



2-ШІ БАСЫЛЫМ
ЖАҢАРТЫЛҒАН ЖӘНЕ
КЕҢЕЙТИНДІ

DATA DRIVEN CONSTRUCTION

ҚҰРЫЛЫС ИНДУСТРИЯСЫНДАҒЫ
ДЕРЕКТЕР ДӘУІРІН ШАРЛАУ

AI ЖӘНЕ LLM ҚОЛДАНУ
ЖАҒДАЙЛАРЫНЫҢ
КӨБІРЕК МЫСАЛДАРЫ

Artem Boiko



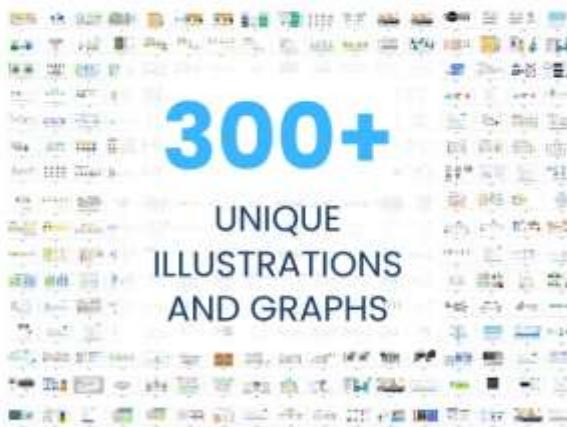
100+

KEY DATA-
RELATED
TOPICS



50+

PRACTICAL
PROBLEMS
SOLVED



300+

UNIQUE
ILLUSTRATIONS
AND GRAPHS

DATA-DRIVEN CONSTRUCTION

Деректер дәуіріндегі навигация
құрылышы саласында

Екінші, түзетілген және толықтырылған басылым

АРТЕМ БОЙКО

“

Бойко *IT* саласындағы Джеймс Карвилл сияқты - соңғысының әйгілі “Экономика, ақымақ” деген сөзін тек бір сөзben алмастыру қажет. “Деректер, ақымақ.” (бағдарламалық қамтамасыз ету емес) Деректер әлемінде жол табу үшін, ежелгі римдіктердің грек дәүіріне жататын бір мақалы бұгінгі күні де өзекті: “Навигация қажет.” Автор оқырмандарын деректер мұхитының тереңдіктері мен таяздықтарында сенімді қолмен және тұрақты компаспен навигациялайды, сонымен қатар кешенді тарихи көзқарасымен, сонымен қатар өте түпнұсқалық графикамен және екінші қарауда ғана байқалатын жақсы әзіл сезімімен. Бойконың кітabyна халықаралық жауаптар қуанышты мақұлдаудан бастап, біршама қышқыл скептицизмге дейінгі аралықта болды, бұл кітаптың екінші неміс басылымына пайдасын тигізді. Бойко - түпнұсқалық және догматикалық емес деректер ойшылы. Ол оқырманға қызықты түсініктер мен әрдайым батыл, тіпті провокациялық тезистерді ұсынады, олар одан әрі ойлануға шабыт береді. Неміс қоғамында латентті консенсуализмнің ауруына тамаша ем. Айтпақшы, жоғарыда аталған латын мақалының толықтыруы бар: “тіршілік ету қажет емес.” Бұл Бойконың деректер әлемінде деген көзқарасына қатысты емес - деректер өмір сүреді және оның өмірі қажет, тіпті шешуші.

— Др. Буркхард Талебитари, фриланс редакторы -
2013 жылдан бері Ernst & Sohn басып шығаратын BIM
журналы үшін.

Артем Бойконың кітабы құрылыш саласындағы цифрандыруды демократияландырудың маңызды кезеңі болып табылады - және шағын және орта кәсіпорындар (ШОК) үшін нағыз ойын өзгеруші. Ерекше жаңашылдық: заманауи ашық кодты, темен кодты және кодсыз құралдарды пайдалану арқылы компаниялар деректерді бизнес процестеріне тиімді интеграциялап, оны пайдамен талдай алады - терең бағдарламалу білімінсіз. Бұл қымбат, ауыр коммерциялық бағдарламалық пакеттерді пайдалануды артық етеді. Бұл кітап - әрекетке шақыру! Ол құрылыш саласындағы цифрлық трансформацияны түсінгісін келетін, сонымен қатар оны белсенді турде қылыштастырысын келетін кез келген адам үшін құнды нұсқаулық - pragmatikaлық, тиімді және болашақта бағытталған. Енді осы білімді бөлісіп, құрылыш саласының өнімділігін тұрақты түрде арттыру үақыты келді.

— Др. Майкл Макс Бюлер, HTWG Констанцтағы Құрылыш Менеджменті профессоры, GemeinWerk Ventures-тің теңисі және DevvStream-да тәуелсіз директор.

DataDrivenConstruction кітабы құрылышылардың, олардың күрделі жобалару және басқару жүйелерінің әдептегі әлемінің шекарасынан асып шығудың алғашқы қадамдарының бірі болып табылады, өйткені, деректердің күрделілігі мен қанықтылыны радикалды оңайлатуға және құрылыш деректермен жұмыс істеуде ашықтықты арттыруға мүмкіндік бермейді. Артем өз кітабында заманауи деректермен жұмыс істеу технологияларының бізге қандай мүмкіндіктер аштының қарапайым тілмен көрсетеді және жұмыссызыда дереву қолдануға болатын накты қадамдарды үсінады. Мен автоматтандыру жүйелерін құрылыш саласында қайда баратынын түсінгісін келетін барлық адамдарды осы кітапты мүккіт зерттеуге шакырамын, себебі құрылыш саласындағы деректер революциясы біздің есігімізді қағып жатыр. Қазір тек технологиктерге қызықты, бірақ бірнеше жылдан кейін, BIM сияқты, мұндай тәсілдер мен бағдарламалық қамтамасыз ету кең таралатын болады!

— Игорь Рогачев, RGD-дағы IMT Компетенция Орталығының жетекшісі, BIM & Цифрлық Трансформация, және InfraBIM.Pro-ның негізін қалаушы.

DataDrivenConstruction кітабын ұсынамын, ол атаянда айтылғандай, AECO үшін деректерге негізделген ақпаратты басқару тәсілін қарастырады. Мен қазір оны әртүрлі топтармен бірнеше талқылауларды бастау үшін пайдаланып жатырмын. Мен оны өте қолжетімді анықтама ретінде талтыйм. AECO-дағы құралдардың тарихи контекстін, деректерді және бірнеше негізгі технологияларды таныстырумен қатар, кітапта деректер көздерінің ауқымын және соңғы пайдалануши артефактілерін улғи жұмыс ағындарымен көрсететін бірнеше өте пайдалы диаграммалар бар. Мениңше, бұл ақпараттық стратегияларды әзірлеу және мониторинг кезінде бізге көбірек қажет диаграммалар.

— Пол Рансли, Acstena-дағы бас консультант және Transport for London-дағы жүйелер интеграциясы инженері.

"Егер "деректер жаңа мұнай" болса, біз оны анықтауды, таңдауды, өндіруді, тазартуды және құнды ету үшін үйренуіміз керек. Мен DataDrivenConstruction кітабын өте ақпараттаныратын және түсінікті деп таптым. Кітап пайдалы тарихи фонды ұсынады және деректермен жұмыс істеуді қаралай тілмен түсіндіреді. Цифрлық трансформациямен қызығатындар үшін деректердің қалай жұмыс істейтінін, қалай құрылымдағанын және қалай пайдаланылатынын жақсы түсінуге мүмкіндік береді."

— Ральф Монтер, ArcDox директоры, BIM Координаторлары Саммитінің директоры және Ирландияның Үлттық Стандарттар Агенттігінде BIM Үлттық Айна Комитетінің тәрағасы.

Кітапта атап өтілгендей, ақпарат құрылыш секторы үшін маңызды актив болып табылады және оны қолжетімді форматтарда ұстау дәл шешім қабылдауды жеңілдетеді және жобаның мерзімдерін жеделдетеді. Кітапта осы деңгээзді шешім қабылдауда пайдалану үшін нейтралды және тиімді тәсіл ұсынылған. Кітапта ұсынылған әдістеме заманауи тәсілді, жасанды интеллектке негізделген бағдарламаларды және қолжетімді ашық кодты құралдарды біріктіреді. Жасанды интеллектті күшін пайдаланып, ашық кодты бағдарламалық қамтамасыз етуді қолдану арқылы, әдістеме автоматтандыруды арттыруға, процестердің оңтайланылуға және салада қолжетімділік пен ынтымақтастықты дамытуға бағытталған. Кітаптың тілі анық және түсінікті.

— Др. Салих Офлуоғлу, Антalia Білім Университетінің Сәулет және Өнер Факультетінің деканы, және Еуразиялық BIM Форумының үйімдастырушысы.

Дарылық айтатыным, WOW! Сіздің тарихты, LLM-ді, графикаларды және жалпы ойларыңызды түсінудің жеңілдігін қалай біріктіргенің шын мәнінде таңгаларлық. Кітаптың ағымы көремет. Бұл кітапта көптеген тамаша аспекттер бар; ол шын мәнінде ойын өзгертуші. Бұл ақпарат көзі ретінде тамаша, және сіздің оған салған күш-жігерізің бен құштарлығыңыз үшін сізді құттықтайдың. Мұндай таңғажайып жұмыс жасағаныңыз үшін құттықтайдың. Мен одан әрі айта алар едім, бірақ жеткілікті, мен өте әсерлендім!

— Наташа Принслу, energylab_ компаниясының Цифрлық Практика Жетекшісі.

Құрылыш саласындағы кез келген адамға, жаңадан бастаушылардан бастап тәжірибелі мамандарға дейін, бұл кітап ойын өзгертуші! Бұл сіздің дәстүрлі, ескірген оқуышың емес — ол түсінкітер, стратегиялар және сізді қызықтырытын юморға толы. Ежелгі деректерді жазу әдістерінен бастап, заманауи цифрлық технологияларға дейін, ол құрылыш саласындағы деректерді пайдаланудың эволюциясын қамтиды. Бұл құрылыш деректерінің эволюциясына үақыт машинасымен саяхат жасау сияқты. Сіз архитектор, инженер, жобаларды басқарушы немесе деректер аналитигі болсаңыз да, бұл кешенде нұсқаулық жобаларының деңгөн көзқарасыныңды революциялайды. Процестердің оңтайланылуға, шешім қабылдауды жақсартуға және жобаларды бұрын-соңды болмағандай басқаруға дайын болыңыз!

— Пьерпаоло Вергати, Римдегі Сapienца университетінің оқытушысы және Fintech компаниясының аға құрылыш жобасының менеджери.

Мен кітапты бір деммен, б сағаттан аз уақыт ішінде оқып шықтым. Кітаптың өндірістік саласы тамаша, тығыз жылтыр қағаз, түс схемалары, жағымды шрифт. Құрылыс саласына тән LLM-мен жұмыс істеу бойынша практикалық мысалдардың көлтігі сізге айтарлықтай уақытты, тіпті жылдардың үнемдейді. Жұмыс мысалдары ете әртүрлі, қаралайымнан күрделігі дейін, күрделі және қынбат бағдарламалық қамтамасыз етуді сатып алуды талап етпейді. Кітап құрылыс саласындағы кез келген бизнес иесіне өз бизнес стратегиясына, цифrlандыруға және даму перспективаларына жаңа көзқарастен қарауға мүмкіндік береді. Ал шағын компанияларға қолжетімді және тегін құралдармен тиімділікті арттыруға көмектеседі.

— Михаил Косарев, TIM-ASG компаниясында құрылыс саласындағы цифрлық трансформация бойынша оқытушы және консультант.

"DATA DRIVEN CONSTRUCTION" кітабы деректер дәүіріндегі құрылыс саласының болашағына қызығушылық танытқан кез келген адам үшін ойын өзгертуші. Артем тек беткі қабатқа ғана нараз аудармайды; ол құрылыс саласындағы қазіргі даму, қындықтар мен перспективалы мүмкіндіктеге терең бойлайды. Бұл кітапты ерекше ететін нәрсе – Артем күрделі идеяларды түсінікті аналогиялар арқылы түсіндіреді, бұл мазмұнды түсініді жеңілдетеді. Мен бұл кітапты өте ақпараттандыратын, бірақ сонымен бірге қызықты деп таптым. Қорытындылай келе, Артем құнды ресурс жасаған, ол тек ақпарат беріл қана қоймай, сонымен қатар шабыттандырады. Сіз тәжірибелі маман болсаңыз да, құрылыска жаңадан келген болсаңыз да, бұл кітап сіздің көзқарасыңызды қеңейтеді және саланың қайда бара жатқанын тереңірек түсінуге көмектеседі. Күшті ұсынылады!

— Моайдад Салех, TMM GROUP Gesamtplanungs GmbH компаниясындағы архитектор және BIM енгізу менеджері.

Мен "Data-Driven Construction" кітабы университеттерде оқытылатын оқулық ретінде лайықты деп айтуды керек және ол BIM саласының дамуына құнды үлес қосады. Data-Driven Construction кітабы техникалық гlosсарийді қамтиды, ол үғымдарды өте жақсы түсіндіреді. Түсіндіру өте күнін тақырыптар өте әдемі визуалды тілмен қаралайым және түсінікті етіледі. Визуализацияларда түсіндіру үшін оқырманға қысқаша болса да, визуализацияларда не түсіндіру керек екені көрсетілуі тиіс деп ойлаймын. Кейбір визуализациялардың түсініктілігі, яғни визуалды оку үшін бөлек ақпарат қажет. Мен Артем Бойконың құнды жұмысын университеттердегі дәрістерім мен семинарларында таныстыруға қуаныштымын.

— Др. Эдиз Язиджоғлу, ArchCube компаниясындағы иесі және Стамбул техникалық университетінің және Медипол университетінің Архитектура кафедрасында құрылыс жобаларын басқару бойынша оқытушы.

"Data Driven Construction" құрылыс деректерімен ақпаратқа негізделген жұмыстың негіздерін айқын жеткізеді. Ақпарат ағындары мен негізгі экономикалық концепцияларды қарастыратын кітап, осылайша, басқа BIM кітаптарынан ерекшеленеді, өйткені ол тек бағдарламалық қамтамасыз етушінің көзқарасын ғана емес, сонымен қатар негізгі концепцияларды жеткізуге тырысады. Оқыған және көргенге тұралрлық кітап.

— Яков Хирн, Build Informed GmbH компаниясындағы бас директоры және "On Top With BIM" инновациялық форумының бастамашысы.

"Деректер – жаңа мұнай" деп айтылғандай, сондықтан оның іздеушілері немесе кешілір 21-ші ғасырдың ресурсынан құн алу үшін дұрыс құралдар мен ойлау тәслюне ие болуы керек. Құрылыс саласы тым ұзақ уақыт бойы "3D ақпарат" негізіндегі процестердің тайғақ жолында болды, мұнда жобаны жеткізу басқалардың дайын ақпараттарна (мысалы, олар диаграмманы немесе бағаны алдын ала сызған) негізделген, ал негізгі "деректер" (мысалы, шикі электрондық кесте) одан да кеп нәрсені жеткізу мүмкін, әсіресе көп деректерді біріктіру мен ИИ шексіз мүмкіндіктерді әкеледі. Егер сіз құрылыс саласында жұмыс істеп жатсаңыз (немесе оқытып/зерттеп жатсаңыз), бұл кітап – біз тап болған деректерге негізделген әлемде навигация жасау үшін ең жақсы және әзірле жалғыз ресурс.

— Др. Зулфикар Адаму, Лондон қаласындағы LSBU университетінде құрылыс саласындағы стратегиялық IT бойынша қауымдастырылған профессор.

"Data-Driven Construction" Артем Бойконың технологиялар мен ақпарат мүмкіндіктері үнемі өсіп келе жатқан кезеңінде құрылыс саласына беретін мықты негізін ұсынатын әсерлі жұмысы. Бойко күрделі тақырыптарды түсінікті түрде ұсынуға және сонымен бірге көреген идеяларды енгізуге қабілетті. Кітап қазіргі даму үрдістерін ғана емес, сонымен қатар болашақ инновацияларға да шолу жақын жақсы ойластырылған жинақ. Деректерге негізделген құрылыс жоспарлау мен орындауды түсінгісі келетін кез келген адам үшін жоғары ұсынылады.

— Маркус Эйбергер, Штутгарт қолданбалы ғылымдар университетінің оқытушысы, аға жобалар менеджері және Konstruktionsgruppe Bauen компаниясындағы орынбасарлық филиал менеджері, BIM кластері Баден-Вюртемберг асоциациясындағы басқарма мүшесі.



Екінші басылым, 2025 жылдың наурызы. © 2025 | Артем Бойко | Карлсруэ

ISBN 978-3-912002-12-6



Артем Бойко Авторлық құқық

boikoartem@gmail.com
info@datadrivenconstruction.io

Бұл кітаптың ешбір бөлігі электронды немесе механикалық құралдармен, фотокөшірмелеге, жазу немесе ақпаратты сақтау және іздеу жүйелері арқылы, автордың жазбаша рұқсатының, қайта шығарылмайды немесе берілмейді – коммерциялық емес тарату жағдайынан басқа. Кітап тегін таратылады және авторлық құқықты және түпнұсқаға сілтемелерді сақтай отырып, басқа пайдаланушыларға жеке, білім беру немесе зерттеу мақсатында еркін берілуі мүмкін. Автор мәтінге қатысты барлық мүліктік емес құқықтарды сақтайды және ешқандай айқын немесе жанама кепілдіктер бермейді. Кітапта аталған компаниялар, өнімдер мен атаулар ойдан шығарылған немесе мысал ретінде қолданылуы мүмкін. Автор берілген ақпаратты пайдаланудың кез келген салдары үшін жауп бермейді. Кітапта қамтылған ақпарат "барынша" ұсынылған, толықтыры немесе өзектілігіне кепілдіктерсіз. Автор берілген ақпарат, код немесе бағдарламаларды пайдаланумен байланысты кездейсоқ немесе жанама шығындар үшін жауп бермейді. Кітапта ұсынылған код мысалдары тек білім беру мақсатында арналған. Оқырмандар оларды өз жауапкершілігімен пайдаланады. Автор өндірістік ортада қолданар алдында барлық бағдарламалық шешімдерді тексеруді ұсынады. Мәтінде аталған барлық сауда маркалары мен өнім атаулары тиісті компаниялардың сауда маркалары, тіркелген сауда маркалары немесе қызмет көрсету белгілері болып табылады және олардың тиісті иелерінің меншігі. Кітапта осы атауларды қолдану олардың иелерімен қандай да бір қатынастарды немесе олардың тарапынан мақұлдауды білдірмейді. Ушінші тарап өнімдері немесе қызметтерін атау ұсыныс емес және олардың қолдауын білдірмейді. Мысалдарда қолданылған компаниялар мен өнімдердің атаулары олардың иелерінің сауда маркалары болуы мүмкін. Сыртқы веб-сайттарға сілтемелер ынғайлыштық үшін берілген және автордың осы сайттардағы ақпаратты мақұлдайтынын білдірмейді. Барлық көлтірлген статистика, дәйексөздер мен зерттеулер кітап жазылған кезде өзекті болды. Деректер уақыт ете келе өзгері мүмкін.

Бұл кітап Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) лицензиясымен таратылады. Сіз оны авторлық құқықты сақтай отырып, өзгеріссіз коммерциялық емес мақсатта көшіріп, тарата аласыз.



CC BY-NC-ND 4.0

© 2024 Артём Бойко. Бірінші басылым. © 2025 Артём Бойко. Екінші басылым, қайта өндөлген және толықтырылған. Барлық құқықтар қорғалған.

ЕКІНШІ БАСЫЛЫМГА АРНАЛҒАН ПРЕДИСЛОВИЕ

Бұл кітап кәсіби қауымдастықпен белсенді диалогтың нәтижесі болып табылады. Оның негізі құрылыс саласындағы деректермен жұмыс істеу мәселелері бойынша әртурлі кәсіби аландар мен әлеуметтік платформаларда өткізілген көптеген кәсіби талқылауларға негізделген. Бұл талқылаулар мақалалар, жарияланымдар және кәсіби қауымдастықта кеңінен жауап алған визуалды материалдар үшін негіз болды. Автордың материалдары жыл сайын әртурлі платформалар мен тілдерде миллиондаған қаралымдарды тартады, құрылыс цифрландыруы саласындағы мамандарды біріктіреді.

Бірінші басылым шыққаннан кейінгі жыл ішінде кітапты 50-ден астам елдің мамандары – Бразилия мен Перудан Маврикий мен Жапонияға дейін – тапсырыс берді. Сіз қолыңызда ұстап отырған кітаптың екінші басылымы сарапшылардың пікірлері, бірінші басылымға қатысты сынекертпелері мен кәсіби ортадағы талқылаулар негізінде қайта өндөліп, толықтырылған. Пікірлердің арқасында екінші басылым едәуір кеңейтілді: CAD (BIM) технологиялары мен тиімді ETL-процестерін құру туралы жаңа тараулар қосылды. Сондай-ақ, практикалық мысалдар мен жағдайлардың саны айтарлықтай артты. Құрылыс саласындағы көшбасшылар, консалтингтік компаниялар және ірі IT-компанияларынан алынған кері байланыс ерекше құнды, олар авторға цифрландыру мен интероперабельділік мәселелері бойынша бірінші басылым шыққанға дейін және кейін жүргінді. Олардың көпшілігі кітапта сипатталған тәсілдерді қолдануда немесе жақын арада қолдануды жоспарлауда.

Сіз қолыңызда пікірталастар мен белсенді пікір алмасу нәтижесінде жасалған кітапты ұстап отырысыз. Прогресс диалогта, көзқарастардың қақтығысында және жаңа тәсілдерге ашықтықта туындейді. Осы диалогтың бір бөлігі болғаның үшін раҳмет. Сіздің конструктивті сынныңыз – болашақтағы жақсартулардың негізі. Егер мәтінде қателер табылса немесе идеялар мен ұсыныстармен бөлісуге ниетіңіз болса, кез келген кері байланыс құпталады. Байланыс деректері кітаптың соңында көрсетілген.

НЕЛІКТЕҢ КІТАП ТЕГІН?

Бұл кітап құрылыс саласындағы деректерді басқарудың заманауи тәсілдерін таратуға бағытталған ашық білім беру ресурсы ретінде ойластырылған. Кітаптың бірінші нұсқасы кәсіби қауымдастықтан пікірлер мен ұсыныстар жинау үшін негіз болды, бұл материалдың құрылымын және мазмұнын жақсартуға мүмкіндік берді. Барлық ескертпелер, ұсыныстар мен идеялар мүқият талданып, осы өндөлген нұсқада ескерілді. Кітаптың мақсаты – құрылыс саласының мамандарына деректермен жүйелі, саналы және ұзақ мерзімді ақпараттың құндылығын ескере отырып жұмыс істеудің маңыздылығын түсінуге көмектесу. Автор цифрландыру саласындағы 10 жылдан астам жұмыс тәжірибесінде жинақталған мысалдар, иллюстрациялар мен практикалық байқауларды жинады. Бұл материалдардың көпшілігі нақты жобалар, инженерлер мен әзірлеушілермен пікірталастар, халықаралық бастамаларға қатысу және оқу семинарларын өткізу барысында пайда болды. Кітап – жинақталған тәжірибелі құрылымдаға және оны қолжетімді түрде бөлісуге тырысу. Егер сіз кітап идеяларын әрі қарай таратуға қолдау көрсеткініз келсе және мысалдар мен визуалды материалдармен жұмыс істеуге ыңғайлы формат алғызыз келсе, баспа нұсқасын сатып ала аласыз.

МАТЕРИАЛДАРДЫ ПАЙДАЛАНУ ҚҰҚЫҚТАРЫ

Бұл кітаптың барлық материалдары, иллюстрациялары мен фрагменттері міндетті түрде авторлық құқықты: Артем Бойко авторлығын және «Data-Driven Construction» кітабының атаянын көрсете отырып, кез келген форматта және кез келген тасымалдаышта reproduced, cited немесе пайдаланылуы мүмкін. Еңбекті құрметтеп, білімді тарату үшін раҳмет.

Шын жүректен алғысымды білдіре отырып, бұл кітапты отбасымен арнаймын, олар маған жастайымнан құрылышқа терең сүйіспеншілік дарытты, туған шахтерлік қалама – өмірлік төзімділік сабактары үшін және менің маркшейдер әйеліме, оның тұрақты қолдауы менің үнемі шабыт көзі болды.

БҮЛ КІТАП КІМГЕ АРНАЛҒАН

Қолжетімді тілмен жазылған бұл кітап құрылымы саласындағы кең ауқымды оқырмандарға арналған - заманауи құрылым процестерінің негіздерін мен өзара кеңінен қолданылады. Бұл кітап архитектор, инженер, прораб, құрылым менеджері немесе деректер аналитигі болсаңыз да, бұл кешенді нұсқаулық көптеген бірегей иллюстрациялар мен графиктермен бірге бизнес процестерін оңтайландыру және автоматтандыру, шешім қабылдауды жақсарту және құрылымын өртүрлі деңгейде басқару үшін деректерді қалай пайдалануға болатындығы туралы құнды ақпарат ұсынады.

Кітап деректерді басқару әдістерін құрылым процестеріне интеграциялау бойынша теориялық негіздер мен практикалық ұсыныстарды біріктіретін кешенді нұсқаулық болып табылады. Кітап ақпаратты стратегиялық пайдалану арқылы операциялық қызметті оңтайландыру, процестерді автоматтандыру, шешім қабылдауды жетілдіру және заманауи цифровалық құралдарды қолдану арқылы жобаларды тиімді басқаруға назар аударады.

Бұл кітаптың миссиясы - құрылымы саласындағы ақпаратпен жұмыс істеудің теориялық және практикалық аспекттерін көздейді. Мәселе-лерді параметрлеу, талаптарды жинау, құрылымдаған және өртүрлі форматтағы деректерді өндереу және оларды құрылым компаниялары үшін тиімді шешімдерге айналдыру әдістемесі егжей-тегжейлі мысалдар арқылы зерттеледі.

Оқырман талаптарды қалыптастыру мен деректердің базалық модельдерін әзірлеуден бастап, өртүрлі ақпарат көздерін интеграциялау, ETL-процестерін құру, ақпараттық Pipeline және машиналық оқыту модельдерін құруға дейінгі жолды кезең-кезеңімен өтеді. Кезеңдік тәсіл бизнес процестерін үйімдастыру мен автоматтандыру механизмдерін және құрылымы саласындағы шешім қабылдау жүйелерін айқын көрсетуге мүмкіндік береді. Кітаптың әрбір бөлімі практикалық тараумен аяқталады, онда алынған білімдерді нақты жобаларда деруе қолдануға мүмкіндік беретін қадамдық нұсқаулар беріледі.

КІТАП БӨЛІМДЕРІНІҢ ҚЫСҚАША ШОЛУЫ

Бұл кітап деректердің құндылық тізбегінде (value chain) трансформациялау концепциясы айналасында құрылды: олардың жинау мен сапасын қамтамасыз етуден аналитикалық өңдеу мен заманауи құралдар мен әдістемелердің қолдану арқылы құнды практикалық шешімдерді алуына дейін.

1-бөлім: Құрылдың саласындағы цифрлық эволюция - деректердің басқарудың тарихи трансформациясын балшықтан жасалған тақталардан заманауи цифрлық жүйелерге дейін қадағалайды, модульдік жүйелердің пайда болуын және ақпараттың цифрандыруының маңыздылығының артуын өнеркәсіптік революциялар контекстінде талдайды.

2-бөлім: Құрылдың индустриясындағы ақпараттық қындықтар - деректердің фрагментациясы, "акпараттық силостар", HiPPO тәсілінің шешім қабылдауға әсері және меншікті форматтардың шектеулері мәселелерін зерттейді, AI және LLM әкімшілдегі әкпараттық технологиялар көшүмдігін қарастыруды ұсынады.

3-бөлім: Құрылдың саласындағы деректердің жүйеленеу - құрылдың деректерінің типологиясын қалыптастырады, олардың үйімдастырылу әдістерін, корпоративтік жүйелермен интеграциясын сипаттайтын және ақпараттық процестерді стандарттау үшін құзырет орталықтарын құруды талқылады.

4-бөлім: Деректердің сапасын қамтамасыз ету - шашыраңқы ақпаратты сапалы, құрылымдаған деректерге айналдыру әдістемелерін ашады, әртүрлі көздерден деректерді алу, валидациялау және LLM қолдану арқылы модельдеуді қамтиды.

5-бөлім: Баға мен уақытты есептеу - баға есептеулерін және жоспарлауды цифрандыруға, CAD- (BIM-) модельдерден көлемдерді автоматты түрде алуға, 4D-8D модельдеу технологияларына және құрылдың жобаларының ESG көрсеткіштерін есептеуге арналған.

6-бөлім: CAD және BIM - жобалау технологияларының эволюциясын, жүйелердің үйлесімділік мәселелерін, деректердің ашық форматтарына көшу тенденцияларын және жобалауда жасанды интеллектінің қолдану перспективаларын сыйни түрфыдан талдайды.

7-бөлім: Деректер аналитикасы және автоматтандыру - ақпаратты визуализациялау принциптерін, негізгі тиімділік көрсеткіштерін, ETL процестерін, жұмыс процестерін оркестрациялау құралдарын және рутиндік тапсырмаларды автоматтандыру үшін тілдік модельдердің қолдануды қарастырады.

8-бөлім: Деректердің сақтау және басқару - деректердің сақтау форматтарын, қоймалардың және деректер көлдерінің концепцияларын, деректердің басқару принциптерін және векторлық деректер базалары мен DataOps және VectorOps әдістемелерін қоса алғанда, жаңа тәсілдерді зерттейді.

9-бөлім: Үлкен деректер және машиналық оқыту - тарихи деректер негізінде объективті талдауға, құрылдың аландаудағы заттар интернетіне және жобалардың құны мен мерзімдерін болжай үшін машиналық оқыту алгоритмдерін қолдануға арналған.

10-бөлім: Цифрлық деректер дәуіріндегі құрылдың саласы - құрылдың саласының болашағына көзқарас ұсынады, себеп-салдарлы талдаудан корреляциялармен жұмыс істеуге көшу, құрылдың "уберизациясы" концепциясы және цифрлық трансформация стратегияларын талдайды.

What is meant by **data-driven construction** ?



KІRІСПЕ

Сіздің компанияңыз технологиялардың жылдам дамып жатқан әлемінде, мерзімдер мен бағаларды есептеуден бастап, тәуекелдерді талдауға дейін, машиналық оқыту модельдері арқылы автоматтандырылатын бизнесіңіздің әр аспектісінде басекеге қабілеттілігін қанша уақыт сақтай алады?

Адамзат тарихымен бірге өмір сүріп келе жатқан құрылыш саласы революциялық өзгерістердің алдында тұр, олар дәстүрлі құрылыш туралы түсінігімізді түбекейлі өзгертуге үеде береді. Басқа экономикалық секторларда цифрландыру қазіргі уақытта тек әдеттегі ережелерді өзгертіп қана қоймай, жаңа деректерді өңдеу шарттарына бейімделе алмаған және шешім қабылдау жылдамдығын арттыра алмаған компанияларды нарықтан аяусыз ығыстырып жатыр (Сурет 1).



Сурет 1 Құрылыш саласындағы шешім қабылдау жылдамдығы басқа салаларға қарағанда адам факторының әсеріне көбірек тәуелді.

Банктік сектор, бөлшек сауда, логистика және агроенеркәсіптік кешен толық цифрландыруға жылдам бет бұрып жатыр, мұнда дәлсіздіктер мен субъективті пікірлерге орын жоқ. Заманауи алгоритмдер орасан зор деректер массивтерін талдай алады және клиенттерге дәл болжамдар ұсынады - бұл не кредитті қайтару ықтималдығы, не жеткізудің оңтайлы маршруттары, не тәуекелдерді болжау.

Құрылыш саласы жоғары ақы төленетін мамандардың пікірлеріне негізделген шешімдерден деректерге негізделген шешімдерге өтудің қажеттілігі туындаған соңғы салалардың бірі болып табылады. Бұл өтудің себебі тек жаңа технологиялық мүмкіндіктермен ғана емес, сонымен қатар нарық пен тапсырыс берушілердің ашықтық, дәлдік және жылдамдыққа деген талаптарының артуымен де байланысты.

Роботтандыру, процестерді автоматтандыру, ашық деректер және олардың негізінде болжамдар –

Бұлардың барлығы тек мүмкіндіктер емес, міндеттілікке айналды. Құрылыш саласындағы көптеген компаниялар, жақында клиенттер алдында көлемдер, құн, жобалардың мерзімі мен сапа бақылауы үшін жауап беріп келген, енді негізгі шешімдер қабылдамайтын, тек бүйректарды орындаушыларға айналуы мүмкін (Рис. 2).

Есептеу қуаттарының, машиналық оқыту алгоритмдерінің дамуы және деректерге қол жеткізудің демократиялануы әртүрлі көздерден деректерді автоматтый түрде біркітіруге мүмкіндік берді, бұл процестерді тереңірек талдауға, тәуекелдерді болжауға және құрылыш жобасын талқылау кезеңінде шығындарды оңтайландыруға мүмкіндік береді. Бұл технологиялар бүкіл секторда тиімділікті радикалды түрде арттыру және шығындарды азайту әлеуетін қалыптастырады.

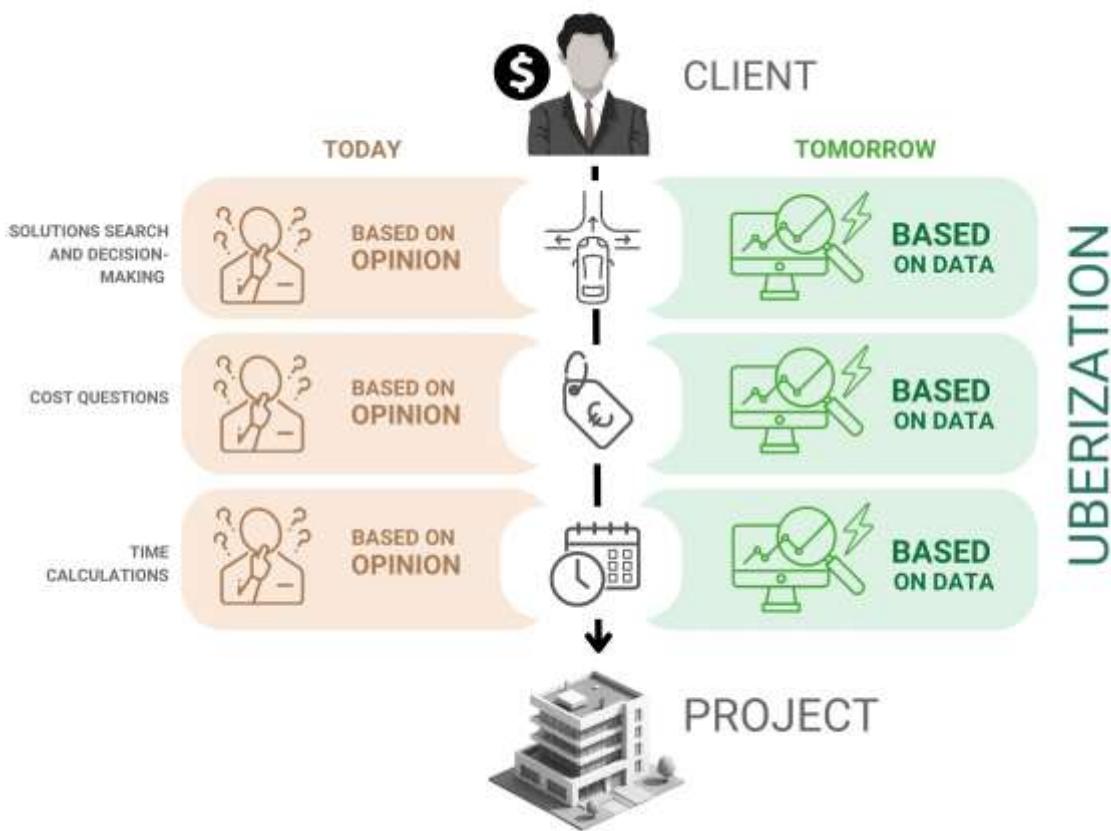


Рис. 2 Клиент өз жобасын жүзеге асыруда артық адам факторының болуына қызықпайды.

Жаңа құралдар мен концепциялардың барлық артықшылықтарына қарамастан, құрылыш саласы жаңа технологияларды енгізу бойынша басқа экономикалық секторлардан едәуір артта қалып отыр.

„IT Metrics Key Data 2017“ есебіне сәйкес, құрылыш саласы 19 басқа экономикалық сектордың арасында IT шығындары бойынша соңғы орында тұр [1].

Деректер көлемінің жылдам өсуі мен процестердің күрделілігі компаниялардың басшылығы үшін бас ауруына айналуда, ал жаңа технологияларды қолданудағы басты проблема – деректердің, олардың молдығына қарамастан, бөлшектелген, құрылымдалмаған және көбінесе әртүрлі жүйелер мен бағдарламалық өнімдер арасында үйлесімсіз болып қалуында. Сондықтан бүгінгі күні құрылыш секторының көптеген компаниялары деректер сапасы мәселелерімен алаңдаулы,

оларды тек тиімді, автоматтандырылған басқару және аналитика жүйелерін енгізу арқылы шешуге болады.

2023 жылы KPMG® жүргізген құрылыш компанияларының менеджерлері арасында жүргізілген сауалнамаға сәйкес, жобалардағы инвестициялардың рентабельдігін артыру үшін ақпараттық басқару жүйелері (PMIS), алдыңғы және базалық деректер аналитикасы және ғимараттардың ақпараттық моделін жасау (BIM) ең үлкен әлеуетке ие (Рис. 3).

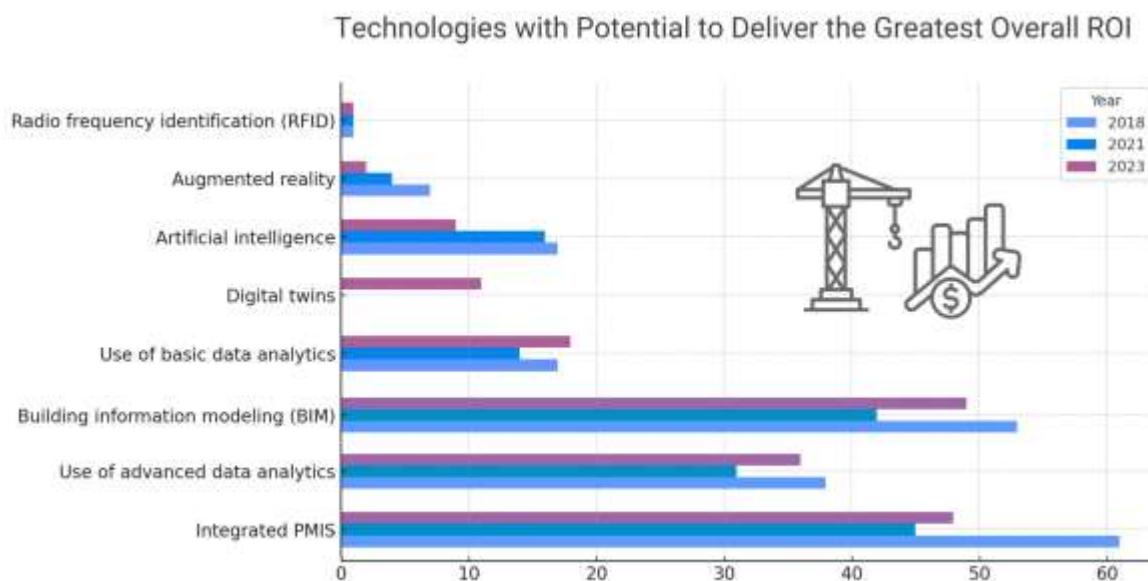


Рис. 3 Құрылыш компанияларының басшылары арасында: қандай технологиялар капиталды жобалардағы инвестициялардың ең үлкен қайтарымын қамтамасыз етеді? (материалдар бойынша [2]).

Деректерді бизнес-процестерге интеграциялау мәселелерін шешу жоғары сапалы ақпаратты қамтамасыз ету, деректердің тиісті форматтарын пайдалану және деректерді жасау, сақтау, талдау және өндіре үшін тиімді әдістерді қолдануда жатыр.

Деректердің құндылығын түсіну әртүрлі салаларды бөлшектелген қосымшалар мен күрделі бюрократиялық басқару құрылымдарынан бас тартуға мәжбүрлейді. Оның орнына ақпараттық архитектураға жаңа тәсілдер жасауға назар аударылады, бұл компанияларды деректермен басқарылатын заманауи кәсіпорындарға айналдырады. Ерте ме, кеш пе, бұл қадам құрылыш саласында да жүзеге асырылып, кезең-кезеңімен цифрлық эволюциядан нағыз цифрлық революцияға өтеді, бұл барлық компанияларды қамтиды.

Деректерге негізделген бизнес-процестерге көшу оңай болмайды. Көптеген компаниялар қыындықтарға тап болады, себебі басшылар хаостық деректер массивтерін тиімділік пен бизнестің өсуін арттыру үшін қалай пайдалану керектігін әрдайым түсінбейді.

Бұл кітапта біз деректер әлеміне терең үнілеміз, мұнда ақпарат стратегиялық ресурс ретінде тиімділік пен бизнес-процестердің тұрақтылығын анықтайды. Ақпарат көлемінің жедел өсуі

жағдайында компаниялар жаңа қыындықтарға тап болады. Цифрлық трансформация тек сәнді терминге айналмай, қажеттілікке айналады.



Рис. 4 Деректер мен процестер құрылыш негізі болып табылады.

Трансформацияны түсіну – бұл күрделі нәрсені қарапайым сөздермен түсіндіре білу. Сондықтан кітап қолжетімді тілмен жазылған және негізгі ұғымдарды визуализациялау үшін арнайы жасалған авторлық иллюстрациялармен сүйемелденеді. Бұл схемалар, графіктер мен визуализациялар қабылдау кедергілерін жоюға және материалды бұрынғыдай күрделі деп санайтындарға да түсінікті етуге арналған. Кітаптағы барлық иллюстрациялар, схемалар мен графіктер автор тара-пынан жасалған және мәтінде сипатталған негізгі концепцияларды визуализациялау үшін арнайы әзірленген.

Бір сурет мың сөзге тең.

– Фред Р. Барнард, ағылшын иллюстраторы, 1927

Теорияны практикамен байланыстыру үшін біз жасанды интеллект құралдарын (әсіресе, тілдік модельдерді) қолданамыз, олар терең бағдарламалау білімінсіз шешімдер әзірлеуге мүмкіндік береді. Егер сіз практикалық материалға бағытталған болсаңыз және деректермен практикалық жұмыс сізді көбірек қызықтырса, бірінші кіріспе бөлімді өткізіп, кітаптың екінші бөліміне, нақты мысалдар мен кейстердің сипаттамасы басталатын жерге өтуінізге болады.

Дегенмен, ИИ (Жасанды Интеллект), машиналық оқыту және LLM (Үлкен Тіл Модельдері) құралдана артық күтім артпау керек. Сапалы бастапқы деректер мен пәндік сала туралы терең түсінік болмаса, тіпті ең озық алгоритмдер де сенімді және мағыналы нәтижелерді қамтамасыз ете алмайды.

Microsoft-тың бас директоры Сатья Наделла 2025 жылдың басында жасанды интеллект саласында пузырь пайда болу қаупі туралы ескертеді, қазіргі ажиотажды "дотком пузырімен" салыстыра отырып. Ол AGI (Жасанды Жалпы Интеллект) кезеңдеріне жету туралы мәлімдемелердің негізсіз екенін "мағынасыз көрсеткіштермен манипуляциялау" деп атап өтті. Наделла ИИ-дің нақты жетістігі әлемдік ЖІӨ-нің өсуіне қосқан үлесімен өлшенуі керек, ал даңқты мәлімдемелерге артық назар аударылмауы тиіс деп санайды.

Барлық жаңа технологиялар мен концепцияларға қатысты дауысты сөздердің артында деректердің сапасын қамтамасыз ету, бизнес-процестерді параметрлеу және құралдарды нақты міндеттерге бейімдеу бойынша құрделі және мұқият жұмыс жатыр.

Деректерге негізделген тәсіл – бұл тек жүктеп алуға немесе сатып алуға болатын өнім емес. Бұл стратегия, оны құру қажет. Ол бар процестер мен мәселелерге жаңа көзқарастан басталады, содан кейін таңдалған бағытта тәртіпті түрде қозғалысты талап етеді.

Бағдарламалық қамтамасыз етуді әзірлеушілер мен қосымша жеткізушилері құрылыш индустрисындағы өзгерістердің локомотиві болмайды, өйткені олардың көпшілігі деректерге негізделген тәсіл бизнес-моделінің қалыптасқан құрылымына қауіп төндіреді.

Басқа салалар [құрылыш саласынан айырмашылығы], мысалы, автомобиль өнеркәсібі, радикалды және жойқын өзгерістерге ұшырады, және олардың цифрлық трансформациясы толық қарқынмен жүріп жатыр. Құрылыш компанияларына тез және шешуші әрекет ету қажет: икемді компаниялар үлкен пайдаға ие болады, ал тербелмелі компаниялар үшін қауіптер айтарлықтай болады. Цифрлық фотографияның осы индустрияда тузызған дурбелеңін еске алыңыз [5].Дүниежүзілік экономикалық форумның "Құрылыш болашағын қалыптастыру" есебі, 2016

Уақытында жаңа тәсілдердің мүмкіндіктері мен артықшылықтарын түсінетін компаниялар тұрақты бәсекелестік артықшылыққа ие болып, ірі вендорлардың шешімдеріне тәуелді болмай да-мып, өсетін болады.

Бұл сіздің ақпараттық цифрандырудың алдағы дауылын тек күтіп қана қоймай, сонымен қатар оны бақылауға алу мүмкіндігіңіз. Кітапта сіз саладағы қазіргі жағдайды талдаумен қатар, процестеріңіз беріңізде қайта ойластыру және қайта құрылымдау бойынша нақты ұсыныстар табасыз, бұл сізді құрылыш саласындағы жаңа дәуірдің көшбасшысы етуге және кәсіби тәжірибеліздейтін арттыруға мүмкіндік береді.

Цифрлық құрылыш болашағы – бұл тек жаңа технологиялар мен бағдарламаларды пайдалану емес, деректермен жұмыс істеу мен бизнес-модельдерді түбәгейлі қайта қарастыру.

Сіздің компанияңыз осы стратегиялардың өзгерістерге дайын ба?

БАСҚАРУ ТІЗІМІ

KІРІСПЕ.....	1
БАСҚАРУ ТІЗІМІ.....	I
I БӨЛІМ ГЛИНА ТАБЛИЦАЛАРЫНАН ЦИФРЛЫҚ РЕВОЛЮЦИЯГА: ҚҰРЫЛЫС САЛАСЫНДАҒЫ АҚПАРАТТЫҢ ЭВОЛЮЦИЯСЫ.....	2
ТАРАУ 1.1. ДЕРЕКТЕРДІ ПАЙДАЛАНУ ЭВОЛЮЦИЯСЫ ҚҰРЫЛЫС САЛАСЫНДА.....	3
Деректер дәүірінің құрылышы саласындағы тууы	3
Балшықтан және папирустан цифрлыш технологияларға дейін	4
Процесс деректермен басқарылатын тәжірибе құралы ретінде.	5
Құрылыш процесінің ақпаратын цифрландыру.....	7
ТАРАУ 1.2. ЗАМАНАУИ ҚҰРЫЛЫС САЛАСЫНДАҒЫ ТЕХНОЛОГИЯЛАР МЕН БАСҚАРУ ЖҮЙЕЛЕРИ.....	11
Цифрлыш революция және модульдік MRP/ERP жүйелерінің пайда болуы	11
Деректерді басқару жүйелері: деректерді жинаудан бизнес-мәселелеріне дейін.....	13
Корпоративті мицелий: деректер бизнес-процестерге қалай бірігеді.....	17
ТАРАУ 1.3. ЦИФРЛЫҚ РЕВОЛЮЦИЯ ЖӘНЕ ДЕРЕКТЕРДІҢ ЖАРЫЛЫСЫ	20
Деректер көлемінің өсуінің басталуы: эволюциялық толқын.....	20
Қазіргі компанияда өндірілетін деректер көлемі.....	22
Деректерді сақтау құны: экономикалық аспект	23
Деректерді жинақтаудың шекаралары: массадан мағынаға	25
Келесі қадамдар: деректер теориясынан практикалық өзгерістерге	27
II БӨЛІМ ҚҰРЫЛЫС БИЗНЕСІ ДЕРЕКТЕР ХАОСЫНДА ҚАЛУ	28
ТАРАУ 2.1. ДЕРЕКТЕРДІҢ ФРАГМЕНТАЦИЯСЫ ЖӘНЕ СИЛОСЫ.....	29
Құралдардың саны көп болған сайын, бизнес тиімдірек пе?	29
Деректер силостары және олардың компания тиімділігіне әсері	31
Дубликация және деректер сапасының болмауы: бөлінудің салдары	35
HiPPO немесе шешім қабылдаудағы пікірлердің қауіптілігі	36
Бизнес-процестердің күрделілігі мен динамикасының тұрақты артуы	39
Төртінші өнеркәсіптік революция (Индустрія 4.0) және бесінші өнеркәсіптік революция (Индустрія 5.0) құрылышы саласында	42
ТАРАУ 2.2. ХАОСЫНЫҢ ТӘРТІПКЕ АЙНАЛУЫ ЖӘНЕ КҮРДЕЛІЛІКТІ ТӨМЕНДЕТУ	45
Артық код және жабық жүйелер: өнімділікті арттырудың кедергі	45

Сиолостардан деректердің біртұтас қоймасына.....	47
Интеграцияланған сақтау жүйелері AI агенттерін пайдалануға мүмкіндік береді	48
Деректерді жинаудан шешім қабылдауға: автоматтандыруға жол.....	51
Келесі қадамдар: хаосты басқарылатын жүйеге айналдыру	53
III БӨЛІМ ҚҰРЫЛЫС БИЗНЕС-ПРОЦЕСТЕРІНДЕГІ ДЕРЕКТЕР ҚҰРЫЛЫСЫ.....	55
ТАРАУ 3.1. ҚҰРЫЛЫС САЛАСЫНДАҒЫ ДЕРЕКТЕР ТҮРЛЕРІ.....	56
Құрылыш саласындағы ең маңызды деректер түрлері	56
Структураланған деректер.....	60
Реляциялық деректер базалары RDBMS және SQL сұраныс тілі.....	62
Деректер базаларындағы SQL сұраныстары және жаңа трендтер	64
Структураланбаған деректер	66
Мәтіндік деректер: структура мен структураланбаған хаос арасында	67
Полуструктураланған және әлсіз структураланған деректер	68
Геометриялық деректер және олардың қолданылуы	69
CAD деректері: жобалаудан деректерді сақтауға.....	72
BIM (ВОМ) концепциясының пайда болуы және CAD-ды процестерде пайдалану.....	75
ТАРАУ 3.2. ДЕРЕКТЕРДІ БІРІКТІРУ ЖӘНЕ СТРУКТУРАЛАУ	81
Құрылыш саласындағы жүйелерді деректермен толтыру.....	81
Деректерді трансформациялау: заманауи бизнес-талдаудың критикалық негізі	84
Деректер модельдері: деректердегі қатынастар мен элементтер арасындағы байланыстар .	87
Проприетарлық форматтар және олардың цифрлық процестерге әсері	92
Ашық форматтар цифрландыруға көзқарасты өзгертеуді	95
Параидигманың өзгеруі: Ашық Код бағдарламалық қамтамасыз етуші вендорлардың үстемдік дәүірінің аяқталуы.....	97
Структурленген ашық деректер: цифрлық трансформацияның негізі	99
ТАРАУ 3.3. LLM ЖӘНЕ ОЛАРДЫҢ ДЕРЕКТЕРДІ ӨНДЕУ ЖӘНЕ БИЗНЕС-ПРОЦЕСТЕРДЕГІ РӨЛІ.	102
LLM чаттары: ChatGPT, LlaMa, Mistral, Claude, DeepSeek, QWEN, Grok деректерді өндеу процестерін автоматтандыру үшін	102
Үлкен тілдік модельдер LLM: бұл қалай жұмыс істейді	103
Компанияның сезімтал деректері үшін жергілікті LLM пайдалану	106
Компанияда ИИ-ға толық бақылау және өз LLM-ді іске қосу жолдары	108
RAG: Корпоративтік деректерге қол жеткізетін интеллектуалды LLM-ассистенттер	110
ТАРАУ 3.4. LLM ҚОЛДАУЫ БАР IDE ЖӘНЕ БАҒДАРЛАМАЛАУДАҒЫ БОЛАШАҚ ӨЗГЕРИСТЕР	113
IDE таңдау: LLM эксперименттерінен бизнес шешімдеріне	113

LLM қолдауымен IDE және бағдарламалаудағы болашақ өзгерістер	115
Python Pandas: деректермен жұмыс істеудің алмастырылмайтын құралы	116
DataFrame: кестелік деректердің әмбебап форматы	120
Келесі қадамдар: тұрақты деректер каркасын құру	123
IV БӨЛІМ ДЕРЕКТЕР САПАСЫ: ҰЙЫМДАСТЫРУ, СТРУКТУРЛАУ, МОДЕЛДЕУ	125
ТАРАУ 4.1. ДЕРЕКТЕРДІ СТРУКТУРЛАНГАН ФОРМАҒА АЙНАЛДЫРУ	126
Құжаттарды, PDF, суреттер мен мәтіндерді құрылымдық форматтарға айналдыруды үйрену	126
PDF құжатын кестеге айналдыру мысалы	127
JPEG, PNG суреттерін құрылымдық формаға айналдыру	131
Мәтіндік деректерді құрылымдық формаға айналдыру	134
CAD (BIM) деректерін құрылымдық формаға аудару	137
CAD шешімдерінің вендорлары құрылымдық деректерге көшуде	142
ТАРАУ 4.2. СЫНЫПТАУ ЖӘНЕ ИНТЕГРАЦИЯ: ҚҰРЫЛЫС ДЕРЕКТЕРІНІҢ БІРДЕЙ ТІЛ	145
Шешім қабылдау жылдамдығы деректер сапасына байланысты	145
Деректерді стандарттау және интеграциялау	146
Цифрлық үйлесімділік талаптардан басталады	149
Құрылыш үшін бірдей тіл: цифрлық трансформациядағы классификаторлардың рөлі	151
Masterformat, OmniClass, Uniclass және CoClass: классификациялық жүйелердің эволюциясы	154
ТАРАУ 4.3. ДЕРЕКТЕРДІ МОДЕЛДЕУ ЖӘНЕ ОЗЫҚ ТӘЖІРИБЕ ОРТАЛЫҒЫ	159
Деректерді моделдеу: концептуалдық, логикалық және физикалық модель	159
Құрылыш контекстінде практикалық деректер моделдеуі	163
LLM көмегімен деректер базасын құру	166
Деректер моделдеуі бойынша Озық Тәжірибе Орталығы (CoE)	168
ТАРАУ 4.4. ТАЛАПТАРДЫ ЖҮЙЕЛЕУ ЖӘНЕ АҚПАРATTЫ ВАЛИДАЦИЯЛАУ	172
Талаптарды жинау және талдау: коммуникацияларды құрылымдық деректерге айналдыру	172
Процесстердің блок-схемалары және концептуалдық схемалардың тиімділігі	176
Құрылымдық талаптар және RegEx тұрақты өрнектері	178
Тексеру процесі үшін деректерді жинау	183
Деректерді тексеру және тексеру нәтижелері	185
Тексеру нәтижелерін визуализациялау	190
Деректер сапасын тексеруді адамның өмірлік қажеттіліктерімен салыстыру	192

Келесі қадамдар: деректерді дәл есептеулер мен жоспарларға айналдыру	194
V БӨЛІМ ШЫҒЫН ЖӘНЕ УАҚЫТ ЕСЕПТЕУЛЕРІ: ДЕРЕКТЕРДІ ҚҰРЫЛЫС ПРОЦЕСТЕРИНЕ ЕҢГІЗУ	196
ТАРАУ 5.1. ҚҰРЫЛЫС ЖОБАЛАРЫНЫң ШЫҒЫНДАРЫ ЖӘНЕ СМЕТАЛАРЫН ЕСЕПТЕУ	197
Құрлыс негіздері: көлем, шығын және уақытты бағалау	197
Жобалардың сметалық құнын есептеу әдістері	199
Құрлыс саласындағы сметалар мен есептеулерді ресурс әдісімен жасау	199
Құрлыс ресурстарының деректер базасы: құрлыс материалдары мен жұмыстар каталогы	200
Ресурс базасына негізделген жұмыстардың құнын есептеу және сметаларды жасау	202
Жобаның жалпы құнын есептеу: сметалардан бюджетке дейін.....	206
ТАРАУ 5.2. КӨЛЕМДІ АЛУ ЖӘНЕ СМЕТАЛАР МЕН КҮНТІЗБЕ ЖОСПАРЛАРЫН АВТОМАТТАНДЫРУ	211
3D-ден 4D және 5D-ге өту: көлемдік және сандық параметрлерді пайдалану	211
5D атрибуттары және CAD-тан атрибуттардың көлемдерін алу	211
QTO Көлемді алу: жобалық деректерді атрибуттар бойынша топтастыру	215
QTO автоматтандыруы LLM және құрлыымдық деректерді пайдалану арқылы	220
Excel кестесіндегі топтар үшін ережелерді пайдалана отырып, жобаның QTO есебі	224
ТАРАУ 5.3. 4D, 6D-8D ЖӘНЕ CO ₂ ГАЗЫНЫң ШЫҒАРЫЛУЫН ЕСЕПТЕУ	230
4D-модель: құрлыс сметаларына уақытты интеграциялау.....	230
Құрлыс графигі және оны калькуляция деректері негізінде автоматтандыру	231
6D-8D кеңейтілген атрибуттық қабаттары: энергия тиімділігінен қауіпсіздікті қамтамасыз етуге дейін.....	233
CO ₂ бағалау және құрлыс жобаларындағы көмірқышқыл газының шығарындыларын есептеу	236
ТАРАУ 5.4. ҚҰРЫЛЫС ERP ЖӘНЕ PMIS ЖҮЙЕЛЕРІ.....	240
Құрлыс ERP жүйелері есептеулер мен сметалар мысалында	240
PMIS: ERP мен құрлыс алаңы арасындағы аралық буын	245
Алыпсатарлық, пайда, жақындық және болмауыERP және PMIS-тегі мәлдірлік	246
Жабық ERP/PMIS дәүірінің соңы: құрлыс саласы жаңа тәсілдерге мұқтаж	249
Келесі қадамдар: жобалық деректерді тиімді пайдалану	251
VI БӨЛІМ CAD ЖӘНЕ BIM: МАРКЕТИНГ, ШЫНДЫҚ ЖӘНЕ ҚҰРЫЛЫС САЛАСЫНДАҒЫ ЖОБАЛАУ ДЕРЕКТЕРІНІҢ БОЛАШАҒЫ	254
ТАРАУ 6.1. ҚҰРЫЛЫС САЛАСЫНДАҒЫ BIM-КОНЦЕПТЕРДІҢ ПАЙДА БОЛУЫ	255
BIM және open BIM-нің CAD жеткізушилері тарапынан маркетингтік концепт ретінде пайда	

булуның тарихы	255
ВІМ шындығы: интеграцияланған деректер базаларының орнына жабық модульдік жүйелер	258
Құрылым саласында IFC ашық форматының пайда болуы	260
IFC форматының геометриялық ядроға тәуелділігі мәселесі.....	262
Құрылым саласында семантика және онтология тақырыбының пайда болуы	265
Неліктен семантикалық технологиялар құрылым саласында күтілген нәтижелерді ақтамайды	267
ТАРАУ 6.2. ЖАБЫҚ ЖОБА ФОРМАТТАРЫ ЖӘНЕ ИНТЕРОПЕРАБЕЛІК МӘСЕЛЕЛЕР	271
Жабық деректер және өнімділіктің төмендеуі: CAD (BIM) саласының дағдарысы	271
CAD жүйелері арасындағы интероперабельділік туралы миф	273
USD және гранулирленген деректерге көшу	277
ТАРАУ 6.3. ҚҰРЫЛЫС САЛАСЫНДАҒЫ ГЕОМЕТРИЯ: СЫЗБАЛАРДАН КУБОМЕТРЛЕРГЕ	281
Қашан сызықтар ақшаға айналады немесе құрылышшыларға геометрия не үшін қажет	281
Сызықтардан көлемдерге: қалайша алаң мен көлем деректерге айналады	282
MESH, USD және полигондарға көшу: геометрия үшін тесселяцияны пайдалану	284
LOD, LOI, LOMD – CAD (BIM) жүйесіндегі детализацияның ерекше классификациясы	285
CAD (BIM) жаңа стандарттары - AIA, BEP, IDS, LOD, COBie	288
ТАРАУ 6.4. ЖОБАЛАУДЫҢ ПАРАМЕТРЛЕНДІРЛУІ ЖӘНЕ CAD МЕН ЖҰМЫС ИСТЕУ ҮШІН LLM ПАЙДАЛАНУ	293
CAD (BIM) деректерінің бірегейлігі иллюзиясы: аналитика мен ашық форматтарға жол	293
Параметрлер арқылы жобалау: CAD және BIM-нің болашағы.....	296
CAD жобалық деректерін өңдеу процестерінде LLM-нің пайда болуы	299
LLM және Pandas көмегімен DWG файлдарын автоматтандырылған талдау	302
Келесі қадамдар: жабық форматтардан ашық деректерге көшу	308
VII БӨЛІМ ДЕРЕКТЕРГЕ НЕГІЗДЕЛГЕН ШЕШІМ ҚАБЫЛДАУ, АНАЛИТИКА, АВТОМАТТАНДЫРУ ЖӘНЕ МАШИНАЛЫҚ ОҚУ	310
ТАРАУ 7.1. ДЕРЕКТЕР АНАЛИТИКАСЫ ЖӘНЕ ДЕРЕКТЕРГЕ НЕГІЗДЕЛГЕН ШЕШІМ ҚАБЫЛДАУ	311
Шешім қабылдауда деректер ресурсы	311
Деректерді визуализациялау: түсіну мен шешім қабылдаудың кілті.....	314
Нәтижелердің тиімділігінің көрсеткіштері KPI және ROI	316
Ақпараттық панельдер мен дашбордтар: тиімді басқару үшін көрсеткіштерді визуализациялау	318
Деректерді талдау және сұрақ қою өнері.....	320

ТАРАУ 7.2. ДЕРЕК АГЫНЫ ҚОЛДАН БОЛМАЙ: ETL НИГЕ ҚАЖЕТ	323
ETL АВТОМАТТАНДЫРУ: ШЫҒЫНДАРДЫ ТӨМЕНДЕТУ ЖӘНЕ ДЕРЕКТЕРМЕН ЖҰМЫС ИСТЕУДІ ЖЫЛДАМДАТУ	323
ETL Extract: деректерді жинау	327
ETL Transform: тексеру ережелерін қолдану және түрлендіру	329
ETL Load: НӘТИЖЕЛЕРДІ ГРАФИКТЕР МЕН ДИАГРАММАЛАР ТҮРІНДЕ КӨРСЕТУ	332
ETL Load: PDF ҚҰЖАТТАРЫН АВТОМАТТЫ ТҮРДЕ ЖАСАУ	337
ETL Load: FPDF ПЕН ҚҰЖАТТАРДЫ АВТОМАТТЫ ГЕНЕРАЦИЯЛАУ	338
ETL Load: ЕСЕПТЕРДІ ҚҰРАСТАРЫУ ЖӘНЕ БАСҚА ЖҮЙЕЛЕРГЕ ЖҮКТЕУ	342
LLM АРҚЫЛЫ ETL: PDF ҚҰЖАТТАРЫНАН ДЕРЕКТЕРДІ КӨРСЕТУ	343
ТАРАУ 7.3. АВТОМАТТЫ ETL КОНВЕЙЕРІ (PIPELINE)	348
Pipeline: Автоматты ETL деректер конвейері	348
Pipeline-ETL ДЕРЕКТЕРДІ ТЕКСЕРУ ПРОЦЕСІ LLM АРҚЫЛЫ	352
Pipeline-ETL: CAD (BIM) ЖОБА ЭЛЕМЕНТТЕРІНІҢ ДЕРЕКТЕРІН ЖӘНЕ АҚПАРАТЫН ТЕКСЕРУ ..	354
ТАРАУ 7.4. ETL ЖӘНЕ ЖҰМЫС ПРОЦЕСТЕРИН ОРКЕСТРАЦИЯЛАУ: ПРАКТИКАЛЫҚ ШЕШІМДЕР	
360	
DAG ЖӘНЕ Apache Airflow: ЖҰМЫС ПРОЦЕСТЕРИН АВТОМАТТАНДЫРУ ЖӘНЕ ОРКЕСТРАЦИЯЛАУ	360
Apache Airflow: ETL АВТОМАТТАНДЫРУ БОЙЫНША ПРАКТИКАЛЫҚ ҚОЛДАНУ	361
Apache NiFi ДЕРЕКТЕРДІ ЖОЛДАНДЫРУ ЖӘНЕ ТҮРЛЕНДІРУ ҮШІН	366
n8n ТӨМЕН КОД, НОЛЬ КОД ПРОЦЕСТЕРДІ ОРКЕСТРАЦИЯЛАУ	367
КЕЛЕСІ ҚАДАМДАР: ҚОЛДАН ОПЕРАЦИЯЛАРДАН АНАЛИТИКАЛЫҚ ШЕШІМДЕРГЕ ӨТУ	370
VIII БӨЛІМ ДЕРЕКТЕРДІ САҚТАУ ЖӘНЕ БАСҚАРУ САЛАСЫНДА	372
ТАРАУ 8.1. ДЕРЕКТЕР ИНФРАСТРУКТУРАСЫ: САҚТАУ ФОРМАТТАРЫНАН ЦИФРЛЫҚ ҚОЙМАЛАРҒА ДЕЙІН	373
ДЕРЕК АТОМДАРЫ: АҚПАРATTЫ ТИІМДІ БАСҚАРУДЫҢ НЕГІЗІ	373
АҚПАРАТ САҚТАУ: ФАЙЛДАР НЕМЕСЕ ДЕРЕКТЕР	374
ҮЛКЕН ДЕРЕКТЕРДІ САҚТАУ: ТАҢЫМАЛ ФОРМАТТАРДЫ ТАЛДАУ ЖӘНЕ ОЛАРДЫҢ ТИІМДІЛІГІ	376
Apache Parquet ПЕН ДЕРЕКТЕРДІ САҚТАУДЫ ОПТИМАЛДАНДЫРУ	379
DWH: Деректер қоймасының (Data Warehouse) САҚТАУ ОРНЫ	381
Data Lake - ETL-ден ELT-ге ЭВОЛЮЦИЯ: ДӘСТҮРЛІ ТАЗАЛАУДАН ИНДИВИДУАЛДЫ ӨНДІРІСКЕ	383
Data Lakehouse АРХИТЕКТУРАСЫ: САҚТАУ ОРЫНДАРЫ МЕН ДЕРЕК ОЗЕРІНІҢ СИНЕРГИЯСЫ	385

CDE, PMIS, ERP НЕМЕСЕ DWH ЖӘНЕ Data Lake.....	387
ТАРАУ 8.2. ДЕРЕКТЕРДІ САҚТАУ ЖӘНЕ ХАОС ТУДЫРУДЫ БОЛДЫРМАУ	391
ВЕКТОРЛЫҚ ДЕРЕК БАЗАЛАРЫ ЖӘНЕ БОКС ШЕГІ	391
ДЕРЕКТЕРДІ БАСҚАРУ (Data Governance), ДЕРЕКТЕРДІҢ МИНИМАЛИЗМІ (Data Minimalism) ЖӘНЕ ДЕРЕКТЕРДІҢ БАТПАҒЫ (Data Swamp)	394
DataOps ЖӘНЕ VectorOps: ДЕРЕКТЕРМЕН ЖҮМЫС ІСТЕУДІҢ ЖАҢА СТАНДАРТТАРЫ	397
КЕЛЕСІ ҚАДАМДАР: ХАОТТЫ САҚТАУДАН ҚҰРЫЛЫМДЫ САҚТАУ ОРНЫНА ӨТУ	399
IX БӨЛІМ УЛКЕН ДЕРЕКТЕР, МАШИНАЛЫҚ ОҚЫ ЖӘНЕ БОЛЖАУ	401
ТАРАУ 9.1. УЛКЕН ДЕРЕКТЕР ЖӘНЕ ОЛАРДЫҢ ТАЛДАNUY.....	402
Улken деректер құрылыс саласында: интуициядан болжамдылық	402
Улken деректердің тиімділігі туралы мәселе: корреляция, статистика және деректерді іріктеу.	403
Улken деректер: Сан-Францискодағы миллион құрылыш рұқсатын талдау.....	406
CAD (BIM) деректеріне негізделген үлken деректердің мысалы	412
IoT (Заттар Интернеті) және смарт-келісімшарттар	416
ТАРАУ 9.2. МАШИНАЛЫҚ ОҚЫТУ ЖӘНЕ БОЛЖАМДАР	420
Машиналық оқыту мен жасанды интеллект біздің құрылышымызды өзгертеді	420
Субъективті бағалаудан статистикалық болжамға	422
Титаник деректер жиыны: Деректер аналитикасы мен үлken деректер әлеміндегі Сәлем Әлем.	424
Машиналық оқыту iс-әрекетте: "Титаник" жолаушыларынан жобаларды басқаруға.....	429
Тарихи деректер негізінде болжамдар мен предсказаниялар	433
Машиналық оқытудың негізгі ұғымдары	435
ТАРАУ 9.3. МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ ҚҰРЫЛЫС ҚҰНЫ ЖӘНЕ МЕРЗІМДЕРДІ БОЛЖАУ	438
Машиналық оқытуды жобаның құнын және мерзімдерін анықтау үшін пайдалану мысалы.	438
Жобаның құны мен уақытын болжау үшін сзықтық регрессияны пайдалану	440
Жоба құны мен уақытының болжамдары K-ең жақын көршілер алгоритмы (k-NN) арқылы	443
Келесі қадамдар: сақтау, талдау және болжау	447
X БӨЛІМ ЦИФРЛЫҚ ДЕРЕКТЕР ДӘУІРІНДЕГІ ҚҰРЫЛЫС САЛАСЫ. МУМКІНДІКТЕР МЕН СЫН-ТЕГЕУРІНДЕР	450
ТАРАУ 10.1. ТҮРМЫС СТРАТЕГИЯЛАРЫ: БӘСЕКЕЛЕСТИК АРТЫҚШЫЛЫҚТАРДЫ ҚАЛЫПТАСТАРЫ	451
Корреляциялар есептеулердің орнына: құрылыш аналитикасының болашағы	451
Деректерге негізделген тәсіл құрылыш саласында: жаңа деңгейдегі инфрақұрылым	454

Келесі ұрпақтың цифрлық оғисі: AI жұмыс кеңістігін қалай өзгертеді.....	456
Ашық деректер мен убериизация – бұл қазіргі құрылыш бизнесіне қауіп тәндіреді.	458
Нерешілген мәселелерді убериизациялау соңғы мүмкіндік ретінде уақытты трансформациялау үшін пайдалану.	461
ТАРАУ 10.2. ДЕРЕККЕ НЕГІЗДЕЛГЕН ЖАҚЫНДЫ ЖУЗЕГЕ АСЫРУ БОЙЫНША ПРАКТИКАЛЫҚ НҰСҚАУЛЫҚ	466
Теориядан практикаға: құрылыш саласындағы цифрлық трансформацияның жол картасы.	466
Цифрлық негізді қалау: 1-5 қадам цифрлық жетілуге.....	468
Деректердің әлеуетін ашу: Цифрлық жетілуге 5-10 қадам.	473
Жол картасы трансформация: хаостан деректерге негізделген компанияға.	479
Индустрія 5.0-да құрылыш: қалай табыс табуға болады, егер жасыру мүмкін болмаса.....	483
ҚОРЫТЫНДЫ	485
АВТОР ТУРАЛЫ	488
КЕРІ БАЙЛАНЫС	490
АУДАРМАҒА ПІКІР	490
БАСҚА ДаҒДЫЛАР ЖӘНЕ КОНЦЕПЦИЯЛАР	491
ГЛОССАРИЙ	495
ӘДЕБИЕТТЕР ТІЗІМІ ЖӘНЕ ONLINE-МАТЕРИАЛДАР	502
ТАҚЫРЫПТЫҚ ІЗДЕУ	518

БАСПА НҰСҚАСЫМЕН МАКСИМАЛДЫ ҚОЛАЙЛЫЛЫҚ

Сіздер Data-Driven Construction-ның тегін цифрлық нұсқасын ұстап отырсыздар. Материалдарға жылдам қол жеткізу және ыңғайлы жұмыс үшін баспа басылымына назар аударуды ұсынамыз:



- Эрдайым қол астында: баспа форматындағы кітап сенімді жұмыс құралы болып табылады, қажетті визуализациялар мен схемаларды кез келген жұмыс жағдайында жылдам табуға және пайдалануға мүмкіндік береді.
- Суреттердің жоғары сапасы: баспа басылымындағы барлық бейнелер мен графиктер максималды сапада ұсынылған.
- Ақпаратқа жылдам қол жеткізу: ыңғайлы навигация, ескертпелер жасау, бетбелгілер қою және кітаппен кез келген жерде жұмыс істеу мүмкіндігі.



Кітаптың толық баспа нұсқасын сатып алу арқылы сіз ақпаратпен ыңғайлы және тиімді жұмыс істеуге арналған құрал аласыз: күнделікті тапсырмаларда визуалды материалдарды жедел пайдалану, қажетті схемаларды жылдам табу және ескертпелер жасау мүмкіндігі. Сонымен қатар, сатып алуыңыз ашық білімнің таралуын қолдайды.

Кітаптың баспа нұсқасын мына жерден тапсырыс беруге болады:
datadrivenconstruction.io/books



I БӨЛІМ

ГЛИНА ТАБЛИЦАЛАРЫНАН ЦИФРЛЫҚ РЕВОЛЮЦИЯҒА: ҚҰРЫЛЫС САЛАСЫНДАҒЫ АҚПАРАТТЫҢ ЭВОЛЮЦИЯСЫ

Кітаптың бірінші бөлімінде құрылыш саласындағы деректерді басқарудың тарихи эволюциясы қарастырылады – физикалық тасымалдаушылардағы примитивті жазбалардан қазіргі заманғы цифрлық әкімшілікке дейін. Ақпаратты басқару технологияларының трансформациясы, ERP жүйелерінің пайда болуы және деректердің фрагментациясының бизнес-процестердің тиімділігіне әсері талданады. Ақпаратты цифрландыру процесіне және субъективті сараптамалық бағалаулардың орнына объективті талдаудың өсіп келе жатқан маңыздылығына ерекше назар аударылады. Қазіргі құрылыш саласы кездесетін ақпарат көлемінің экспоненциалды өсуі және корпоративтік жүйелер үшін туындастырылған қыындықтар егжей-тегжейлі қарастырылады. Құрылыш индустриясының төртінші және бесінші өнеркәсіптік революциялар контекстіндегі позициясы, сондай-ақ жасанды интеллект пен деректерге негізделген тәсілдерді пайдалану әлеуеті тұрақты бәсекелестік артықшылықтарды қалыптастыру үшін зерттеледі.

ТАРАУ 1.1.

ДЕРЕКТЕРДІ ПАЙДАЛАНУ ЭВОЛЮЦИЯСЫ ҚҰРЫЛЫС САЛАСЫНДА

Деректер дәуірінің құрылыш саласындағы тууы

Шамамен 10 000 жыл бұрын, неолит дәуірінде, адамзат өзінің дамуында революциялық өтпелі кезеңді бастаң кешірді, көшпелі өмір салтынан отырықшылыққа көшіп, алғашқы примитивті құрылыш материалдары – саз, ағаш және тас – пайда болды [6]. Осы сәттен бастап құрылыш индустриясының тарихы басталады.

Цивилизациялар дамыған сайын, архитектура барған сайын күрделене түсті, бұл алғашқы ритуалдық храмдар мен қоғамдық ғимараттардың пайда болуына әкелді. Архитектуралық жобалардың күрделенуі ежелгі инженерлер мен басқарушылардан алғашқы жазбалар мен есептеулер жасауды талап етті. Алғашқы жазбалар саздан жасалған тақталар мен папирустарда жиі қажетті құрылыш материалдарының санын, олардың құнын және орындалған жұмыстың ақысының есебін сипаттайтын логиканы қамтыды [7]. Осылайша, құрылыш саласында деректерді пайдаланудың дәүірі басталды – қазіргі заманғы цифрлық технологиялардың пайда болуынан әлдеқайда бұрын (Рис. 1.11).

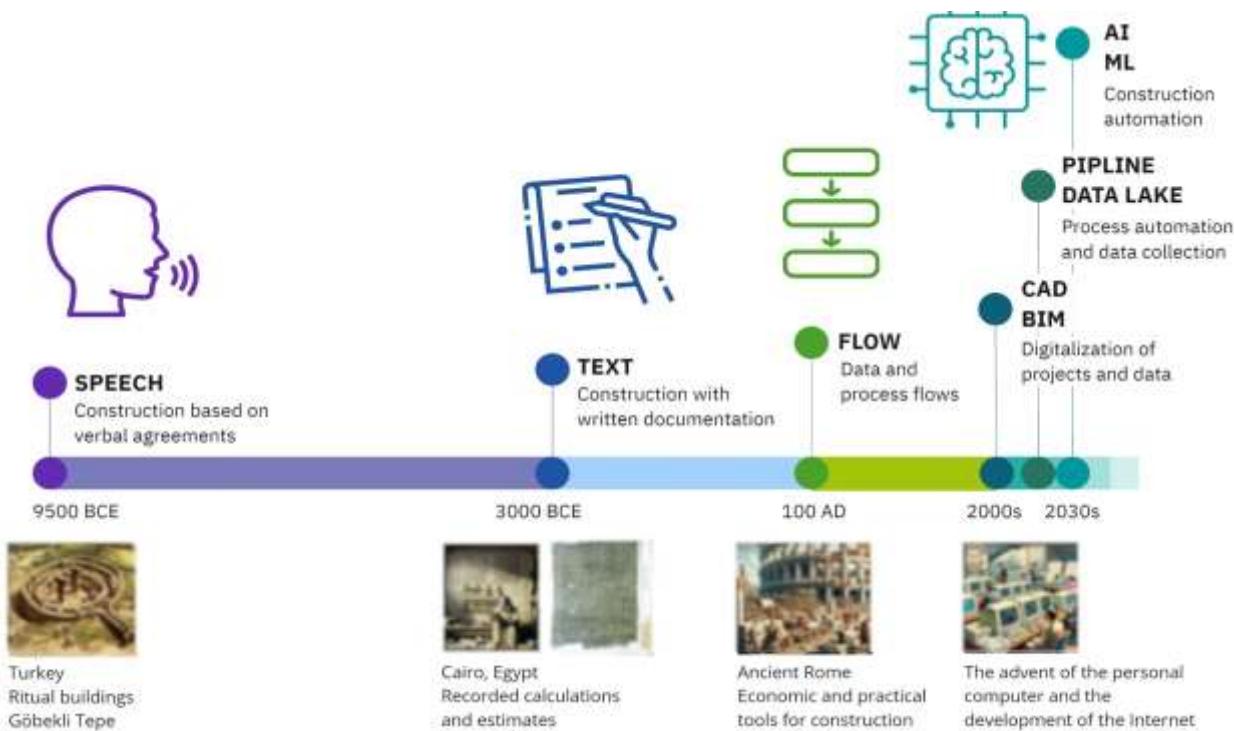


Рис. 1.11 Құрылыш саласындағы ақпараттық технологиялардың даму хронологиясы: ауызша ақпараттан жасанды интеллектке дейін.

Балшықтан және папиустан цифрлық технологияларға дейін

Құрылыш саласындағы алғашқы құжаттық дәлелдер пирамидаларды салу кезеңіне, шамамен 3000–4000 жылдар б.з.д. жатады. Содан бері жазбаша жазбаларды жүргізу құрылыш саласындағы прогресті жеңілдетіп, сүйемелден отырды, бұл 10 000 жыл бойы құрылыш әдістері мен архитектурада елеулі инновацияларды жинақтап, жүйелеуге мүмкіндік берді.

Құрылыш саласындағы алғашқы физикалық тасымалдаушылар, мысалы, керамикалық тақталар, мыңжылдық папирус (Рис. 1.12) немесе 1980-жылдардағы “A0” форматындағы қағаз, деректерді жазу үшін бастапқыда жаңа жобаларда қолдану мақсатында болмады. Мұндай жазбалардың негізгі мақсаты жобаның ағымдағы жағдайын, қажетті материалдар мен жұмыстардың құнын есептеуді егжей-тегжейлі сипаттау болды. Дәл сол сияқты, қазіргі әлемде цифрлық жобалық деректер мен модельдердің болуы болашақ жобаларда олардың қолданылуын әрдайым қамтамасыз етпейді және көбінесе құрылыш материалдарының қажеттіліктері мен құнын есептеуге арналған ақпарат ретінде қызымет етеді.

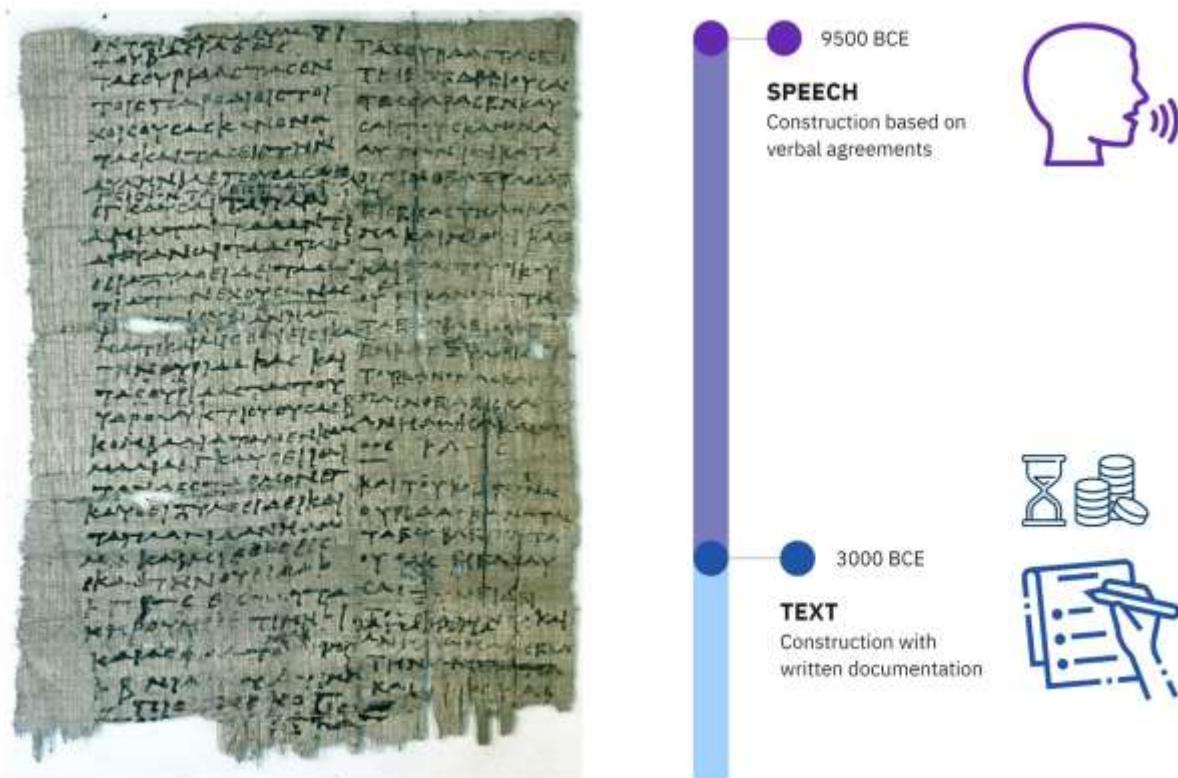


Рис. 1.12 Б.з.д. III ғасырдағы папирус, корольдік сарайдағы әртүрлі төрлөрінің сурет салу құнын әнкаустика техникасымен сипаттайды.

Адамзатқа құрылыш жобаларын басқаруда ауызша әңгімелерден жазбаша құжаттарға өту үшін шамамен 5 000 жыл қажет болды, сондай-ақ қағаз тасымалдаушылардан цифрлық деректерге негізгі жоспарлау және бақылау ресурсы ретінде өту үшін де сонша уақыт қажет болды.

Сауда және ақша қатынастарының дамуы жазудың пайда болуына және даулы мәселелерді шешетін алғашқы заңгерлердің пайда болуына ықпал еткен сияқты, құрылыш саласындағы материалдардың құны мен жұмыс көлемі туралы алғашқы жазбалар құрылыш саласындағы алғашқы мәнеджерлердің пайда болуына әкелді, олардың міндеттеріне жобаның мерзімдері мен құны туралы негізгі ақпаратты құжаттау, мониторинг жүргізу және жауапкершілік жүктеу кірді.

Бұғынгі күні деректер әлдеқайда маңызды рөл атқарады: олар тек қабылданған шешімдерді тіркепп қана қоймай, сонымен қатар болашақты болжай мен модельдеудің құралына айналады. Осы негізде жобаларды басқарудағы заманауи процестік тәсіл қалыптасады – жинақталған тәжірибелі құрылымдалған және тексерілген деректерге негізделген шешім қабылдау жүйесіне айналдыру.

Процесс деректермен басқарылатын тәжірибе құралы ретінде.

Әрбір процестің негізінде өткен тәжірибелі болашақты жоспарлау құралына айналдыру жатыр. Заманауи түсінікте тәжірибе – бұл құрылымдалған деректер жиынтығы, оның талдауы негізделген болжамдар жасауға мүмкіндік береді.

Тарихи деректер болжаудың негізі болып табылады, себебі олар орындалған жұмыстардың нәтижелерін айқын көрсетіп, осы нәтижелерге әсер ететін факторлар туралы түсінік береді.

Конкретті мысалды монолитті құрылыш саласынан қарастырайық: әдетте, жұмыстардың мерзімдерін жоспарлағанда бетон көлемі, конструкцияның күрделілігі және ауа райы жағдайлары ес-керіледі. Мысалы, белгілі бір прорабтың құрылыш алаңында немесе компанияның тарихи деректері соңғы үш жылда (2023–2025) жаңбырлы ауа райында 200 м² алаңдағы монолитті конструкцияны құюға 4,5-тен 6 күнге дейін уақыт қажет болғанын көрсетеді (Рис. 1.13). Міне, осындай жинақталған статистика болашақ жобаларда үқсас жұмыстарды жоспарлауда мерзімдерді болжай және ресурстарды калькуляциялау үшін негіз болады. Осы тарихи деректер негізінде прораб немесе сметчик, 2026 жылы үқсас жағдайларда болашақта осындай жұмыстарды орындауға қажетті уақыт туралы негізделген болжам жасай алады.

Бұл жағдайда уақытты бағалау – аналитикалық процесс бөлшектелген деректерді құрылымдалған тәжірибеге, содан кейін – дәл жоспарлау құралдарына айналдыру механизмі ретінде қызмет етеді. Деректер мен процестер – бұл біртұтас экожүйе, мұнда біреуі екіншісіз өмір сүре алмайды.

Санаққа келетін нәрсelerdі санаңыз, өлшеуге болатын нәрсelerdі өлшеңіз, ал өлшеуге келмейтін нәрсelerdі өлшеуге жарамды етіңіз. Галилео Галилей

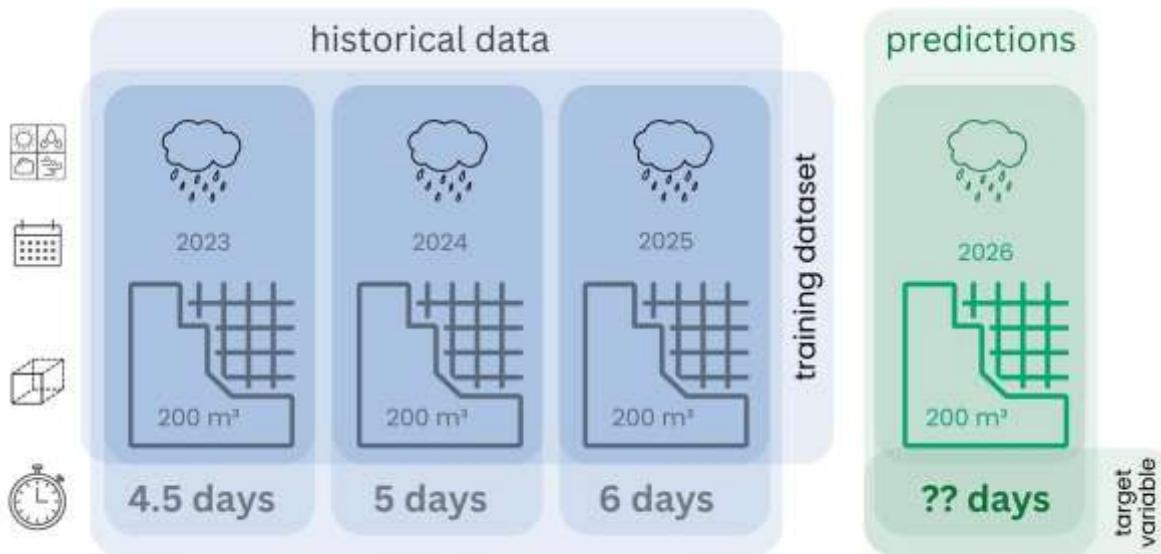


Рис. 1.13 Тарихи деректер болашақта бір көрсеткішті болжау үшін оқу деректерінің жиынтығы ретінде қызмет етеді.

Заманауи бизнес-ландшафта деректерді талдау жобаларды тиімді басқару, процестерді оңтайландыру және стратегиялық шешім қабылдау үшін критикалық маңызды компонентке айналуда. Құрылыш индустриясы төрт негізгі аналитика деңгейін біртіндеп игеруде, әрқайсысы нақты сұраққа жауап беріп, ерекше артықшылықтар ұсынады (Сурет 1.14):-

- Сипаттамалық аналитика – «не болды?» деген сұраққа жауап береді және тарихи деректер мен өткен оқынадар мен нәтижелер туралы есептерді ұсынады: соңғы үш жылда (2023–2025) жаңбырлы аяа-райында 200 м² көлеміндегі монолитті құрылымды құю үшін 4,5-тен 6 күнге дейін үақыт қажет болды.
- Диагностикалық аналитика «неліктен бұл болды?» деген сұраққа жауап береді, проблемалардың туындау себептерін анықтайды: талдау монолитті құрылымның құйылу үақытының жаңбырлы аяа-райы салдарынан бетонның қату процесінің баяулағанын көрсетеді.
- Болжамдық аналитика - болашаққа бағытталған, ықтимал қауіптер мен жұмыстардың орындалу мерзімдерін болжайды, «не болады?» деген сұраққа жауап береді: тарихи деректер негізінде 2026 жылы жаңбырлы аяа-райында 200 м² алаңдағы ұқсас монолитті құрылымды құю үшін барлық белгілі факторлар мен тенденцияларды ескере отырып, шамамен 5,5 күн қажет болады.
- Прескриптивті аналитика автоматтандырылған ұсыныстар береді және «не істеу керек?» деген сұраққа жауап береді, компанияларға оңтайлы әрекеттерді таңдауға мүмкіндік береді: Мысалы, жұмыстарды оңтайландыру үшін мынадай ұсыныстар жасалады: ылғалдылығы жоғары жағдайда бетонның қатуын жеделдететін арнайы қоспаларды пайдалану; жауын-шашынның ықтималдығы ең төмен кезеңдерде құюды жоспарлау; құрылымды үақытша жапқыштармен жабдықтау, бұл қолайсыз аяа райы жағдайында жұмыстардың үақытын 4-4,5 күнге дейін қысқартуға мүмкіндік береді.

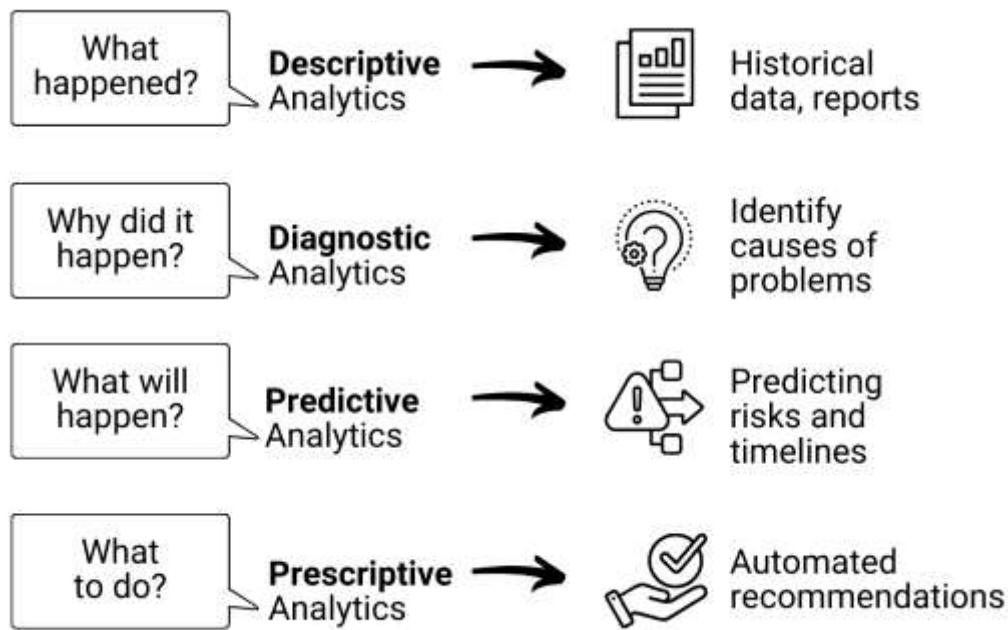


Рис. 1.14 Негізгі аналитика түрлері: өткенді сипаттаудан автоматтандырылған шешім қабылдау процесіне дейін.

Толыққанды цифрлық трансформация, жүйелі аналитика мен деректер негізінде басқаруға көшу, тек сыртқы мердігерлерді тартуды емес, ішкі білікті команданы қалыптастыруды талап етеді. Мұндай команданың негізгі қатысушыларына өнім менеджерлері, деректер инженерлері, аналитиктер және әзірлеушілер кіруі тиіс, олар бизнес бөлімдерімен тығыз ынтымақтастықта жұмыс істейді (Рис. 4.39). Бұл ынтымақтастық аналитикалық сұрақтарды дұрыс қою және шешім қабылдау бизнес міндеттерін тиімді параметрлеу үшін қажет. Ақпараттық қоғам жағдайында деректер тек көмекші құрал емес, болжau мен оңтайландырудың негізі болып табылады.-

Құрылыш саласында цифрлық трансформация жобалау, басқару және объектілерді пайдалану тәсілдерін түбекейлі өзгертерді. Бұл процесс ақпаратты цифрландыру деп аталады – құрылыш процесінің барлық аспекттерін талдауға қолайлы цифрлық формада ауыстыру.

Құрылыш процесінің ақпаратын цифрландыру.

Мыңжылдықтар бойы құрылыш саласында жазылған ақпарат көлемі шамамен өзгермеген, бірақ соңғы онжылдықтарда ол қарқынды өсті (Рис. 1.15).

PwC® зерттеуіне сәйкес «Деректермен басқару. Студенттерге тез өзгеретін бизнес әлемінде табысқа жету үшін не қажет» (2015) [9], әлемдегі барлық деректердің 90%-ы соңғы екі жылда жасалған (2015 жылғы деректер бойынша). Алайда, көптеген компаниялар бұл деректерді толық көлемде пайдаланбайды, себебі олар не бөлшектелген жүйелерде қалады, не шын мәнінде талданбай архивтеледі.

Соңғы жылдары деректер көлемінің артуы тек жеделдеді, 2015 жылы 15 зеттабайттан 2025 жылы

181 зеттабайтқа дейін екі есеге өсті [10]. Күн сайын құрылыш және жобалау компанияларының сервверлері жобалық құжаттамамен, жұмыс кестелерімен, есептеулермен және калькуляциялармен, қаржылық есептермен толтырылады. 2D/3D сызбалары үшін DWG, DXF және DGN форматтары, ал 3D модельдері үшін RVT, NWC, PLN және IFC™ форматтары қолданылады. Мәтіндік құжаттар, кестелер және презентациялар DOC, XLSX және PPT форматтарында сақталады. Құрылыш алаңынан видео және суреттер MPG және JPEG форматтарында, ал IoT компоненттерінен, RFID® белгілерінен (идентификация және бақылау) және BMS (мониторинг және бақылау) басқару жүйелерінен деректер нақты уақыт режимінде түседі.

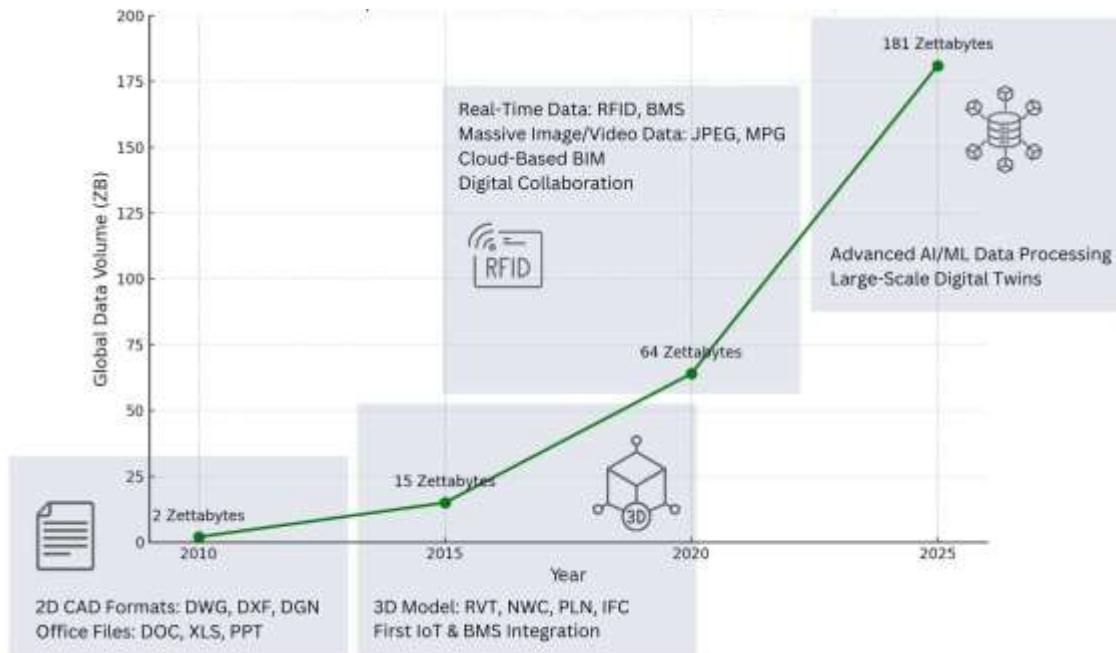


Рис. 1.15 Деректердің параболалық өсуі 2010–2025 (материалдар [10] бойынша).

Ақпарат көлемінің қарқынды өсуі жағдайында құрылыш саласы деректерді жинау мен сақтау қажеттілігімен ғана емес, сонымен қатар олардың тексерілуі, верификациялануы, өлшенуі және аналитикалық өндөлеуімен де бетпе-бет келеді. Бүгінгі күні индустрия ақпаратты цифрландырудың белсенді кезеңін бастап кешіруде – құрылыш қызыметінің барлық аспектілерін талдауға, интерпретациялауға және автоматтандыруға қолайлы цифрлық формага жүйелі түрде түрлендіру.

Ақпаратты цифрландыру құрылыш жобасының және құрылыш процесінің барлық элементтері мен объектілері туралы ақпарат алу дегенді білдіреді - бұрын ақпарат деп есептелмегендерді де қоса алғанда - және оны деректер форматына түрлендіру, осылайша ақпаратты сандық түрде өлшеуге және талдауға ыңғайлы ету.

Құрылыш контекстінде бұл жобалардың барлық элементтері мен процестерінің ақпаратын цифрлық түрде фиксациялау және көрсету дегенді білдіреді – құрылыш алаңындағы техника мен адамдардың қозғалысынан бастап, құрылыш алаңындағы ая-райы мен климаттық жағдайларға, материалдардың ағымдағы бағаларына және орталық банктің пайыздық мөлшерлемесіне дейін – аналитикалық модельдерді қалыптастыру мақсатында.

Егер сіз айтып отырған нәрсені өлшей алсаңыз және оны сандармен білдіре алсаңыз, онда сіз бұл тақырып туралы белгілі бір білімге ие боласыз. Бірақ егер сіз оны сандық түрде білдіре алмасаңыз, онда сіздің біліміңіз шектеулі және қанағаттанарлықсыз. Бұл, мүмкін, бастапқы кезең, бірақ бұл нағыз ғылыми білім деңгейі емес. У. Томсон (Лорд Кельвин), 1824–1907, британдық ғалым

Цифрландыру ақпаратты жинаудың дәстүрлі тәсілдерінен әлдеқайда асып түседі, мұнда тек базалық көрсеткіштер – адам-сағаттар немесе материалдардың нақты шығындары сияқты – тіркеletтін. Бүгінгі күні кез келген оқиға терең талдау үшін қолайлы деректер ағынына айналдырылуы мүмкін, бұл үшін озық аналитика құралдары мен машиналық оқыту әдістері қолданылады. Құрылыш индустрисында қағаз сызбалардан, Excel сметаларынан және ауызша нұсқаулардан – цифрлық жүйелерге (Сурет 1.24) көшу бойынша принципті өзгеріс орын алды, мұнда обьектінің әрбір элементі деректер көзіне айналады. Тіпті қызметкерлер – инженерлерден бастап құрылышшарға дейін – енді цифрлық айнымалылар мен деректер жиынтығы ретінде қарастырылады.

КРМГ-дің «Таныс мәселелер - жаңа тәсілдер: 2023 жылғы жаһандық құрылыш шолуы» деректері бойынша, цифрлық егіздер, жасанды интеллект (AI) және Үлкен деректер (Big Data) жобалардың рентабельділігін арттырудың негізгі қозғаушы күштері болып табылады.

Заманауи технологиялар ақпарат жинауды жеңілдетіп қана қоймай, деректерді сақтау құнын радикалды түрде төмөндөтеді. Нәтижесінде, компаниялар таңдамалы тәсілден бас тартып, болашақта талдау үшін барлық ақпарат массивін сақтауды жөн көреді, бұл процестерді оңтайландыру үшін әлеуетті мүмкіндіктерді ашады.

Ақпаратты цифрландыру және сандық трансформация жасырын, бұрын пайдаланылмаған ақпараттың құндылығын анықтауға мүмкіндік береді. Деректерді дұрыс ұйымдастырған жағдайда, олар екінші өмірге ие болады: олар қайта пайдаланылып, қайта қаралып, жаңа қызметтер мен шешімдерге интеграциялануы мүмкін.

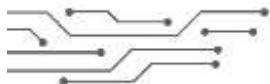
Болашақта ақпараттың цифрануы, ықтимал түрде, құжат айналымының толық автоматтандырылуына, өзін-өзі басқарушы құрылыш процестерінің енгізілүіне және жаңа мамандықтардың пайда болуына әкеледі – құрылыш деректерінің аналитиктері, ЖИ-мен жобаларды басқару бойынша сарапшылар және цифрлық инженерлер. Құрылыш нысандары динамикалық ақпарат көздеріне айналады, ал шешім қабылдау интуицияға немесе субъективті тәжірибеге емес, сенімді және қайталанатын цифрлық фактілерге негізделеді.

Ақпарат – 21 ғасырдың мұнайы, ал аналитика – ішкі жану қозғалтқышы. Питер Сондергаард, аға вице-президенті Gartner®

IoT Analytics 2024 жылғы деректеріне сәйкес, жаһандық деректерді басқару және аналитикаға жұмысалатын шығындардың 2023 жылы 185,5 миллиард доллардан 2030 жылы 513,3 миллиард долларға дейін күтілуде, орташа жылдық өсу қарқыны 16% құрайды. Дегенмен, барлық компоненттер бірдей қарқынмен өспейді: аналитика жылдам дамып келеді, ал деректерді сақтау жүйелерінің өсуі баяулайды. Аналитика деректерді басқару экожүйесінде ең жылдам өсуді қамтамасыз етеді: болжам бойынша, оның көлемі 2023 жылы 60,6 миллиард доллардан 2030 жылы 227,9 миллиард долларға дейін өседі, бұл орташа жылдық өсу қарқыны 27% құрайды.

Ақпараттың цифрлану процесінің жеделдеуі және ақпарат көлемінің тез өсуімен құрылыш жобалары мен компаниялардың басшылығы әртүрлі, көбінесе гетерогенді деректерді жүйелі түрде сақтау, талдау және өндеде қажеттілігімен бетпе-бет келеді. Осыған жауап ретінде, 1990-шы жылдардың ортасынан бастап, сала электронды құжаттарды жасау, сақтау және басқаруға жаппай көшуді бастады – кестелер мен жобалық есептеулерден бастап сызбалар мен келісімшарттарға дейін.

Қолмен жазылған дәстүрлі құжаттар, қол қоюды, физикалық сақтауды, тұрақты қайта қарауды және шкафттарда архивтеуді талап ететін, біртіндеп құрылымдалған түрде деректерді сақтайтын цифрлық жүйелермен ығыстырылуда – арнайы қосымшалардың дерекқорларында.



ТАРАУ 1.2.

Заманауи құрылымдық саласындағы технологиялар мен басқару жүйелері

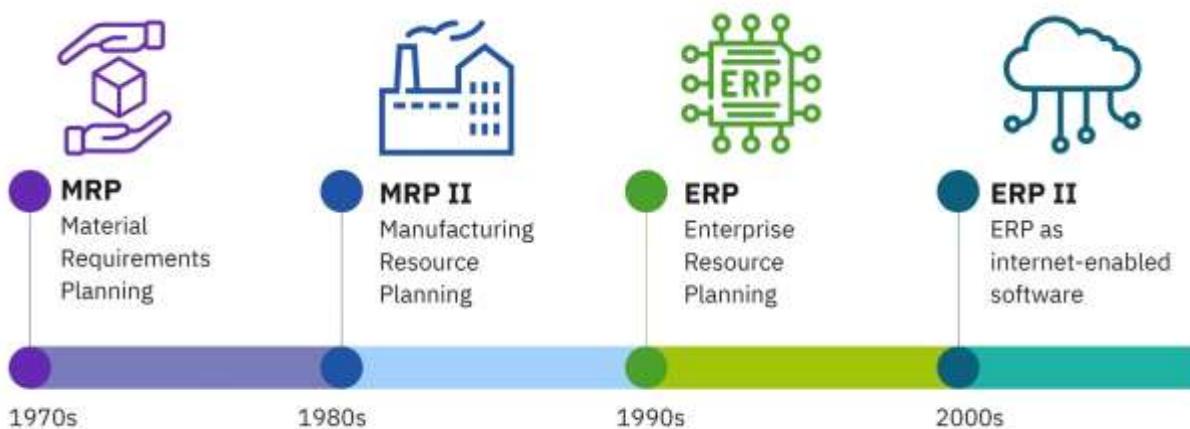
Цифрлық революция және модульдік MRP/ERP жүйелерінің пайда болуы

Цифрлық форматта заманауи деректерді сақтау және өндіреуде дәуірі 1950-ші жылдары магниттік ленталардың пайда болуымен басталды, бұл үлкен көлемдегі ақпаратты сақтау және пайдалану мүмкіндігін ашты. Келесі серпілік дискілік жинақтаушылардың пайда болуы болды, олар құрылымдық саласындағы деректерді басқару тәсілін түбекейлі өзгерти.

Деректер қоймаларының дамуы нарыққа көптеген компаниялардың шығуына әкелді, олар деректерді жасау, сақтау, өндіреуде және рутиндік тапсырмаларды автоматтандыру үшін модульдік бағдарламалық қамтамасыз етуді өзірлеуді бастады.

Ақпарат пен құралдардың экспоненциалды өсуі интеграцияланған модульдік шешімдерді өзірлеу қажеттілігіне әкелді, олар жеке файлдармен жұмыс істемей, әртүрлі процестер мен жобалар аясында деректер ағындарын басқаруға және бақылауға көмектеседі.

Алғашқы кешенді платформалық құралдар тек құжаттарды сақтап қана қоймай, сонымен қатар процестердегі өзгерістер мен операцияларға барлық сұрауларды құжаттауы тиіс болды: кім бастады, сұраудың көлемі қандай болды және нәтижесінде қандай мән немесе атрибут ретінде жазылды. Осы мақсаттар үшін дәл есептеулер мен қабылданған шешімдерді бақылауға мүмкіндік беретін жүйе қажет болды. Мұндай платформалар 1990-шы жылдардың басынан бастап танымал болған алғашқы MRP (Material Requirements Planning) және ERP (Enterprise Resource Planning) жүйелері болды.



Деректерді сақтау технологиясындағы жетістіктер 1980-ші жылдары ERP жүйелерінің пайда болуына әкелді.

Алғашқы MRP және ERP жүйелері бизнес-процесстер мен құрылыш жобаларын басқаруда цифрландыру дәүірінің негізін қалады. Бастапқыда негізгі бизнес-процесстерді автоматтандыру үшін арналған модульдік жүйелер уақыт өте келе қосымша, икемді және бейімделгіш бағдарламалық шешімдермен интеграциялана бастады.

Бұл қосымша шешімдер деректерді өңдеу және жобалардың мазмұнын басқару үшін арналған, олар не үлкен жүйелердің белгілі бір модульдерін алмастыруды, не тиімді түрде толықтырып, бүкіл жүйенің функционалдығын кеңейтті.

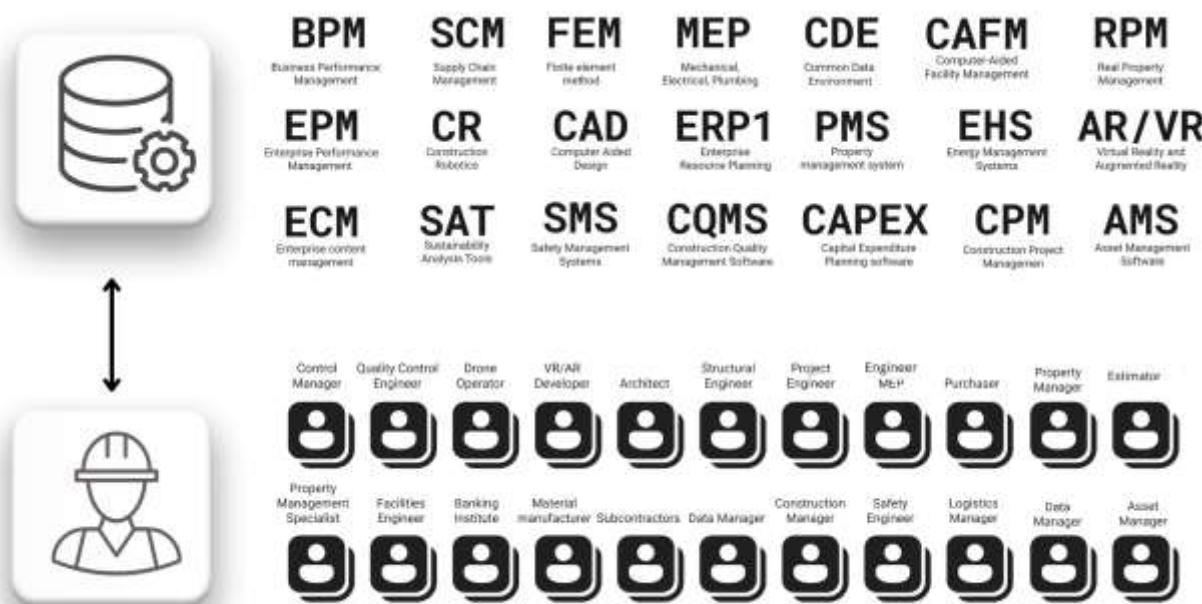


Рис. 1.22 Жаңа бағдарламалық шешімдер бизнеске деректер ағындарын басқару үшін менеджерлердің үлкен армиясын тартты.

Соңғы онжылдықтарда компаниялар модульдік жүйелерге елеулі қаражат инвестициялады [15], оларды ұзақ мерзімді интеграцияланған шешімдер ретінде қабылдай отырып.

2022 жылғы Software Path есебінің деректері бойынша [16], ERP жүйесінің бір пайдаланушысына орташа бюджет 9 000 АҚШ долларын құрайды. Орташа алғанда, компания қызметкерлерінің шамамен 26%-ы осындан жүйелерді пайдаланады. Мысалы, 100 пайдаланушысы бар ұйым үшін ERP енгізуудің жалпы шығындары шамамен 900 000 долларды құрайды.

Проприетарлық, жабық модульдік шешімдерге инвестициялар қазіргі заманғы, икемді және ашық технологиялардың қарқынды дамуы аясында барған сайын аз негізделген болып келеді. Егер мұндай инвестициялар жасалған болса, бар жүйелердің рөлін объективті қайта бағалау маңызды: олар ұзақ мерзімді перспективада шынымен қажет пе, әлде олардың функцияларын қайта қараяу және тиімді әрі ашық түрде жүзеге асыру мүмкін бе.

Қазіргі заманғы модульдік деректер өңдеу платформаларының негізгі мәселелерінің бірі – деректерді жабық қосымшалар ішінде орталықтандыру. Нәтижесінде, компанияның негізгі активі – деректер – нақты бағдарламалық шешімдерге тәуелді болады, ал керісінше емес. Бұл ақпаратты

қайта пайдалануды шектейді, миграцияны қынданатады және бизнеске тез өзгеретін цифрлық ландшафта икемділікті төмендетеді.

Егер жабық модульдік архитектуралық маңыздылығы немесе сұранысы болашақта төмөндөйтін болса, бүгінгі күні жұмсалған шығындарды қайтарымсыз деп тану және ашық, масштабталатын және бейімделетін цифрлық экожүйеге стратегиялық көшуге назар аудару орынды.

Проприетарлық бағдарламалық қамтамасыз ету әзірлеуші компанияның бастапқы код пен пайдалануши деректеріне ерекше бақылауымен сипатталады, олар осындаид шешімдерді пайдалану барысында жасалады. Ашық бастапқы кодты бағдарламалардан айырмашылығы, пайдаланышылар қосымшаның ішкі құрылымына қол жеткізе алмайды және оны өз қажеттіліктеріне қарай қарап, өзгерте немесе бейімдей алмайды. Оның орнына, олар лицензияларды сатып алуға міндетті, бұл бағдарламалық қамтамасыз етуді жеткізуі белгілеген шектерде пайдалануға құқық береді.

Деректерге бағытталған қазіргі заманғы тәсіл басқа парадигманы ұсынады: деректер стратегиялық актив ретінде қарастырылуы тиіс — тәуелсіз, ұзақ мерзімді және нақты бағдарламалық шешімдерден бөлек. Қосымшалар, өз кезегінде, деректермен жұмыс істеу үшін тек құралдар болып табылады, оларды алмастыру кезінде маңызды ақпаратты жоғалту қаупінсіз еркін ауыстыруға болады.

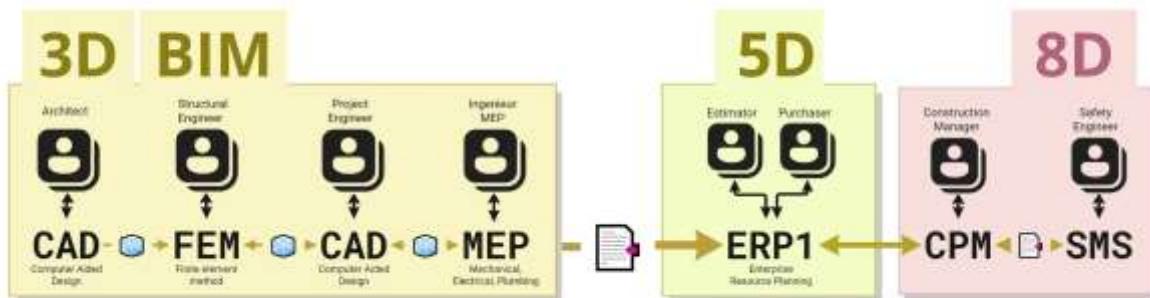
1990 жылдардағы ERP және MRP жүйелерінің дамуы (Рис. 1.21) бизнеске процестерді басқару үшін қуатты құралдар ұсынды, алайда күтпеген салдармен — ақпараттық ағындарды қолдаумен айналысадын қызметкерлер санының елеулі артуымен бірге жүрді. Автоматтандыру мен операциялық тапсырмаларды жеңілдетудің орнына, мұндай жүйелер жиі жаңа күрделілік, бюрократия және ішкі IT ресурстарына тәуелділікті тудырды.

Деректерді басқару жүйелері: деректерді жинаудан бизнес-мәселелеріне дейін

Қазіргі заманғы компаниялар деректерді басқарудың көптеген жүйелерін интеграциялау қажеттілігімен бетпе-бет келеді. Деректерді басқару жүйелерін таңдау, оларды тиімді басқару және әртүрлі дереккөздерді интеграциялау бизнес тиімділігі үшін критикалық маңызды міндетке айналады.

2020-жылдардың ортасында орташа құрылымы компанияларында жүздеген (ал ірі компанияларда мындаған) әртүрлі жүйелер табуға болады, олар құрылымы процесінің барлық аспектілерінің тегіс және үйлесімді өтуі үшін бірлесіп жұмыс істеуі тиіс.

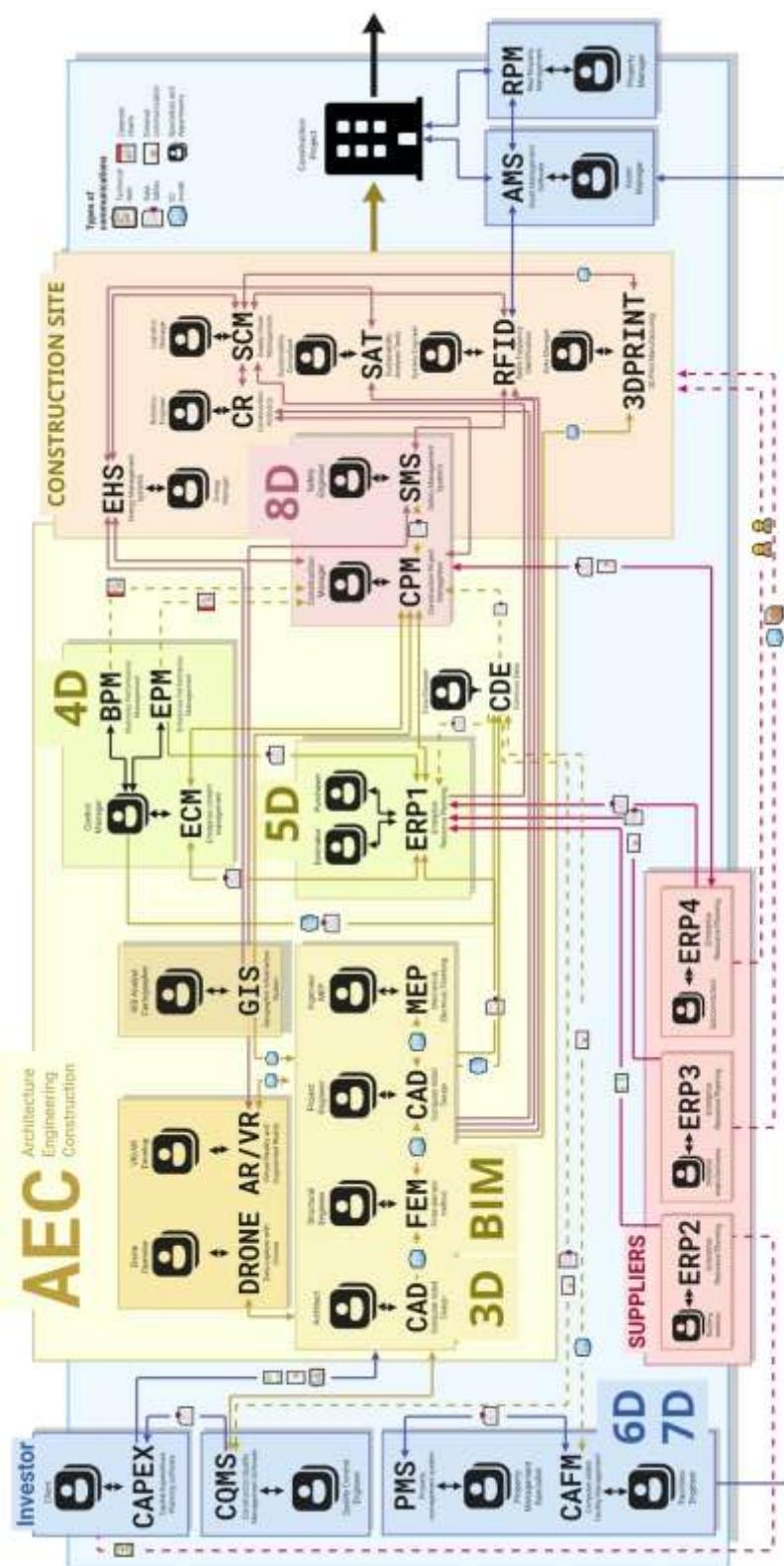
Deloitte®-тың 2016 жылғы «Деректерге негізделген басқару цифрлық капиталдық жобаларда» зерттеуіне сәйкес, орташа құрылымы маманы күн сайын 3,3 бағдарламалық қосымшаны пайдаланады, алайда олардың тек 1,7-сі бір-бірімен интеграцияланған.



Әрбір бизнес-жүйе деректерді сапалы басқару үшін кәсіби команда мен жауапты менеджерді талап етеді.

Орташа және ірі құрылыш компанияларында құрылыш жобаларын тиімді басқару үшін қолданылатын танымал жүйелердің тізімі төменде келтірілген:

- ERP (Enterprise Resource Planning) – бухгалтерлік есеп, сатып алу және жобаларды басқару сияқты бизнес-процесстердің интеграциясын қамтамасыз етеді.
- CAPEX (Capital Expenditure Planning Software) – құрылыш жобаларындағы қаржылық инвестицияларды жоспарлау және басқару үшін қолданылады, негізгі құралдар мен үзақ мерзімді активтерге инвестициялардың құнын анықтауға көмектеседі.
- CAD (Computer-Aided Design) және BIM (Building Information Modeling) – жобалардың егжей-тегжейлі және дәл техникалық сызбаларын және 3D модельдерін жасау үшін қолданылады. Бұл жүйелерде геометриялық ақпаратпен жұмыс істеуге басты назар аударылады.
- MEP (Mechanical, Electrical, Plumbing) – механикалық, электрлік және сантехникалық компоненттерді қамтитын инженерлік жүйелер, жобаның ішкі "қан тамырлары" жүйесін егжей-тегжейлі сипаттайды.
- ГИС (географиялық ақпараттық жүйелер) – картография және кеңістіктік талдау сияқты жерді талдау және жоспарлау үшін қолданылады.
- CQMS (құрылыш сапасын басқару бағдарламалық қамтамасыз етуі) – құрылыш процестерінің белгіленген стандарттар мен нормаларға сәйкестігін қамтамасыз етеді, ақауларды жоюға көмектеседі.
- СРМ (құрылыш жобаларын басқару) – құрылыш процестерін жоспарлау, үйлестіру және бақылауды қамтиды.
- CAFM (Computer-Aided Facility Management) – ғимараттарды басқару және қызмет көрсету жүйелері.
- SCM (Supply Chain Management) – материалдар мен ақпараттың жеткізушилер мен құрылыш алаңы арасындағы ағынын оңтайландыру үшін қажет.
- EPM (Enterprise Performance Management) – бизнес-процестер мен өнімділікті жақсартуға бағытталған.
- AMS (Asset Management Software) – активтердің өмірлік циклінің барлық кезеңінде жабдықтар мен инфрақұрылымды пайдалану, басқару және қызмет көрсету үшін қолданылады.
- RPM (Real Property Management) – ғимараттар мен жер участкерін, сондай-ақ онымен байланысты ресурстар мен активтерді басқару және пайдалану міндеттері мен процестерін қамтиды.



Жүйелердің өзара байланысы, компанияның процестерін әртүрлі бөлімшелер арасындағы ақпарат ағысымен байланыстырады.

- CAE (Computer-Aided Engineering) – компьютерлік инженерия, есептеу және симуляциялық жүйелерді, мысалы, шекті элементтер әдісі (FEA) және есептеу гидродинамикасын (CFD) қамтиды.
- CFD (Computational Fluid Dynamics) – сұйықтықтар мен газдардың ағындарын модельдеу. CAE-нің бір категориясы.
- CAPP (Компьютерлік Процесс Жоспарлау) – технологиялық процестерді компьютерлік жоспарлау. Маршруттық және технологиялық карталарды жасау үшін қолданылады.
- CAM (Компьютерлік Жасау) – автоматтандырылған өндіріс, ЧПУ станоктарына басқару бағдарламаларын генерациялау.
- PDM (Өнім Деректерін Басқару) – өнім туралы деректерді басқару, техникалық құжаттаманы сақтау және басқару жүйесі.
- MES (Өндірістік Атқару Жүйесі) – өндірістік процестерді нақты уақыт режимінде басқару жүйесі.
- PLM (Өнімнің Өмірлік Цикліне Басқару) – жобаның элементінің өмірлік цикліне басқару, PDM, CAPP, CAM және басқа жүйелерді біріктіріп, өнімді өзірлеуден жоюға дейін толық бақылауды қамтамасыз етеді.

Бұл және басқа да көптеген жүйелер, әртүрлі бағдарламалық шешімдерді қамти отырып, қазіргі құрылыш саласының ажырамас бөлігіне айналды (Рис. 1.24). Мұндай жүйелер, негізінен, жобалау және құрылыш кезеңдерінде ақпаратты тиімді енгізу, өңдеу және талдау үшін интуитивті интерфейстермен қамтамасыз етілген мамандандырылған деректер базалары болып табылады. Цифрлық құралдардың интеграциясы жұмыс процестерін оңтайландыруға ғана емес, сонымен қатар қабылданатын шешімдердің дәлдігін едәуір арттыруға ықпал етеді, бұл жобалардың мерзімі мен сапасына оң әсер етеді.

Алайда, интеграцияның жартысында жоқ. Статистикаға сәйкес, тек әрбір екінші қосымша немесе жүйе басқа шешімдермен интеграцияланған [17]. Бұл цифрлық ортадағы фрагментацияның сақталуын көрсетеді және құрылыш жобасы аясында үздіксіз ақпарат алмасуды қамтамасыз ету үшін ашық стандарттар мен унификацияланған интерфейстерді дамыту қажеттілігін атап көрсетеді.

Интеграциядағы басты қындықтардың бірі қазіргі компаниялар үшін цифрлық жүйелердің жоғары күрделілігі мен ақпаратты тиімді іздеу және интерпретациялау үшін қажетті пайдаланушы біліктілігі талаптары болып табылады. Әрбір бизнес жүйесін енгізу үшін мамандар командасы құрылып, оны негізгі менеджер басқарады (Рис. 1.22).

Негізгі жүйелік менеджер деректер ағынның қажетті бағытына жауап береді және соңғы ақпараттың сапасына жауапты, мындаған жылдар бүрын алғашқы менеджерлер папиrustар немесе балшық тақталарындағы жазбалар үшін жауап бергендей.

Ақпарат ағындарын басқару құралдарына айналдыру үшін жүйелік интеграция және деректерді басқару қабілеті қажет. Бұл архитектурада менеджерлер біртұтас желінің элементтері ретінде әрекет етуі тиіс – компанияның жеке бөліктерін байланыстыратын мицелий сияқты, бейімделіп, дамуға қабілетті тұтас тірі организм.

Корпоративті мицелий: деректер бизнес-процестерге қалай бірігеді.

Деректерді қосымшалар мен деректер базаларына интеграциялау процесі әртүрлі бөлімдер мен мамандардан алынған ақпаратты агрегациялауға негізделеді (Рис. 1.24). Мамандар қажетті деректерді іздең, оларды өндеп, әрі қарай пайдалану үшін өз жүйелері мен қосымшаларына береді.

Компанияның әрбір жүйесі, құралдар, технологиялар және деректер базаларынан тұратын – тарихи деректерге тамырын жайып, жаңа шешімдер: құжаттар, есептеулер, кестелер, графіктер және панельдер түрінде жаңа өнімдер әкелетін білімдер ағашы. Компаниядағы жүйелер, орманнның белгілі бір бөлігіндегі ағаштар сияқты, бір-бірімен әрекеттесіп, байланысып, маман менеджерлермен қолдау көрсетілетін және басқарылатын құрделі әрі жақсы құрылымдалған жүйені құрайды.

Компаниядағы ақпаратты іздеу және беру жүйесі ағаштардан (жүйелер) және санырауқұлақтардың мицелийінен (менеджерлер) тұратын құрделі орман желісі ретінде жұмыс істейді, олар ақпараттың берілуін және қажетті жүйелерге түсін қамтамасыз ететін бағыттаушылар мен өндеушілер рөлін атқарады. Бұл компания ішіндегі деректердің сау және тиімді ағынын және таралуын қолдауға көмектеседі.

Сарапшылар, тамырлар сияқты, жобаның бастапқы кезеңдерінде шикі деректерді сіңіріп, оларды корпоративтік экожүйе үшін қоректік заттарға айналдырады. Деректер мен контентті басқару жүйелері (Рис. 1.24 - ERP, CRM, BIM және т.б.) бұл білімдердің компанияның барлық деңгейлері арқылы айналымын қамтамасыз ететін қуатты ақпараттық магистральдар рөлін атқарады.

Табиғаттағыдай, экожүйенің әр элементі өз рөлін атқарса, компанияның бизнес-ландшафтында процесс қатысушылары - инженерлерден аналитиктерге дейін - ақпараттық ортаның өсуі мен құнарлылығына өз үлестерін қосады. Бұл жүйелік "деректер ағаштары" (Рис. 1.25) ақпарат жинау механизмдері ғана емес, сонымен қатар компанияның тұрақты дамуын қамтамасыз ететін бәсекелестік артықшылық болып табылады.-

Орман экожүйелері цифрлық корпоративтік құрылымдарды үйімдастыру принциптерін таңқалаңық дәлдікпен көрсетеді. Орманнның көп деңгейлі құрылымына - тәменгі қабаттан ағаштардың шындарына дейін - корпоративтік басқару жауапкершілік деңгейлері мен функционалдық бөлімдер бойынша міндеттерді бөледі.

Ағаштардың терең және тармақталған тамырлары тұрақтылықты және қоректік заттарға қолжетімділікті қамтамасыз етеді. Сол сияқты, мықты үйімдік құрылым мен сапалы деректермен жұмыс істей процестері компанияның ақпараттық экожүйесін қолдайды, оның тұрақты өсуі мен дамуына ықпал етеді, тіпті нарықтық тұрақсыздық пен дағдарыс кезеңдерінде (күшті жел) де.

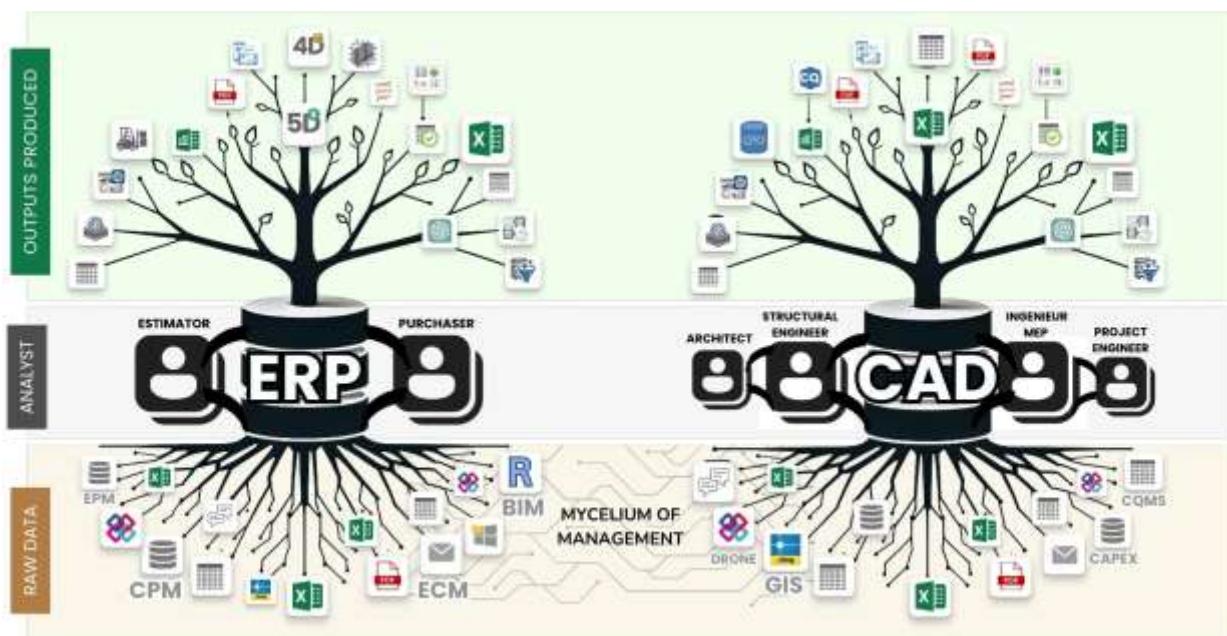


Рис. 1.25 Деректерді әртүрлі жүйелер арқылы интеграциялау мен менеджерлер мен мамандарды біртұтас ақпараттық желіге біркіттертін мицелийге үксас.

Бизнес саласындағы қазіргі түсінік эволюцияға ұшырады. Бұгінгі күні компанияның құндылығы тек оның көрінетін бөлімі - "кроналары" - соңғы құжаттар мен есептер түрінде ғана емес, сонымен қатар сапалы жинақталған және жүйелі түрде өндөлген деректердің "тамыр жүйесінің" тереңдігімен анықталады. Неғұрлым көп ақпарат жинап, өндесе, бизнес құндылығы соғұрлым жоғары болады. "Компост" ретінде қайта өндөлген деректерді жүйелі түрде жинақтап, одан пайдалы инсайттарды шығара алатын компаниялар стратегиялық артықшылыққа ие болады.

Тарихи ақпарат жаңа капитал түріне айналып, өсу, процестерді оңтайландыру және бәсекелестік артықшылықты қамтамасыз етеді. Деректерге бағытталған әлемде женетіндер - көп билетіндер, көп билетіндер емес.

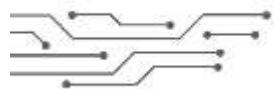
Құрылыш саласы үшін бұл жобаларды нақты уақыт режимінде басқаруға көшу дегенді білдіреді, мұнда барлық процестер - жобалаудан сатып алуға дейін, мердігерлерді үйлестіруге дейін - күнделікті жаңартылатын деректерге негізделеді. Әртүрлі көздерден (ERP жүйелері, CAD модельдері, құрылыш алаңындағы IoT датчиктері, RFID) ақпаратты интеграциялау дәл болжаулар жасауға, өзгерістерге жедел жауап беруге және өзекті деректердің болмауынан туындастырын кідірістерді болдырмауға мүмкіндік береді.

McKinsey & Company® (2022 [18]) жүргізген "Деректермен басқарылатын 2025 жылғы кесіпорын" зерттеуі бойынша, болашақтағы табысты компаниялар өз қызметінің барлық негізгі аспектілерінде – стратегиялық шешімдерден бастап операциялық өзара әрекеттерге дейін – деректерге сүйенетін болады.

Деректер тек аналитикалық құрал болып қалмай, барлық бизнес-процесстердің ажырамас бөлігіне айналады, басқаруды ашықтық, бақылау және автоматтандырумен қамтамасыз етеді. Деректерге

негізделген тәсіл үйімдарға адам факторының әсерін минимизациялауға, операциялық тәуекелдерді төмендетуге және шешім қабылдаудың ашықтығы мен тиімділігін арттыруға мүмкіндік береді.

XXI ғасыр экономикалық парадигманы өзгертуде: бұрын мұнай "қара алтын" деп аталып, механизмдер мен көліктерді қозғау қабілеті үшін бағаланса, бүгінгі күні тарихи деректер уақыт қысымымен жаңа стратегиялық ресурсқа айналып, бизнес қозғалысын қамтамасыз ететін шешім қабылдау алгоритмдерін қоректендіреді.



ТАРАУ 1.3.

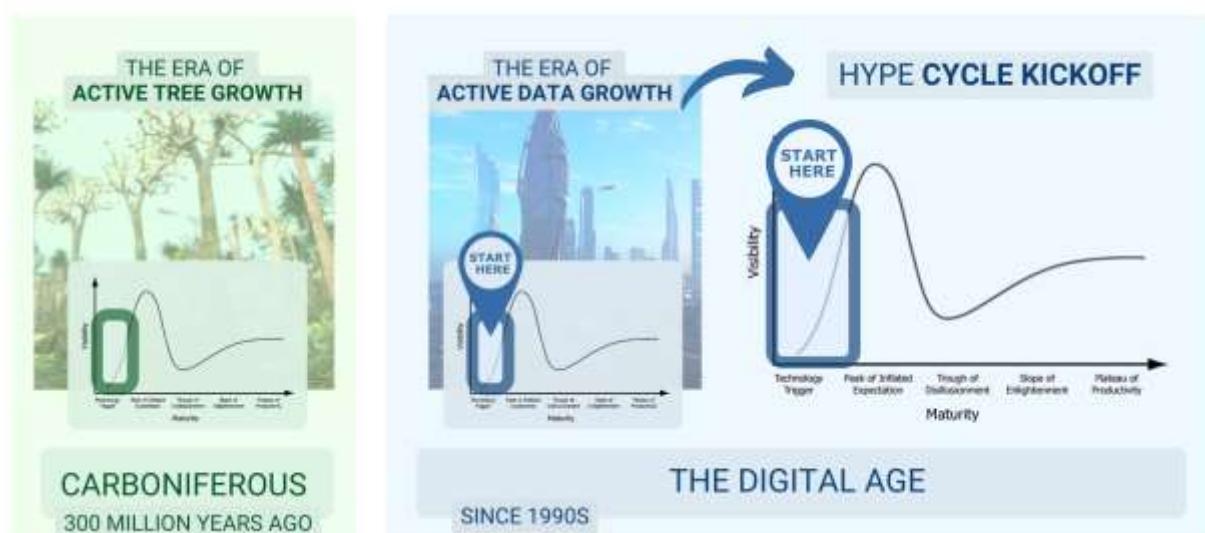
ЦИФРЛЫҚ РЕВОЛЮЦИЯ ЖӘНЕ ДЕРЕКТЕРДІҢ ЖАРЫЛЫСЫ

Деректер көлемінің өсуінің басталуы: эволюциялық толқын

Құрылыш саласы бұрын-соңды болмаған ақпараттық жарылышты бастаң кешіруде. Бизнесті деректермен қоректенетін білім ағашы ретінде елестетсек, қазіргі цифрландыру кезеңін тас көмір дәуіріндегі өсімдіктердің қарқынды өсуімен салыстыруға болады – бұл кезеңде Жердің биосферасы биомассаның жедел жинақталуы арқасында өзгерді.-

Глобалдық цифрлық даму жағдайында құрылыш саласындағы ақпарат көлемі жыл сайын екі есе артуда. Заманауи технологиялар деректерді фоновый режимде жинауға, оларды нақты уақытта талдауға және бұрын мүмкін емес масштабта пайдалануға мүмкіндік береді.

Intel®-дің соучредителі Гордон Мурдың формулировкасы бойынша, Мур заңы интегралды схемалардың тығыздығы мен құрделілігі, сондай-ақ өндөлеттің және сақталатын деректер көлемі шамамен әр екі жыл сайын екі есе артады [19].



Цифрландырудың басталуы деректердің экспоненциалды өсуіне әкелді, бұл тас көмір дәуіріндегі өсімдіктердің жарылышына үқсас.

Егер ежелгі мегалитикалық құрылымдар, мысалы, Гёбекли-Тепе (Түркия), қайталанатын пайдалануға жарамды құжатталған білім қалдырmasa, бұғынгі цифрлық технологиялар ақпаратты жинақтау және қайта пайдалану мүмкіндігін береді. Бұл споралы өсімдіктерден тұқымды өсімдіктерге (ангиосперма) эволюциялық өтпелі кезеңмен салыстыруға болады: тұқымның пайда болуы планетада өмірдің кеңінен таралуына серпін берді.

Сол сияқты, өткен жобалардан алынған деректер "цифрлық тұқымдар" ретінде қызмет етеді – масштабтауға және жаңа жобалар мен өнімдерде пайдалануға болатын ДНҚ білімінің тасымалдаушылары. Заманауи жасанды интеллект құралдарының – машиналық оқыту мен үлкен тілдік модельдердің (LLM), мысалы, ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok – пайда болуы

деректерді автоматты түрде шығарып, интерпретациялауға және жаңа контексттерде қолдануға мүмкіндік береді.

Тұқымдар бастапқыда тіршіліксіз планетада өмірдің таралуын революциялағандай, "деректер тұқымдары" жаңа ақпараттық құрылымдар мен білімдердің автоматты түрде пайда болуына негіз болады, бұл цифрлық экожүйелердің өздігінен дамуына және пайдаланушылардың өзгермелі талаптарына бейімделуіне мүмкіндік береді.

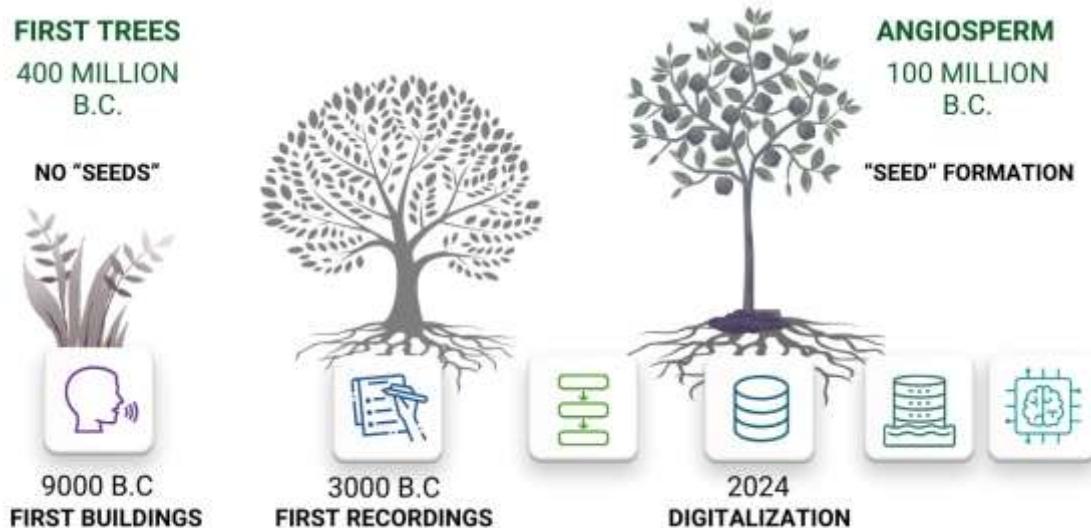


Рис. 1.32 Цифрлық "деректер тұқымдары" ангиоспермдер – гүлді өсімдіктердің экожүйені трансформациялаудағы эволюциялық рөлін атқарады.

Біз құрылыш саласында жаңа дәуірдің алдында тұрмыз, мұнда деректердің жарылышы және "деректер тұқымдарының" – өткен және қазіргі жобалардан алынған құрылымдық ақпараттық белсенді таралуы – саланың цифрлық болашағының негізін қалыптастырады. Оларды үлken тілдік модельдер (LLM) арқылы "тозаңдандыру" тек цифрлық өзгерістерді бақылап қана қоймай, өзін-өзі үйрететін, бейімделгіш экожүйелерді құруға белсенді қатысуға мүмкіндік береді. Бұл эволюция емес – бұл цифрлық революция, мұнда деректер жаңа шындықтың негізгі құрылыш материалына айналады.

Құрылыш саласындағы деректер көлемі жобалардың өмірлік циклі бойында әртүрлі пәндерден келетін ақпараттың арқасында күрт артып келеді. Бұл деректердің үлken жинақталуы құрылыш индустриясын Үлken Деректер дәуіріне итермеледі. – Профессор Ханг Ян, азаматтық құрылыш және архитектура факультеті, Ухань технологиялық университеті, Ухань, Қытай

Ақпарат көлемінің өсуі ақпараттық дәуірде табиғаттағы эволюциялық процестерді еске салады: ормандардың дамуы планетаның ежелгі ландшафтың өзгергендей, қазіргі ақпараттық жарылыш құрылыш саласының ландшафтың өзгертуде.

Қазіргі компанияда өндірілетін деректер көлемі

Соңғы екі жылда әлемдегі барлық деректердің 90%-ы жасалды. 2023 жылға қарай әр адам, соның ішінде құрылыш саласының мамандары, секундына шамамен 1,7 мегабайт деректерді генерациялайды, ал әлемдегі жалпы деректер көлемі 2023 жылы 64 зеттабайтқа жетеді және 2025 жылға қарай 180 зеттабайттан, яғни 180×10^{15} мегабайттан асып кететіні болжануда.

Бұл ақпараттық жарылыш XV ғасырда Иоганн Гутенбергтің баспа машинасын ойлап табуымен тарихи прецедентке ие. Оның пайда болғанына небәрі елу жыл өткен соң, Еуропадағы кітаптар саны екі есе артты: бірнеше онжылдықта қолмен жасалған кітаптардан 1200 жыл бойы шығарылған кітаптар санына тең көлемде кітаптар басылып шықты. Бүгін біз одан да жылдам өсуді байқаймыз: әлемдегі деректер көлемі үш жыл сайын екі есе артып отырады.

Деректердің қазіргі өсу қарқынын ескере отырып, құрылыш саласы келесі бірнеше онжылдықта өзінің бұрынғы тарихында жинақталған ақпарат көлеміне тең деректерді генерациялауы мүмкін.



Рис. 1.33 Әр қызметкердің компания серверлерінде күнделікті деректерді сақтау түрақты деректер көлемінің өсуіне ықпал етеді.

Қазіргі құрылыш бизнесінде тіпті шағын компаниялар күн сайын үлкен көлемде әртүрлі форматтағы ақпаратты генерациялайды, ал шағын құрылыш компаниясының цифрлық ізі күн сайын ондаған гигабайтқа жетуі мүмкін – модельдер мен сыйбалардан бастап фотофиксация мен объектідегі датчиктерге дейін. Егер әр маман орташа есеппен секундына шамамен 1,7 МБ деректерді жасайтынын ескерсек, бұл күніне шамамен 146 ГБ, немесе жылына 53 ТБ-ға тең (Рис. 1.33).

Егер 10 адамнан тұратын ұжым күніне 3 сағат бойы белсенді жұмыс істесе, күніне генерацияланатын ақпарат көлемі 180 гигабайтқа жетеді (Рис. 1.34).

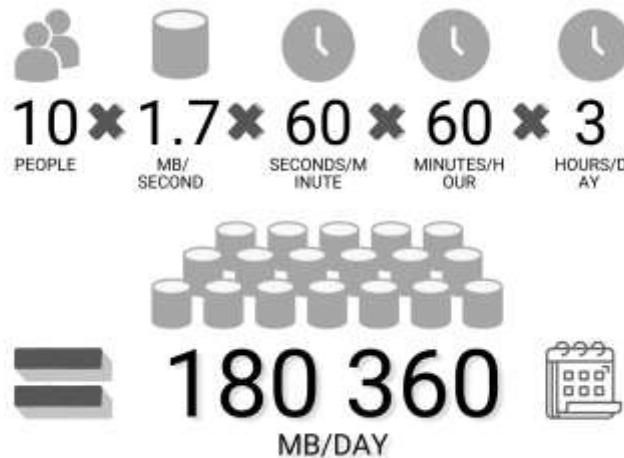


Рис. 1.34 10 адамнан тұратын компания құніне шамамен 50–200 гигабайт деректердің генерациялайды.

Егер 30% жұмыс деректері жаңа деп есептесек (қалғаны қайта жазылады немесе жойылады), онда 10 адамнан тұратын компания айына бірнеше жұз гигабайт жаңа деректерді жасай алады (нақты көрсеткіштер компанияның қызмет түріне байланысты).

Осылайша, біз тек деректерді көбейтіп қана қоймай, олардың тиімді басқарулына, сақталуына және ұзақ мерзімді қолжетімділігіне деген өсіп келе жатқан қажеттілікпен бетпе-бет келеміз. Егер бұрын деректер жергілікті серверлерде ерекше шығынсыз «жатса», цифрлық трансформация жағдайында көптеген компаниялар ақпараттық инфрақұрылымының негізі ретінде бұлттық шешімдерді қолдана бастайды.

Деректердің сақтау құны: экономикалық аспект

Соңғы жылдары көптеген компаниялар деректердің сақтау үшін бұлттық қызметтерге жүгінуде. Мысалы, егер компания өз деректерінің жартысын бұлтта орналастыrsa, орташа бағасы 0,015 доллар гигабайт үшін айына, оның сақтау шығындары айына 10–50 доллар [25] артуы мүмкін.

Кішігірім компания үшін деректердің генерациялау бойынша типтік модельдермен бұлттық сақтау шығындары айына жүздеген доллардан, бірнеше жылдан кейін мүмкін 1000 доллардан асуы мүмкін (Рис. 1.35), бұл әлеуетті түрде елеулі қаржылық жүктеме туғызады.

Forrester зерттеуіне сәйкес «Кәсіпорындар деректердің сақтау жүйелерін аутсорсингке береді, себебі күрделілік артуда» [26], онда технологиялық инфрақұрылым бойынша шешім қабылдайтын 214 басшы қатысқан, үйымдардың үштен бірі деректер операцияларының өсіп келе жатқан көлемі мен күрделілігіне байланысты деректердің сақтау жүйелерін аутсорсингке беруді таңдайды, ал екі үшінші кәсіпорын жазылым негізіндегі модельді қалайды.



Рис. 1.35 Деректердің бүлтқа көшіру айна 2000 долларға дейінгі сақтау шығындарын арттыруы мүмкін, тіпті 10 қызметкере бар компания үшін.

Жағдай CAD (BIM), CAFM, PMIS және ERP жүйелері сияқты бүлттық технологияларға жедел көшу арқылы құрделене түсуде, бұл деректердің сақтау және өндеу шығындарын қосымша арттырады. Нәтижесінде, компаниялар шығындарды оңтайландыру және бүлттық провайдерлерге тәуелділікті азайту жолдарын іздеуге мәжбүр.

2023 жылдан бастап, үлкен тілдік модельдердің (LLM) белсенді дамуы аясында деректердің сақтау тәсілдері өзгеруде. Қоғытеген компаниялар деректерді бақылауды қайтару туралы ойлана бастады, себебі ақпаратты өз серверлерінде өндеу қауіпсіз ері тиімдірек болуда.

Осы контексте тек қажетті деректердің сақтау және өндеуден бас тарту тенденциясы корпоративтік LLM және ИИ шешімдерін жергілікті орналастыруға қарай алға шығады. Microsoft-тың бас директоры өз сұхбаттарының бірінде [27] атап өткендей, әртүрлі тапсырмаларды орындау үшін бірнеше жеке қосымшаларға немесе SaaS бүлттық шешімдерге сүйенудің орнына, ИИ агенттері деректер базаларындағы процестерді басқаруға, әртүрлі жүйелердің функцияларын автоматтандыруға мүмкіндік береді.

Ескі тәсіл осы мәселені [деректерді өндеу] шешу үшін мынадай болды: әртүрлі бизнес-қосымшалары интеграциямен қалай айналысқанын еске алсақ, олар коннекторларды пайдаланды. Компаниялар осы коннекторларға лицензиялар сатты, және осының айналасында бизнес-модель қалыптасты. SAP [ERP] – классикалық мысалдардың бірі: SAP деректеріне қол жеткізу тек тиісті коннектор болған жағдайда мүмкін болды. Сондықтан менің ойымша, агенттердің [ЖИ] өзара әрекеттесуі жағдайында да осындай нәрсе пайда болады. Біз қабылдайтын тәсіл, ең болмағанда, бизнес-қосымшаларының болуы концепциясы, мүмкін, [ЖИ] агенттер дәүірінде құлайды. Себебі, егер ойласақ, олар негізінен бизнес-логикамен толтырылған деректер базалары болып табылады.— Сатья Наделла, Microsoft компаниясының бас директоры, BG2 арнасына берген сұхбат, 2024 ж.

Осы парадигмада деректерге негізделген тәсіл LLM-дің қолдану классикалық жүйелерден асып

түседі. Жасанды интеллект пайдалануши мен деректер арасында делдал болып, көптеген делдал интерфейстердің қажеттілігін жояды және бизнес-процестердің тиімділігін арттырады. Деректермен жұмыс істеудің осы тәсілі туралы біз «Киындықты жеңу және күрделілікті азайту» бөлімінде толығырақ сөйлесеміз.-

Бұғаңға таңда болашақ архитектурасы тек қалыптасып жатыр, компаниялар өткен шешімдердің салдарымен бетпе-бет келуде. Соңғы онжылдықтардағы жаппай цифрландыру, бөлшектелген жүйелердің енгізілуімен және бақылаусыз деректердің жиналудымен бірге жаңа проблема – ақпараттық жүктеме туындағы.

Деректердің жинақтаудың шекаралары: массадан мағынаға

Заманауи компания жүйелері басқарылатын өсу жағдайында сәтті дамып, жұмыс істейді, мұнда деректер көлемі мен қосымшалар саны ИТ бөлімдерінің және менеджерлердің мүмкіндіктерімен теңдестірілген. Алайда, соңғы онжылдықтардағы цифрландыру деректердің көлемі мен күрделілігінің бақылаусызы өсуіне әкелді, бұл компаниялардың ақпараттық экожүйесінде артық жүктеме әсерін тудырды.

Бұғаңға таңда серверлер мен сақтау орындары өндөлмеген және әртүрлі форматтағы ақпараттың бұрын-соңды болмаған ағынына ұшырауда, ол компостқа айналуға үлгермейді және тез арада өзектілігін жоғалтады. Компанияның шектеулі ресурстары осы ағынмен күрсес алмайды, ал деректер оқшауланған сақтау орындарында (так называемые «силосы») жиналады, олардан пайдалы ақпаратты алу үшін қолмен өңдеу қажет.

Нәтижесінде, плющпен жабылған және зеңмен жабылған орманға үқсас, қазіргі заманғы компанияларды басқару жүйелері жиі ақпараттық жүктемеден зардап шегеді. Корпоративтік экожүйенің негізінде қоректік ақпараттық гумус орнына әртүрлі форматтағы деректердің оқшауланған участкерлері қалыптасады, бұл бизнес-процестердің жалпы тиімділігінің төмендеуіне әкеледі.

Соңғы 40 жылда байқалған деректер көлемінің экспоненциалды өсу кезеңі міндетті түрде қанығу және кейінгі салқындау фазасына ауысады. Сақтау орындары шегіне жеткенде, сапалық өзгеріс орын алады: деректер тек сақтау объектісі болмай, стратегиялық ресурсқа айналады.

Жасанды интеллект пен машиналық оқытуудың дамуы компанияларға ақпаратты өңдеу шығындарын азайтуға және сандық өсуден деректерді сапалы пайдалануға көшуге мүмкіндік береді. Алдағы онжылдықта құрылыштың саласы назарын жаңа деректер массивтерін жасаудан, олардың құрылымдылығын, тұтастырының және аналитикалық құндылығын қамтамасыз етуге ауыстыруы тиіс.

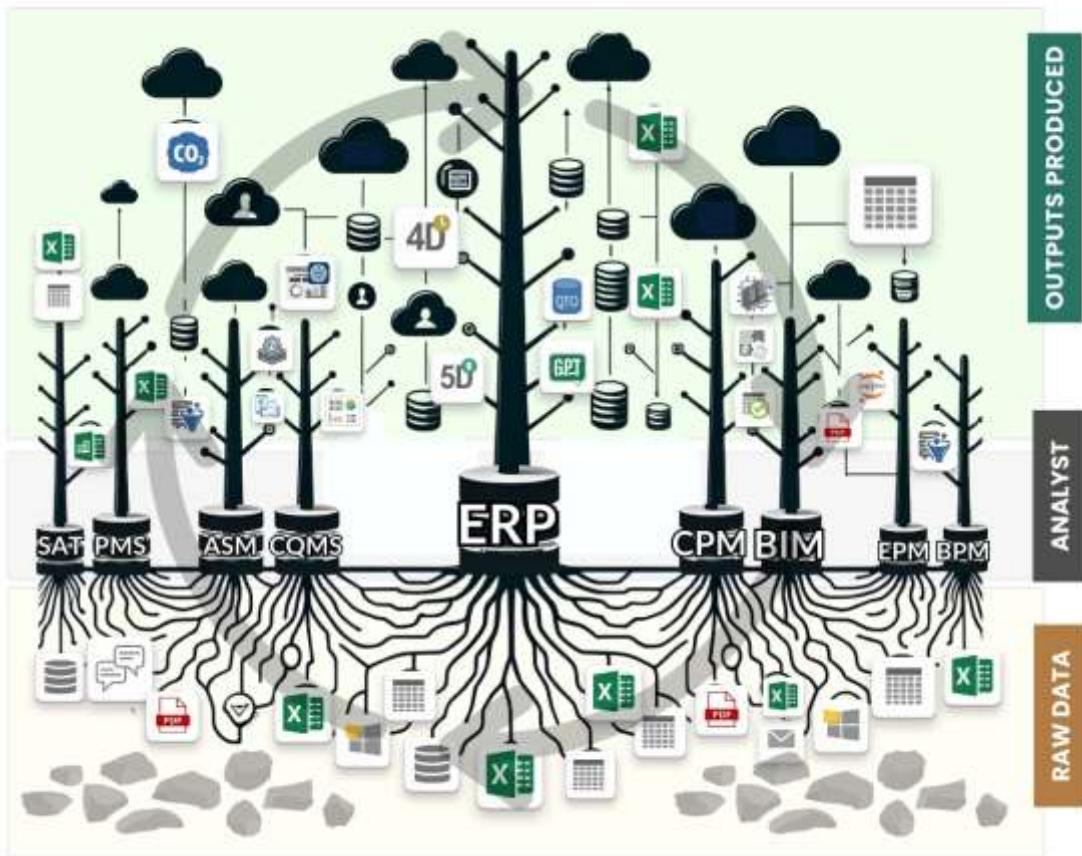


Рис. 1.36 Изолирленген дереккөздер ақпараттың жүйелер арасында алмасуына кедергі келтіреді.

Негізгі құндылық енді ақпарат көлемінде емес, оны автоматты түрде интерпретациялау және басқарушылық шешімдер қабылдау үшін пайдалы қолданбалы білімге айналдыру қабілетінде. Деректер шын мәнінде пайдалы болуы үшін оларды дұрыс басқару қажет: жинау, тексеру, құрылымдау, сақтау және нақты бизнес міндеттер контекстінде талдау.

Компаниядағы деректерді талдау процесі ормандағы ағаштардың өмір циклі мен ыдырау процесіне және жаңа жас ағаштардың пайда болуына ұқсас: жетілген ағаштар өлеңді, шіриді және жаңа ескіндер үшін қоректік ортаға айналады. Дайын және аяқталған процестер өз қолданылуынан кейін компанияның ақпараттық әкімшілікке түсіп, ақырында жаңа жүйелер мен деректердің өсуіне қоректік гумусқа айналады.

Алайда, тәжірибеде бұл цикл жиі бұзылады. Органикалық жаңарудың орнына, қабаттасқан хаос қалыптасады – геологиялық қабаттарға ұқсас, мұнда жаңа жүйелер ескі жүйелердің үстіне қабатталып, терең интеграция мен құрылымдаусыз қалыптасады. Нәтижесінде білімнің айналымына кедергі келтіретін және деректерді басқаруды қындараттын бөлшектенген ақпараттық «силостар» пайда болады.

Келесі қадамдар: деректер теориясынан практикалық өзгерістерге

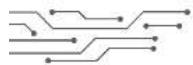
Құрылыш саласындағы деректердің эволюциясы – балшықтан жасалған тақталардан қазіргі заманғы модульдік платформаларға дейінгі жол. Бүгінгі күні шақыру ақпаратты жинауда емес, әртүрлі және форматталған деректерді стратегиялық ресурсқа айналдыратын құрылымды құруда. Сіздің рөліңіз – компания басшысы немесе қаралайым инженер болсын – деректердің құндылығын түсіну және олармен жұмыс істеу болашақта негізгі кәсіби дағдыға айналатын болады.

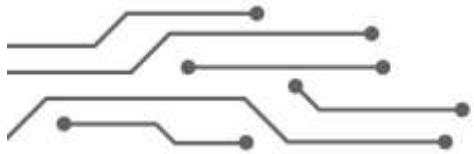
Осы бөлімнің қорытындысында, қарастырылған тәсілдердің күнделікті міндеттеріндеңде қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

- Ақпараттық ағындарды жеке аудит жүргізіңіз.
- Сіз күнделікті жұмыс істейтін барлық жүйелер мен қосымшалардың тізімін жасаңыз.
- Деректердің іздеуге немесе қайта тексеруге көп уақыт жұмсайтын жерлерді белгілеңіз.
- Негізгі ақпарат көздерінде анықтаңыз.
- Қосымшалардың қазіргі ландшафтын артықшылықтар мен функциялардың дублиренуі түрғысынан талдаңыз.
- Аналитикалық жетілу деңгейлері бойынша процестерде алға жылжуға ұмтылыңыз.
- Міндеттеріндеңде сипаттамалық аналитикадан бастаңыз (не болды?).
- Постепенно диагностикалық (неге бұл болды?) енгізіңіз.
- Процестерде болжамдық (не болады?) және ұсыныстық (не істеу керек?) аналитикаға қалай көшуге болатынын ойластырыңыз.
- Жұмыс деректеріндеңде құрылымдауды бастаңыз.
- Жиі пайдаланатын файлдар мен папкалардың біртұтас атаяу жүйесін енгізіңіз.
- Жиі қолданылатын құжаттар мен есептер үшін шаблондар жасаңыз.
- Тұрақты түрде аяқталған жобаларды айқын құрылыммен архивтеңіз.

Егер сіз өз командаңыздағы немесе компанияңыздағы ақпараттық инфрақұрылымды толық өзгерте алмасаңыз да, өз процестеріңіз бер күнделікті жұмысыңыздағы шағын жетілдірулерден бастаңыз. Деректердің шынайы құндылығы олардың көлемінде емес, олардан практикалық пайда алу қабілетінде екенін есте сақтаңыз. Тіпті шағын, бірақ дұрыс құрылымдалған және талданған ақпарат массивтері шешім қабылдау процестеріне интеграцияланған жағдайда елеулі әсер етуі мүмкін.

Келесі бөлімдерде біз деректермен жұмыс істеудің нақты әдістері мен құралдарына өтеміз, құрылымдалмаған ақпаратты құрылымдалған жиынтықтарға айналдыру тәсілдерін қарастырамыз, талдауды автоматтандыру технологияларын зерттейміз және құрылыш компаниясында тиімді аналитикалық әкожүйені қалай құруды егжей-тегжейлі талқылаймыз.





ІІ БӨЛІМ ҚҰРЫЛЫС БИЗНЕСІ ДЕРЕКТЕР ХАОСЫНДА ҚАЛУ

Екінші бөлім құрылым компанияларының өсіп келе жатқан деректер көлемдерімен жұмыс істеу кезінде кездесетін мәселелерді сынни талдауға арналған. Ақпараттық фрагментациясының салдары мен "силос түріндегі деректер" феномені, тиімді шешім қабылдауға кедергі келтіретін мәселелер егжей-тегжейлі қарастырылады. НiPPO тәсілінің (Ең жоғары жалақы алатын адамның пікірі) проблематикасы және оның құрылымы жобаларындағы басқарушылық шешімдердің сапасына әсері зерттеледі. Бизнестің динамикалық процестерінің және олардың өсіп келе жатқан күрделілігінің ақпараттық ағындарға және операциялық тиімділікке әсері бағаланады. Артық күрделілік жүйелерінің шығындарды арттырып, ұйымдардың икемділігін төмендететін нақты мысалдары келтіріледі. Проприетарлық форматтардың шектеулеріне және құрылым саласында ашық стандарттарды қолдану перспективаларына ерекше назар аударылады. Артық күрделілікті және техникалық кедергілерді минимизациялайтын ИИ және LLM негізіндегі бағдарламалық экожүйелерге көшу концепциясы ұсынылады.

ТАРАУ 2.1.

ДЕРЕКТЕРДІҢ ФРАГМЕНТАЦИЯСЫ ЖӘНЕ СИЛОСЫ

Құралдардың саны көп болған сайын, бизнес тиімдірек пе?

Бір қарағанда, сандық құралдардың санының артуы тиімділіктің өсуіне әкеледі деп ойлауға болады. Алайда, тәжірибеде жағдай басқаша. Әр жаңа шешіммен, ол бүлттық қызмет, ескірген жүйе немесе кезекті Excel есебі болса да, компания өз цифрлық ландшафтында тағы бір қабат қосады – бұл қабат көбінесе қалғандарымен интеграцияланбайды.

Деректердің көмірмен немесе мұнаймен салыстыруға болады: олар жылдар бойы қалыптасып, хаос, қателіктер, құрылымдалмаған процестер мен ұмытылған форматтардың қабаттары астында «сығылып» жатады. Олардың ішінен шын мәнінде пайдалы ақпаратты алу үшін компанияларға ескірген шешімдер мен цифрлық шудың қабаттарынан өтуге тұра келеді.

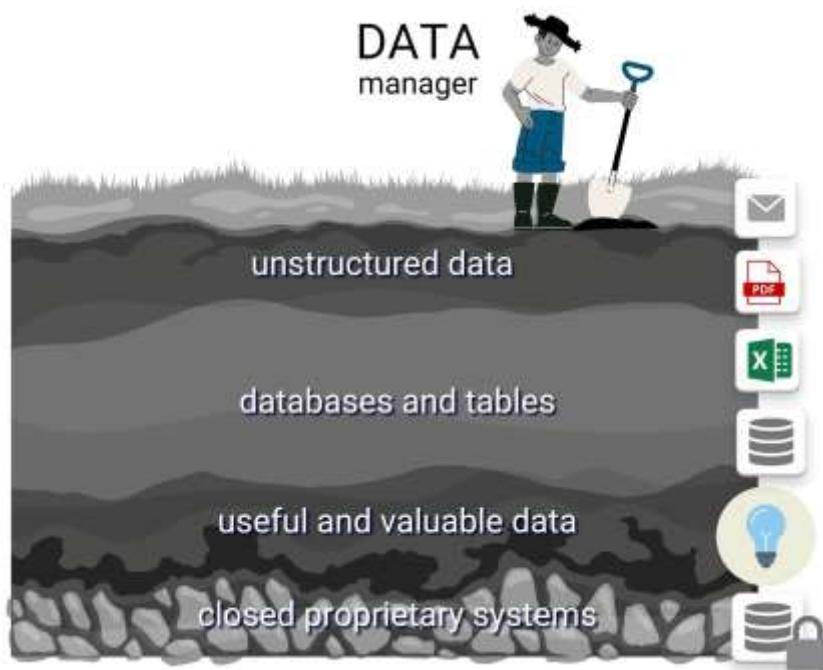


Рис. 2.11 Әр түрлі форматтағы деректер оқшауланған қабаттардың құрайды – тіпті «алтын» инсайттар жүйелік күрделіліктің геологиялық қабаттарында жоғалып кетеді.

Әрбір жаңа қосымша өзінен кейін із қалдырады: файл, кесте немесе серверде толық оқшауланған «силос». Бір қабат – бұл саз (ескірген және ұмытылған деректер), екінші – құм (бөлек кестелер мен есептер), үшінші – гранит (жабық меншікті форматтар, интеграцияға көнбейтін). Уақыт өте келе компанияның цифрлық ортасы бақылаусыз ақпарат жинақталуымен пластикалық қоймаға үқсас болып барады, мұнда құндылық компания серверлерінің тереңдігінде жоғалады.

Әр жаңа жоба мен жаңа жүйе арқылы инфрақұрылым ғана емес, сонымен қатар пайдалы сапалы деректерге жету жолы да күрделене түседі. Құнды «таужынысына» жету үшін терең тазалау

жүргізу, ақпаратты құрылымдау, «бөліп», оны мағыналы фрагменттерге топтастыру және аналитика мен деректер моделдеу арқылы стратегиялық маңызды мәліметтерді алу қажет.

Деректер - құнды нәрсе, және олар деректерді өндөйтін жүйелерден үзак үақыт қызмет етеді.

— Тим Бернерс-Ли, Галамтордың әкесі және алғашқы сайттың жасаушысы

Деректер «құнды нәрсеге» айналып, шешім қабылдаудың сенімді негізі болу үшін, олар мұқият дайындықтан өтуі тиіс. Дұрыс алдын ала өңдеу шашыраңқы мәліметтерді құрылымдалған тәжірибеге, пайдалы ақпараттық гумусқа айналдырады, ол кейіннен болжау мен онтайландыру құралына айналады.

Анализді бастау үшін деректердің мінсіз таза болуы қажет деген түсінік бар, алайда практикада «лас» деректермен жұмыс істей білу — процестің ажырамас бөлігі.

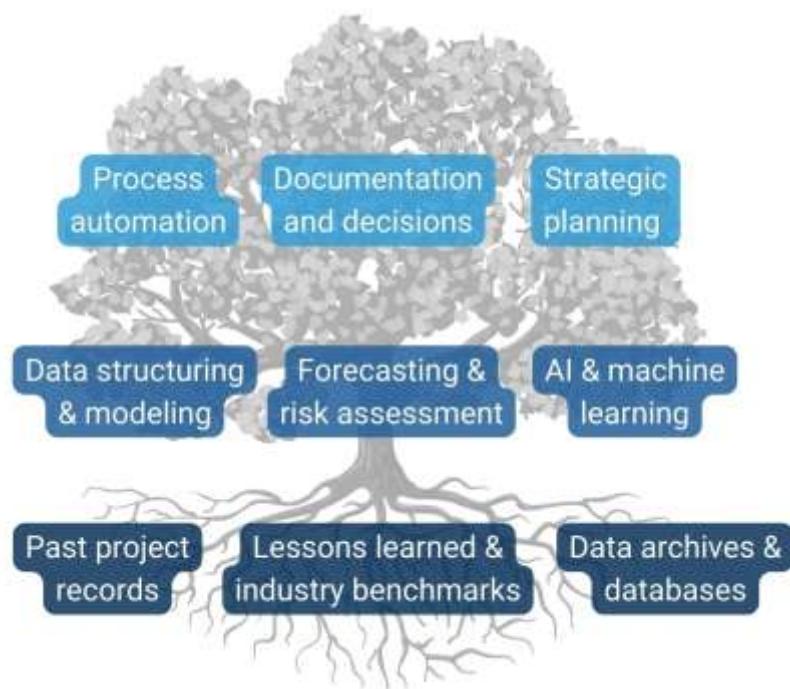


Рис. 2.12 Деректер бизнес негізі мен тамырлық жүйесі болып табылады, ол өз кезегінде шешім қабылдау процестеріне негізделеді.

Технологиялар тоқтап қалмайтындықтан, сіздің бизнесіңіз де алға жылжып, деректерден құн жасауға үйренуі тиіс. Мұнай және көмір компаниялары пайдалы қазбаларды өндіру үшін инфрақұрылым құрастыратын болса, бизнес те өз серверлерінде жаңа ақпарат ағындарын тиімді басқаруды және пайдаланылмаған, әртүрлі форматтағы және ескірген деректерден құнды мәліметтерді алып,

оларды стратегиялық ресурсқа айналдыруды үйренуі керек.

Деректер қоймаларын (мәліметтер қоймаларын) құру – бұл алғашқы қадам. Ең қуатты құралдар да деректердің оқшаулануы мен әртүрлі форматтағы деректер мәселеңін шешпейді, егер компаниялар бөлшектелген жүйелерде жұмыс істей берсе. Деректер бір-бірінен бөлек, ақпарат алмаспай, өмір сүрген кезде, бизнес «деректер силосы» эффектісіне тап болады. Біртұтас, келісілген инфрақұрылымның орнына компаниялар деректерді біріктіру мен синхрондауға ресурстарын жұмсауға мәжбүр болады.

Деректер силостары және олардың компания тиімділігіне әсері

Сіз тұрғын үй кешенін салып жатқанда, әр бригадада өз жобасы бар деп елестатіңіз. Біреулері қабырғаларды тұрғызыса, басқалары коммуникацияларды жүргізеді, үшіншілері – жолдарды төсейді, бір-бірімен тексермей. Нәтижесінде, құбырлар қабырғалардағы тесіктермен сәйкес келмейді, лифт шахталары қабаттармен сәйкес келмейді, ал жолдарды қайтадан бұзып, қайта төсеуге тұра келеді.

Мұндай жағдай – тек гипотетикалық сценарий емес, қазіргі заманғы көптеген құрылыш жобаларының шындығы. Әртүрлі жүйелермен жұмыс істейтін көптеген бас мердігерлер мен субмердігерлердің көптігі, бірегей үйлестіру орталығының болмауы, процесті шексіз келісімдер, қайта өңдеулер мен қақтығыстар тізбегіне айналдырады. Мұның бәрі елеулі кешіктірuler мен жобалардың бірнеше есе қымбаттауына әкеледі.

Құрылыш алаңында пайда болатын классикалық жағдай: опалубка дайын, бірақ арматура уақтылы жеткізілмеді. Әртүрлі жүйелердегі ақпаратты тексеру кезінде байланыс келесідей жүреді:

- Құрылыш алаңындағы прораб 20-шы күні жоба менеджеріне: «Біз опалубканы орнатуды аяқтадық, арматура қайда?» деп жазады.
- Жобаның менеджері (PMIS) жабдықтау бөлімінде: – «Опалубка дайын. Менің жүйемде [PMIS] арматура 18-ші күні келуі тиіс деп көрсетілген. Арматура қайда?»
- Жабдықтау маманы (ERP): – «Біздің ERP-де жеткізу 25-ші күні болады деп көрсетілген».
- Деректер инженері немесе ИТ бөлімі (интеграцияларға жауап береді): – PMIS-те 18-ші күн, ERP-де 25-ші күн көрсетілген. ERP мен PMIS арасында OrderID бойынша байланыс жоқ, сондықтан деректер синхрондалмаған. Бұл ақпараттық алшақтықтың типтік мысалы.
- Жобаның менеджері бас директорға – «Арматураның жеткізілуі кешіктірілуде, алаң тоқтап тұр, ал кім жауапты – түсініксіз».

Инциденттің себебі деректердің әртүрлі жүйелерде оқшаулануы болды. Деректер көздерін интеграциялау және унификациялау, біртұтас ақпарат қоймасын құру, сондай-ақ ETL құралдары (Apache NiFi, Airflow немесе n8n) арқылы автоматтандыру жүйелер арасындағы алшақтықты жоюға мүмкіндік береді. Бұл және басқа әдістер мен құралдар кітаптың келесі бөлімдерінде егжей-тегжейлі қарастырылады.

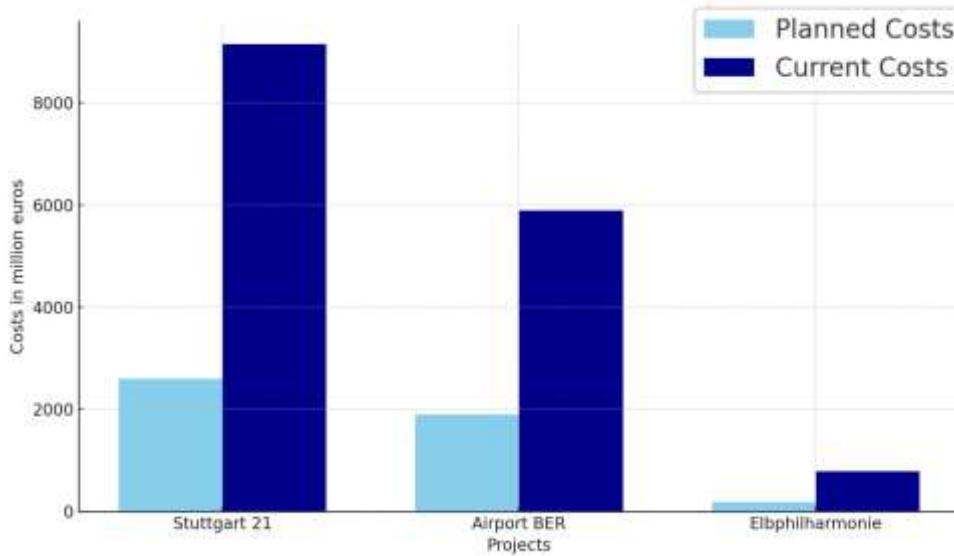


Рис. 2.13 Германиядағы ірі инфрақұрылымдық жобалар бойынша жоспарланған және нақты шығындарды салыстыру.

Корпоративтік жүйелермен де дәл осындай жағдай орын алады: алдымен оқшауланған шешімдер жасалады, содан кейін оларды интеграциялау және келісін үшін үлкен бюджеттер жұмсалады. Егер бастапқыда деректер модельдері мен байланыстарын ойластырган болса, интеграция қажеттілігі мүлдем туындаламас еді. Оқшауланған деректер цифрлық әлемде хаос тудырады, құрылыс процесінің үйлесімсіздігіне үқсас.

KPMG-нің 2023 жылғы «Cue construction 4.0: Уақыт жасау немесе бұзу» зерттеуіне сәйкес, тек 36% компаниялар бөлімдер арасында деректерді тиімді алмасады, ал 61% оқшауланған «силос» деректерімен байланысты елеулі проблемаларға тап болады [30].

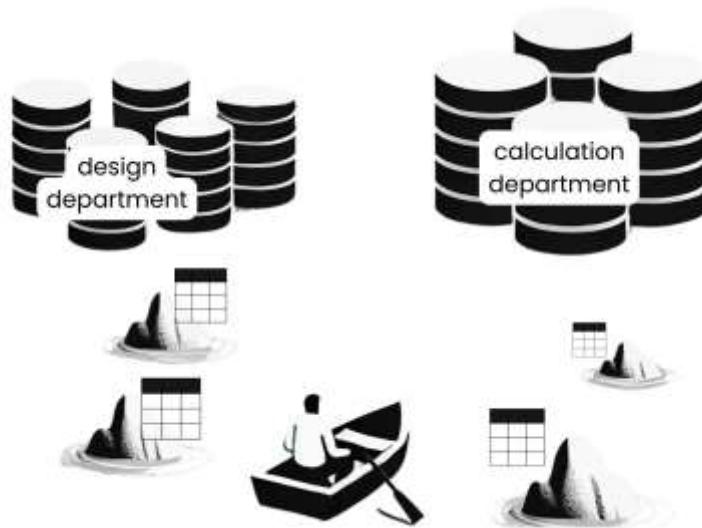


Рис. 2.14 Жылдар бойы жинақталған, бірақ қыындықпен алынатын деректер оқшауланған «силос» қоймаларында жиналышп, ешқашан пайдаланылмау қаупіне үшірайды.

Компанияның деректері оқшауланған жүйелерде, ландшафтта шашыраған жеке ағаштар сияқты сақталады. Әрқайсысы құнды ақпаратты қамтиды, бірақ олардың арасында байланыс болмауы біртұтас, өзара байланысты экожүйені құруға кедергі келтіреді. Мұндай алшақтық деректердің ағынына кедергі келтіреді және үйлемнің толық көріністі көру қабілетін шектейді. Бұл синостарды байланыстыру — менеджмент деңгейінде ақпараттың жеке бөліктерін жүйелер арасында беру үшін мицелий өсіру сияқты ұзақ және күрделі процесс.

WEF-тің 2016 жылғы зерттеуіне сәйкес, цифрлық трансформацияның негізгі кедергілерінің бірі деректердің біртұтас стандарттарының болмауы және фрагментация болып табылады.

Құрылым саласы әлемдегі ең фрагменттелген салалардың бірі болып табылады және құн тізбегінің барлық қатысушыларының үйлесімді өзара әрекеттесуіне тәуелді [5].

— Дүниежүзілік экономикалық форум 2016: Құрылым болашағын қалыптастыру.

Дизайнерлер, менеджерлер, координаторлар және әзірлеушілер көбінесе автономды жұмыс істеуді қалайды, координацияның қындықтарынан қашады. Бұл табиғи үмттылыс ақпараттық "синостардың" пайда болуына әкеледі, мұнда деректер жеке жүйелер ішінде оқшауланады. Мұндай оқшауланған жүйелердің саны артқан сайын, олардың өзара әрекеттесуін орнату қындейтін түседі. Уақыт өте келе, әр жүйе өз деректер базасына және менеджерлерден тұратын арнайы қолдау бөлімшесіне ие болады, бұл интеграцияны одан әрі қынданады (Рис. 1.24).

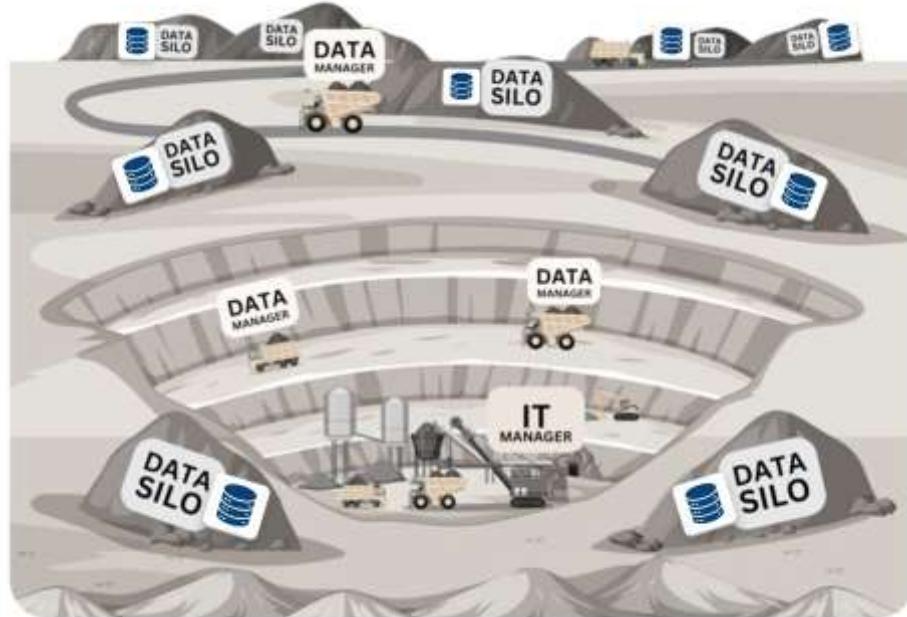


Рис. 2.15 Әр жүйе өзіне тән деректер силюстрирует, что они находятся в изолированных системах [31].

Корпоративтік жүйелердегі тұйық шеңбер былай көрінеді: компаниялар күрделі оқшауланған шешімдерге инвестиция салады, содан кейін олардың интеграциясына жоғары шығындармен

бетпе-бет келеді, ал әзірлеушілер жүйелерді біріктірудің қыындығын түсініп, өздерінің тұйық экожүйелерінде жұмыс істеуді қалайды. Бұл IT-ландшафттың фрагментациясын арттырып, жаңа шешімдерге көшу процесін қыындаатады (Рис. 2.15). Менеджерлер, ақырында, деректердің разрозналығын сынға алады, бірақ оның себептерін және алдын алу жолдарын сирек талдайды. Басшылар ескірген IT-жүйелеріне шағымданады, бірақ олардың ауыстырылуы елеулі инвестицияларды талап етеді және сирек қутілетін нәтижені береді. Нәтижесінде, тіпті осы мәселені шешуге бағытталған әрекеттер көбінесе жағдайды одан әрі нашарлатады.

Разрозналықтың негізгі себебі – қосымшалардың деректерден басым болуы. Компаниялар алдымен жеке жүйелерді әзірлейді немесе вендорлардан дайын шешімдерді сатып алады, содан кейін оларды біріктіруге тырысады, бұл қайталанатын және бір-бірімен үйлеспейтін сақтау орындары мен деректер базаларын қалыптастырады.

Фрагментация мәселесін шешу түбекейлі жаңа тәсілді талап етеді – деректердің қосымшалардан басымдылығы. Компаниялар алдымен деректерді басқару стратегияларын және деректер модельдерін әзірлеуі керек, содан кейін біртұтас ақпарат жиынтығымен жұмыс істейтін жүйелерді құруы немесе шешімдерді сатып алуы тиіс, жаңа кедергілерді қалыптастырмай.

Біз деректер бағдарламалық қамтамасыз етуден маңыздырақ болуы мүмкін жаңа әлемге қадам басудамыз. – Тим О'Reilly, O'Reilly Media, Inc. бас директоры.

McKinsey Global Institute-тің «Құрылыш саласын қайта ойлау: өнімділікті арттыру жолы» (2016) зерттеуі құрылыш саласының цифрлық трансформацияда басқа секторлардан артта қалғанын көрсетеді [32]. Есепке сәйкес, автоматтандырылған деректерді басқару мен цифрлық платформаларды енгізу өнімділікті айтартықтай арттырып, процестердің үйлесімсіздігімен байланысты шығындарды азайтуы мүмкін. Цифрлық трансформация қажеттілігін Иганның (Ұлыбритания, 1998) есебі де атап етеді [33], ол құрылыш саласындағы интеграцияланған процестер мен бірлескен тәсілдің маңызды рөлін көрсетеді.

Нәтижесінде, егер соңғы 10 000 жылда деректер менеджерлерінің негізгі проблемасы деректердің жетіспеушілігі болса, деректер мен деректерді басқару жүйелерінің лавинообразды өсуімен пайдаланушылар мен менеджерлер зандау және сапалы ақпаратты іздеуде қыындықтарға тап болды - деректердің артықшылығы.

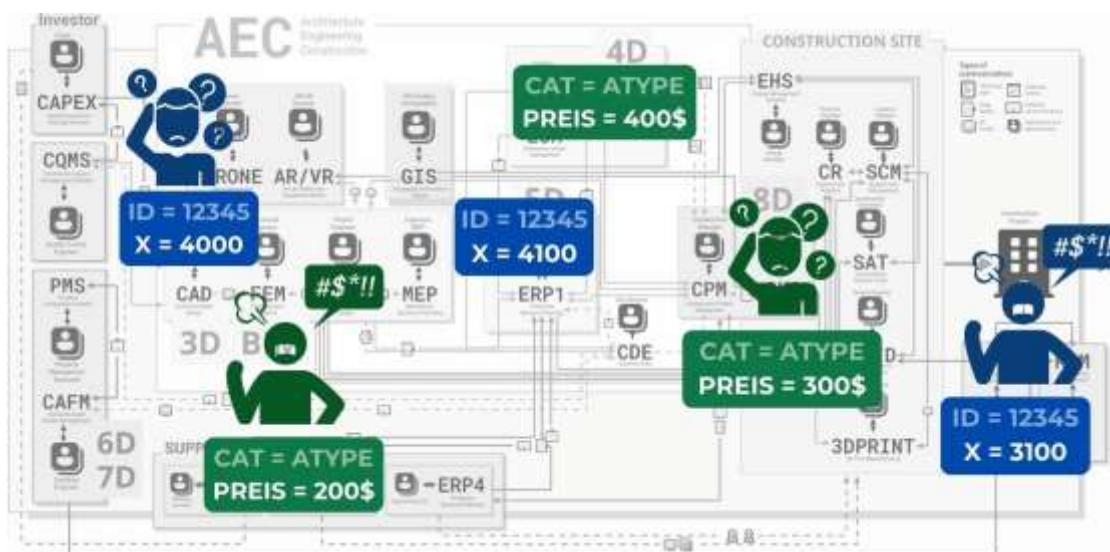
Деректер силостарының разрозналығы сөзсіз деректер сапасының төмендеуіне әкеледі. Көптеген тәуелсіз жүйелер болған кезде, бірдей деректер әртүрлі нұсқаларда, жиі қарама-қайшы мәндермен болуы мүмкін, бұл пайдаланушыларға қандай ақпараттың өзекті және сенімді екенін анықтауды қыындаатады.

Дубликация және деректер сапасының болмауы: бөлінудің салдары

Деректер сипаттық мәселесі салдарынан менеджерлер деректерді іздеуге және тексеруге едәүір уақыт жұмысайға мәжбүр. Сапа мәселелерінен сақтану үшін компаниялар ақпаратты басқарудың күрделі құрылымдарын құрады, онда менеджерлер вертикалінің міндегі деректерді іздеу, верификациялау және келісу болып табылады. Алайда, бұл тәсіл тек бюрократияны арттырып, шешім қабылдау процесін баяулатады. Деректер саны артқан сайын, оларды талдау және интерпретациялау қындағы түседі, әсіресе сақтау және өндеге стандарттары болмаған жағдайда.

Соңғы онжылдықта көптеген бағдарламалық қосымшалар мен жүйелердің пайда болуымен, деректер сипаттық және сапасыз деректер мәселесі соңғы пайдаланушылар үшін барған сайын өзекті болуда. Бірдей деректер, бірақ әртүрлі мәндермен, енді әртүрлі жүйелер мен қосымшаларда кездеседі. Бұл соңғы пайдаланушылар үшін көптеген қолжетімді нұсқалардың арасында қайсысы өзекті және дұрыс деректер екенін анықтауда қындықтарға әкеледі. Нәтижесінде, талдау кезінде қателіктер туындалап, ақырында шешім қабылдауда қындықтар пайда болады.-

Қажетті деректерді іздеуден сақтану үшін компания басшылары менеджерлерден тұратын көденгейлі бюрократия құрады. Олардың міндегі – деректерді жылдам табу, тексеру және қажетті деректерді кестелер мен есептер түрінде жіберу, әртүрлі жүйелердің лабиринтінде бағдарлану.

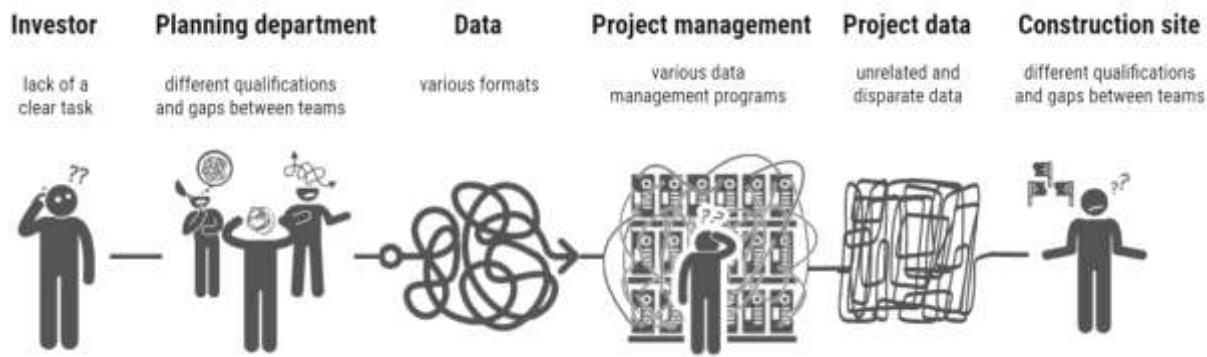


Деректерді іздеу кезінде менеджерлер әртүрлі жүйелер арасындағы деректердің сапасы мен заңдылығын қамтамасыз етуі керек.

Алайда, тәжірибеде бұл модель жаңа қындықтарды тудырады. Деректерді басқару қолмен жүзеге асырылғанда, ал ақпарат көптеген байланыссыз шешімдер арасында шашырап кеткенде, дәл және өзекті ақпаратты алу әрекеті жауапты тұлғалар пирамидасы арқылы (Рис. 2.17) уақыттың көп алатын және қателіктерге ұшырайтын тар орынға айналады.

Цифрлық шешімдердің лавинообразды өсуі жағдайды одан әрі қынданады. Бағдарламалық қамтамасыз ету нарығы жаңа, уәде беретін құралдармен толығуды жалғастыруды. Алайда, деректерді басқару бойынша нақты стратегиясыз, бұл шешімдер біртұтас жүйеге интеграцияланбайды, керісінше, қосымша күрделілік пен дублирлеу қабаттарын тудырады. Нәтижесінде, процестердің жеңілдетудің орнына, компаниялар одан әрі фрагменттелген және хаостық ақпараттық ортада

қалады.



Жүйелердің күрделілігі мен деректер форматтарының әртүрлілігі құрылыш процесінде келісімділіктің жоғалуына әкеледі.

Көптеген байланыссыз шешімдерді басқарумен байланысты барлық аталған мәселелер, ерте ме, кеш пе, компания басшылығын маңызды бір түсінікке әкеледі: мәселе деректер көлемінде немесе оларды өңдеудің «универсалды» құралын іздеуде емес. Негізгі себеп деректердің сапасында және үйымның оларды қалай жасайтынында, алатынында, сақтайтынында және пайдаланатынында жатыр.

Тұрақты табыстың кілті – жаңа «сиқырлы» қосымшаларды іздеуде емес, компания ішінде деректермен жұмыс істеу мәдениетін қалыптастыруды. Бұл деректерді стратегиялық актив ретінде қарастыруды білдіреді, ал олардың сапасы, тұтастығы және өзектілігі мәселелері үйымның барлық деңгейлерінде басымдыққа ие болады.

«Сапа мен санның» дилеммасын шешу біртұтас деректер құрылымын құруда жатыр, ол қайталауды жояды, қарама-қайшылықтарды жояды және ақпараттық ағындарды біріктіреді. Мұндай архитектура негізінде негізделген, сенімді деректер көзі қалыптастыруға мүмкіндік береді, соның негізінде негізделген, дәл және уақытылы шешімдер қабылданады.

Әйтпесе, қазіргі уақытта жиі кездесетіндей, компаниялар сенімді фактілердің орнына HiPPO сарапшыларының субъективті пікірлері мен интуитивті бағалауларына сүйенуді жалғастыруды. Құрылыш саласында, мұнда дәстүрлі түрде сараптамалық тәжірибе маңызды рөл атқарады, бұл әсіресе айқын көрінеді.

HiPPO немесе шешім қабылдаудағы пікірлердің қауіптілігі

Дәстүр бойынша құрылыш саласында негізгі шешімдер тәжірибе мен субъективті бағалаулар негізінде қабылданады. Уақытылы және сенімді деректер болмаған жағдайда, компания басшылары ең жоғары жалақы алатын қызметкерлердің интуициясына сүйенуге мәжбүр, объективті фактілердің орнына (Рис. 2.18).-

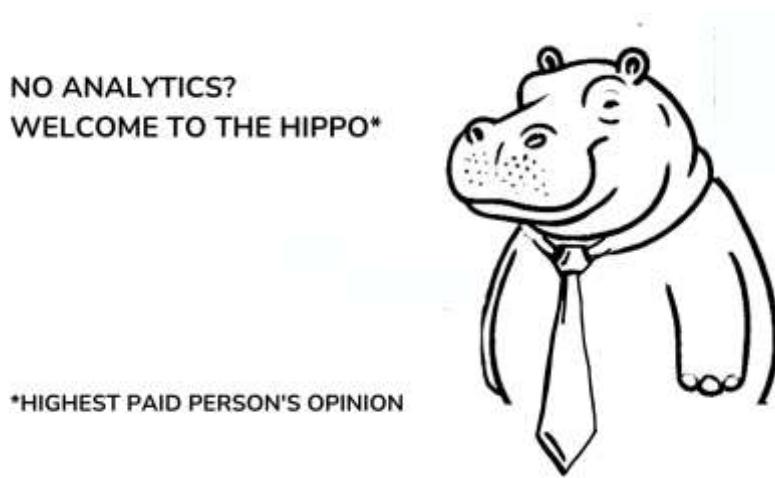


Рис. 2.18 Аналитика болмаған жағдайда бизнес тәжірибелі мамандардың субъективті пікіріне тәуелді.

Мұндай тәсіл, мүмкін, тұрақтылық пен баяу өзгерістер жағдайында ақталған, бірақ цифрлық трансформация дәүірінде бұл елеулі қауіпке айналады. Интуиция мен болжамдарға негізделген шешімдер бүрмалануға ұшырайды, көбінесе расталмаған гипотезаларға негізделеді және деректерде көрсетілген кешенді көріністі ескермейді.

Компанияда шешім қабылдау деңгейінде ақылға қонымын пікірталастар ретінде ұсынылатын нәрсе көбінесе нақты нәрсеге негізделмейді. Компанияның табысы сарапшылардың беделіне және жалақы деңгейіне емес, деректермен тиімді жұмыс істей білу, заңдылықтарды анықтау және салмақты шешімдер қабылдау қабілетіне байланысты болуы тиіс.

Бедел немесе тәжірибе автоматты түрде шешімнің дұрыстығын білдіретін концепциядан бас тарту маңызды. Деректерге негізделген тәсіл ойын ережелерін өзгертеді: енді шешім қабылдаудың негізі деректер мен аналитика, ал лауазым мен жалақы емес. Үлкен деректер, машиналық оқыту және визуалды аналитика заңдылықтарды анықтауға және болжамдарға емес, фактілерге сүйенуге мүмкіндік береді (Рис. 1.14).-

Деректерсіз сіз жай ғана пікір білдіруші адамсыз. – У. Эдвардс Деминг, ғалым және менеджмент бойынша консультант

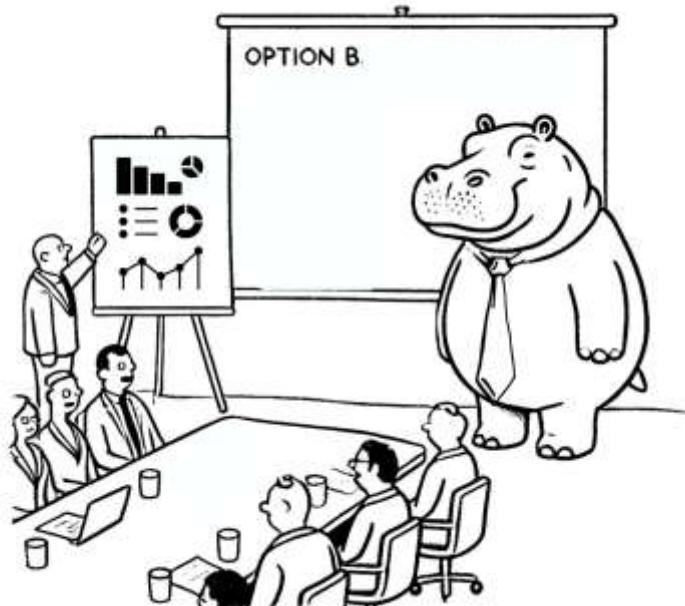
Заманауи деректерді басқару әдістері компаниядағы білімнің сабактастығын қамтамасыз етеді. Айқын сипатталған процестер, автоматтандыру және жүйелі тәсіл кілтті рөлдерді тиімділікті жоғалтпай беруге мүмкіндік береді.

Алайда, деректерге соқыр сенім де елеулі қателіктерге әкелуі мүмкін. Деректер өздігінен тек сандар жиынтығы. Дұрыс талдау, контекст және заңдылықтарды анықтау қабілеті болмаса, олар құнсыз және процестерді басқаруға қабілетсіз. Табыс кілті Hippo интуициясы мен аналитика арасында таңдау жасауда емес, шашыраңқы ақпаратты басқарылатын, негізделген шешімдерге айналдыратын интеллектуалды құралдарды құруда.

Цифрлық құрылымың жағдайында табыстың шешуші факторлары еңбек өтілі мен иерархиядағы орны емес, реакция жылдамдығы, шешімдердің дәлдігі және ресурстарды тиімді пайдалану болып табылады.

Деректер – бұл құрал, абсолютті шындық емес. Олар адам ойлаудың толықтыруы тиіс, алмастыруды емес. Аналитиканың барлық артықшылықтарына қарамастан, деректер адам интуициясы мен тәжірибесін толықтай алмастыра алмайды. Олардың рөлі – дәл және саналы шешімдер қабылдауға көмектесу.

Бәсекелестік артықшылық стандарттарға сәйкес болумен ғана емес, барлық үшін бірдей ресурстарды тиімді пайдалануда бәсекелестерден асып тұсу қабілетімен қамтамасыз етіледі. Болашақта деректермен жұмыс істеу дағдысы, бұрынғыдан, сауаттылық немесе математикамен танысумен бірдей маңызды болады. Деректерді талдай және интерпретациялай алатын мамандар, тек жеке тәжірибесіне сүйенетіндерден гөрі, дәл шешімдер қабылдай алады.



Шешімдер объективті талдауға негізделуі тиіс, ең жоғары жалақы алатын қызыметкердің пікірімен емес.

Менеджерлер, мамандар және инженерлер деректер аналитиктері ретінде әрекет етеді, жобалардың құрылымын, динамикасын және негізгі көрсеткіштерін зерттейді. Адам ресурстары деректер негізінде максималды тиімділікке қол жеткізу үшін икемді түрде реттеуді талап ететін жүйенің элементтері болады.

Дұрыс емес деректерді пайдалану кезінде қателіктер, деректердің жоқтығына қарағанда, әлдеқайда аз. Чарльз Бэббидж, алғашқы аналитикалық есептеу машинасының ойлап табушысы.

Улken деректердің пайда болуы және LLM (Улken Тіл Модельдері) енгізілуі тек талдау тәсілдерін ғана емес, шешім қабылдау табиғатын да түбекейлі өзгерту. Бұрын назар аударылған себептілік (неге бір нәрсе болды – диагностикалық аналитика) болса, бүгінгі таңда болашақты болжай қабілеті (прогноздық аналитика) алдыңғы қатарға шығады, ал болашақта – прескриптивтік аналитика, мұнда машиналық оқыту мен ИИ шешім қабылдау процесінде ең жақсы таңдау жасауға көмектеседі.-

SAP™ жаңа зерттеуі бойынша «Жаңа зерттеу көрсеткендей, басшылардың жартысынан көбі жасанды интеллектке өздерінен гөрі көбірек сенеді» 2025 жылы, 44% жоғары басшылар ИИ ұсыныстарына негізделген бұрын қабылданған шешімдерін өзгертуге дайын, ал 38% ИИ-ге өз атынан бизнес шешімдер қабылдауға сенер еді. Сонымен қатар, 74% басшылар ИИ кеңестеріне достары мен отбасыларына қарағанда көбірек сенетіндерін айтты, ал 55% жылдық табысы 5 миллиард доллардан асатын ұйымдарда ИИ арқылы алынған инсайттар дәстүрлі шешім қабылдау әдістерін алмастыратынын немесе жиі айналып өтетінін атап өтті. Бұдан басқа, респонденттердің 48% генеративті ИИ құралдарын күн сайын пайдаланады, олардың 15% – бірнеше рет күн сайын.

LLM және автоматтандырылған деректер басқару жүйелерінің дамуы жаңа мәселені туындалады: ақпаратты тиімді пайдалану, оның құндылығын үйлесімсіз форматтар мен әртүрлі көздер хаосында жоғалтпай. Бұл бизнес процестерінің өсіп келе жатқан күрделілігі мен динамикасымен тоғызырылады.

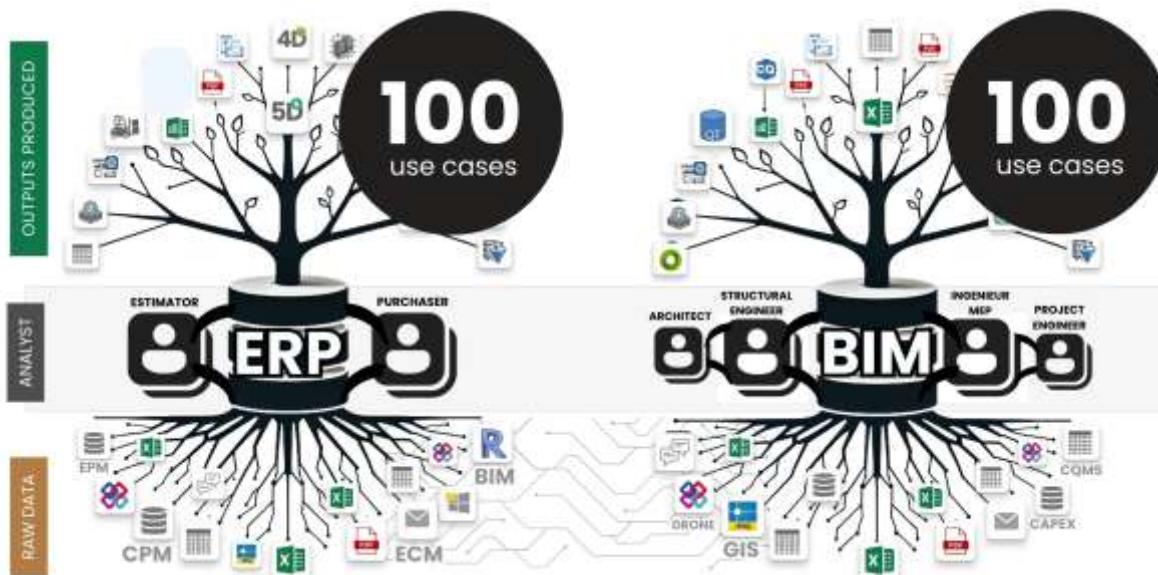
Бизнес-процестердің күрделілігі мен динамикасының тұрақты артуы

Құрылыс саласы бүгінгі күні деректер мен процестерді басқаруда маңызды қыындықтарға тап болуда. Негізгі қыындықтар – ақпараттық жүйелердің бөлшектенуі, артық бюрократия және цифрлық құралдар арасында интеграцияның болмауы. Бұл мәселелер бизнес-процестердің технологиялардың, клиенттердің өзгермелі талаптарының және жаңартылған нормативтердің әсерінен барған сайын күрделене түсуімен күштейе түсude.

Құрылыс жобаларының бірегейлігі тек техникалық ерекшеліктерімен ғана емес, әртүрлі елдердің ұлттық стандарттары мен реттеуші талаптарындағы айырмашылықтармен де анықталады (Сур. 4.210, Сур. 5.17). Бұл әр жобаға икемді, жекелендірілген тәсілді талап етеді, ал бұл дәстүрлі модульдік басқару жүйелері шеңберінде жүзеге асыру қын. Процестердің күрделілігі мен деректердің үлкен көлемі себепті көптеген компаниялар мамандандырылған шешімдер ұсынатын вендорларға жүгінеді. Алайда нарық артық жүктелген – көптеген стартаптар ұқсас өнімдерді ұсынып, тар міндеттерге назар аударады. Нәтижесінде, деректерді басқарудағы кешенді тәсіл жиі жоғалып кетеді.-

Үздіксіз жаңа технологиялар мен нарық талаптарына бейімделу бәсекеге қабілеттіліктің критикалық факторы болып табылады. Алайда, барша патентtelgen қосымшалар мен модульдік жүйелердің бейімделу қабілеті төмен – кез келген өзгерістер жиңізу үшін жаңа қымбат қайта өндіріледі тарап етеді, ал әзірлеушілер құрылыш процестерінің ерекшеліктерін әрдайым түсінбейді.

Компаниялар технологиялық артта қалушылықтың тұтқынына айналып, жаңа жаңартуларды құтудің орнына инновациялық интеграцияланған тәсілдерді жедел енгізуге дайын емес. Нәтижесінде құрылыш үйымдарының ішкі құрылымы көбінесе өзара байланысты иерархиялық және жиңіз жабық жүйелердің курделі экокожүесін құрайды, олардың арасындағы координация көп деңгейлі менеджерлер желісі арқылы жүзеге асырылады.



Сур. 2.110 Компаниялар өзара байланысты жүйелерден тұрады, олардың бірігүі автоматтандыруды талап ететін процестерді қалыптастырады.

Канададағы 2021 жылды Канадалық құрылыш асоциациясы мен KPMG компаниясы жүргізген зерттеуге сәйкес, тек 25% компаниялар бәсекелестеріне қарағанда технологияларды немесе цифрлық шешімдерді енгізу бойынша елеулі немесе тамаша жағдайда екендіктерін есептейді. Тек 23% респонденттер өз шешімдерінің елеулі немесе үлкен дәрежеде деректерге негізделгенін хабарлады. Сонымен қатар, сауалнамаға қатысушылардың басым бөлігі басқа технологияларды қолдануды тек эксперименттік деп сипаттады немесе оларды мұлдем қолданбайтындарын мойындағы.

Бұл технологиялық эксперименттерге қатысуға деген ықылассызық әсіресе ірі инфрақұрылымдық жобаларда айқын көрінеді, мұнда қателіктер миллиондаған доллар тұруы мүмкін. Тіпті ең озық технологиялар – цифрлық егіздер, предиктивті аналитика – жиңіз қарсылықта тап болады, себебі олардың тиімділігіне емес, нақты жобалардағы дәлелденген сенімділіктің болмауына байланысты.

Дүниежүзілік экономикалық форумның (WEF) "Құрылым болашағын қалыптастыру" есебіне сәйкес, құрылым саласындағы жаңа технологияларды енгізу тек техникалық қыындықтарға ғана емес, тапсырыс берушілер тарапынан психологиялық кедергілерге де тап болады. Көптеген клиенттер заманауи шешімдерді қолданудың өз жобаларын эксперименттік алаңға айналдырып, оларды "тәжірибелік қояндарға" айналдыруынан қорқады, ал болжанбаған салдар қосымша шығындар мен тәуекелдерге әкелуі мүмкін.

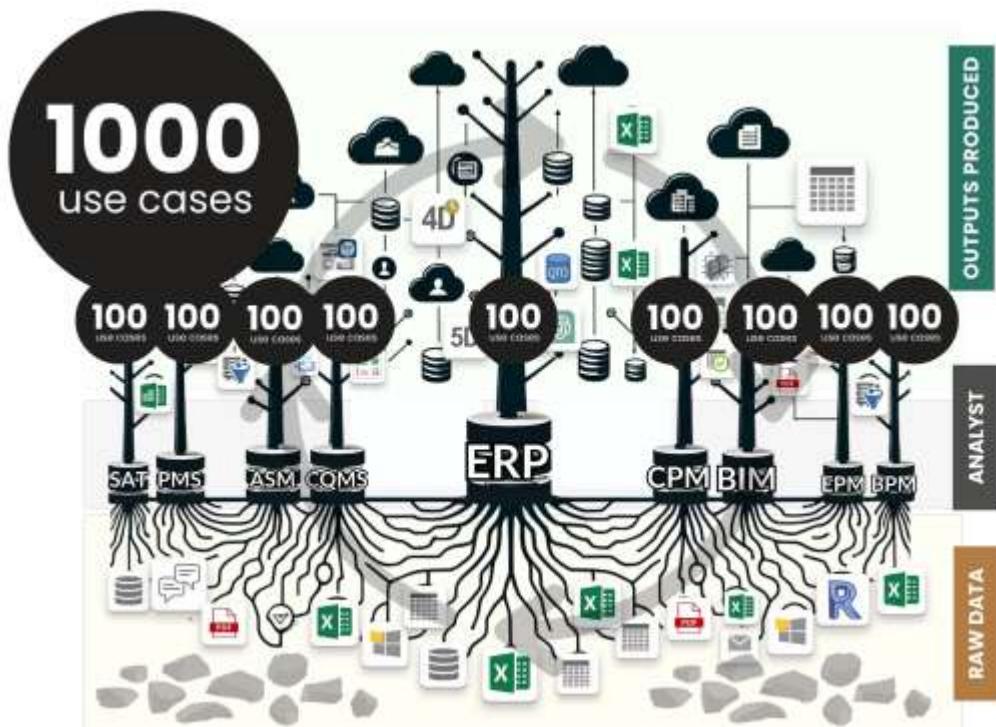


Рис. 2.111 Әрбір қолдану жағдайы үшін шешімдер нарығы процестерді оңтайландыру және автоматтандыру үшін қосымшаларды ұсынады.

Құрылым саласы өте әртүрлі: әртүрлі жобалардың әртүрлі талаптары, аймақтық ерекшеліктер, заңнамалық нормалар (Рис. 4.210), калькуляция стандарттары (Рис. 5.17) және т. б. Сондықтан барлық осы талаптар мен жобалардың ерекшеліктеріне толық сәйкес келетін патентtelген әмбебап қосымша немесе жүйе жасау іс жүзінде мүмкін емес.--

Жүйелердің күрделілігі мен бағдарламалық қамтамасыз етушілерге тәуелділікті жену үшін, деректерді тиімді басқарудың кілті тек ашықтық пен стандарттау ғана емес, сонымен қатар процестердің архитектурасын женілдету екендігіне барған сайын көбірек назар аударылады. Бизнес-процесстердің есіп келе жатқан күрделілігі мен динамикасы жаңа тәсілдерді талап етеді, мұнда басымдық деректерді жинаудан олардың құрылымдалуы мен тәртіпке келтірілуіне ауысады. Дәл осы өзгеріс құрылым саласының дамуының келесі кезеңі болады, бағдарламалық қамтамасыз етушілердің үстемдік дәүірінің аяқталуын және ақпаратты мағыналы үйымдастыру дәүірінің басталуын білдіреді.

Әмбебап шешімдердің шектеулі екендігін және күрделіліктің өсуіне осалдықты түсіну базымдықтардың өзгеруіне әкеледі: жабық платформалар мен деректерді жинаудан – ашықтыққа, бейімделгіштікке және ақпаратпен құрылымдық жұмысқа. Бұл ойлау өзгерісі жаһандық экономика мен технологиядағы кең ауқымды өзгерістерді көрсетеді, олар "өнеркәсіптік революциялар" призмасы арқылы сипатталады. Құрылыш саласының қайда бағыт алып жатқанын және оның болашақ бағдарларын түсіну үшін, саланың Төртінші және Бесінші өнеркәсіптік революциялардың контекстінде орнын қарастыру қажет – автоматтандыру мен цифрандырудан деректердің персонализациясы, ашық стандарттар мен қызмет көрсету моделіне дейін.

Төртінші өнеркәсіптік революция (Индустрия 4.0) және бесінші өнеркәсіптік революция (Индустрия 5.0) құрылыш саласында

Технологиялық және экономикалық құрылымдар – қоғам мен экономиканың даму кезеңдеріндегі эволюциясын сипаттау және талдау үшін қолданылатын теориялық концепциялар. Сонымен қатар, әртүрлі зерттеушілер мен сарапшылар оларды әртүрлі түрде интерпретациялауы мүмкін.

- Төртінші өнеркәсіптік революция (4IR немесе Industry 4.0) ақпараттық технологиялармен, автоматтандырумен, цифрандырумен және глобализациямен байланысты. Оның негізгі элементтерінің бірі – патентtelgen бағдарламалық шешімдерді жасау, яғни нақты міндеттер мен компанияларға арналған мамандандырылған цифрлық өнімдерді әзірлеу. Бұл шешімдер жиі ИТ-инфрақұрылымының маңызды бөлігіне айналады, бірақ қосымша модификацияларды қажет етпей, нашар масштабталады.
- Бесінші өнеркәсіптік революция (5IR) бүгінгі таңда 4IR-дан концептуализация және даму кезеңінде ертерек түр. Оның негізгі принциптері өнімдер мен қызметтердің жеке дараланын арттыруды қамтиды. 5IR – бұл экономикалық қызметтің бейімделгіш, икемді және жеке дараланған түріне көшу, жеке даралануға, консалтингке және қызметке бағытталған модельдерге акцент жасайды. Бесінші экономикалық жүйенің негізгі аспектің шешім қабылдау үшін деректерді пайдалануды қамтиды, бұл ашық деректер мен ашық құралдарды қолданбай жүзеге асырылуы мүмкін емес (Рис. 2.112).-

**FOURTH INDUSTRIAL REVOLUTION
(4IR OR INDUSTRY 4.0)**



**FIFTH INDUSTRIAL REVOLUTION
(5IR)**



Рис. 2.112 Төртінші жүйе шешімдерге, ал бесінші – жеке даралануға және деректерге назар аударады.

Құрылыс саласындағы компаниялар үшін он немесе жүз ұйымға арналған қосымша жасау, оны басқа компанияларда, аймақтарда немесе елдерде сәтті масштабтауға кепілдік бермейді, елеулі модификациялар мен жетілдірулерсіз. Мұндай шешімдердің сәтті кеңею ықтималдығы төмен, себебі әр ұйымның ерекше процестері, талаптары мен жағдайлары бар, олар жеке дараланған бейімдеулерді талап етуі мүмкін.

Бұгінгі таңда технологиялық шешімдердің сәтті интеграциясы әр процесс, жоба және компанияяға терең жеке дараланған тәсілді талап ететінін түсіну маңызды. Бұл, тіпті әмбебап құрылым, құрал немесе бағдарлама әзірленгеннен кейін, әрбір нақты компания мен нақты жоба үшін ерекше талаптар мен жағдайларға сәйкес оны егжей-тегжейлі бейімдеу мен баптауды қажет ететінін білдіреді.

PwC-дің «Бесінші өнеркәсіптік революцияны декодтау» [38] есебіне сәйкес, биыл әртүрлі салалардағы жоғары басшылықтың шамамен 50%-ы озық технологиялар мен адам тәжірибесін интеграциялауға ставка жасауда. Мұндай тәсіл өнім дизайны немесе тапсырыс берушінің талаптарына жедел бейімделуге мүмкіндік береді, жеке дараланған өндірісті қалыптастырады.

Әр процесс үшін ерекше функция немесе қосымша әзірлеу қажет, бұл әлемдік құрылыс саласының ауқымын және жобалардың әртүрлілігін ескере отырып, әрдайым ерекше Pipeline логикасын білдіретін бизнес-кейстердің үлкен санын тудырады (Рис. 2.113). Әрбір осындай кейстің өз ерекшеліктері бар және жеке тәсілді талап етеді. Біз машиналық оқыту мен «Титаник» деректер жиынтығын талдау тақырыбында әртүрлі тәсілдердің контекстінде бірдей аналитикалық тапсырманың әртүрлі шешімдерін егжей-тегжейлі қарастырамыз (Рис. 9.29).--

Цифрлық процестер контекстінде Pipeline – бұл жобаның өмірлік циклі кезеңдерінде автоматтандырылған немесе құрылымдық деректер мен жұмыстардың ағынын қамтамасыз ететін әрекеттер, процестер мен құралдардың тізбегі.

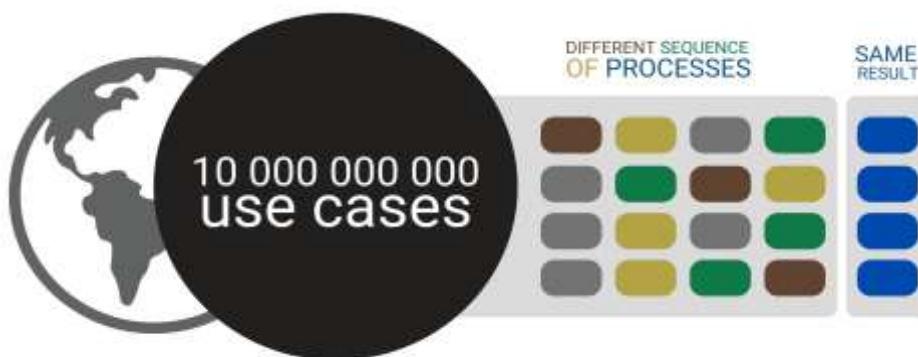


Рис. 2.113 Бизнес-кейстердің даралығы мен вариативлігі масштабталатын жабық платформалар мен құралдарды жасауды мүмкін емес етеді.

Біздің өміріміз цифрлық трансформацияның әсерінен көп өзгерді, және бұгінгі таңда құрылыс саласының экономикалық дамуында жаңа кезеңнің басталғаны туралы айтуға болады. Осы «жаңа экономикада» бәсекелестік басқа ережелер бойынша ұйымдастырылады: қоғамдық білім мен

ашық деректердің сұранысқа ие өнімдер мен қызметтерге тиімді айналдыра алатындар бесінші индустриалды революция жағдайында негізгі артықшылыққа ие болады.

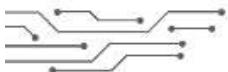
Экономист Кейт Маскус 2012 жылы «Жеке құқықтар мен қоғамдық мәселелер: XXI ғасырдағы зияткерлік меншік глобалдық экономикасы» атты кітабында атап өткендей, біз білім экономикасында өмір сүріп жатырымыз, және болашақ ғылыми жаңалықтарды тауарға айналдыру қабілетіне ие адамдарға тиесілі.

Бесінші экономикалық құрылымға көшу жабық IT-шешімдерден ашық стандарттар мен платформаларға ауысады білдіреді. Компаниялар дәстүрлі бағдарламалық өнімдерден деректер, ал проприетарлық технологиялар емес, негізгі актив ретінде қызмет көрсетуге бағытталған модельдерге көшуді бастайды.

2024 жылы Гарвард бизнес мектебінің зерттеуі ашық бастапқы кодты бағдарламалық қамтамасыз етудің (Open Source Software, OSS) экономикалық құндылығының зор екенін көрсетеді. Зерттеуге сәйкес, OSS барлық бағдарламалық кодтардың 96%-ында бар, ал кейір коммерциялық бағдарламалық қамтамасыз ету 99,9%-ға OSS-компоненттерінен тұрады. OSS болмаса, компаниялар бағдарламалық қамтамасыз етуге 3,5 есе көп қаржат жүмсайтын еді.

Құрылыс компанияларының экожүйелері, әлемдік тенденцияларға сәйкес, бесінші экономикалық парадигмаға біртіндеп өтеді, мұнда деректерге негізделген аналитикалық және консалтингтік қызметтер оқшауланған жабық шешімдерден гөрі басымдыққа ие болады.

Цифрландыру дәуірі саладағы күштер теңгерімін өзгертерді: компаниялар вендорлардың шешімдеріне тәуелділіктен деректерді тиімді пайдалану қабілетіне негізделген бәсекеге қабілеттілікті құра бастайды. Нәтижесінде құрылыс индустриясы ескірген қатты жүйелерден икемді, бейімделгіш экожүйелерге өтеді, мұнда ашық стандарттар мен үйлесімді құралдар жобаларды басқарудың негізі болады. Қолданбалы вендорлардың үстемдік дәуірінің аяқталуы жаңа жағдайларды туыннатады, мұнда құндылық жабық кодқа және арнайы коннекторларға ие болумен емес, деректерді стратегиялық артықшылыққа айналдыру қабілетімен анықталады.



ТАРАУ 2.2.

ХАОСЫНЫҢ ТӘРТІПКЕ АЙНАЛУЫ ЖӘНЕ КҮРДЕЛІЛІКТІ ТӨМЕНДЕТУ

Артық код және жабық жүйелер: өнімділікті арттырудың кедергі

Соңғы онжылдықтарда IT-сферадағы технологиялық өзгерістер негізінен бағдарламалық қамтамасыз етуші жеткізушилермен анықталды. Олар даму векторын белгілеп, компанияларға қандай технологияларды енгізу керектігін, ал қайсысын қалдыру керектігін анықтады. Жеке шешімдерден орталықтандырылған деректер базалары мен интеграцияланған жүйелерге көшу дәуірінде вендорлар лицензияланатын өнімдерді алға тартты, қолжетімділік пен масштабтауды бақылауды қамтамасыз етті. Кейінірек, бұлтты технологиялар мен Software as a Service (SaaS) модельдерінің кеңуімен, бұл бақылау жазылым моделіне трансформацияланды, пайдаланушыларды цифрлық қызметтердің тұрақты клиенттері ретінде бекітті.

Мұндай тәсіл парадокс туғызыды: жасалған бағдарламалық кодтың бұрын-соңды болмаған көлемдеріне қарамастан, шын мәнінде тек оның аз бөлігі ғана пайдаланылады. Мүмкін, кодтың саны қажеттіліктен жүздеген немесе мындаған есе көп, өйткені бірдей бизнес-процестер әртүрлі бағдарламаларда, тіпті бір компания ішінде де әртүрлі сипатталады және қайталанады. Сонымен қатар, әзірлеуге бұрыннан төленген, және бұл шығындар қайтарылмайды. Дегенмен, индустрия бұл циклды жалғастыра береді, жаңа өнімдер шығара отырып, соңғы пайдаланушы үшін минималды қосымша құнмен, көбінесе нарықтық күтілімдердің қысымымен, нақты қажеттіліктерден гөрі.

Бағдарламалық қамтамасыз етуді әзірлеу құнның бағалау жөніндегі нұсқаулыққа сәйкес, АҚШ-тың Қорғаныс сатып алулары университеті (DAU) әзірлеген, бағдарламалық қамтамасыз етуді әзірлеу құнны бірнеше факторларға, соның ішінде жүйенің күрделілігі мен таңдалған технологияларға байланысты айтарлықтай өзгеруі мүмкін. Тарихи түрғыдан алғанда, 2008 жылға арналған әзірлеу құнны шамамен \$100 код жолы (SLOC) құраса, ал қызмет көрсету құнны \$4 000-ға дейін өсуі мүмкін.

CAD-қосымшаларының тек бір компоненті – геометриялық ядро – ондаған миллион код жолдаурын қамтуы мүмкін. ERP-жүйелерінде де осындай жағдай байқалады, олардың күрделілігіне біз кітаптың бесінші бөлімінде қайта ораламыз. Алайда, тереңірек қарағанда, бұл кодтың едәуір бөлігі қосымша құнды жасамайды, тек деректерді дерекқор, API, пайдаланушы интерфейсі және жүйенің басқа кестелері арасында механикалық түрде тасымалдайды. Бизнес-логиканың маңызды екендігі туралы танымал мифке қарамастан, қазіргі код базалары ескірғен шаблондық блоктарға (легаси код) толы, олардың жалғыз мақсаты – кестелер мен компоненттер арасында деректерді беру, шешім қабылдауға немесе бизнес тиімділігін арттыруға әсер етпей.-

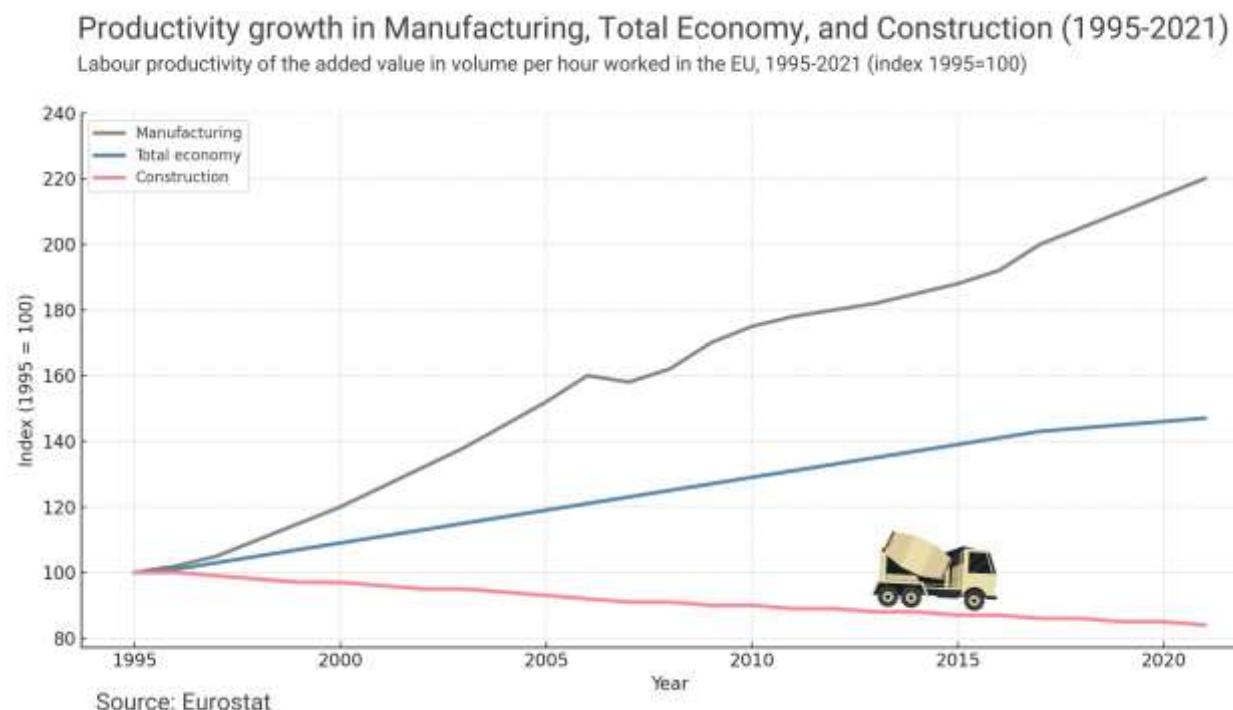
Нәтижесінде, әртүрлі дереккөздерден деректерді өңдеумен айналысатын жабық шешімдер, сөзсіз, күрделі "спагетти-екожүйелерге" айналады. Мұндай күрделі, бір-бірімен байланысқан жүйелерді тек менеджерлердің үлкен армиясы ғана жартылай қолмен жұмыс істей алады. Мұндай деректерді басқару үйімін ресурстар түрғысынан тиімді емес, сонымен қатар бизнес-процестерде критикалық осалдық нүктелерін тудырады, компанияны осы технологиялық лабиринттің қалай жұмыс істейтін түсінетін тар шеңбердегі мамандарға тәуелді етеді.

Код көлемінің, қосымшалар санының және вендорлар ұсынатын концепциялардың күрделенуі ақпараттық технологиялар экожүйесінің күрделілігінің өсуіне әкелді. Бұл салада цифрландыруды

практикалық жүзеге асыруды қосымшалардың санының артуы арқылы тиімділігін төмендедті. Пайдаланушылардың қажеттіліктеріне жеткілікті назар аударылмай әзірленген бағдарламалық өнімдер, әдетте, енгізу мен қолдауға едәүір ресурстарды талап етеді, бірақ күтілетін нәтиже бермейді.

McKinsey компаниясының «Құрылыш өнімділігін арттыру» зерттеуіне сәйкес, соңғы екі онжылдықта құрылыш саласындағы еңбек өнімділігінің жаһандық өсуі жылына орташа 1%-ды құрады, ал әлемдік экономикада өсу 2,8%-ды, өндөу өнеркәсібінде 3,6%-ды құрады. Америка Құрама Штаттарында құрылыш саласындағы еңбек өнімділігі 1960-жылдардан бері екі есе төмендеді.

Жүйелердің күрделенуі, деректердің оқшаулануы мен жабықтығы мамандар арасындағы коммуникацияны нашарлатты, бұл құрылыш саласын ең аз тиімді секторлардың бірі етті.



Жабықтық пен деректердің күрделілігі, соның салдарынан мамандар арасындағы нашар коммуникация құрылыш саласын экономиканың ең аз тиімді секторларының біріне алып келді.

McKinsey (2024) зерттеуінде «Құрылыш өнімділігін қамтамасыз ету енді міндettі емес» деп атап көрсетілгендей, ресурстардың тапшылығы мен саланың өсу қарқының екі есе арттыруға ұмтылу жағдайында құрылыш қазіргі өнімділік деңгейінде қалуға мүмкіндік бермейді. 2023 жылы жаһандық құрылыш шығындарының 13 трлн доллардан 2040 жылға қарај 22 трлн долларға дейін өсетіні болжануда, бұл тиімділік мәселесін тек өзекті емес, сонымен қатар критикалық маңызды етеді.

Эффективтілікті арттырудың негізгі тәсілдерінің бірі – деректерді өндайтін қосымшалар мен экожүйе құрылымдарын міндettі түрде унификациялау және оңайлату. Мұндай рационализация тәсілі корпоративтік жүйелерде жылдар бойы жинақталған артық абстракция қабаттарынан және

қажетсіз күрделіліктен арылуға мүмкіндік береді.

Силостардан деректердің біртұтас қоймасына

Ұйым жинақтаған деректер неғұрлым көп болса, олардан нақты пайда алу соғұрлым қыын болады. Ақпараттың фрагментациясы изоляцияланған силостарда сақталуымен қазіргі компаниялар бизнес-процестерінде мындаған әртүрлі қоймаларда сақталған материалдардан зәулім ғимарат салуға тырысатын құрылышшыларға ұқсайды. Ақпараттың артықтығы заңды маңызы бар мәліметтерге қол жеткізуіді қыындағы қана қоймай, шешім қабылдауды да баяулатады: әр қадамды бірнеше рет тексеріп, растауға тұра келеді.

Әрбір тапсырма немесе процесс жеке кесте немесе деректер базасына қатты байланған, ал жүйелер арасында деректер алмасу күрделі интеграцияларды талап етеді. Бір жүйедегі қателіктер мен сәйкесіздіктер басқаларында тізбекті ақауларды тудыруы мүмкін. Дұрыс емес мәндер, кешіктірілген жаңартулар және ақпараттың дубликациясы қызметкерлерді деректерді қолмен тексеруге және келісуге көп уақыт жұмсауға мәжбүр етеді. Нәтижесінде ұйым фрагментацияның салдарын жоюға көбірек уақыт жұмсайды, ал процестерді дамыту мен оңтайландыруға аз уақыт бөледі.

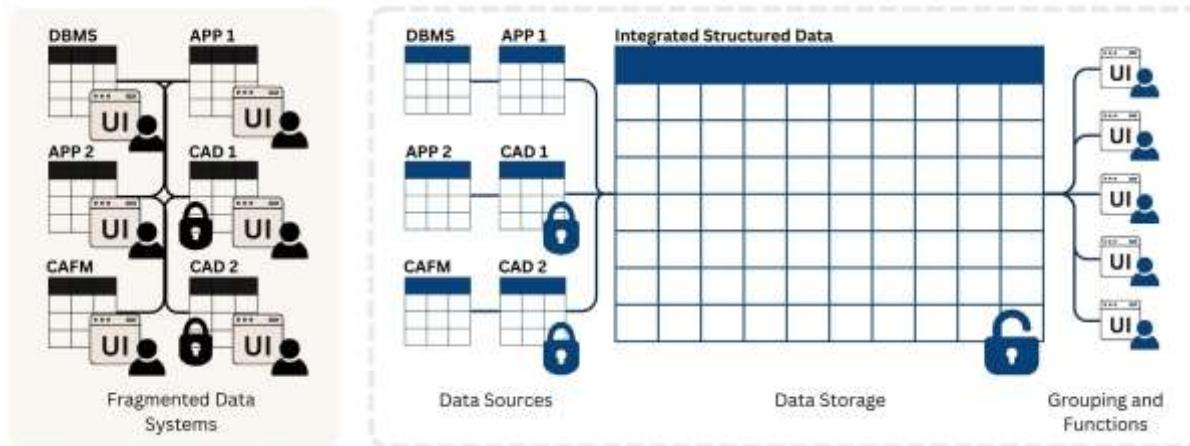
Бұл мәселе әмбебап: кейбір компаниялар хаоспен құресуді жалғастыруды, ал басқалары интеграцияда – ақпараттық ағындарды орталықтандырылған сақтау жүйесіне көшіруде шешім табуда. Мұны тапсырмалар, жобалар және объектілермен байланысты кез келген мәндерді сақтауға болатын бір үлкен кесте ретінде елестетіңіз. Онда ондаған бөлшектенген кестелер мен форматтардың орнына біртұтас байланысты сақтау пайда болады, бұл:-

- деректердің жоғалуын минимизациялауға;
- ақпаратты тұрақты келісуді жоюға;
- деректердің қолжетімділігі мен сапасын жақсартуға;
- аналитикалық өндеуді және машиналық оқытуды жеңілдетуге мүмкіндік береді.

Деректерді бір стандартқа келтіру, дереккөзге қарамастан, ақпараттың унификацияланған және машинамен оқылатын форматқа түрлендірілуін білдіреді. Мұндай деректерді ұйымдастыру олардың тұтастығын тексеруге, нақты уақытта талдауға және басқарушылық шешімдер қабылдау үшін жедел пайдалануға мүмкіндік береді.

Интеграцияланған сақтау жүйелерінің концепциясы мен олардың аналитика мен машиналық оқудағы қолданылуы туралы «Үлкен деректерді сақтау және машиналық оқу» бөлімінде толығырақ сөйлесеміз. Деректерді модельдеу және құрылымдау тақырыптары «Деректерді құрылымдық форматта түрлендіру» және «Стандарттар ойын ережелерін қалай өзгертерді: кездейсоқ файлдардан

ойластырылған деректер моделіне» бөлімдерінде егжей-тегжейлі қарастырылады.



Интеграцияланған деректер оқшауланды жояды, ақпаратқа қолжетімділікті жақсартады және бизнес-процестерді оңтайландырады.

Деректерді құрылымдау және біріктіруден кейінгі логикалық қадам олардың тексерілуі болып табылады. Біртұтас интеграцияланған қойма болған жағдайда, бұл процесс айтарлықтай жеңілдейді: енді көптеген келісілменеген схемалар, дублирленген құрылымдар мен кестелер арасындағы күрделі өзара байланыстар жоқ. Барлық ақпарат біртұтас деректер моделіне келтірілген, бұл ішкі қарама-қайшылықтарды жояды және валидация процесін жеделдетеді. Деректерді тексеру және сапасын қамтамасыз ету – барлық бизнес-процестердің негізі аспектілері, және біз оларды кітаптың тиісті тарауларында толығырақ қарастырамыз.

Соңғы кезеңде деректер топтастырылады, сүзгіден өткізіледі және талданады. Оларға әртүрлі функциялар қолданылады: агрегаттау (қосу, көбейту), кестелер, бағандар немесе жолдар арасындағы есептеулер. Деректермен жұмыс істеу кезеңдер тізбегіне айналады: жинау, құрылымдау, тексеру, түрлендіру, аналитикалық өндеу және соңғы қосыншаларға шығару, мұнда ақпарат практикалық міндеттерді шешу үшін пайдаланылады. Мұндай сценарийлерді құру, кезеңдерді автоматтандыру және өндеу ағындарын құру туралы біз ETL-процестер мен data pipeline тәсіліне арналған тарауларда сөйлесеміз.

Осылайша, цифрлық трансформация – бұл ақпаратпен жұмыс істеуді жеңілдету ғана емес. Бұл деректерді басқарудағы артық күрделіліктен бас тарту, хаостан болжамдылыққа, көптеген жүйелерден басқарылатын процесста көшу. Архитектураның күрделілігі неғұрлым төмен болса, соғұрлым қолдауға қажетті код мөлшері азаяды. Ал болашақта код мүлдем жойылып, оның орнына деректерді өздігінен талдайтын, жүйелейтін және түрлендіретін интеллектуалды агенттер пайда болуы мүмкін.

Интеграцияланған сақтау жүйелері AI агенттерін пайдалануға мүмкіндік береді

Деректер мен жүйелердің күрделілігі неғұрлым аз болса, соғұрлым код жазу және қолдау қажет

болады. Ал кодты жазуды үнемдеудің ең қарапайым тәсілі – кодтан мұлдем бас тарту, оны деректермен алмастыру. Қолданба кодының дамуы кодтан деректер модельдеріне ауысқанда, деректерге негізделген (data-driven) тәсілге қарай ауысу міндетті түрде пайда болады, өйткені бұл концепциялар мұлдем басқа ойлау тәсілін білдіреді.

Адам деректерді орталықтандырылған түрде өңдеу жолын таңдағанда, олардың рөлін басқаша көреді. Деректер тек қосымшалар үшін «шикізат» болып қалмайды – енді бұл архитектура, логика және өзара әрекеттестік негізі болып табылады.

Деректерді басқарудың дәстүрлі тәсілі әдетте қосымшалар деңгейінде басталады және құрылыш жағынан ауыр бюрократиялық жүйені еске салады: көп деңгейлі келісімдер, қолмен тексерулер, тиісті бағдарламалық өнімдер арқылы шексіз құжат нұсқалары. Цифрлық технологиялардың дамуы барысында көптеген компаниялар минимализм принципіне көшуге мәжбүр болады – тек шын мәнінде қажет және пайдаланылатын нәрсені сақтау және пайдалану.

Минимизация логикасын вендорлар қабылдады. Деректерді сақтау және өңдеу процестерін жеңілдешту үшін пайдаланушылардың жұмысы оффлайн қосымшалар мен құралдардың функционалышын бұлтты қызметтер мен SaaS шешімдеріне ауыстырылуда.

SaaS (Software as a Service, немесе "қызмет ретінде бағдарламалық қамтамасыз ету") концепциясы қазіргі IT-инфрақұрылымдарындағы негізгі бағыттардың бірі болып табылады, пайдаланушыларға интернет арқылы қосымшаларға қол жеткізуге мүмкіндік береді, бұл өз компьютерлерінде бағдарламалық қамтамасыз етуді орнату және қолдауды қажет етпейді.

Бір жағынан, SaaS масштабтауды, нұсқаларды басқаруды жеңілдетіп, қолдау мен қызмет көрсету шығындарын азайты, бірақ екінші жағынан, пайдалануышыны нақты қосымшаның логикасына ғана емес, сонымен қатар провайдердің бұлттық инфрақұрылымына толық тәуелді етті. Егер қызмет істен шықса, деректер мен бизнес-процестерге қолжетімділік уақытша немесе тіпті ұзақ мерзімге бұғатталуы мүмкін. Сонымен қатар, SaaS қосымшаларын пайдалану барысында барлық пайдаланушы деректері провайдердің серверлерінде сақталады, бұл қауіпсіздік және реттеуші талаптарға сәйкестік түрғысынан қауіптерді туындалады. Тарифтер немесе пайдалану шарттарының өзгеруі де шығындардың артуына немесе шұғыл миграция қажеттілігіне әкелуі мүмкін.

AI, LLM-агенттерінің және деректерге негізделген тәсілдің дамуы дәстүрлі қосымшалардың және SaaS орындауларының болашағына құмән келтірді. Бұрын қосымшалар мен қызметтер бизнес-логиканы басқару және деректерді өңдеу үшін қажет болса, AI-агенттердің келуімен бұл функциялар тікелей деректермен жұмыс істейтін интеллектуалды жүйелерге ауысуы мүмкін.

Дәл осы себепті IT бөлімдерінде және менеджмент деңгейінде гибридті архитектуралар, яғни AI-агенттер мен жергілікті шешімдер бұлтты қызметтерді толықтырып, SaaS платформаларына тәуелділікті азайту мәселелері жиі талқылануда.

Біз үстанатын тәсіл дәстүрлі бизнес-қосымшалар немесе SaaS-қосымшалардың агенттер дәуірінде түбегейлі өзгеруі мүмкін екенін мойындаиды. Бұл қосымшалар, негізінен, бизнес-логикасы бар CRUD [жасау, оқу, жаңарту және жою] дәрекқорлары болып табылады. Алайда, болашақта бұл логика ИИ агенттеріне аудиосатын болады.

Сатыя Наделла, Microsoft компаниясының бас директоры, 2024 ж.

Деректерге негізделген тәсіл мен ИИ/LLM агенттерін пайдалану артық процестерді қысқартуға мүмкіндік береді, демек, қызметкерлердің жұктемесін азайтады. Деректер дұрыс үйімдастырылған кезде, оларды талдау, визуализациялау және шешім қабылдау үшін қолдану оқыярақ болады. Шексіз есептер мен тексерулердің орнына мамандар бірнеше рет басу арқылы немесе LLM агенттері арқылы автоматты түрде дайын құжаттар мен дэшбордтар түрінде өзекті ақпаратқа қол жеткізеді.

Деректермен жұмыс істеуде бізге жасанды интеллект (ЖИ) құралдары мен үлкен тілдік модельдер (YTM) чаты көмектеседі. Соңғы жылдары дәстүрлі CRUD операцияларынан (жасау, оқу, жаңарту, жою) үлкен тілдік модельдерді (YTM) деректерді басқару үшін пайдалану тенденциясы байқалуда. YTM табиғи тілдерді интерпретациялап, автоматты түрде деректер базасына сәйкес сұрауларды генерациялауға қабілетті, бұл деректерді басқару жүйелерімен өзара әрекеттесуді жеңілдетеді.-

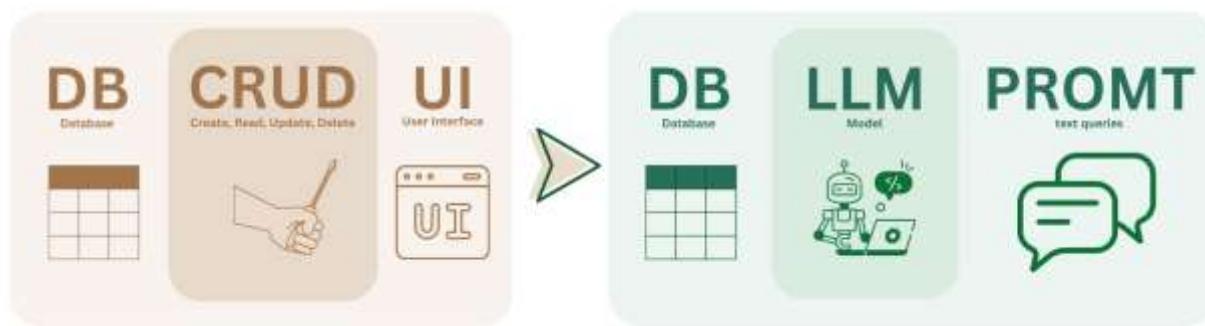


Рис. 2.23 Жасанды интеллект деректер қоймалары мен базаларымен жұмыс істеу және оларды интеграциялау бойынша шешімдерді алмастырады, дәстүрлі қосымшалар мен CRUD операцияларын біртіндеп ығыстыра отырып.

Келесі 3–6 айда ЖИ 90% кодты жазатын болады, ал 12 айдан кейін дерлік барлық код ЖИ арқылы генерациялануы мүмкін. Дарио Амодей, LLM Anthropic компаниясының бас директоры, наурыз 2025 ж.

Жасанды интеллект (AI) құралдарының жылдам дамуына қарамастан (мысалы, GitHub Copilot), 2025 жылы әзірлеушілер әлі де осы процесте маңызды рөл атқарады. AI-агенттері барған сайын пайдалы көмекшілерге айналуда: олар пайдаланушы сұрауларын автоматты түрде интерпретациялайды, SQL және Pandas сұрауларын генерациялайды (бұдан әрі – келесі тарауларда) немесе де-

ректерді талдау үшін код жазады. Осылайша, жасанды интеллект дәстүрлі қолданушы интерфейстерін біртіндеп алмастыруда.

Жасанды интеллект модельдерінің, мысалы, тілдік модельдердің таралуы гибридті архитектуралардың дамуын ынталандырады. Біз толықтай бұлттық шешімдер мен SaaS өнімдерінен бас тартудың орнына, бұлттық қызметтерді жергілікті деректерді басқару жүйелерімен интеграциялауды көре аламыз. Мысалы, федеративті оқыту (federated learning) сезімтал деректерді бұлтқа жібермей-ақ қуатты AI модельдерін пайдалануға мүмкіндік береді. Осылайша, компаниялар өз деректерін бақылауда ұстай отырып, заманауи технологияларға қол жеткізе алады.

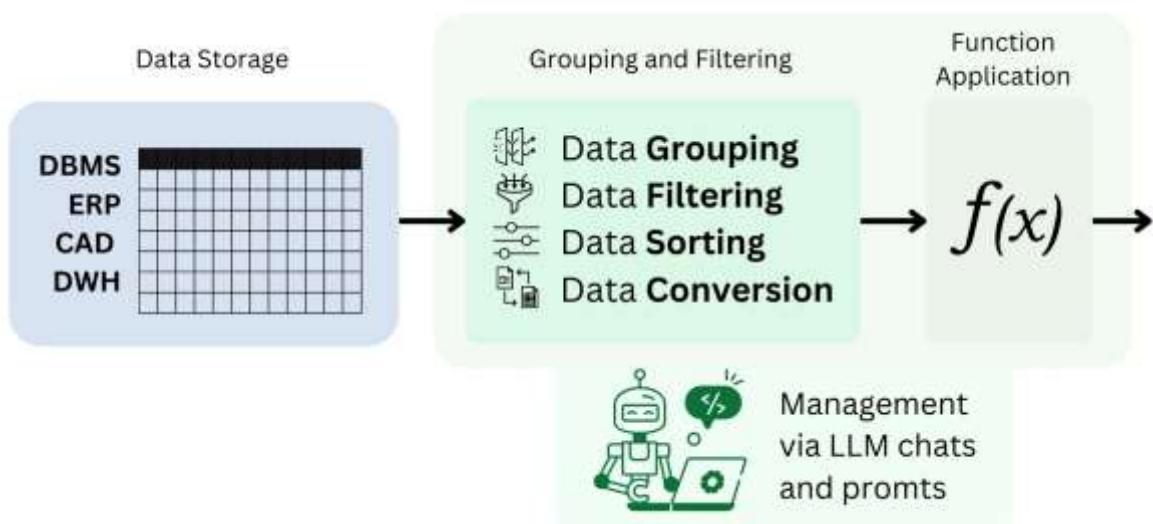


Рис. 2.24 Негізгі топтастыру, сүзу және сұрыптау операцияларын LLM чаттары жүзеге асырады.

Құрылыш саласының болашағы жергілікті шешімдер, бұлттық қуаттар және интеллектуалды модельдердің тиімді және қауіпсіз деректерді басқару жүйелерін құру үшін бірге жұмыс істеуіне негізделеді. LLM пайдаланушыларға терең техникалық білімсіз деректер базалары мен қоймала-рына табиғи тілде сұраулар формулоривкасын жасау арқылы өзара әрекеттесуге мүмкіндік береді. LLM және AI агенттері, сондай-ақ олардың қалай жұмыс істейтіні туралы біз «LLM агенттері және құрылымдық деректер форматтары» тарауында толығырақ сейлесеміз.

Дұрыс үйімдастырылған деректер мен LLM қолдауымен қарапайым, ыңғайлы аналитикалық құралдар ақпаратпен жұмыс істеуді жеңілдетіп қана қоймай, қателіктерді азайтуға, тиімділікті арттыруға және процестерді автоматтандыруға көмектеседі.

Деректерді жинаудан шешім қабылдауға: автоматтандыруға жол

Кітаптың келесі бөлімдерінде мамандардың өзара әрекеттесуі мен деректердің шешім қабылдау, автоматтандыру және жұмыс тиімділігін арттыру негізі ретінде қалай қызмет ететінін егжей-тегжейлі қарастырамыз. Рис. 2.25 деректерді өңдеу кезеңдерінің тізбегін көрсететін схеманың мысалын ұсынады, бұл схема деректерге негізделген тәсілдің үздіксіз жетілдіру контурын (Continuous Improvement Pipeline) иллюстрациялайды, оның кейбір бөліктері кітапта кейінірек егжей-тегжейлі

Қарастырылады.

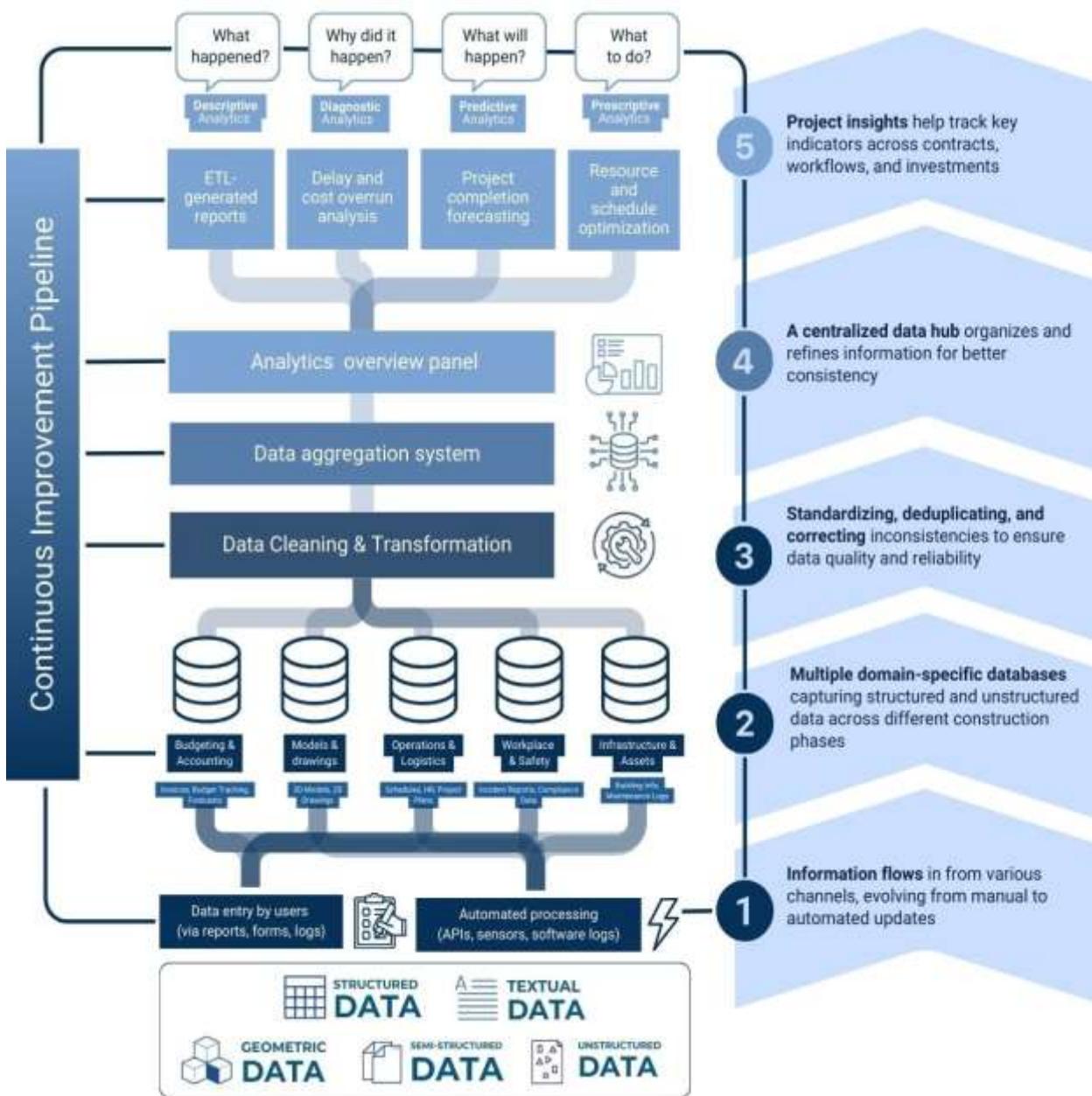


Рис. 2.25 Деректердің үздіксіз жетілдіру конвейері: құрылыш жобаларындағы деректерді өңдеу және талдау ағыны.

Орташа компанияның бизнес-процесстерін сипаттайтын жүйе көп деңгейлі принципке негізделеді. Оған деректерді жинау, тазалау, агрегация, аналитикалық өңдеу және алынған нәтижелер негізінде шешім қабылдау кіреді. Біз осы кезеңдердің барлығын кітапта әрі теориялық контексте, әрі практикалық мысалдар арқылы зерттейміз:

- Бірінші деңгейде деректер енгізіледі (Рис. 3.11). Ақпарат қолмен (есептер, формалар, журналдар арқылы) және автоматтандырылған түрде (API, датчиктер, бағдарламалық жүйелер арқылы) түседі. Деректер әртүрлі құрылымда болуы мүмкін: геометриялық, мәтіндік, құрылымсыз. Осы кезеңде ақпарат ағындарын стандарттау, құрылымдау және унификациялау қажеттілігі туындейды.
- Келесі деңгей – деректерді өңдеу және түрлендіру. Бұл кезең деректерді тазалау, дубликаттарды жою, қателерді түзету және ақпаратты әрі қарай талдауға дайындау процестерін қамтиды. Бұл кезең аналитиканың сапасы деректердің тазалығы мен дәлдігіне тікелей байланысты болғандықтан, өте маңызды.
- Содан кейін деректер функционалдық бағыттар бойынша бөлінген арнайы кестелерге, деректер фреймдеріне немесе базаларға түседі: бюджеттеу және есеп, модельдер мен сызбалар, логистика, қауіпсіздік және инфрақұрылым. Мұндай бөліну ыңғайлы қолжетімділікті үйімдастыруға және ақпаратты кросс-талдау мүмкіндігін қамтамасыз етуге мүмкіндік береді.
- Одан кейін деректер агрегатталып, аналитикалық панельде (витрина) көрсетіледі. Мұнда сипаттамалық, диагностикалық, болжаушы және ұсынымдық аналитика әдістері қолданылады. Бұл негізгі сұрақтарға жауап беруге мүмкіндік береді: не болды, неге бұл болды, болашақта не болады және қандай әрекеттер жасау қажет. Мысалы, жүйе кідірістерді анықтап, жобалардың аяқталуын болжай алады немесе ресурстарды оңтайландыра алады.-
- Соңында, соңғы деңгейде аналитикалық қорытындылар мен негізгі көрсеткіштер қалыптасады, олар келісімшарттардың орындалуын бақылауға, инвестицияларды басқаруға және бизнес процестерін жақсартуға көмектеседі. Бұл ақпарат компанияның шешім қабылдау және даму стратегиясының негізі болады.

Осылайша, деректер жинаудан стратегиялық басқаруда пайдалануға дейінгі жолдан өтеді. Кітаптың келесі бөлімдерінде біз әр кезеңді егжей-тегжейлі қарастырамыз, деректердің түрлеріне, деректерді өңдеу әдістеріне, аналитика құралдарына және құрылышы саласындағы осы тәсілдерді қолданудың нақты мысандарына ерекше назар аударамыз.

Келесі қадамдар: хаости басқарылатын жүйеге айналдыру

Осы бөлімде біз ақпараттық синостар мәселелерін зерттеп, жүйелердің артық құрделілігінің бизнес тиімділігіне әсерін қарастырдық, деректердің, ал қосымшалардың емес, орталық рөл атқаратын бесінші өнеркәсіптік революцияға көшүін талдадық. Біз бөлшектелген ақпараттық жүйелер білім алмасуға кедегілер жасайтынын, ал ИТ ландшафтың одан әрі құрделендіру құрылышы саласындағы өнімділікті төмендетіп, инновацияларды тежейтінін көрдік.

Осы бөлімнің қорытындысында, қарастырылған тәсілдердің күнделікті міндеттерінде қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

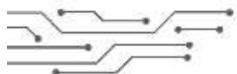
- Ақпараттық ландшафтыңызды визуализациялаңыз
- Сіз үнемі жұмыс істейтін деректер көздерінің визуалды картасын жасаңыз
- Осы картада сіздің жұмысыңызды пайдаланатын жүйелер мен қосымшаларды қосыңыз
- Потенциалды дублирленген функционалдылықтар мен артық шешімдерді анықтаңыз
- Жүйелер арасында деректерді беру кезінде жоғалту немесе бүрмалау орын алуы мүмкін

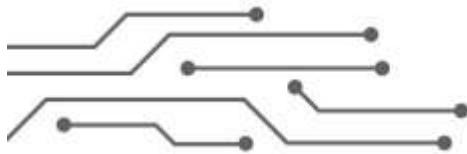
kritikalық нүктелерді анықтаңыз

- Деректерді басқарудың жеке тәжірибелерін енгізіңіз
- Процестерде деректерді негізгі актив ретінде қолдану фокусын өзгертіңіз
- Транспаренттілікті қамтамасыз ету үшін деректер көздері мен өндөр әдістемесін құжаттаңыз
- Деректердің сапасын бағалау және арттыру механизмдерін әзірлеңіз
- Деректер бір рет енгізіліп, бірнеше рет пайдаланылатында болуға үмтүлінген – бұл процестерді тиімді үйымдастырудың негізі
- Өз командаңызда деректерге негізделген (data-driven) тәсілді алға жылжытыңыз
- Әріптестер арасында деректерді алмасудың стандартталған және унификацияланған форматтарын пайдалануды ұсыныңыз
- Деректердің сапасы мен қолжетімділігіне қатысты мәселелерді командалық кездесулерде тұрақты түрде көтеріңіз.
- Сіз қолданатын құралдардың Open Source баламаларымен танысыңыз.

Кішкентайдан бастаңыз – жұмысының үшін маңызды бір нақты процесс немесе деректер жиынтығын таңдаңыз да, оған деректерге негізделген тәсілді қолданыңыз, құралдардан деректерге назар аударыңыз. Бір пилоттық тапсырмада табысқа жеткеннен кейін, сіз тек практикалық тәжірибе ғана емес, сонымен қатар жаңа әдістемелердің артықшылықтарын командаңызға айқын көрсету мүмкіндігіне ие боласыз. Көптеген осы қадамдарды орындау барысында сұрақтар туындаған жағдайда, кез келген заманауи LLM-ге түсініктемелер мен көмек сұрауға болады.

Кітаптың келесі бөлімдерінде деректерді құрылымдау және унификациялау әдістерін тереңірек қарастырамыз және әртүрлі ақпаратты интеграциялаудың практикалық тәсілдерін зерттейміз. Бөлшектелген қоймалардан деректердің біртұтас экожүйелеріне көшуге ерекше назар аударылады, бұл құрылыштың саласындағы цифрлық трансформацияда маңызды рөл атқарады.





III БӨЛІМ ҚҰРЫЛЫС БИЗНЕС- ПРОЦЕСТЕРІНДЕГІ ДЕРЕКТЕР ҚҰРЫЛЫСЫ

Ушінші бөлімде құрылыс саласындағы деректер типологиясы мен оларды тиімді үйімдастыру әдістері туралы кешенді түсінік қалыптасады. Құрылымдалған, құрылымдалмаған, жартылай құрылымдалған, мәтіндік және геометриялық деректермен жұмыс істеудің сипаттамалары мен ерекшеліктері құрылыс жобалары контекстінде талданады. Салада қолданылатын әртүрлі жүйелер арасында ақпарат алmasу үшін заманауи сақтау форматтары мен протоколдары қарастырылады. Әртүрлі форматтағы деректерді біртұтас құрылымдалған ортаға айналдырудың практикалық құралдары мен әдістері, соның ішінде CAD (BIM) деректерін интеграциялау тәсілдері сипатталады. Деректердің сапасын стандарттау және валидациялау арқылы қамтамасыз ету тәсілдері ұсынылады, бұл құрылыс есептеулерінің дәлдігі үшін маңызды. Заманауи технологияларды (Python Pandas, LLM-модельдері) қолданудың практикалық аспектілері, құрылыс индустриясындағы типтік тапсырмаларды шешу үшін код мысалдарымен егжей-тегжейлі талданады. Ақпаратты басқару тәсілдерін координациялау және стандарттау үшін үйімдастыруышылық құрылым ретінде құзырет орталығын (CoE) құрудың құндылығы негізделеді.

ТАРАУ 3.1.

ҚҰРЫЛЫС САЛАСЫНДАҒЫ ДЕРЕКТЕР ТҮРЛЕРИ

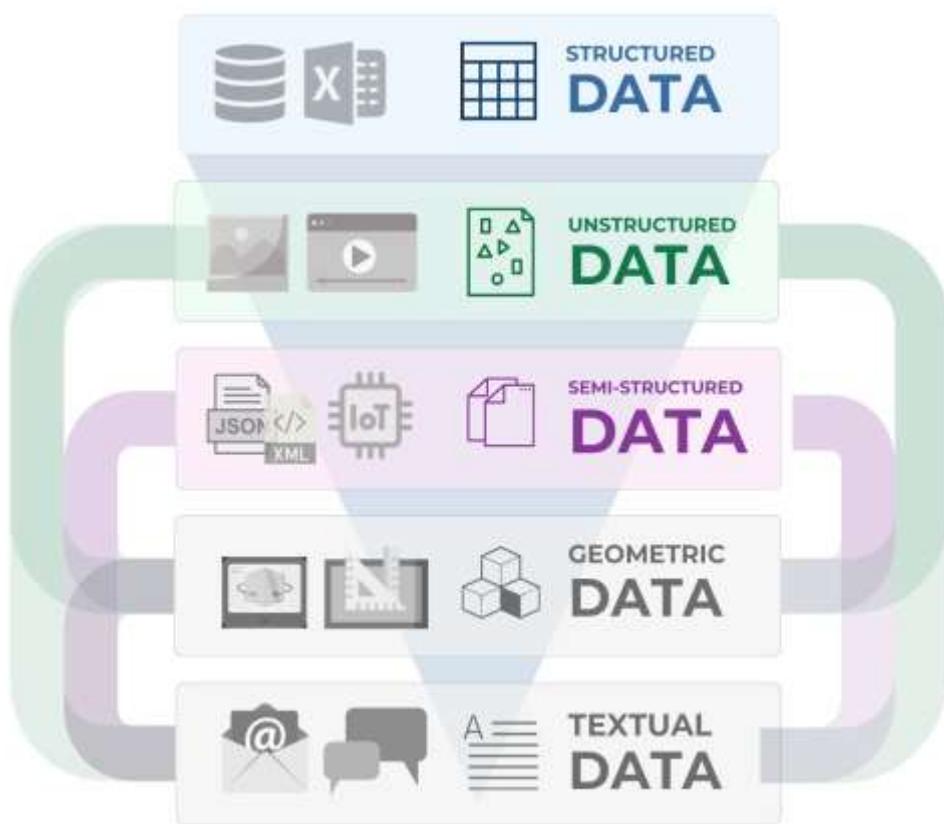
Құрылыш саласындағы ең маңызды деректер түрлері

Қазіргі құрылыш саласында компаниялардың жүйелері, қосымшалары және деректер қоймалары әртүрлі типтегі және форматтағы ақпаратпен белсенді түрде толтырылады. Құрылыш саласында жұмыс істейтін заманауи компанияның ақпараттық ландшафттың қалыптастыратын негізгі деректер типтерін егжей-тегжейлі қарастырамыз:-

- Құрылымдалған деректер: бұл деректерде нақты үйімдастырушылық құрылым бар, мысалы, электрондық кестелер мен реляциялық деректер базалары.
- Құрылымдалмаған деректер: бұл ақпарат қатаң ережелерге сәйкес үйімдастырылмаған. Мұндай деректерге мәтін, бейне, фотосуреттер және аудиозапистар жатады.
- Жартылай құрылымдалған деректер: бұл деректер құрылымдалған және құрылымдалмаған деректер арасында аралық орын алады. Олар құрылым элементтерін қамтиды, бірақ бұл құрылым әрдайым анық емес немесе жиі әртүрлі схемалар арқылы сипатталады. Құрылыш саласындағы жартылай құрылымдалған деректердің мысалдары: техникалық спецификациялар, жобалық құжаттама немесе орындалған жұмыстар туралы есептер.
- Мәтіндік деректер: ауызша және жазбаша коммуникациялар нәтижесінде алынған барлық нәрсені қамтиды, мысалы, электрондық пошта, жиналыстар мен көздесулердің стенограммалары.
- Геометриялық деректер: бұл деректер CAD бағдарламаларынан келеді, онда мамандар жобаның элементтерінің геометриялық деректерін визуализациялау, көлем мәндерін растау немесе қақтығыстарды тексеру үшін жасайды.

Геометриялық және мәтіндік (әріп-цифрлық) деректердің жеке категория емес екенін атап өткен жөн, олар үш деректер түрінде де болуы мүмкін. Мысалы, геометриялық деректер құрылымдалған деректердің (параметрлік CAD форматтары) бөлігі ретінде, сондай-ақ құрылымдалмаған деректердің (сканерленген сыйбалар) бөлігі ретінде болуы мүмкін. Мәтіндік деректер де, сәйкесінше, деректер базаларында (құрылымдалған деректер) үйімдастырылуы мүмкін, сондай-ақ нақты құрылымы жоқ құжаттар түрінде болуы мүмкін.

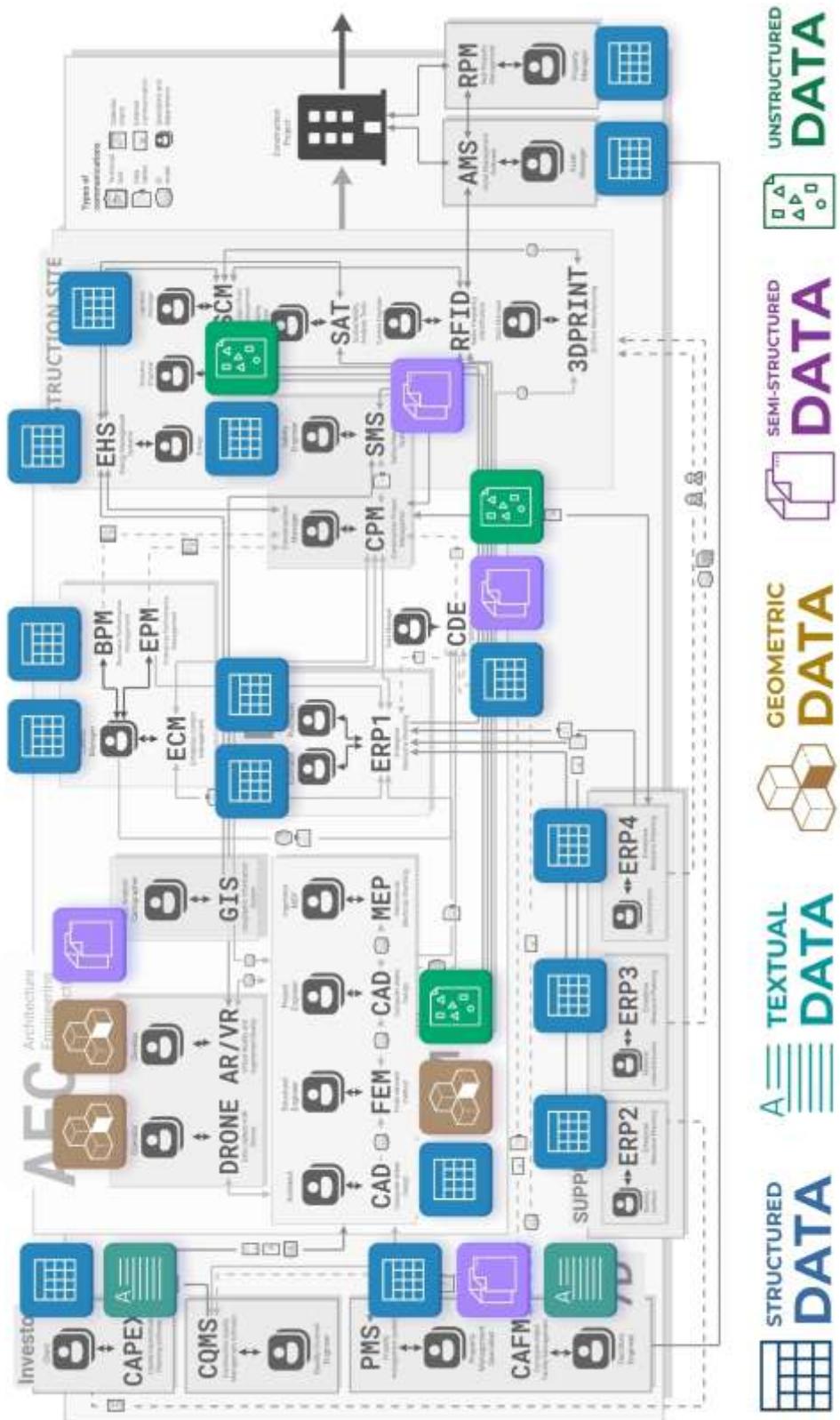
Құрылыш компаниясындағы әрбір деректер түрі – компанияның ақпараттық активтер мозаикасындағы бірегей элемент. Құрылыш алаңдарынан алынған суреттер мен жиналыстардың аудиозапистары сияқты құрылымдалмаған деректерден, кестелер мен деректер базаларын қамтитын құрылымдалған жазбаларға дейін – әрбір элемент компанияның ақпараттық ландшафтының қалыптасуында маңызды рөл атқарады.



Инженерлер мен деректермен жұмыс істейтін менеджерлер құрылыс саласында қолданылатын барлық деректер түрлерімен жұмыс істеуді үйренуі тиіс.

Міне, құрылыс саласында қолданылатын кейбір жүйелер мен оларға қатысты деректер түрлерінің тізімі (Рис. 3.12): -

- ERP (Enterprise Resource Planning) - әдетте құрылымдалған деректерді өндейді, кәсіпорын ресурстарын басқаруға және әртүрлі бизнес процестерін интеграциялауға көмектеседі.
- CAD (Computer-Aided Design) BIM (Building Information Modeling) жүйесімен бірге - құрылыс жобаларын жобалау және модельдеу үшін геометриялық және жартылай құрылымдалған деректерді пайдаланады, жобалау кезеңінде ақпараттың дәлдігі мен үйлесімділігін қамтамасыз етеді.
- GIS (Geographic Information Systems) - картографиялық деректер мен кеңістіктік қатынастарды жасау және талдау үшін геометриялық және құрылымдалған деректермен жұмыс істейді.
- RFID (Radio-Frequency Identification) - құрылыс алаңында материалдар мен жабдықтарды тиімді бақылау үшін радиожиілік идентификациясы арқылы жартылай құрылымдалған деректерді пайдаланады.
- ECM (Engineering Content Management) – техникалық сызбалар мен жобалық құжаттама сияқты жартылай құрылымдалған және құрылымдалмаған деректерді, инженерлік деректер мен құжаттаманы басқару жүйесі.



Әр түрлі форматтар мен деректер әр түрлі жүйелерді толтырады, кешенді интеграция үшін қолайлы формага аударуды талап етеді.

Бұл және көптеген басқа жүйелер компанияның құрылымдалған кестелік деректерден бастап күрделі геометриялық модельдерге дейінгі кең ауқымды деректерді басқаруға мүмкіндік береді, жобалау, жоспарлау және құрылыш басқару процестерінде интеграцияланған өзара әрекеттесуді қамтамасыз етеді.

Қаралайым диалог мысалында (Рис. 3.13) құрылыш жобасының мамандары әртүрлі деректер түрлерімен алмасады:

- Архитектор: "Клиенттің тілектерін ескере отырып, мен шатырда демалыс аймағын қостым. Жаңа дизайнмен танысыңыз" (геометриялық деректер - модель).
- Инженер-конструктор: "Жоба алынды. Мен жаңа демалыс аймағы үшін шатырдың көтеріш қабілетін есептеймін" (құрылымдалған және жартылай құрылымдалған деректер - есептеу кестелері).
- Сатып алу менеджері: "Демалыс аймағы үшін материалдардың спецификациялары мен саны қажет, сатып алушы үйімдастыру үшін" (мәтіндік және жартылай құрылымдалған деректер - тізімдер мен спецификациялар).
- Еңбек қорғау және қауіпсіздік инженері: "Жаңа аймақ туралы мәліметтер алдым. Тәуекелдерді бағалап, қауіпсіздік жоспарын жаңартып жатырмын" (жартылай құрылымдық деректер - құжаттар мен жоспарлар).
- ВІМ-моделдеу маманы: "Жұмыс құжаттамасын түзету үшін жобаның жалпы моделіне өзгерістер енгізу" (геометриялық деректер және жартылай құрылымдық деректер).
- Жоба жетекшісі: "Демалыс аймағын жұмыс кестесіне енгіземін. Жобаны басқару жүйесінде кестелер мен ресурстарды жаңартып жатырмын" (құрылымдық және жартылай құрылымдық деректер - кестелер мен жоспарлар).
- Нысандарды басқару маманы (FM): "Мен демалыс аймағын болашақта қызмет көрсету үшін деректерді дайындал, оларды мүлікті басқару жүйесіне енгіземін" (құрылымдық және жартылай құрылымдық деректер - техникалық қызмет көрсету нұсқаулары мен жоспарлары).

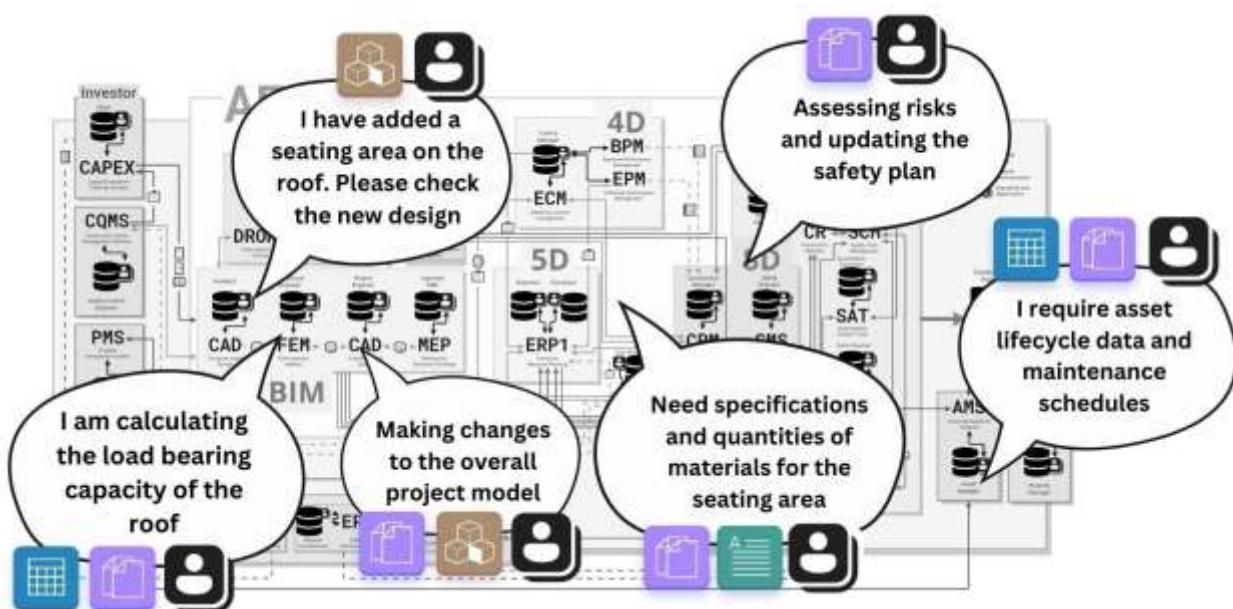


Рис. 3.13 Мамандар арасындағы байланыс мәтін деңгейінде де, деректер деңгейінде де жүзеге асырылады.

Әрбір маман әртүрлі деректер түрлерімен жұмыс істейді, бұл команда арасындағы тиімді өзара әрекеттесуді және жобаның сәтті орындалуын қамтамасыз етеді. Құрылымдық, жартылай құрылымдық және құрылымдық емес деректер арасындағы айырмашылықтарды түсіну, олардың цифрлық бизнес-процестердегі ерекше рөлін түсінуге мүмкіндік береді. Әр түрлі деректер формаларының бар екенін білу ғана емес, сонымен қатар оларды қалай, қайда және не үшін қолдану көрктігін түсіну маңызды.

Жақында осындағы әртүрлі деректердің біріктіру идеясы амбициялық, бірақ қыын жүзеге асырылатын болып көрінген. Бұғынгі күні бұл құнделікті практикаға айналды. Әртүрлі схемалар мен құрылымдардың деректерін интеграциялау қазіргі ақпараттық жүйелер архитектурасының ажырамас бөлігіне айналды.

Келесі тарауларда құрылымдық, жартылай құрылымдық және құрылымдық емес деректердің біртұтас келісілген көрініске біріктіруге мүмкіндік беретін негізгі стандарттар мен тәсілдерді ежей-тегжейлі қарастырамыз. Құрылымдық деректер мен реляциялық деректер базаларына ерекше назар аударылады - құрылымдық саласындағы ақпаратты сақтау, өңдеу және талдау үшін негізгі механизмдер.

Структуранған деректер

Құрылымдық саласында ақпарат көптеген көздерден - сызбалар, спецификациялар, кестелер мен есептерден келеді. Бұл ағынды тиімді басқару үшін деректердің құрылымдау қажет. Құрылымдық деректер ақпаратты ыңғайлы, оқуға жеңіл және қолжетімді түрде үйімдастыруға мүмкіндік береді.

JB Knowledge компаниясының 5-ші жыл сайынғы құрылымдық технологиялары туралы есебіне сәйкес [17], құрылымдық жобаларын басқару мамандарының 67%-ы жұмыстардың тиімділігін қолмен немесе электрондық кестелер арқылы бақылап, бағалайды.

Құрылымдық деректердің ең кең тараулған форматтарының бірі XLSX және CSV болып табылады. Олар электрондық кестелерде ақпаратты сақтау, өңдеу және талдау үшін кеңінен қолданылады. Мұндай кестелерде деректер жолдар мен бағандар түрінде ұсынылған, бұл оларды оқу, өңдеу және талдау үшін ыңғайлыш етеді.

Microsoft компаниясы жасаған XLSX форматы XML құрылымдарын пайдалану негізінде құрылған және ZIP алгоритмімен архивтеледі. Форматтың негізгі ерекшеліктері:

- Күрделі формулаларды, диаграммаларды және макростарды қолдау.
- Деректердің әртүрлі параллельдерда сақтау және ақпаратты форматтау мүмкіндігі.
- Microsoft Excel ортасында жұмыс істеуге оңтайланылған, бірақ басқа офис пакеттерімен де үйлесімді.

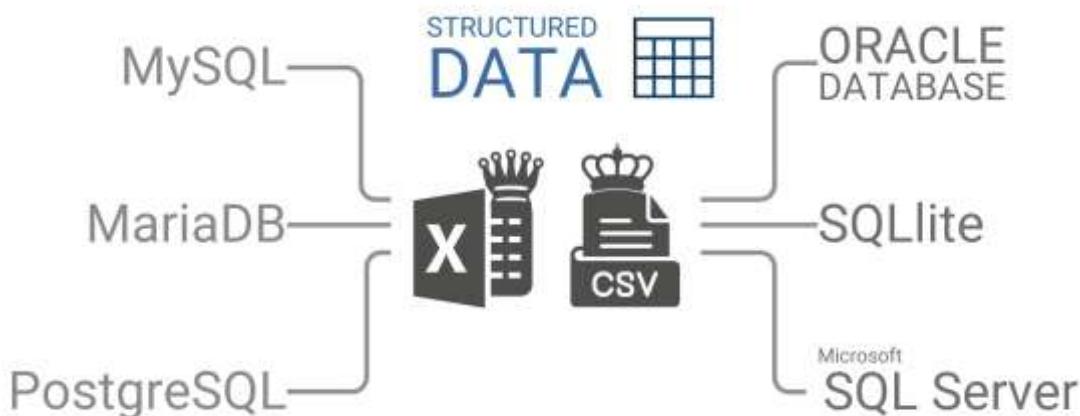
CSV форматы - бұл мәндер үтірлермен, нұкте-үтірлермен немесе басқа бөлгіш символдармен бөлінген қаралайым мәтіндік файл. Негізгі артықшылықтары:

- Әртүрлі бағдарламалар мен операциялық жүйелермен әмбебап үйлесімділік.
- Деректер базалары мен аналитикалық жүйелерге импорт/экспорт үшін ыңғайлыштың.

■ Тіпті мәтіндік редакторларда өңдеудің жеңілдігі.

Дегенмен, CSV формулалар мен форматтауды қолдамайды, сондықтан оның негізгі қолданылуы - жүйелер арасында деректер алмасу және ақпаратты жаппай жаңарту. Өзінің әмбебаптығы мен платформалардан тәуелсіздігі арқасында CSV гетерогенді ИТ-орталарында деректерді беру үшін танымал құралға айналды.

XLSX және CSV форматтары құрылымдық деректермен жұмыс істейтін әртүрлі жүйелер арасында байланыстыруши буын ретінде қызмет етеді. Олар әсіресе оқылымдылық, қолмен өңдеу және негізгі үйлесімділік маңызды тапсырмаларда пайдалы.-



XLSX және CSV форматтары құрылымдық деректермен жұмыс істейтін әртүрлі жүйелер арасында байланыстыруши буын болып табылады.

Платформадан тәуелсіздік CSV-ны гетерогенді ИТ-орталарында және жүйелерде деректерді беру үшін ең танымал формат етеді.

Дегенмен, XLSX және CSV жоғары өнімді есептеулер немесе үлкен көлемдегі деректерді ұзақ мерзімді сақтау үшін арналмаған. Мұндай мақсаттар үшін Apache Parquet, Apache ORC, Feather, HDF5 сияқты заманауи құрылымдық форматтар қолданылады. Бұл форматтарды біз «Улken деректерді сақтау: танымал форматтарды талдау және олардың тиімділігі» атты кітаптың тоғызыншы бөлімінде толығырақ қарастырамыз.

Практикада Excel XLSX форматымен көбінесе шағын тапсырмалар мен рутиналық процестерді автоматтандыру үшін қолданылады. Құрделі сценарийлер деректерді басқару жүйелерін, мысалы, ERP, PMIS CAFM, CPM, SCM және басқаларын пайдалануды талап етеді. Дәл осы жүйелерде компанияның ақпараттық ағындарын үйімдастыру мен басқаруға негіз болатын құрылымдық деректер сақталады.-

Құрылыш саласында қолданылатын заманауи деректерді басқару ақпараттық жүйелері деректерді кесте түрінде үйімдастырылған құрылымдық деректерге сүйенеді. Улken ақпарат көлемін сенімді,

масштабталатын және тұтас басқару үшін қосымша және жүйе әзірлеушілері реляциялық деректер базасын басқару жүйелеріне (RDBMS) жүгінеді.

Реляциялық деректер базалары RDBMS және SQL сұраныс тілі

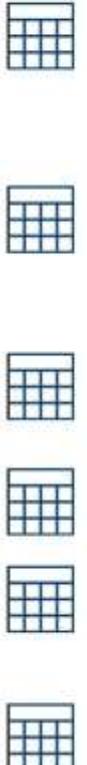
Деректерді тиімді сақтау, өңдеу және талдау үшін реляциялық деректер базалары (RDBMS) қолданылады - бұл ақпаратты кестелерде белгілі бір қатынастармен ұйымдастырытын деректерді сақтау жүйелері.

Деректер, RDBMS-те ұйымдастырылған, тек цифрлық ақпаратты білдірмейді; олар транзакциялар мен әртүрлі жүйелер арасындағы өзара әрекеттестіктің негізі болып табылады.

Ең көп таралған реляциялық деректер базасын басқару жүйелерінің бірнешеуі:-

- MySQL (Open Source) - LAMP (Linux, Apache, MySQL, PHP/Perl/Python) стекінің құрамына кіретін ең танымал RDBMS-тердің бірі. Веб-дамытуда қарапайымдылығы мен жоғары өнімділігі арқасында кеңінен қолданылады.
- PostgreSQL (Open Source) - сенімділігі мен кеңейтілген мүмкіндіктерімен танымал қуатты объектілік-реляциялық жүйе. Қурделі корпоративтік шешімдер үшін қолайлы.
- Microsoft SQL Server – Microsoft компаниясының коммерциялық жүйесі, корпоративтік ортада кеңінен қолданылып, компанияның басқа өнімдерімен интеграциясы мен жоғары қауіпсіздік деңгейімен ерекшеленеді.
- Oracle Database – ірі кәсіпорындар мен критикалық маңызды қосымшаларда қолданылатын ең қуатты және сенімді СУБД-лардың бірі.
- IBM DB2 – ірі корпорацияларға бағытталған, жоғары өнімділік пен сенімділікті қамтамасыз етеді.
- SQLite (Ашық код) – мобильді қосымшалар мен автономды жүйелер, мысалы, CAD (BIM) жобалау бағдарламалары үшін тамаша жарамды жеңіл встраиваемая база деректері.
- Құрылыш бизнесінде танымал деректер базасын басқару жүйелері - MySQL, PostgreSQL, Microsoft SQL Server, Oracle® Database, IBM® DB2 және SQLite - құрылымдалған деректермен жұмыс істейді. Бұл СУБД-лар бизнес-процестер мен қосымшаларды басқару үшін қуатты және икемді шешімдерді ұсынады, шағын веб-сайттардан бастап ауқымды корпоративтік жүйелерге дейін (Рис. 3.21). -

Statista [48] деректері бойынша, 2022 жылы реляциялық деректер базасын басқару жүйелері (RDBMS) жалпы қолданылатын DBMS-тің шамамен 72%-ын құрады.



	Rank			DBMS	Database Model	Open Source vs Commercial
	Mar2025	Feb2025	Mar2024			
1.	1.	1.	Oracle®	Relational, Multi-model	Commercial	
2.	2.	2.	MySQL	Relational, Multi-model	Open Source	
3.	3.	3.	Microsoft® SQL Server	Relational, Multi-model	Commercial	
4.	4.	4.	PostgreSQL	Relational, Multi-model	Open Source	
5.	5.	5.	MongoDB	Document, Multi-model	Open Source	
6.	7.	9.	Snowflake®	Relational	Commercial	
7.	6.	6.	Redis®	Key-value, Multi-model	Open Source	
8.	8.	7.	Elasticsearch®	Multi-model	Open Source	
9.	9.	8.	IBM Db2	Relational, Multi-model	Commercial	
10.	10.	10.	SQLite	Relational	Open Source	
11.	11.	12.	Apache Cassandra®	Multi-model	Open Source	
12.	12.	11.	Microsoft Access®	Relational	Open Source	
13.	13.	17.	Databricks®	Multi-model	Commercial	
14.	14.	13.	MariaDB	Relational, Multi-model	Open Source	
15.	15.	14.	Splunk	Search engine	Commercial	
16.	16.	16.	Amazon DynamoDB	Multi-model	Commercial	
17.	17.	15.	Microsoft Azure SQL	Relational, Multi-model	Commercial	

Рис. 3.15 DBMS рейтингінде құрылымдаған деректер базаларын қолданудың танымалдығы (көкпен белгіленген) (по материалам [49]).

Ашық кодты деректер базаларын орнату өте оңай – тіпті терең техникалық білімсіз. PostgreSQL, MySQL немесе SQLite сияқты ашық кодты жүйелер тегін қолжетімді және Windows, macOS және Linux операциялық жүйелерінің көпшілігінде жұмыс істейді. Бар болғаны, жобаның ресми сайтына кіріп, орнатушыны жүктеп алғып, нұсқауларды орындау қажет. Көп жағдайда орнату 10–15 минуттан аспайды. Біз осындай деректер базасын төртінші бөлімде модельдеп, жасаймыз (Рис. 4.38).-

Егер сіздің компанияңызда бүлтты қызметтер (мысалы, Amazon Web Services, Google Cloud немесе Microsoft Azure) қолданылса, онда деректер базасын бірнеше рет басу арқылы орналастыруға болады – платформа сізге орнату үшін дайын шаблондарды ұсынады. Кодтың ашықтығы арқасында, мұндаидарды өз міндеттеріңізге оңай бейімдеуге болады, ал үлкен пайдаланушылар қауымдастыры кез келген мәселені шешуге көмектеседі.

RDBMS көптеген бизнес-қосымшалар мен аналитикалық платформалардың негізі болып табылады (Рис. 3.16), олар компанияларға деректерді тиімді сақтау, өңдеу және талдау мүмкіндігін береді – демек, негізделген және уақтылы шешімдер қабылдауға мүмкіндік береді.-

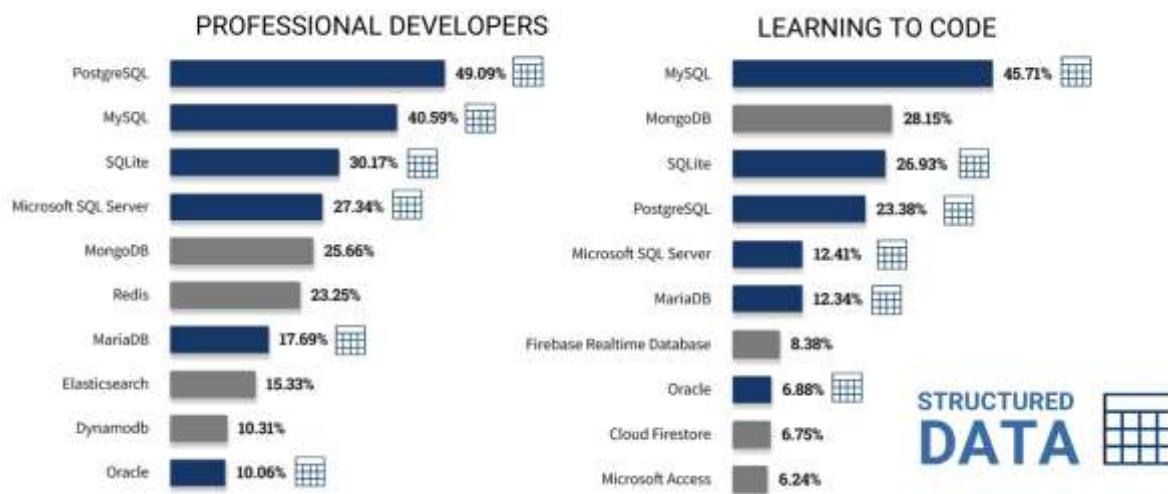


Рис. 3.16 StackOverFlow (ең ірі IT форум) сайтында әзірлеушілер арасында өткен жылы қандай деректер базаларын қолданғаны және келесі жылы қандай деректер базаларын қолданғысы келетіні туралы сауалнама (RDBMS көк түспен ерекшеленген) (по материалам [50]).

RDBMS деректердің сенімділігін, келісімділігін, транзакцияларды қолдауды қамтамасыз етеді және SQL (Structured Query Language) қуатты сұраныс тілін пайдаланады, ол аналитикада жиі қолданылады және деректер базасында сақталған ақпаратты оңай алуға, өзгертуге және талдауға мүмкіндік береді. SQL реляциялық жүйелерде деректермен жұмыс істеудің негізгі құралы болып табылады.

Деректер базаларындағы SQL сұраныстары және жаңа трендтер

Реляциялық деректер базаларында жиі қолданылатын SQL тілінің негізгі артықшылығы, басқа ақпаратты басқару әдістерімен (мысалы, классикалық электрондық кестелер Excel арқылы) салыстырғанда, өте үлкен деректер базаларын жоғары сұраныс өңдеу жылдамдығымен қолдауда жатыр.

Структурланған сұрау тілі (SQL) – реляциялық деректер базаларында ақпаратты сақтау, өңдеу және талдау үшін арналған мамандандырылған бағдарламалу тілі. SQL деректерді жасау, басқару және оларға қол жеткізу үшін қолданылады, бұл ақпаратты тиімді түрде табуға, сұзуге, біріктіруге және агрегаттауға мүмкіндік береді. Ол деректерге қол жеткізудің негізгі құралы болып табылады, ақпарат сақтау орындарымен өзара әрекеттесудің ыңғайлы және формализденген тәсілін қамтамасыз етеді.

Эволюция SEQUEL-SQL жүйелерінің маңызды өнімдері мен компаниялары, мысалы, Oracle, IBM DB2, Microsoft SQL Server, SAP, PostgreSQL және MySQL арқылы өтеді, және SQLite пен MariaDB-нің пайда болуымен аяқталады. SQL кестелермен жұмыс істеу мүмкіндіктерін ұсынады, олар Excel-де жоқ, бұл деректермен жұмысты масштабалатын, қауіпсіз және автоматтандыруға ыңғайлы етеді.

- Деректер құрылымдарын жасау және басқару (DDL): SQL-де дереккорда кестелерді жасау,

өзгерту және жою, олардың арасындағы байланыстарды орнату және деректерді сақтау құрылымдарын анықтау мүмкіндігі бар. Ал Excel-де жұмыс фиксиленген парапттар мен ұяшықтармен жүргізіледі, парапттар мен деректер жиынтықтары арасында нақты байланыстар жоқ.

- Деректермен операциялар (DML): SQL үлкен көлемдегі деректерді жаппай қосуға, өзгертуге, жоюға және жылдам шығарып алуға мүмкіндік береді, күрделі сұрауларды фильтрациялау, сұрыптау және кестелерді біріктіру арқылы орындау (Сурет 3.17). Excel-де үлкен ақпарат көлемін өңдеу қолмен әрекеттерді немесе арнайы макростарды талап етеді, бұл процессті баяулатып, қателіктердің ықтималдығын арттырады.-
- Қол жеткізуі басқару (DCL): SQL деректерге қол жеткізу құқықтарын әртүрлі пайдаланушылар арасында шектеуге мүмкіндік береді, ақпаратты өңдеу немесе қараша мүмкіндігін шектейді. Ал Excel-де қол жеткізу не жалпы (файлды беру кезінде), не бұлттық қызметтер арқылы құқықтарды бөлу үшін күрделі параметрлеуді талап етеді.

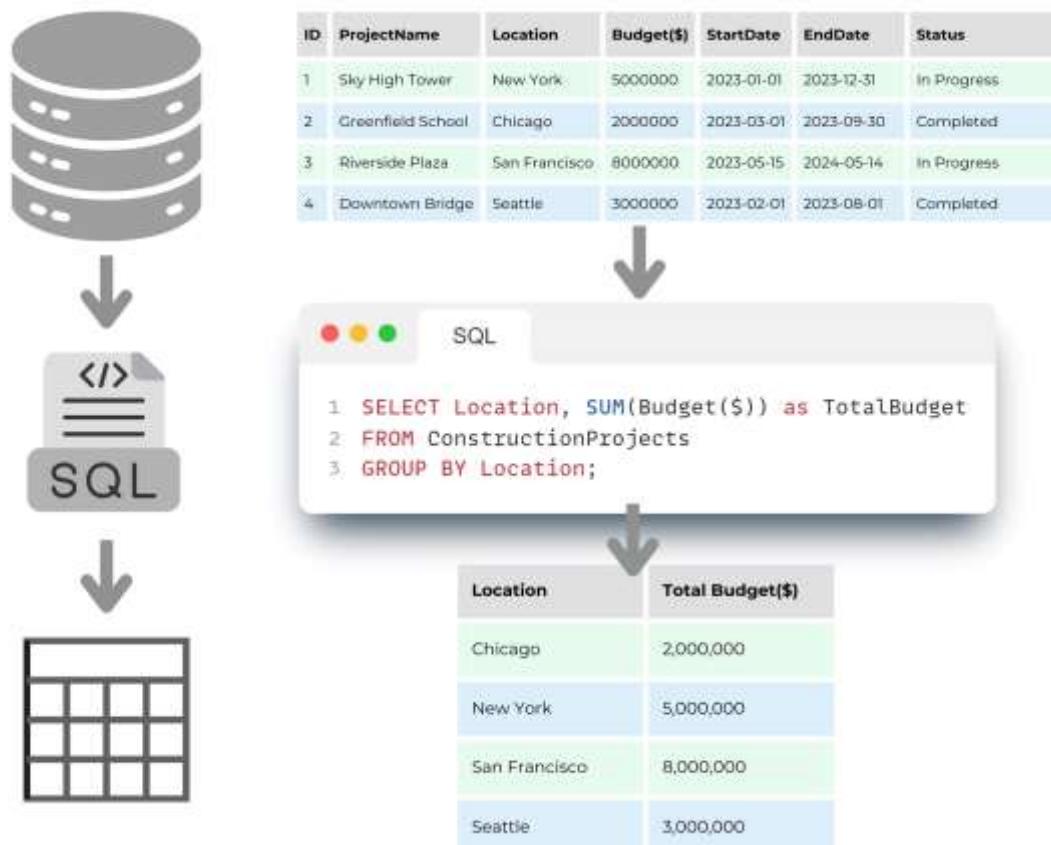


Рис. 3.17 DML мысалы SQL: деректерді автоматты өңдеу үшін бірнеше код жолдарын пайдалана отырып, жылдам өңдеу, топтау және агрегаттау.

Excel деректермен жұмыс істеуді визуалды және интуитивті құрылымымен жеңілдетеді. Алайда, деректер көлемі артқан сайын Excel-дің өнімділігі төмендейді. Excel сонымен қатар сақталатын деректер көлеміне қатысты шектеулерге тап болады – максимум бір миллион жол, және өнімділік осы шекке жетпей-ақ төмендейді. Сондықтан, Excel шағын деректер көлемдерімен визуализация және манипуляция жасау үшін қолайлыш болып көрінгенімен, үлкен деректер массивтерімен жұмыс

істей үшін SQL әлдеқайда тиімдірек.

Келесі кезең құрылымдық деректерді дамытуда бағандық деректер базаларының (Columnar Databases) пайда болуы болды, олар дәстүрлі реляциялық деректер базаларына балама болып табылады, әсіресе едәүір үлкен деректер көлемдері мен аналитикалық есептеулер жүргізу кезінде. Жолдық СУБД-лерден айырмашылығы, мұнда деректер жол бойынша емес, баған бойынша сақталады. Классикалық деректер базаларымен салыстырғанда, бұл мүмкіндік береді:

- Деректердің бағандарындағы біртекті мәліметтерді тиімді сұғу арқылы сақтау көлемін азайту.
- Аналитикалық сұраныстарды жеделдету, себебі тек қажетті бағандар оқылады, ал бүкіл кесте емес.
- Big Data және деректер қоймалары, мысалы, Data Lakehouse архитектурасында жұмысын оңтайландыру.

Колоночалық дерекқорлар, Pandas DataFrame, Apache Parquet, HDF5 және олардың негізінде деректерді талдау мен өңдеу мақсатында Big Data сақтау қоймаларын құру туралы біз осы кітаптың келесі тарауларында – «DataFrame: әмбебап кестелік деректер форматы» және «Деректерді сақтау форматтары мен Apache Parquet-пен жұмыс: DWH деректер қоймалары мен Data Lakehouse архитектурасы» – толығырақ сөйлесетін боламыз.

Структураланбаған деректер

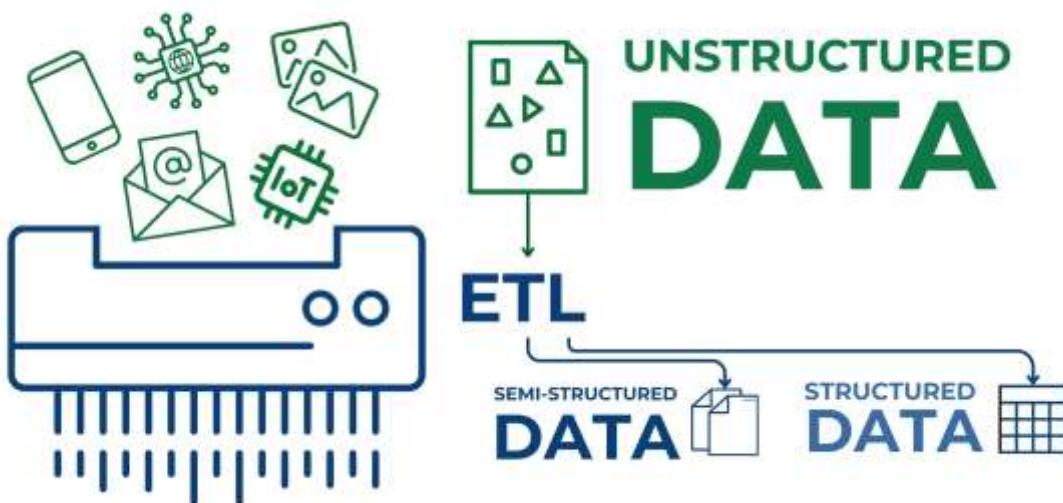
Қолданбалар мен ақпараттық жүйелерде пайдаланылатын деректердің басым бөлігі құрылымдық формада болса да, құрылыс саласында өндірілетін ақпараттың көп бөлігі құрылымдалмаған деректер түрінде ұсынылады – суреттер, бейнeler, мәтіндік құжаттар, дыбыс жазбалары және басқа контент түрлері. Бұл әсіресе құрылыс, пайдалану және техникалық бақылау кезеңдерінде, визуалды және мәтіндік ақпараттың басым болуымен байланысты.

Құрылымдалмаған деректер – бұл алдын ала белгіленген модель немесе құрылымға ие емес, дәстүрлі жолдар мен бағандарда ұйымдастырылмаған ақпарат, мысалы, деректер базасы немесе кестелерде.

Жалпы түрде құрылымдалмаған деректерді екі категорияға бөлуге болады:

- Адамдармен өндірілген құрылымдалмаған деректер, оған әртүрлі адам жасаған контент түрлері жатады: мәтіндік құжаттар, электрондық хаттар, суреттер, бейнeler және т.б.
- Машиналармен өндірілген құрылымдалмаған деректер құрылғылар мен датчиктер арқылы жасалады: бұл журнал файлдары, GPS деректері, Internet of Things (IoT) нәтижелері және, мысалы, құрылыс алаңынан алғынған басқа телеметриялық ақпарат.

Құрылымдалған деректерден айырмашылығы, олар кестелер мен деректер базасында ыңғайлы ұйымдастырылған, құрылымдалмаған деректер ақпараттық жүйелерге интеграцияланар алдында қосымша өңдеу кезеңдерін талап етеді. Автоматтандырылған жинау, талдау және мұндай деректерді түрлендіру технологияларын пайдалану құрылыс тиімділігін арттыру, қателіктерді азайту және адам факторының әсерін минимизациялау үшін жаңа мүмкіндіктер ашады.



Құрылымдалмаған деректерді өңдеу полуструктураланған және құрылымдалған деректерге түрлендіруден басталады.

Құрылымдалмаған деректер барлық ақпараттың 80%-на дейінгі бөлігін құрайды, сондықтан біз олардың түрлері мен өңдеуін кітаптың келесі тарауларында мысалдармен егжей-тегжейлі қарастырамыз.

Талқылауды жеңілдету үшін мәтіндік деректерді жеке категорияға бөліп алайық. Олар құрылымдалмаған деректердің бір түрі болғанымен, олардың құрылыс саласындағы маңыздылығы мен таралуы ерекше назар аударуды талап етеді.

Мәтіндік деректер: структура мен структураланбаған хаос арасында

Құрылыш саласындағы мәтіндік деректер әртүрлі форматтар мен ақпарат түрлерін қамтиды, қағаз құжаттардан бастап, құрылыш алаңындағы хат алмасу, әңгімелер, жұмыс хаттамалары және ауызша кездесулер сияқты бейресми байланыс әдістеріне дейін. Барлық осы мәтіндік деректер құрылыш жобаларын басқару үшін маңызды ақпаратты қамтиды - жобалық шешімдер мен жоспарлардағы өзгерістерден бастап, қауіпсіздік мәселелерін талқылау және мердігерлер мен клиенттермен келіссөздер жүргізуге дейін.-



Мәтіндік деректер, жобаның қатысушылары арасындағы байланыс үшін ең танымал ақпарат түрлерінің бірі.

Мәтіндік ақпарат формализован және құрылымдалмаған болуы мүмкін. Формализован деректерге Word (.doc,.docx), PDF форматындағы құжаттар, сондай-ақ жиналыс хаттамаларының мәтіндік файлдары (.txt) жатады. Неформализован деректер мессенджерлер мен поштадағы хат алмасуды, жиналыстардың стенограммаларын (Teams, Zoom, Google Meet), сондай-ақ талқылаулардың дыбыс жазбаларын (.mp3,.wav) қамтиды, олар мәтінге түрлендіруді талап етеді.

Бірақ ресми сұраулар, келісім шарттардың шарттары және электрондық хабарламалар сияқты жазбаша құжаттар әдетте белгілі бір құрылымға ие болса, ауызша хабарламалар мен жұмыс хат алмасулары жиі құрылымсыз болып қалады, бұл олардың талдауын және жобаларды басқару жүйелеріне интеграциясын қынданады.

Мәтіндік деректерді тиімді басқарудың кілті – оларды құрылымдалған форматқа түрлендіру. Бұл өндөлген ақпаратты құрылымдалған деректермен жұмыс істейтін қолданыстағы жүйелерге автоматты түрде интеграциялауға мүмкіндік береді.



Рис. 3.110 Мәтіндік контентті құрылымдалған деректерге түрлендіру.

Мәтіндік ақпаратты тиімді пайдалану үшін оны автоматты түрде құрылымдалған формаға түрлендіру қажет (Рис. 3.110). Бұл процесс әдетте бірнеше кезеңді қамтиды:-

- Мәтінді тану (OCR) – құжаттар мен сыйбалардың кескіндерін машиналық оқуға жарамды форматқа түрлендіру.
- Мәтінді талдау (NLP) – жобамен байланысты негізгі параметрлерді (күндер, сомалар және сандар) автоматты түрде анықтау.
- Деректерді класификациялау – ақпаратты категорияларға (қаржы, логистика, тәуекелдерді басқару) бөлу.

Танылған және класификацияланғаннан кейін құрылымдалған деректер базаларға интеграциялаңып, автоматтандырылған есеп беру және басқару жүйелерінде пайдаланылуы мүмкін.

Полуструктуранған және әлсіз структуранған деректер

Полуструктурленген деректер белгілі бір деңгейде ұйымдастыруды қамтиды, бірақ қатаң схема немесе құрылымға ие емес. Мұндай ақпарат құрылымдалған элементтерді (мысалы: күндер, қызметкерлердің аттары және орындалған тапсырмалар тізімдері) қамтығанымен, ұсыныс форматы әртүрлі жобаларда немесе тіпті жеке қызметкерлерде айтартылғанда өзгеруі мүмкін. Мұндай

деректерге жұмыс уақытының есеп журналдары, орындалған жұмыстар туралы есептер және әртүрлі форматта ұсынылуы мүмкін графиктер жатады.

Полуструктурленген деректерді талдау құрылымсыз деректерге қарағанда жеңіл, алайда оларды стандартталған жобаларды басқару жүйелеріне интеграциялау үшін қосымша өндөуді талап етеді.

Полуструктурленген деректермен жұмыс істеу, құрылымның тұрақты өзгеруімен сипатталатын, елеулі қыындықтарды туыннатады. Бұл деректер құрылымының өзгергіштігі әрбір полуструктурленген дереккөзді өндөу және талдау үшін жеке тәсілдерді талап етеді.

Бірақ құрылымсыз деректермен жұмыс істеу үлкен күш-жігерді талап етсе, полуструктурленген деректерді өндөу салыстырмалы түрде қарапайым әдістер мен құралдар арқылы жүзеге асырылуы мүмкін.

Әлсіз құрылымдалған деректер – минималды немесе толық емес құрылыммен сипатталатын жалпы термин. Көбінесе бұл мәтіндік құжаттар, чаттар, электрондық хаттар, мұнда жеке метадеректер (мысалы, күн, жіберуші) кездеседі, бірақ ақпараттық көп бөлігі хаостық түрде ұсынылады.

Құрылыш саласында әлсіз құрылымдалған деректер әртүрлі процестерде кездеседі. Мысалы, оларға мынашынан жатады:

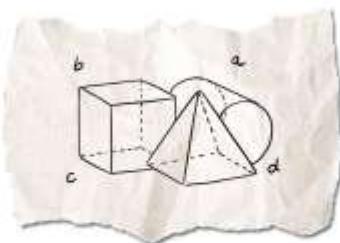
- Сметалық есептеулер мен коммерциялық ұсыныстар – материалдар, көлемдер және бағалар туралы деректерді қамтитын, бірақ біртекті форматқа ие емес кестелер.
- Сызбалар мен инженерлік схемалар – мәтіндік аннотациялар мен метадеректерді қамтитын PDF немесе DWG файлдары, бірақ қатаң бекітілген құрылымсыз.
- Жұмыстарды орындау графиктері – MS Project, Primavera P6 немесе басқа жүйелерден алынған деректер, экспорт құрылымы әртүрлі болуы мүмкін.
- CAD (BIM-модельдер) – құрылым элементтерін қамтиды, бірақ деректерді ұсыну бағдарламалық қамтамасыз ету мен жобаның стандартына байланысты.

CAD жүйелерімен өндірілген геометриялық деректер жартылай құрылымдалған деректер ретінде класификациялануы мүмкін. Дегенмен, біз геометриялық CAD (BIM) деректерін жеке тип ретінде бөліп көрсетеміз, себебі олар, мәтіндік деректер сияқты, компания процестерінде жиі жеке деректер типі ретінде қарастырылады.

Геометриялық деректер және олардың қолданылуы

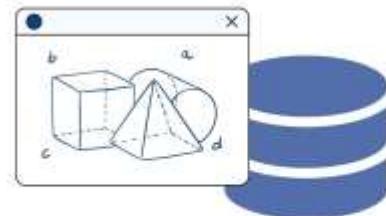
Жоба элементтері туралы метадеректер көбінесе кестелер түрінде, құрылымдалған немесе әлсіз құрылымдалған форматтарда сақталса, жоба элементтерінің геометриялық деректері көбінесе арнайы CAD құралдары арқылы жасалады, олар жобаның элементтерін сыйықтар (2D) немесе геометриялық (3D) түрінде егжей-тегжейлі визуализациялауға мүмкіндік береді.-

3000 BCE - 1960s



physical medium
(artefact)

1960s to present day

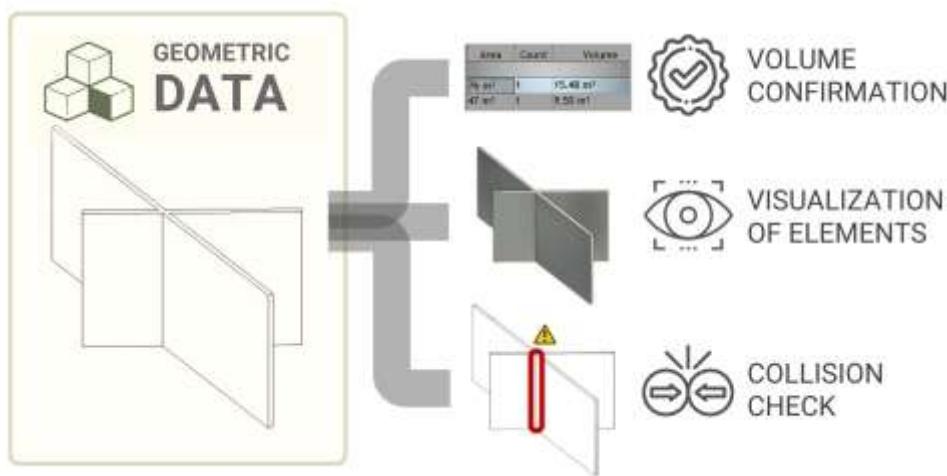


digital project data
(CAD data)

CAD құралдары физикалық тасымалдаушылардан геометриялық ақпаратты деректер базасы форматына көшіруге көмектесті.

Құрылыш және архитектурада геометриялық деректермен жұмыс істегендеге, геометриялық деректердің үш негізгі қолдану саласын бөліп көрсетуге болады:-

- Көлемдерді растау: CAD (BIM) бағдарламаларының ішінде арнайы геометриялық ядролар арқылы генерирленген геометриялық деректер жобаның элементтерінің көлемдері мен өлшемдерін автоматты және дәл анықтау үшін қажет. Бұл деректер жоспарлау, бюджет құру және ресурстар мен материалдарды тапсырыс беру үшін қажетті автоматты түрде есептелген алаңдар, көлемдер, ұзындықтар және басқа да маңызды атрибуттарды қамтиды.
- Жобаны визуализациялау: жобаға қандай да бір өзгерістер енгізілген жағдайда, элементтердің визуализациясы әртүрлі жазықтықтарда жаңартылған сыйбаларды автоматты түрде генерациялауға мүмкіндік береді. Жобаның бастапқы кезеңдерінде визуализация барлық қатысушылар арасында өзара түсіністікті жеделдетуге көмектеседі, құрылыш процесінде уақыт пен ресурстарды үнемдеуге мүмкіндік береді.
- Қақтығыстарды тексеру: бірнеше элемент категориялары (мысалы, құбырлар мен қабырғалар) арасындағы "геометриялық қақтығыстар" болмауы критикалық маңызды болған күрделі құрылыш және инженерлік жобаларда, қақтығыстарды тексеру маңызды рөл атқарады. Қақтығыстарды анықтау бағдарламалық қамтамасыз етуді пайдалану жобаның элементтері арасындағы әлеуетті геометриялық қақтығыстарды алдын ала анықтауға мүмкіндік береді, құрылыш процесінде қымбатқа түсетін қателіктерді болдырмауға көмектеседі.
- Инженерлік-конструкторлық бюролардың пайда болған сәтінен бастап, алғашқы күрделі конструкциялардың құрылышынан бері, инженерлер геометриялық ақпаратты сыйбалар, сыйықтар және жазық геометриялық элементтер түрінде ұсынды, олар негізінде прорабтар мен сметашылар, мындаған жылдар бойы, сыйғыштар мен транспортирлердің көмегімен атрибуттық көлемдер немесе элементтер мен элементтер топтарының санын жинады.-



Геометрия элементтердің көлемдік параметрлерін алу үшін негіз болып табылады, олар кейін жобаның құнын және мерзімдерін есептеу үшін пайдаланылады.

Бұғынгі қуні бұл қолмен және еңбек-intensive тапсырма заманауи CAD (BIM) құралдарындағы көлемдік моделдеу арқасында толық автоматтандыру арқылы шешіледі, бұл арнайы геометриялық ядроның көмегімен кез келген элементтің көлемдік атрибуттарын автоматты түрде алуға мүмкіндік береді, көлемдік параметрлерді қолмен есептеудің қажеті жоқ.

Заманауи CAD құралдары жобаның элементтерін классификациялауға және категориялауға мүмкіндік береді, бұл жобаның дерекқорынан спецификациялар кестелерін шығарып, оларды бағалау, графиктер құру немесе CO2 есептеу сияқты әртүрлі жүйелерде пайдалануға мүмкіндік береді (Рис. 3.113). Спецификацияларды, QTO кестелерін және көлемдерді алу, сондай-ақ практикалық мысалдар туралы "Көлемдерді алу және сандық есептеу" бөлімінде талқылаймыз. -



Рис. 3.113 CAD (BIM) құралдары деректердің басқа жүйелермен интеграциялау және өзара әрекеттесу үшін арналған дерекқорларда сақтайты.

CAD ортасында қолданылатын дерекқорлар мен форматтардың жабықтығына байланысты, CAD шешімдерінде жасалған геометриялық деректер фактілі түрде жеке ақпарат түріне айналды. Бұл элементтердің геометриясын және арнайы файлдар мен форматтарға енгізілген метаақпаратты (құрылымдалған немесе жартылай құрылымдалған) біріктіреді.

CAD деректері: жобалаудан деректерді сақтауға

Заманауи CAD және BIM жүйелерінде деректер көбінесе меншікті, жиі патенттелген форматтарда сақталады: DWG, DXF, RVT, DGN, PLN және басқалар. Бұл форматтар объектілердің екі өлшемді және үш өлшемді көріністерін қолдайды, сонымен қатар геометриямен қатар объектілермен байланысты атрибуттарды сақтайды. Олардың ең кең таралғандары:

- DWG – екі өлшемді (және сирек үш өлшемді) жобалық деректер мен метадеректерді сақтау үшін қолданылатын бинарлық файл форматы.
- DXF – CAD жүйелері арасында 2D және 3D сызбаларды алмасу үшін тексттік формат. Геометрияны, қабаттарды және атрибуттық деректерді қамтиды, ASCII және бинарлық көріністі қолдайды.
- RVT – 3D геометрияны, элементтердің атрибуттарын, байланыстарды және жобаның параметрлерін қамтитын CAD модельдерін сақтау үшін бинарлық формат.
- IFC – CAD (BIM) жүйелері арасында құрылымдардың алмасу үшін ашық тексттік формат. Геометрияны, объектілердің қасиеттерін және олардың өзара байланыстары туралы ақпаратты қамтиды.

Одан басқа, PLN, DB1, SVF, NWC, CPIXML, BLEND, BX3, USD, XLSX, DAE сияқты басқа форматтар да қолданылады. Олар тағайындалуы мен ашықтық деңгейі бойынша әртүрлі (Рис. 3.114), бірақ барлығы жобаның ақпараттық моделін әртүрлі формаларда көрсетуі мүмкін. Кешенді жобаларда бұл форматтар жиі параллель қолданылады – сызбадан бастап жобаларды үйлестіруге дейін.

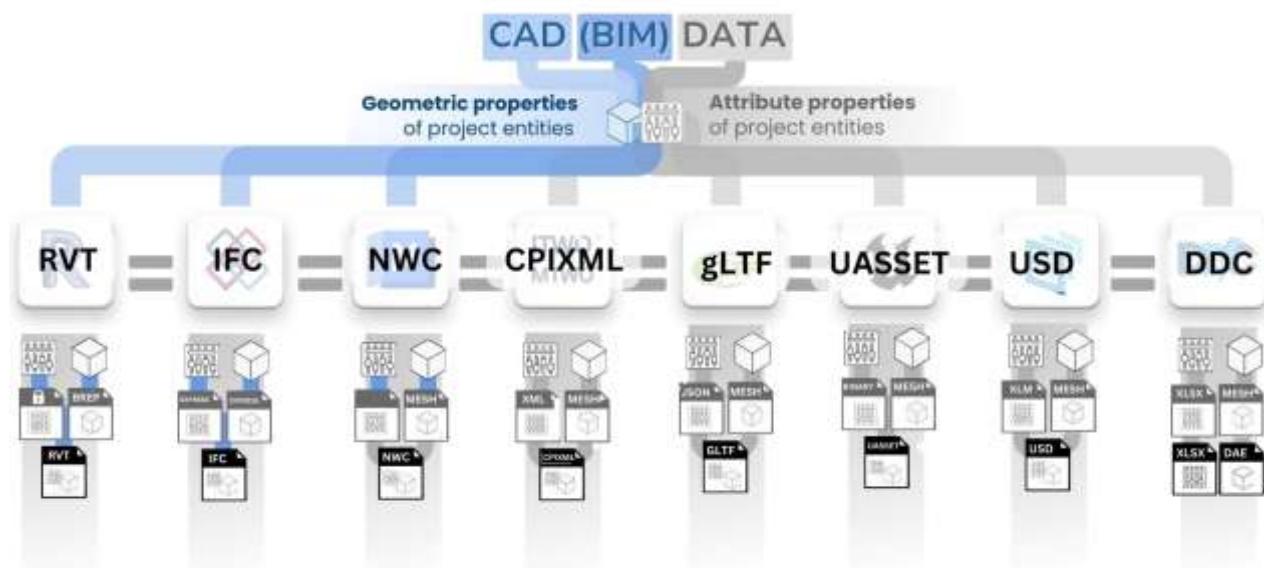


Рис. 3.114 CAD-тан алынған ақпаратты сақтау үшін танымал форматтар BREP немесе MESH параметрлері арқылы геометрияны сипаттайтыны, оларды атрибуттық деректермен толықтырады.

Жоғарыда аталған барлық форматтар құрылымдардың жобасының әр элементі туралы деректерді сақтауға мүмкіндік береді, және барлық аталған форматтарда екі негізгі деректер түрі бар:

- Геометриялық параметрлер – объектінің пішінін, орналасуын және өлшемдерін сипаттайтыны.

Геометрия мен оның қолданылуы туралы толығырақ кітаптың алтыншы бөлімінде, CAD (BIM) шешімдеріне арналған бөлімде талқыланады;

- Атрибуттық қасиеттер – материалдар, элементтердің түрлері, техникалық сипаттамалар, біргей идентификаторлар және жобаның элементтерінде болуы мүмкін басқа қасиеттер туралы әртүрлі ақпаратты қамтиды.

Заманауда жобаларда атрибуттық деректердің маңызы ерекше, себебі олар объектілердің эксплуатациялық сипаттамаларын анықтайды, инженерлік, калькуляциялық есептеулер жүргізуге мүмкіндік береді және жобалау, құрылымың және эксплуатация қатысушылары арасында үздіксіз өзара әрекеттесуді қамтамасыз етеді. Мысалы:

- Терезелер мен есіктер үшін көрсетіледі: конструкция түрі, остекление түрі, ашылу бағыты (Рис. 3.21).
- Қабырғалар үшін материалдар, жылу оқшаулау және акустикалық сипаттамалар туралы ақпарат тіркеледі.
- Инженерлік жүйелер үшін құбырлар, ауа арналар, кабель трассалары және олардың қосылыстарының параметрлері сақталады.

Бұл параметрлер CAD-(BIM)-файлдарының ішінде, сондай-ақ сыртқы деректер базаларында – экспорт, конверсия немесе CAD-тың ішкі құрылымдарына тікелей қол жеткізу нәтижесінде сақталуы мүмкін. Мұндай тәсіл жобалық ақпаратты басқа корпоративтік жүйелермен және платформалармен интеграциялауды жеңілдетеді.

CAD (BIM) контекстіндегі кері инжиниринг – бұл цифрлық модельдің ішкі құрылымын шығару және талдау процесі, оның логикасын, деректер құрылымын және тәуелділіктерін қайта құру мақсатында, бастапқы алгоритмдерге немесе құжаттамаға қол жеткізусіз.

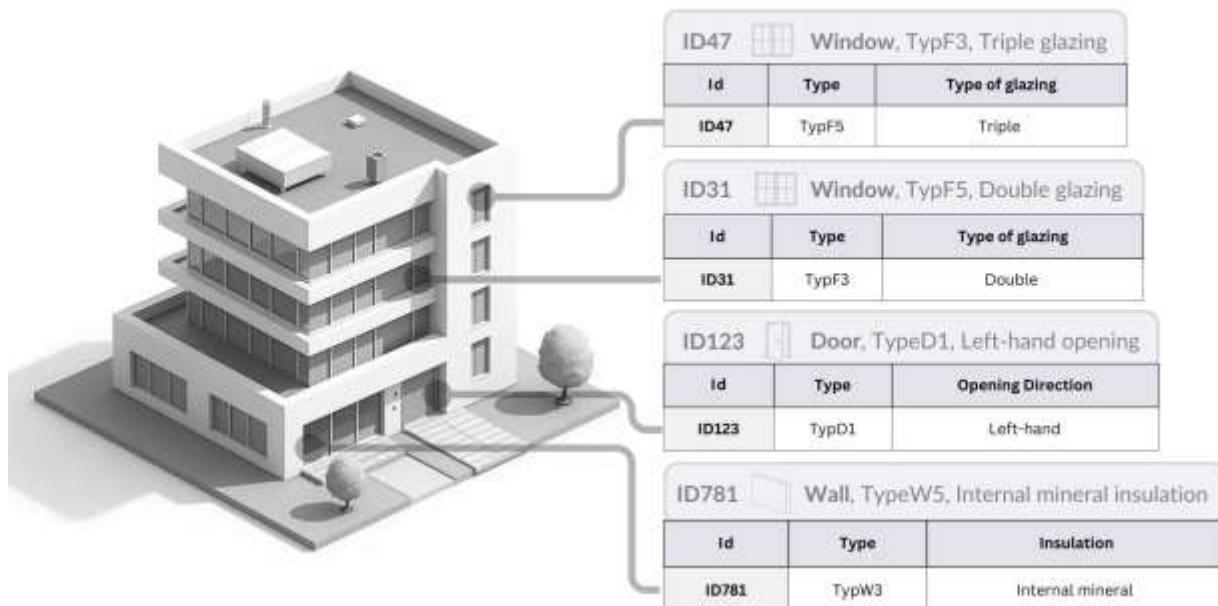


Рис. 3.115 Жобаның элементі параметрлік немесе полигональды геометрияны сипаттаумен қатар, элементтердің параметрлері мен қасиеттері туралы ақпаратты қамтиды.

Нәтижесінде, әр элементтің айналасында бірегей параметрлер мен қасиеттер жиынтығы қалыптасады, ол әр объектінің (мысалы, идентификатор және өлшемдер) бірегей сипаттамаларын, сондай-ақ, элементтер топтары үшін жалпы атрибуттарды қамтиды. Бұл тек жоба элементтерін талдауға ғана емес, сонымен қатар оларды логикалық топтарға біріктіруге мүмкіндік береді, оларды кейін басқа мамандар өз міндеттері мен есептеулері үшін жүйелер мен деректер базаларында пайдалана алады.

Субъект (ағылш. entity) – нақты немесе абстрактілі шындық объектісі, оны біржакты бөліп, сипаттап, деректер түрінде ұсынуға болады.

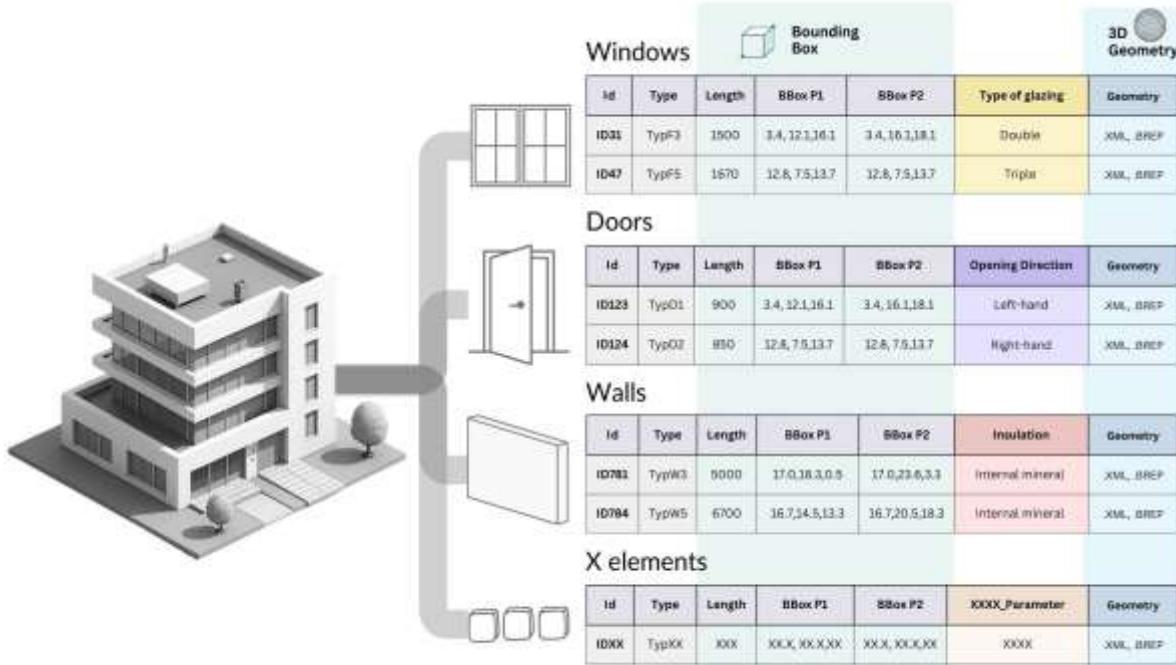


Рис. 3.116 Әр жобаның элементі атрибуттарды қамтиды, олар жобалаушы тараپынан енгізіледі немесе CAD бағдарламасының ішінде есептеледі.

Соңғы онжылдықта құрылыш саласында деректерді жасау, сақтау және беру процесін жеңілдететін көптеген жаңа CAD (BIM) форматтары пайда болды. Бұл форматтар жабық және ашық, кестелік, параметрлік немесе графтық болуы мүмкін. Алайда, олардың әртүрлілігі мен фрагментациясы жобаның өмірлік циклі кезеңдерінде деректерді басқаруды едәуір қынданатады. Құрылыш саласында ақпарат алмасу үшін қолданылатын негізгі форматтардың салыстыру кестесі Рис. 3.117-де берілген (толық нұсқасы QR коды арқылы қолжетімді).-

Интероперабельділік және CAD деректеріне қол жеткізу мәселелерін шешу үшін (BIM) менеджер-

лері мен координаторлары жұмысқа тартылады, олардың міндегі – экспортты бақылау, деректердің сапасын тексеру және CAD (BIM) деректерінің бөліктерін басқа жүйелерге интеграциялау.

Алайда, форматтардың жабықтығы мен күрделілігіне байланысты, бұл процесті автоматтандыру қыын, бұл мамандарды көптеген операцияларды қолмен орындауға мәжбүр етеді, деректерді өңдеу үшін толыққанды ағынды процестерді (pipeline) құру мүмкіндігінсіз.

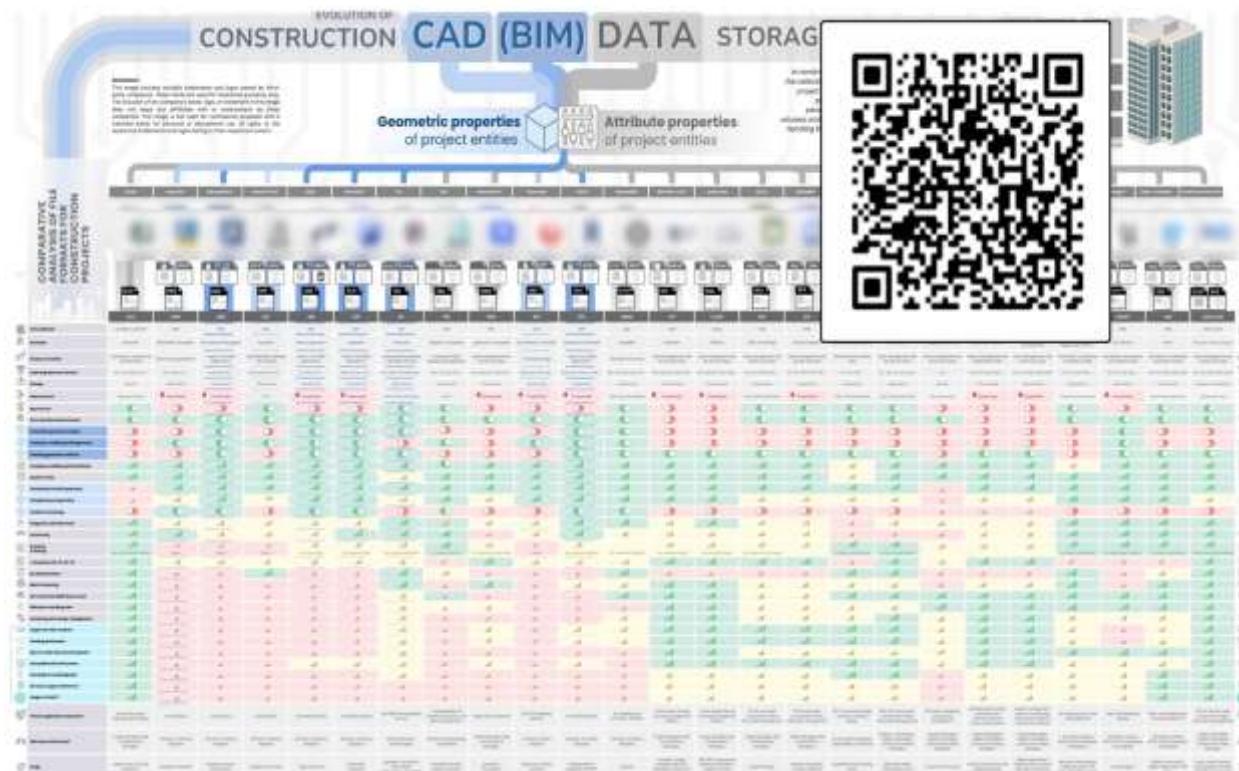


Рис. 3.117 Жобаның элементтері туралы ақпаратты сақтайдын негізгі деректер форматтарының салыстыру кестесі [53].

Неліктен көптеген әртүрлі деректер форматтары бар және олардың көпшілігі жабық екенін түсінү үшін, CAD (BIM) бағдарламаларының ішіндегі процестерге тереңірек үңілу маңызды, бұл кітаптың алтыншы бөлімінде егжей-тегжейлі қарастырылады.

Геометрияға қосылатын қосымша ақпараттық қабат CAD жүйелерінің әзірлеушілері тарапынан BIM (Building Information Modeling) концепциясы түрінде ұсынылды – 2002 жылдан бері құрылыс саласында белсенді түрде насиҳатталатын маркетингтік термин [54].

BIM (BOM) концепциясының пайда болуы және CAD-ды процестерде пайдалану

Ғимараттардың ақпараттық моделдеу концепциясы (BIM), алғаш рет 2002 жылғы Whitepaper BIM құжатында баяндалған, CAD бағдарламалық қамтамасыз ету өндірушілерінің маркетингтік бастамаларының нәтижесінде пайда болды. Бұл концепция, машина жасау саласында жақсы нәтиже

көрсеткен принциптерді құрылыш саласының қажеттіліктеріне бейімдеу мақсатында CAD бағдарламалық қамтамасыз ету әзірлеушілерінің маркетингтік бастамаларының нәтижесінде қалыптасты.

BIM-нің шабыт көзі BOM (Bill of Materials) концепциясы болды – өнімнің құрамын сипаттайтын спецификация, ол өнеркәсіпте 1980-жылдардың сонынан бері белсенді қолданылып келеді. Машина жасауда BOM CAD жүйелерінен PDM (Product Data Management), PLM (Product Lifecycle Management) және ERP жүйелеріне деректерді байланыстыруға мүмкіндік беріп, өнімнің өмірлік циклы бойында инженерлік ақпаратты кешенді басқаруды қамтамасыз етті.-

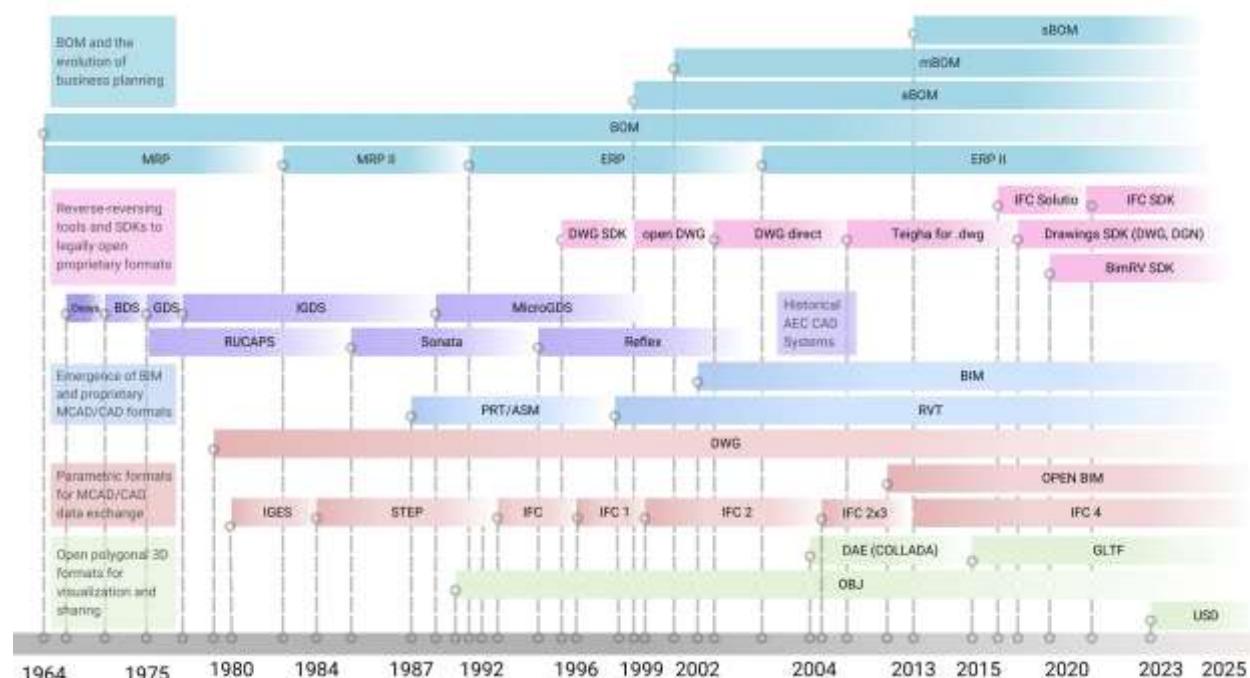


Рис. 3.118 Спецификациялардың (BOM), ақпараттық моделдеудің (BIM) және инженерлік құрылыш саласындағы цифрлық форматтардың эволюциясы.

BOM концепциясының қазіргі даму кезеңі XBOM (Extended BOM) кеңейтілген құрылымының пайда болуына әкелді, ол тек өнімнің құрамын ғана емес, сонымен қатар мінез-құлық сценарийлерін, пайдалану талаптарын, тұрақтылық параметрлерін және болжамдық аналитика үшін деректерді қамтиды. XBOM, негізінен, құрылыш саласындағы BIM-нің рөлін атқарады: екі тәсіл де цифрлық модельді жобаның барлық қатысушылары үшін өмірлік цикл бойында сенімді ақпараттың бірегей көзіне (Single Source of Truth) айналдыруға үмтүлады.

Құрылыш саласында BOM-ның пайда болуындағы негізгі кезең 2002 жылы құрылыш саласына арнайы бейімделген алғашқы параметрліk CAD-тың пайда болуы болды. Оны 1980-жылдардың сонында машина жасау үшін революциялық MCAD жүйесін жасаған команда әзірледі, ол сала стандарттарына айналды.

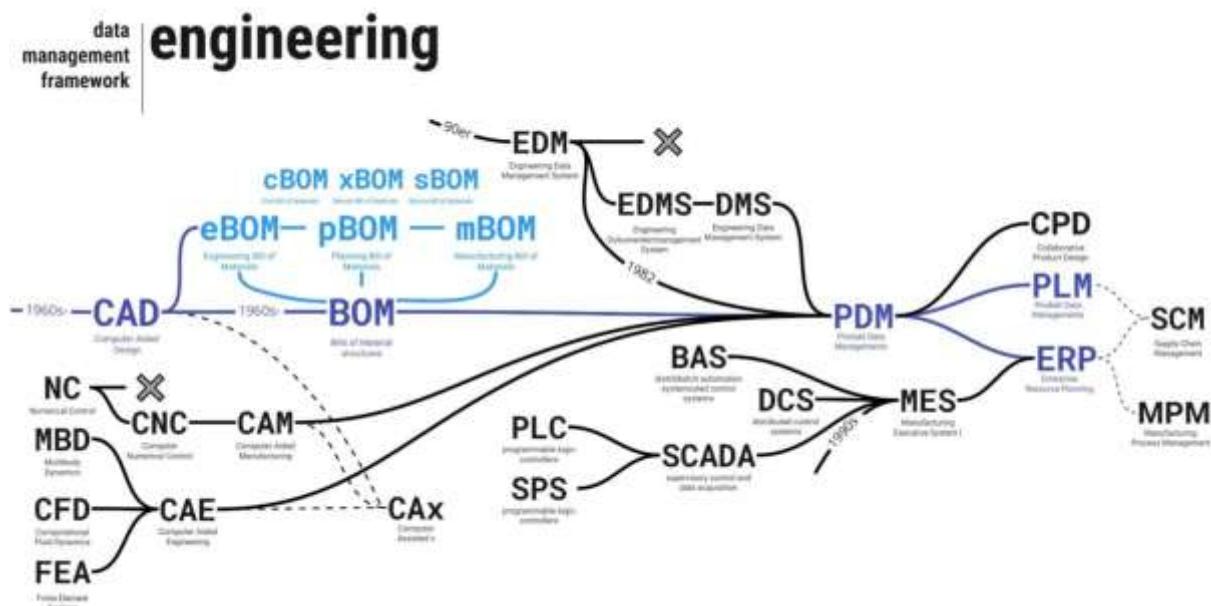
1980-жылдардың сонында мақсат CAD бағдарламаларында бар шектеулерді жою болды. Негізгі міндет – жобаның элементтерінің параметрлерін өзгертуге жұмысалатын еңбек шығындарын азайту және CAD бағдарламаларынан тыс деректер негізінде модельді жаңартуға мүмкіндік беру. Бұл процессте параметризация маңызды рөл атқаруы туис: деректер базасынан автоматты түрде

сипаттамаларды алу және оларды CAD жүйесіндегі модельді жаңарту үшін пайдалану.

Pro-E және оның негізінде жатқан BOM концепциясы CAD және MCAD нарығының да-мұына елеулі әсер етті. 25 жыл бойы бұл модель салада қолданылып, көптеген қазіргі заманғы жүйелер оның концептуалдық мұрагерлері болды.

Мақсат – инженерге әртүрлі конструкцияларды оқай қарастыруға мүмкіндік беретін жүйе құру. Ал жобадағы өзгерістер енгізу шығындары мүмкіндігінше нөлге жақын болуы тиіс. Дәстүрлі CAD/CAM бағдарламалық қамтамасыз ету өзгерістерді енгізуі тек жобалау процесінің ең бастапқы кезеңінде арзан түрде шектеуге мүмкіндік береді.– Самуэль Гейзенберг, Parametric Technology Corporation® компаниясының негізін қалаушы, Pro-E MCAD өнімінің әзірлеушісі және RVT форматындағы CAD өнімінің жасаушысының үстазы.

Машина жасау саласында негізгі платформалар PDM, PLM, MRP және ERP жүйелері болды. Олар деректер мен процестерді басқаруда орталық рөл атқарады, CAx жүйелерінен (CAD, CAM, CAE) ақпарат жинаап, өнім құрылымына (BOM: eBOM, pBOM, mBOM) негізделген жобалау қызметін үйімдастырады (Сурет 3.118). Мұндаидын интеграция қателер санын азайтуға, деректердің дублирленуінен аулақ болуға және жобалаудан өндіріс кезеңдеріне дейінгі барлық сатыларда үздіксіз бақылауды қамтамасыз етуге мүмкіндік береді.



Сурет 3.119 BOM тарихы 1960-шы жылдарда CAx жүйелерінен деректерді құрылымдау және оларды басқару жүйелеріне беру тәсілі ретінде пайда болды.

CAD шешімдерінің жетекші вендорының BOM тәсіліне негізделген бұрынғы Pro-E командасы әзірлеген шешімді сатып алуы 2002-2003 жылдары BIM бойынша сериялы Whitepaper жариялануымен бірден белгіленді [60][61]. 2000-шы жылдардың ортасынан бастап құрылыш саласында BIM концепциясының белсенді түрде ілгерілей бастауы параметрлік бағдарламалық қамтамасыз етуге қызығушылықты айтартылған арттырды. Таңымалдылығы соншалық, Pro-E машина жасау саласынан шыққан параметрлік CAD, осы вендор арқылы ілгерілетілген, архитектуралық және құрылыш жобалауы сегментінде бәсекелестерді факті бойынша ығыстырды (Сурет 3.120). 2020-шы жылдардың басында ол BIM (CAD) нарығында жаһандық үстемдікті факті бойынша бекітті [62]. -



Сурет 3.120 Google-дағы іздеу сұрауының таңымалдылығы (RVT versus IFC): BOM-BIM қолдауымен бұрынғы Pro-E командасы жасаған параметрлік CAD әлемнің көптеген елдерінде таңымалдылыққа ие болды.

Соңғы 20 жылда BIM аббревиатуrasesы көптеген түсіндірмелермен толықты, олардың көпмағыналығы 2000-шы жылдардың басында пайда болған бастапқы маркетингтік концепцияларға тамыр жайды. ISO 19650 стандарты, терминнің таңымалдылығында маңызды рөл атқарған, BIM-ге "ғылыми негізделген" ақпаратты басқару тәсілі мәртебесін факті бойынша бекітті. Алайда, стандарттың мәтінінде, BIM-ді пайдалана отырып, объектілердің өмірлік цикліндегі деректерді басқаруға арналған, BIM аббревиатуrasesы аталғанымен, нақты анықтама берілмейді.

2002 [60] және 2003 [61] жылдары BIM бойынша сериялы Whitepaper жариялаған вендордың ресми сайтында 1990-шы жылдары Pro-E машина жасау бағдарламалық қамтамасыз етуде бұрын қолданылған BOM (Bills of Materials) және PLM (Product Lifecycle Management) концепциялары бойынша маркетингтік материалдар факті бойынша қайта жарияланды [63].

Гимараттардың ақпараттық моделін жасау - 2002 жылы [CAD вендорының атауы] компаниясы ұсынған жобалау, құрылыш және гимараттарды басқарудағы жаңа инновациялық тәсіл, сала мамандарының әлем бойынша технологияларды жобалау, құрылыш және басқаруда қалай қолдануға болатыны туралы көзқарастарын өзгертті. – Whitepaper BIM, 2003 [61]

Осы ерте жарияланымдарда BIM тікелей орталықтандырылған интеграцияланған деректер база-

сымен байланысты болды. 2003 жылғы Whitepaper-де айтылғандай, BIM - бұл ғимарат туралы ақпаратты басқару, мұнда барлық жаңартулар бір ғана қоймада жүзеге асырылады, барлық сызбалар, кесулер және спецификациялардың (BOM – Bills of Materials) синхронизациясын қамтамасыз етеді.

BIM ғимарат туралы ақпаратты басқару ретінде сипатталады, мұнда барлық жаңартулар мен өзгерістер деректер базасында жүзеге асырылады. Осылайша, схемалармен, кесінділермен немесе парақтардағы сызбалармен жұмыс істесеңіз де, бәрі әрқашан үйлестірілген, келісілген және өзекті болып табылады. – CAD вендорының сайты, BIM туралы ақ қағаз, 2003 жыл

Жобалауды біртұтас интеграцияланған деректер базасы арқылы басқару идеясы 1980-жылдары кеңінен талқыланды. Мысалы, Чарльз Истманның BDS концепциясы "деректер базасы" терминін 43 рет қолданды. 2004 жылға қарай BIM материалдарында бұл сан екі есе азайып, 2002 жылғы ақ қағазда 23-ке жетті. Ал 2000-жылдардың ортасына қарай деректер базалары тақырыбы вендорлардың маркетингтік материалдарынан және цифрландыру күн тәртібінен дерлік жоғалып кетті.

Деректер базасы және оған қолжетімділік бастапқыда BIM жүйесінің ядросы ретінде қарастырылғанымен, уақыт өте келе акцент геометрияға, визуализацияға және 3D-ге ауысты. 1994 жылы IFC стандартының тіркеушісі, 2002 жылы BIM туралы ақ қағазды жариялаған вендор, 2000-жылдардың басындағы ақ қағазда IGES, STEP және IFC сияқты нейтралды форматтардың шектеулігін және CAD деректер базаларына тікелей қол жеткізуіндегі қажеттілігін атап өтті:

Әр түрлі қосымшалар үйлесімді болмауы мүмкін, ал қайта енгізілген деректер дәлсіз болуы мүмкін. Традициялық автоматтандырылған жобалаудың нәтижесі: шығындардың артуы, нарыққа шығу уақытын ұлғайту және өнім сапасының төмендеуі. Бұғынгі күні барлық негізгі қосымшалар деректерді төмен деңгейде алмасу үшін стандартты салалық интерфейстерді пайдаланады. IGES немесе жаңа STEP стандарттарын пайдалана отырып, әртүрлі өндірушілердің қосымшалары арасында деректер алмасу арқылы пайдаланушылар өз саласында үздік өнімдер арасында белгілі бір деректер үйлесімділігіне қол жеткізе алады. Алайда IGES және STEP тек төмен деңгейде жұмыс істейді, және олар қазіргі заманғы жетекші қосымшалармен жасалатын бай деректермен алмайды. Бұл және басқа стандарттар күн сайын жетілдірілсе де, олар әрқашан қазіргі өндірушілердің өнімдеріне қатысты деректердің байлығында артта қалатын болады. Қосымша бағдарламалар деректерді алмасу және олардың байлығын сақтау үшін нейтралды трансляторларға, мысалы, IGES, STEP немесе IFC-ге жүгінбей, CAD деректер базасына тікелей қол жеткізуі керек. – CAD вендорының ақ қағазы (IFC, BIM) «Интеграцияланған жобалау және өндіріс: Артықшылықтар мен негіздеме», 2000

Осылайша, 1980-жылдары және 2000-жылдардың басында CAD ортасындағы цифрлық жобалаудың негізгі элементі деректер базасы болып саналды, ал файл форматы немесе нейтралды IFC

форматы емес. Трансляторлардан бас тарту және қосымшаларға деректерге тікелей қол жеткізуді қамтамасыз ету ұсынылды. Алайда, шын мәнінде, 2020-жылдардың ортасына қарай BIM концепциясы «бөліп ал да, биле» стратегиясына ұқсас болып, бағдарламалық қамтамасыз етуші жеткізушилердің мүдделері, жабық геометриялық ядроларды пайдаланатын, ашық ақпарат алмасуды дамыту емес, басымдыққа ие болды.

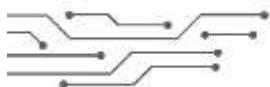
Бүгінгі таңда BIM құрылыш индустрисының ажырамас бөлігі ретінде қабылданады. Алайда соңғы екі онжылдықта деректерді интеграциялау мен жеңілдетілген өзара әрекеттесу уәделері көп жағдайда жүзеге аспады. Көптеген шешімдер әлі де жабық форматтарға немесе нейтралды форматтарға, сондай-ақ мамандандырылған құралдарға тәуелді. BIM-нің пайда болу тарихы, open BIM және IFC, сондай-ақ интероперабельділік пен геометриялық ядролар мәселелерін біз «CAD және BIM: маркетинг, шындық және құрылыш жобалау деректерінің болашағы» атты кітаптың алтыншы бөлімінде қарастырамыз.

Бүгінгі күні салада негізгі сывн-тегеурін – CAD (BIM) дәстүрлі түсінігін модельдеу құралы ретінде пайдалануынан толыққанды деректер базасы ретінде пайдалануына көшу. Бұл ақпаратпен жұмыс істеудің жаңа тәсілдерін, жабық әкімшілдікке тәуелділіктен бас тартуды және ашық шешімдерді енгізуі талап етеді.

CAD деректер базаларына қол жеткізуге мүмкіндік беретін кері инженерия құралдарының дамуы, сондай-ақ Open Source және LLM-технологияларының таралуы арқасында құрылыш саласындағы пайдаланушылар мен әзірлеушілер бағдарламалық қамтамасыз етуші жеткізушилердің анық емес терминдерінен алысталп барады. Оның орнына назар шын мәнінде маңызды нәрсеге: деректерге (деректер базалары) және процестерге ауысады.

Заманауи аббревиатуралар мен визуализациялардың артында деректерді басқарудың стандартты практикалары: сақтау, беру және түрлендіру – яғни классикалық ETL (Extract, Transform, Load) процесі жатыр. Басқа салалардағыдей, құрылыш цифрландыруы тек алмасу стандарттарын ғана емес, сонымен қатар әртүрлі ақпаратпен құрылымды жұмыс істеуді талап етеді.

CAD (BIM) деректерінің әлеуетін толыққанды пайдалану үшін компанияларға ақпаратты басқару тәсілдерін қайта қарастыру кажет. Бұл құрылыш саласының мамандары күн сайын жұмыс істейтін деректерді унификациялау, стандарттау және мағыналы құрылымдау – цифрлық трансформацияның негізгі элементіне әкеледі.



ТАРАУ 3.2. ДЕРЕКТЕРДІ БІРІКТІРУ ЖӘНЕ СТРУКТУРАЛАУ

Құрылыш саласындағы жүйелерді деректермен толтыру

Ірі корпорациялар немесе орташа компаниялар болсын, мамандар күн сайын әртүрлі интерфейстермен бағдарламалық жүйелер мен деректер базаларын әртүрлі форматтағы ақпаратпен толтырумен айналысады, бұл менеджерлер арқылы бір-бірімен үйлесімді әрекет етуі тиіс. Нәтижесінде, осы өзара әрекеттесетін жүйелер мен процестер кешені компанияға табыс пен пайда әкеледі.

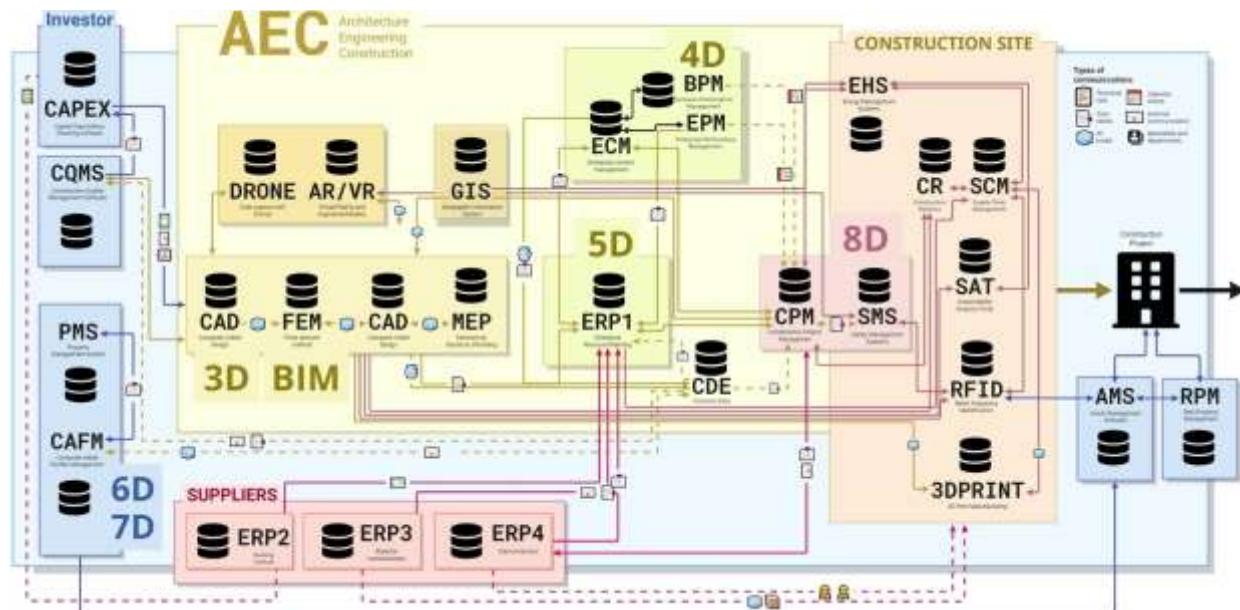


Рис. 3.21 Құрылыш бизнесінде қолданылатын әрбір жүйе немесе қосымша негізінде танымал RDBMS деректер базаларының бірін пайдаланады.

Бұрын айтылған құрылыш саласында қолданылатын жүйелердің әрбір категориясы өздерінің функционалдық рөліне сәйкес деректердің өз типтерімен жұмыс істейді. Абстракттылы деңгейден нақты деңгейге көшу үшін біз деректер типтерінен олардың форматтар мен құжаттар түрінде ұсынылуына етеміз.

Бұрын ұсынылған жүйелер тізіміне (Рис. 1.24) біз енді олар жиі жұмыс істейтін нақты форматтар мен құжаттардың түрлерін қосамыз:-

- Инвестор (CAPEX)
- Қаржылық деректер: бюджеттер, шығындар болжамы (құрылымдалған деректер).
- Нарық тенденциялары туралы деректер: нарықты талдау (құрылымдалған және құрылымдалмаған деректер).
- Заңды және келісімшарттық деректер: келісімшарттар (мәтіндік деректер).
-

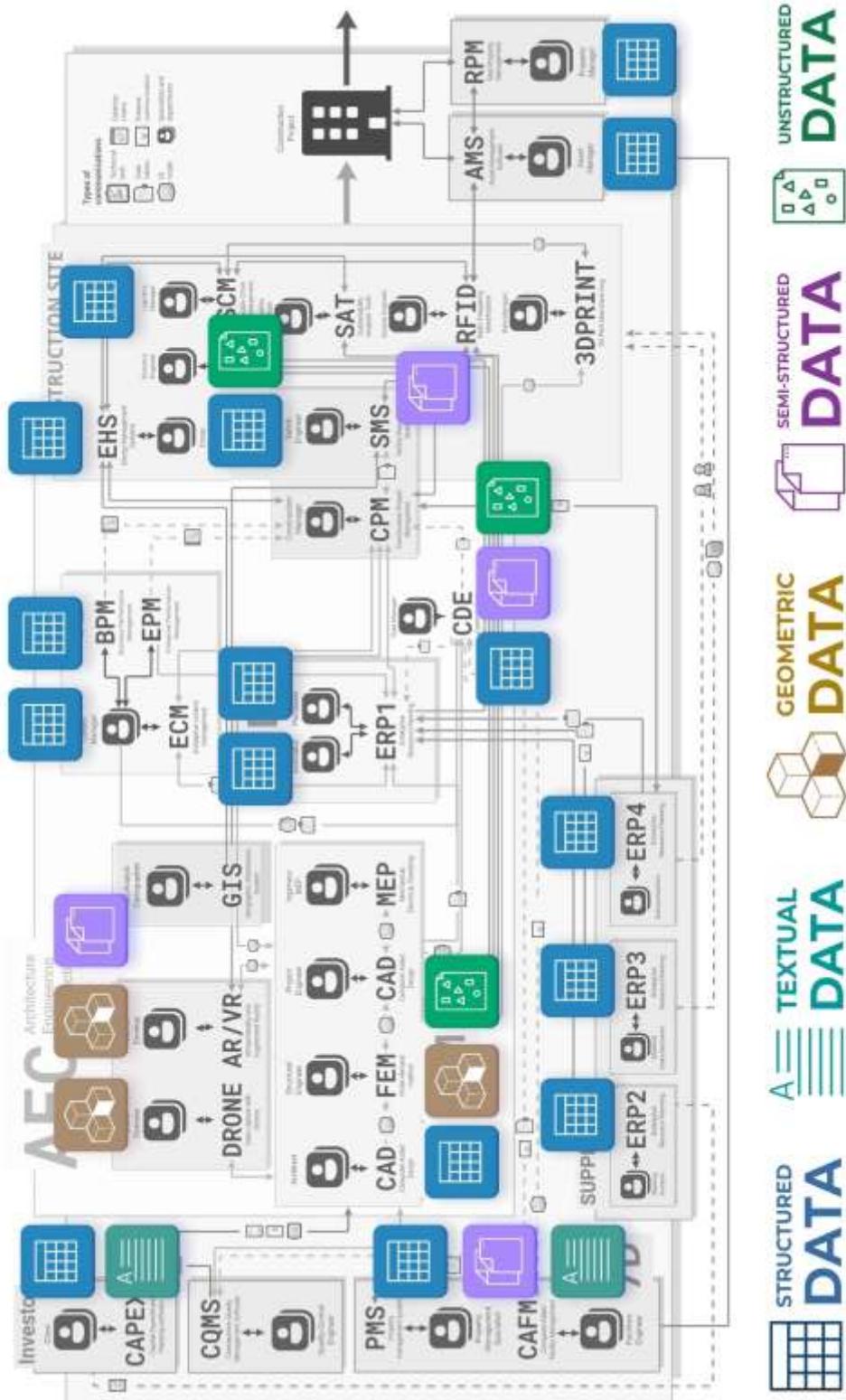


Рис. 3.22 Құрылым саласында әртүрлі интерфейстермен жұмыс істейтін көптеген жүйелер бар, олар әртүрлі деректер типтерімен жұмыс істейді.

■ Басқару жүйелері (PMS, CAFM, CQMS)

- Жоба деректері: графіктер, тапсырмалар (құрылымдалған деректер).
- Нысандың техникалық қызмет көрсету деректері: техникалық қызмет көрсету жоспарлары (мәтіндік және жартылай құрылымдалған деректер).
- Сапа бақылау деректері: стандарттар, тексеру есептері (мәтіндік және құрылымдалмаған деректер).
- CAD, FEM және BIM
- Техникалық сыйбалар: архитектуралық, құрылымдық жоспарлар (геометриялық деректер, құрылымдалмаған деректер).
- Ғимарат модельдері: 3D модельдер, материалдар туралы деректер (геометриялық және жартылай құрылымдалған деректер).
- Инженерлік есептеулер: жүктеме анализі (құрылымдалған деректер).
- Құрылыш алаңын басқару жүйелері (EHS, SCM)
- Қауіпсіздік және денсаулық деректері: қауіпсіздік хаттамалары (мәтіндік және құрылымдалған деректер).
- Жеткізу тізбегі деректері: қорлар, тапсырыстар (құрылымдалған деректер).
- Құнделікті есептер: жұмыс уақыты, өнімділік (құрылымдалған деректер).
- Дрондар, AR/VR, ГИС, 3D басып шығару
- Геодеректер: топографиялық карталар (геометриялық және құрылымдалған деректер).
- Нақты уақыт деректері: бейнелер мен фотосуреттер (құрылымдалмаған деректер).
- 3D басып шығару үшін модельдер: цифровые сыйбалар (геометриялық деректер).
- Қосымша басқару жүйелері (4D BPM, 5D ERP1)
- Уақыт және шығындар туралы деректер: графіктер, сметалар (құрылымдалған деректер).
- Өзгерістерді басқару: жоба бойынша өзгерістер туралы жазбалар (мәтіндік және құрылымдалған деректер).
- Нәтижелер туралы есеп беру: табыс көрсеткіштері (құрылымдалған деректер).
- Деректерді интеграциялау және байланыс (CDE, RFID, AMS, RPM)
- Деректер алмасу: құжат алмасу, деректер модельдері (құрылымдалған және мәтіндік деректер).
- RFID және бақылау деректері: логистика, активтерді басқару (құрылымдалған деректер).
- Мониторинг және бақылау: нысандардағы датчиктер (құрылымдалған және құрылымдалмаған деректер).

Осылайша, құрылыш саласындағы әрбір жүйе – құрылыш алаңын басқару жүйелерінен бастап пайдалану базаларына дейін – өз ақпарат түрімен жұмыс істейді: құрылымдалған, мәтіндік, геометриялық және т.б. Құнделікті жұмыс істейтін мамандардың "деректер ландшафты" өте әртүрлі. Алайда, форматтарды жай ғана тізу ақпаратпен нақты жұмыс істеудің барлық күрделілігін ашпайды.

Практикада компаниялар деректер, тіпті жүйелерден алынған болса да, "осылайша" пайдалануға дайын емес екендігімен бетпе-бет келеді. Бұл әсіресе мәтіндер, суреттер, PDF құжаттары, CAD файлдары және стандартты құралдармен талдауға қызын басқа форматтарға қатысты. Сондықтан келесі маңызды қадам деректерді трансформациялау – бұл өндеу, талдау, визуализация және шешім қабылдауды тиімді автоматтандыру үшін мүмкін емес процесс.

Деректерді трансформациялау: заманауи бизнес-талдаудың критикалық негізі

Бұгінгі таңда көптеген компаниялар парадоксқа тап болуда: олардың күнделікті процестерінің шамамен 80% әлі де классикалық құрылымдалған деректерге – таныс Excel кестелері мен реляциальық деректер базаларына (RDBMS) сүйенеді. Алайда, осы уақытта компаниялардың цифрлық экожүйесіне түсетін жаңа ақпараттың 80% құрылымдалмаған немесе әлсіз құрылымдалған сипатқа ие. Бұл мәтін, графика, геометрия, суреттер, CAD-модельдер, PDF форматындағы құжаттар, аудио- және бейнемазмұн, электрондық хат алмасу және тағы басқалар.-

Бұдан бөлек, құрылымдалмаған деректердің көлемі жылдам өсуде – жыл сайынғы өсім 55–65% аралығында бағаланады. Мұндай динамика жаңа ақпаратты бар бизнес-процестерге интеграциялауда елеулі қыындықтар туғызады. Бұл әртүрлі форматтағы деректер ағынын елемеу ақпараттық бос орындардың пайды болуына және компанияның цифрлық ортасының басқарылу деңгейінің төмендеуіне әкеледі.



Рис. 3.23 Жыл сайынғы құрылымдалмаған деректер көлемінің өсуі бизнес-процестерге ағынды ақпаратты интеграциялауда қыындықтар туғызады.

Күрделі құрылымсыз және жартылай құрылымдалған деректерді автоматтандыру процестерінде елемеу компанияның ақпараттық ландшафтында елеулі олқылықтарға әкеleу мүмкін. Қазіргі әлемде бақылаусыз және қарқынды ақпарат ағында компаниюларға барлық деректер түрлерімен тиімді жұмыс істеуді қамтитын деректерді басқарудың гибридті тәсілін қолдану қажет.

Деректерді тиімді басқарудың кілті әртүрлі "Вавилон" деректерін (құрылымдалмаған, мәтіндік және геометриялық форматтарды, құрылымдалған немесе әлсіз құрылымдалған деректерді) ұйымдастыру, құрылымдау және класификациялауда жатыр. Бұл процесс хаостық деректер жиынтығын ұйымдастырылған құрылымдарға айналдырып, жүйелерге интеграциялауға мүмкіндік береді, осылайша олардың негізінде шешім қабылдауды мүмкін етеді.-

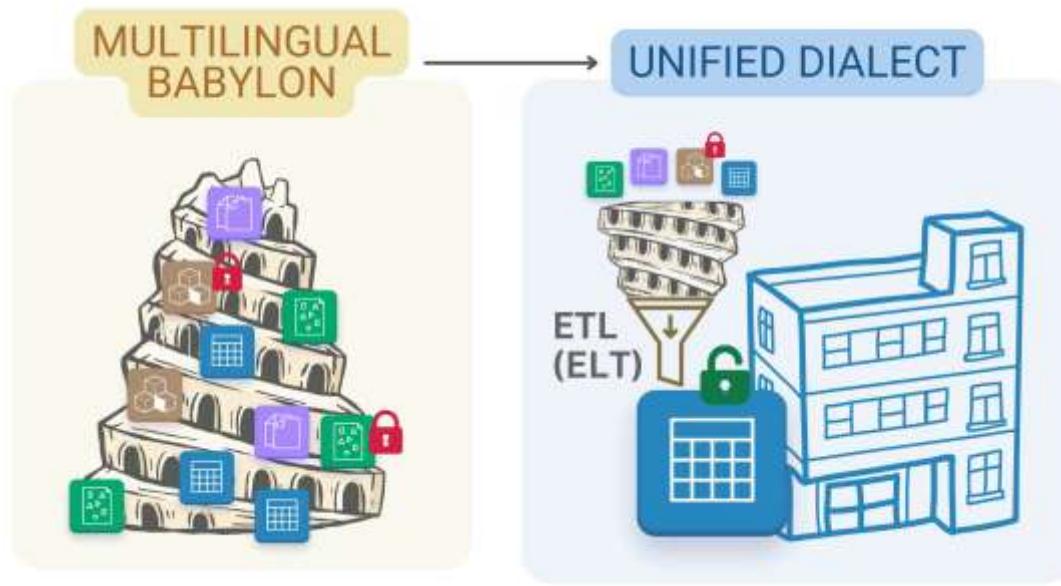


Рис. 3.24 Деректерді басқару бөлімдерінің негізгі міндеті - «Вавилон» әртүрлі және әр форматтағы деректерді құрылымдалған және класификацияланған жүйеге айналдыру.

Осыған үкіс аудиоформаттағы кедергі келтіретін негізгі факторлардың бірі - әртүрлі цифрлық платформалар арасындағы төмен деңгейдегі үйлесімділік, яғни «силос» деп аталатын жүйелер.

Сараптамаға сәйкес, АҚШ-тың ұлттық стандарттар және технологиялар институты (NIST) мәлімдейді, деректердің әртүрлі құрылыш платформалары арасында төменгі үйлесімділігі ақпараттың жоғалуына және елеулі қосымша шығындарға әкеледі. Тек 2002 жылы бағдарламалық қамтамасыз етудің үйлесімділігі мәселелерінен АҚШ-тың капиталдық құрылышында жыл сайын 15,8 миллиард доллар шығын болды, оның екі үшінші бөлігі ғимараттардың иелері мен операторларына, әсіресе пайдалану және қызмет көрсету кезеңінде, тиесілі. Зерттеуде деректер форматтарын стандарттау осы шығындарды азайтуға және объектінің өмірлік циклі кезеңдерінің барлығында жұмыс тиімділігін арттыруға мүмкіндік беретіні атап өтілген.

2016 жылы CrowdFlower жүргізген зерттеуге сәйкес, әлем бойынша 16 мың деректермен жұмыс істейтін маманды қамтыған, негізгі мәселе «лас» және әртүрлі форматтағы деректер болып табылады. Бұл зерттеуге сәйкес, ең құнды ресурс - бұл түпкілікті деректер базалары мен машиналық оқыту модельдері емес, мамандардың ақпаратты дайындауға жұмсайтын уақыты.

Деректерді тазалау, форматтау және үйымдастыру аналитиктер мен деректер менеджерлерінің жұмыс уақытының 60 пайызын алады. Деректер жинау мен іздеу, қажетті деректер жиынтықтарын табу, көбінесе жабық қоймаларда («силос») жасырылған, талдауға қолжетімсіз болып табылады, бұл жұмыс уақытының бесінші бөлігін алады. Тек 9 пайыз уақыт моделдеу, аналитика, болжамдар құру және гипотезаларды тестілеуге жұмсалады. Қалған уақыт коммуникация, визуализация, есеп беру және қосымша ақпарат көздерін зерттеуге кетеді.

Орташа алғанда, деректер менеджерінің жұмысы келесідей бөлінеді (Рис. 3.25): -

- Деректерді тазалау және үйымдастыру (60%): таза және құрылымдалған деректердің болуы аналитиктің жұмыс уақытын едәуір қысқартып, тапсырмаларды орындау процесін

жеделдетуі мүмкін.

- Деректер жинау (19%): деректер ғылымы мамандары үшін негізгі қындықтар релевантты деректер жиынтықтарын табуда. Көп жағдайда компаниялардың деректері хаотикалық түрде үйымдастырылған "силос" түрінде жинақталады, бұл қажетті ақпаратқа қолжетімділікті қынданатады.
- Моделдеу/машиналық оқыту (9%): тапсырыс берушілер тараپынан бизнес мақсаттарының жеткіліксіз анықтылығымен жиі қынданатылады. Міндеттің нақты қойылмауы тіпті ең сапалы модельдің әлеуетін жоюы мүмкін.
- Басқа міндеттер (5%): деректерді өңдеуден басқа, аналитиктер зерттеулермен, деректерді әртүрлі қырынан зерттеумен, нәтижелерді визуализациялар мен есептер арқылы коммуникациялаумен, сондай-ақ процестер мен стратегияларды оңтайландыру бойынша ұсыныстармен айналысусы қажет.

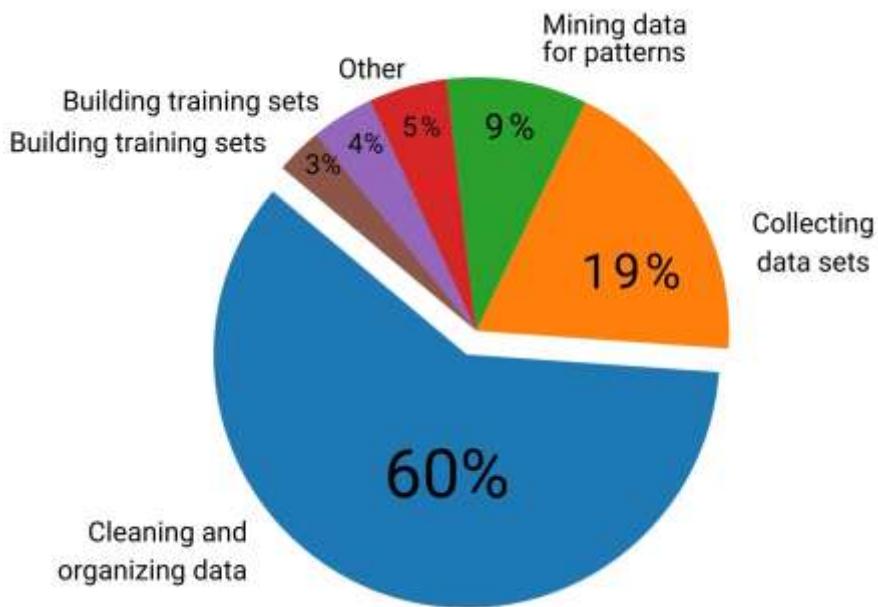


Рис. 3.25 Деректермен жұмыс істейтін менеджерлердің уақытын ең көп жұмсайтын салалар (материалдар бойынша [70]).

Бұл бағалаулар басқа зерттеулермен де расталады. 2015 жылы BizReport-та жарияланған Xplenty зерттеуі бойынша [71], бизнес-аналитика (BI) мамандарының 50%-дан 90%-ға дейінгі уақыты деректерді талдауға дайындауға жұмсалады.

Деректерді тазалау, тексеру және үйымдастыру барлық деректерді өңдеу және аналитика процесстерінің критикалық маңызды негізін құрайды, деректермен жұмыс істейтін мамандардың уақытының 90%-на дейін алады.

Бұл көрінбейтін, бірақ маңызды еңбек, бастапқы деректердегі қателіктер талдау нәтижелерін бұрмалап, адастыруы мүмкін және қымбат басқарушылық қателіктерге әкелуі ықтимал. Сондықтан деректерді тазалау және стандарттау процестері – дубликаттарды жоюдан бастап, бос орындарды толтыруға, өлшем бірліктерін келісуге және жалпы модельге келтіруге дейін – қазіргі цифровық стратегияның негізі болып табылады.

Осылайша, деректерді мүқият трансформациялау, тазалау және стандарттау тек деректермен жұмыс істейтін мамандардың уақытын (деректермен жұмыс істегендегі 80%-ға дейін) алады, сонымен қатар олардың қазіргі бизнес процестерінде тиімді пайдаланылу мүмкіндігін анықтайды. Алайда, деректерді үйымдастыру мен тазалау тек компанияның ақпараттық ағындарын оңтайлы басқару міндетін толық қамтыймайды. Үйымдастыру және құрылымдау кезеңінде ақпаратпен жұмыс істеудің ыңғайлылығы мен тиімділігіне тікелей әсер ететін деректер моделін таңдау маңызды.

Деректер мен бизнес мақсаттары әртүрлі болғандықтан, деректер модельдерінің ерекшеліктерін түсіну және қажетті құрылымды таңдау немесе жасау маңызды. Элементтер арасындағы байланыстарды сипаттау тәсілі мен құрылымдылық дәрежесіне байланысты үш негізгі модель бөлінеді: құрылымдалған, әлсіз құрылымдалған және графтық. Әр модель әртүрлі міндеттер үшін қолайлыша және өз күшті және әлсіз жақтарына ие.

Деректер модельдері: деректердегі қатынастар мен элементтер арасындағы байланыстар

Ақпараттық жүйелердегі деректер әртүрлі үйымдастырылады – сақтау, өңдеу және ақпаратты беру талаптарына байланысты. Деректер модельдерінің түрлері арасындағы негізгі айырмашылық ақпараттың сақталу формасында құрылымдылық дәрежесі мен элементтер арасындағы өзара байланыстарды сипаттау тәсілінде жатыр.

Құрылымдалған деректер нақты және қайталанатын схеманы қамтиды: олар фиксирулған бағандары бар кестелер түрінде үйымдастырылған. Мұндай формат болжамдылықты, өңдеудің қаралайымдылығын және SQL сұрауларын, сүзуді және агрегацияны орындауда тиімділікті қамтамасыз етеді. Мысалдар – деректер базалары (RDBMS), Excel, CSV.

Әлсіз құрылымдалған деректер икемді құрылымды қабылдайды: әртүрлі элементтер әртүрлі атрибуттар жиынтығын қамтуы мүмкін және иерархиялық түрде сақталады. Мысалдар – JSON, XML немесе басқа құжаттық форматтар. Бұл деректер күрделі обьектілер мен олардың арасындағы қатынастарды модельдеу қажет болғанда ыңғайлыш, бірақ, екінші жағынан, деректерді талдау мен стандарттауды қынданатады.-

	Data Model	Storage Format	Example
	Relational	CSV, SQL	A table of doors in Excel
	Hierarchical	JSON, XML	Nested door objects inside a room
	Graph-based	RDF, GraphDB	Relationships between building elements

Модель деректер – деректердің жүйеде қалай ұйымдастырылғанын, сақталуын және өндөлүін сипаттайтын логикалық құрылым.

Қолайлы форматты таңдау міндеттерге байланысты:

- Егер сүзу мен аналитиканың жылдамдығы маңызды болса – реляциялық кестелер (SQL, CSV, RDBMS, колонкалық деректер базалары) қолайлы.
- Егер құрылымның икемділігі қажет болса – JSON немесе XML қолданған жөн.
- Егер деректерде күрделі өзара байланыстар болса – графтық деректер базалары көрнекілік пен масштабталуды қамтамасыз етеді.

Классикалық реляциялық деректер базаларында (RDBMS) әрбір объект (мысалы, есік) жолмен, ал оның қасиеттері кестенің бағандарымен ұсынылады. Мысалы, «Есіктер» категориясындағы элементтер кестесі ID, Білктік, Ені, Отқа төзімділік және бөлмені көрсететін Room ID сияқты өрістерді қамтуы мүмкін.

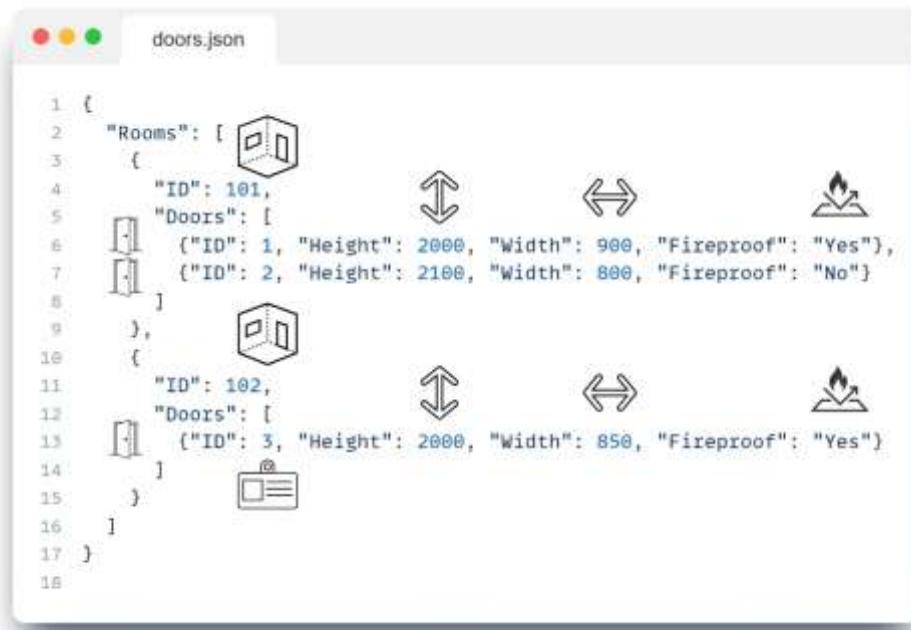
Классикалық реляциялық деректер базаларында (RDBMS) қатынастар кестелер түрінде қалыптасады, мұнда әрбір жазба объектіні, ал бағандар оның параметрлерін білдіреді. Кестелік форматта жобадағы есіктер туралы деректер осылай көрінеді, мұнда әр жол жеке элемент - есік, оның бірегей идентификаторы мен атрибуттары, ал бөлмемен байланыс «Room ID» параметрі арқылы жүзеге асырылады.



Door ID	Room ID	Height (mm)	Width (mm)	Fireproof
ID1001	101	2000	900	Yes
ID1002	101	2100	800	No
ID1003	102	2000	850	Yes

Жобадағы «Есіктер» категориясындағы үш элемент туралы ақпарат кестелік құрылымдалған формада.

Әлсіз құрылымдалған форматтарда, мысалы, JSON немесе XML, деректер иерархиялық немесе кіріктірілген түрде сақталады, мұнда элементтер басқа объектілерді қамтуы мүмкін, ал олардың құрылымы өзгеруі мүмкін. Бұл элементтер арасындағы күрделі байланыстарды модельдеуге мүмкіндік береді. Жобадағы есіктер туралы ақпарат, құрылымдалған түрде (Рис. 3.27) жазылған, әлсіз құрылымдалған форматта (JSON) былайша ұсынылады, олар бөлмелер (Rooms – ID) ішінде кіріктірілген объектілерге айналады, бұл иерархияны логикалық түрде көрсетеді.--



```

1 {
2   "Rooms": [
3     {
4       "ID": 101,
5       "Doors": [
6         {"ID": 1, "Height": 2000, "Width": 900, "Fireproof": "Yes"},
7         {"ID": 2, "Height": 2100, "Width": 800, "Fireproof": "No"}
8       ]
9     },
10    {
11      "ID": 102,
12      "Doors": [
13        {"ID": 3, "Height": 2000, "Width": 850, "Fireproof": "Yes"}
14      ]
15    }
16  ]
17 }
18

```

Жобадағы «Есіктер» категориясындағы элементтер туралы ақпарат JSON форматында.

Графтық модельде деректер түйіндер (нұктелер) және олардың арасындағы байланыстар қабырғалар түрінде ұсынылады. Бұл объектілер мен олардың атрибуттары арасындағы күрделі өзара қатынастарды айқын көрсетуге мүмкіндік береді. Дверьдер мен бөлмелер туралы деректер жобасында графтық ұсыныс келесідей көрінеді:

- Түйіндер (нұктелер) негізгі субъектілерді білдіреді: бөлмелер (Room 101, Room 102) және есіктер (ID1001, ID1002, ID1003)
- Қабырғалар (байланыстар) осы субъектілер арасындағы қатынастарды көрсетеді, мысалы, есіктің белгілі бір бөлмеге тиесілігін
- Атрибуттар түйіндерге байланған және субъектілердің қасиеттерін (биіктік, ені, отқа төзімділік есіктер үшін) қамтиды



Рис. 3.29 Жобадағы есіктердің графтық ұсынысындағы субъектілер туралы ақпарат.

Графтық модельде деректердің сипаттамалары, әр бөлме мен әр есік жеке түйіндер болып табылады. Есіктер бөлмелермен қабырғалар арқылы байланысқан, олар есіктің белгілі бір бөлмеге тиесілігін көрсетеді. Сонымен қатар, есіктердің атрибуттары (биіктік, ені, отқа төзімділік) сәйкес түйіндердің қасиеттері ретінде сақталады. Графтық форматтар мен графтық семантиканың құрылымы саласында қалай пайдаланып туралы «Құрылымы саласында семантика мен онтология тақырыбының пайдалануы» бөлімінде толығырақ сөйлесеміз.

Графтық деректер базалары, деректердің өздері емес, олардың арасындағы қатынастар маңызды болған жағдайда тиімді, мысалы, ұсыныс жүйелерінде, маршруттау жүйелерінде немесе объектілерді басқару жобаларында күрделі өзара қатынастарды модельдеу кезінде. Графтық формат жаңа байланыстарды құруды жеңілдетеді, графқа жаңа деректер түрлерін қосуға мүмкіндік береді, сақтау құрылымын өзгертпей. Алайда, реляциялық кестелер мен құрылымдық форматтармен салыстырғанда, графта деректердің қосынша байланыстырылуы жоқ – екі өлшемді базадан графқа деректерді аудару байланыстардың санын арттырмайды және жаңа ақпарат алуға мүмкіндік бермейді.

Деректердің формасы мен схемасы нақты қолдану жағдайына және шешілтін міндеттерге сәйкес келуі тиіс. Бизнестік процестерде тиімді жұмыс істеу үшін, нәтижені мүмкіндігінше жылдам және оңай алу үшін, сол құралдар мен деректер модельдерін пайдалану маңызды.

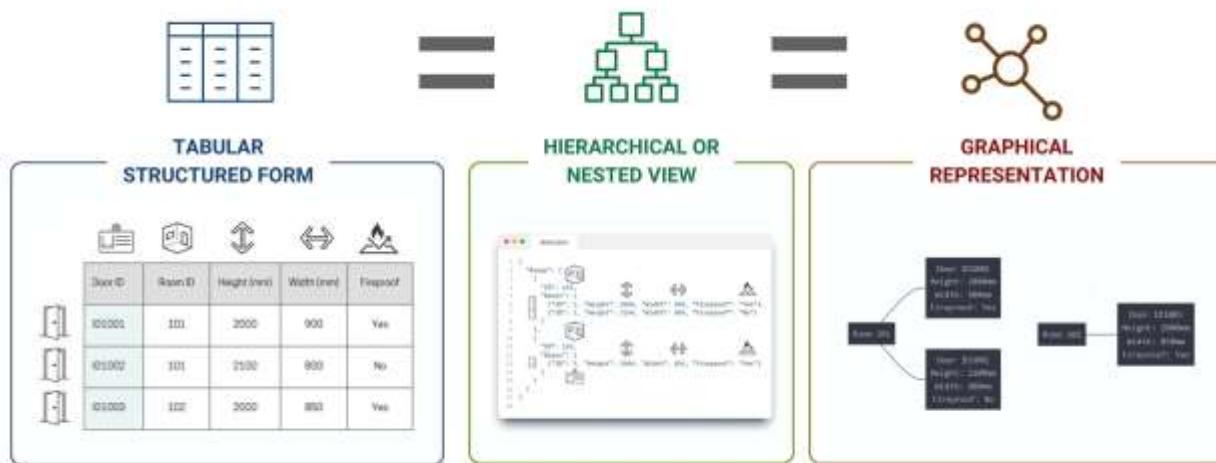


Рис. 3.210 Жобаның элементтері туралы бірдей ақпарат әртүрлі форматтарда әртүрлі деректер модельдері арқылы сақталуы мүмкін.

Бұгінгі күні көптеген ірі компаниялар деректердің артық күрделілігімен бетпе-бет келеді. Эр жүздеген немесе мыңдаған қосымшалар өз деректер модельн пайдаланады, бұл артық күрделілік тудырады — жеке модель жиі қажеттілікке қарағанда ондаған есе күрделі, ал барлық модельдердің жиынтығы — мыңдаған есе. Мұндай артық күрделілік әзірлеушілер мен соңғы пайдаланушылардың жұмысын айтарлықтай қынданатады.

Мұндай күрделілік компанияның жүйелерін дамыту мен қолдауға елеулі шектеулер қояды. Модельдегі әр жаңа элемент қосымша кодты, жаңа логикаларды енгізуі, мүқият тестілеуді және бұрыннан бар шешімдерге бейімдеуді талап етеді. Мұның бәрі шығындарды арттырады және компаниядағы автоматтандыру командасының жұмысын баяулатады, тіпті қарапайым міндеттерді де қымбат және еңбекқор процестерге айналдырады.

Күрделілік деректер архитектурасының барлық деңгейлерін қамтиды. Реляциялық деректер базаларында бұл кестелер мен бағандардың санының артуымен, көбінесе артықшылықтарымен көрінеді. Объектілік-бағытталған жүйелерде күрделілік көптеген сыныптар мен өзара байланысты қасиеттер арқылы артады. XML немесе JSON сияқты форматтарда ауырлық күрделі кіріктірілген құрылымдар, бірегей кілттер мен ретсіз схемалар арқылы көрінеді.

Артық күрделілік деректер модельдерінің жүйелерді тек тиімділігін төмендетіп қана қоймай, сонымен қатар соңғы пайдаланушылар мен болашақтағы үлкен тілдік модельдер мен LLM агенттері үшін түсінуді қынданатады. Деректер модельдерін түсіну және өңдеу мәселесі деректерді жеткілікті қарапайым пайдалануға қалай мүмкіндік беруге болатынын сұрақты көтереді, осылайша олар шын мәнінде жылдам пайда әкелуге бастайды.

Тиісті деректер модельдерін таңдағанның өзінде, деректерге қолжетімділік шектелсе, олардың практикалық пайдасы күрт төмендейді. Проприетарлық форматтар мен жабық платформалар интеграцияны қынданатады, автоматизацияны күрделендіреді және компанияның өз ақпаратын бақылау мүмкіндігін шектейді, жаңа деректердің силосын ғана емес, вендордың рұқсатымен ғана

алуға болатын жабық силосты да құрады. Мәселенің ауқымын түсіну үшін жабық жүйелердің цифрлық процестерге қалай әсер ететінін қарастыру маңызды.

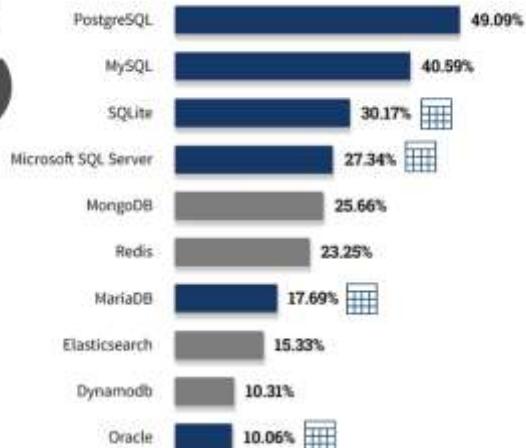
Проприетарлық форматтар және олардың цифрлық процестерге әсері

Құрылыш компанияларының цифрландыру барысында кездесетін негізгі мәселелерінің бірі деректерге шектеулі қолжетімділік болып табылады. Бұл жүйелердің интеграциясын қынданатады, ақпараттың сапасын төмендетеді және тиімді процестерді ұйымдастыруды құрделендіреді. Бұл қындықтардың негізінде көбінесе проприетарлық форматтар мен жабық бағдарламалық шешімдердің қолданылуы жатыр.

Өкінішке орай, құрылыш саласында қолданылатын көптеген бағдарламалар пайдаланушыға деректерді тек өздерінің форматтарында немесе тек шектеулі интерфейстер арқылы қолжетімді болатын бұлттық қоймаларда сақтауға мүмкіндік береді. Сонымен қатар, мұндай шешімдер көбінесе ірі вендорлардың одан да жабық жүйелеріне байланысты құрылады. Нәтижесінде, тіпті ашық архитектураларды ұсынғысы келетін әзірлеушілер де ірі вендорлар белгілеген ережелерді сақтауға мәжбүр.

Қазіргі заманғы құрылыш деректерін басқару жүйелері барған сайын ашық форматтар мен стандарттарды қолдаса да, CAD- (BIM)-құралдарының деректер базалары, сондай-ақ оларға байланысты ERP- және CAFM-жүйелері саладағы цифрлық ландшафта оқшауланған проприетарлық "аралдар" болып қала береді.-

Banking and Finance
Healthcare
E-Commerce
Education
Travel and Hospitality
Telecommunication
Government
Social Media
Real Estate



Construction
Designing



Жабық және проприетарлық деректер сипаты интеграция мен деректерге қолжетімділік үшін кедергілерді туынданатады.

Форматтар мен протоколдардың жабықтығы мен монополизациясы тек құрылыш саласының

мәселесі емес. Көптеген экономиканың секторларында жабық стандарттармен және деректерге шектеулі қолжетімділікпен күрес инновациялардың баяулауымен басталды, жаңа ойыншылардың нарыққа кіруіне жасанды кедергілердің пайда болуымен және ірі жеткізушілерге тәуелділіктің тереңдеуімен сипатталады. Деректердің маңыздылығының жедел өсуі аясында антимонополиялық органдар жаңа цифрлық нарықтармен байланысты шақыруларға жауп беруге үлгермейді, нәтижесінде жабық форматтар мен жабық деректерге қолжетімділік, негізінен, ақпараттың қозғалысын және өсуін тежейтін цифрлық "шекараларға" айналады.

Егер машиналар бізге қажетті барлық нәрсені өндіретін болса, онда біздің жағдайымыз осы иғліктердің қалай бөлінетініне байланысты болады. Эркім өмірден молшылықта ләззат ала алады, тек машиналар өндіретін байлық жалпы меншікте болғанда ғана. Эйтпесе, егер машиналардың иелері байлықты қайта бөлуді сәтті лоббиласа, көшілік адамдар қорқынышты кедейлікте өмір сүрге мәжбүр болады. Қазіргі уақытта, көрініп тұрғандай, жағдай екінші нұсқа бойынша дамып келеді, технологиялар теңсіздікті арттыруда [72].

— Стивен Хокинг, астрофизик, 2015

Monopolies or tight control over critical data formats

Telecommunications:
Proprietary Protocols

1970s-1980s

Computing Industry:
Open Source Movement

1980s

Document Formats:
PDFs and DOCs

Late 1980s to 1990s

Web Browsing:
Browser Wars

Mid-1990s to early 2000s

Media:
Audio and Video Codecs

1990s-2000s

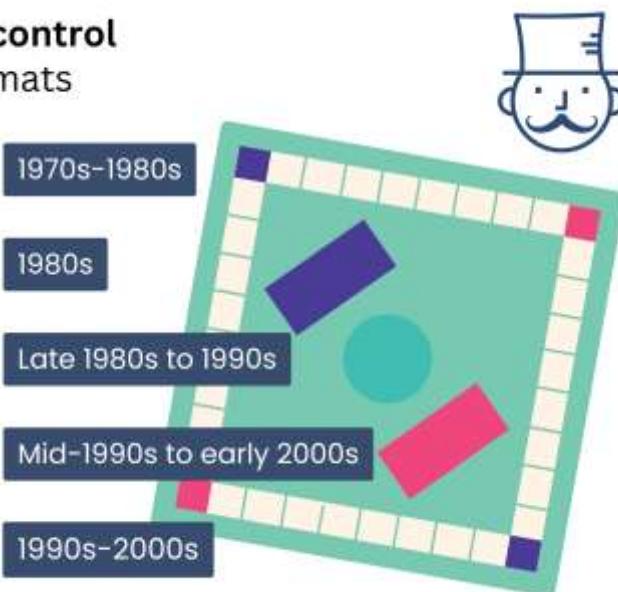


Рис. 3.212 Кілтті деректер форматтары мен протоколдарына монополиялық иелік ету құрылышынан саласының ерекше мәселесі емес.

Нәтижесінде, бағдарламалық базаларға жабық қолжетімділік салдарынан, деректер менеджерлері, аналитиктер, IT мамандары және құрылышынан саласында деректерге қол жеткізу, өндеу және автоматтандыру үшін қосымшалар жасайтын әзірлеушілер көптеген бағдарламалық қамтамасыз етуші жеткізушілерге тәуелділіктермен бетпе-бет келеді (Рис. 3.213). Бұл тәуелділіктер қосымша деңгейдегі қолжетімділікті талап етеді, арнағы API байланыстары мен арнағы құралдар мен бағдарламалық қамтамасыз етуді жасауды қажет етеді.-

API (Қолданбалы Бағдарламалу Интерфейсі) – бұл бір бағдарламаның екіншімен өзара әрекеттесуіне мүмкіндік беретін формализденген интерфейс, деректер мен функционалдылықты алмасу үшін бастапқы кодқа қол жеткізу дің қажеті жоқ. API сыртқы жүйе қандай сұраулар жасай алатынын, олар қандай форматта болуы керектігін және қандай жауап алатынын сипаттайтыны. Бұл бағдарламалық модульдер арасындағы стандартталған «келисім».

Жабық шешімдерге тәуелділіктің көптігі код архитектурасы мен бизнес процестер логикасын «спагетти архитектурасына» айналдырады, бұл бағдарламалық қамтамасыз етуші жеткізу шартынан өзгеруі, деректерге сапалы қолжетімділік көрсету саясатына тәуелді құралдардан тұрады.

Жабық шешімдер мен платформаларға тәуелділік тек икемділікті жоғалтуға ғана емес, сонымен қатар нақты бизнес тәуекелдеріне де әкеледі. Лицензиялау шарттарының өзгеруі, деректерге қолжетімділіктің жабылуы, API форматтарының немесе құрылымының өзгеруі – бұлардың барлығы маңызды процестерді блоктауы мүмкін. Кенеттен бір кестені жаңарту бүкіл интеграциялар мен коннекторлар блогын қайта жасауды талап етеді (Рис. 3.213), ал бағдарламалық қамтамасыз етуді немесе оның API-ін жеткізу шартынан өзгеруі жаңартуы компанияның бүкіл жүйесінің тұрақтылығына әлеуетті қауіп төндіреді.-

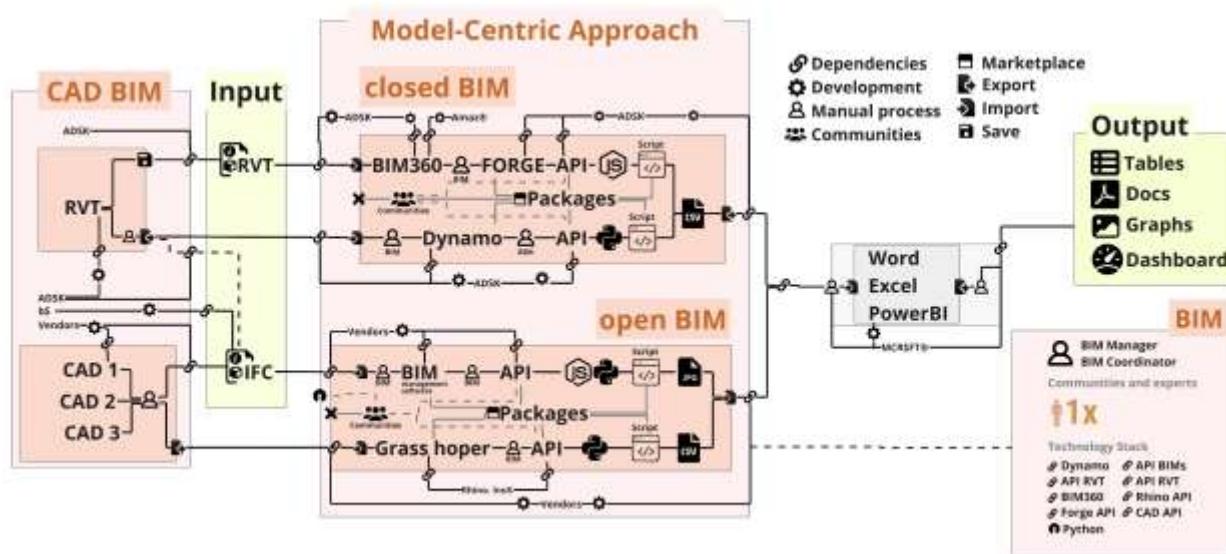


Рис. 3.213 CAD деректерін өңдеудегі көптеген тәуелділіктер құрылышы компаниялары экожүйесінде деректерді интеграциялауға кедергілер туғызады.

Мұндай жағдайда әзірлеушілер мен жүйелік архитектуралар алдын алу үшін емес, өмір сүру үшін жұмыс істеуге мәжбүр. Жаңа шешімдерді енгізу дің орнына – олар бейімделеді. Дамудың орнына – олар үйлесімділікті сақтауға тырысады. Процестерді автоматтандыру мен жылдамдатудың орнына, олар жабық интерфейстерді, API құжаттамасын зерттеуге және кодты қайта құруға уақыт жұмсайды.

Жабық форматтармен және жүйелермен жұмыс істеу – бұл тек техникалық мәселе емес, стратегиялық шектеу. Заманауи автоматтандыру, ИИ, LLM және болжамды аналитика технологиялары ұсынған айқын мүмкіндіктерге қарамастан, көптеген компаниялар олардың әлеуетін толық жүзеге

асыра алмай отыр. Проприетарлық форматтармен (Рис. 3.213) көтерілген кедергілер бизнеске өз деректеріне қол жеткізуі шектейді. Бұл, бәлкім, құрылыш саласындағы цифрлық трансформацияның негізгі ирониясы болып табылады.-

Деректердің ашықтығы мен жүйелердің ашықтығы – бұл сән-салтанат емес, жылдамдық пен тиімділіктің қажетті шарты. Ашықтықтың болмауында бизнес-процестер артық бюрократиямен, көп деңгейлі келісімдер тізбегімен және НiPPO принципіне – ең жоғары жалақы алатын адамның пікірі негізінде шешім қабылдауға тәуелділіктің артуымен толтырылады.

Дегенмен, парадигмалық өзгеріс көкжиегінде қалыптасып келеді. Проприетарлық шешімдердің үстемдігіне қарамастан, көптеген компаниялар Төртінші өнеркәсіп революциясы рухында құрылған архитектуралардың шектеулерін түсіне бастады. Бүтінгі таңда вектор Бесінші революция принциптеріне қарай ауысып, деректер стратегиялық актив ретінде, ашық интерфейстер (API) және жүйелер арасында шынайы интероперабельділік орталықта орналасуда.

Бұл өтпелі кезең жабық әкожүйелерден ашық форматтар, стандарттар және деректер алмасудың ашықтығы бар икемді, модульдік цифрлық архитектураларға көшуін білдіреді.

Ашық форматтар цифрландыруға көзқарасты өзгертеді

Құрылыш саласы деректердің жабықтығы мен проприетарлығына қатысты мәселені соңғы болып көтерді. Басқа экономикалық салалармен салыстырғанда, цифрландыру мұнда баяу дамыды. Себептері – саланың дәстүрлі консерватизмі, бөлшектелген жергілікті шешімдердің басым болуы және қағаз құжат айналымының терең тамырлануы. Онжылдықтар бойы құрылыш саласындағы негізгі процестер физикалық сызбаларға, телефон қонырауларына және синхрондалмаған деректер базаларына сүйенді. Осы контексте жабық форматтар ұзақ уақыт бойы кедергі емес, қалыпты жағдай ретінде қабылданды.

Басқа салалардың тәжірибесі көрсеткендей, жабық деректермен байланысты кедергілерді жою инновациялардың көтерілуіне, дамудың жеделдеуіне және бәсекелестіктің артуына әкеледі. Ғылымда ашық деректер алмасу ашылуларды жеделдетуге және халықаралық ынтымақтастықты дамытуға мүмкіндік береді. Медицинада – диагностика мен емдеудің тиімділігін арттырады. Бағдарламалық инженерияда – бірлескен шығармашылық әкожүйелерін құруға және өнімдерді жылдам жетілдіруге мүмкіндік береді.

McKinsey-дің «Ашық деректер: Ақпараттық ағындар арқылы инновацияларды және өнімділікті ашу» 2013 жылғы есебіне сәйкес, ашық деректер жыл сайын құрылыш, көлік, денсаулық сақтау және энергетика сияқты жеті негізгі экономикалық салада 3-тен 5 трлн долларға дейін құлпын ашуға қабілетті. Сол зерттеуге сәйкес, орталықтандырылмаған деректер әкожүйелері ірі құрылыш компаниялары мен мердігерлерге бағдарламалық қамтамасыз етуді әзірлеу және қолдау шығындарын азайтуға, цифрлық технологияларды енгізуі жеделдетуге мүмкіндік береді.

Ашық архитектураларға көшу, басқа экономикалық секторларда бұрыннан басталған, біртінде құрылыш саласын да қамтыв жатыр. Үлкен компаниялар мен мемлекеттік тапсырыс берушілер, әсіресе құрылыш жобаларына инвестицияларды бақылап отырған қаржы үйімдары, ашық деректерді пайдалану және есептеулер, калькуляциялар мен қосымшалардың бастапқы кодының

қолжетімділігін қамтамасыз ету талаптарын жиі көтеруде. Дамытушылар тек цифрлық шешімдерді жасап, жобаның соңғы сандарын көрсетуі жеткіліксіз — олардан ашықтық, қайта өндіруге мүмкіндік және үшінші тарап қосымшаларына тәуелсіздік күтіледі.

Ашық бастапқы кодпен шешімдерді пайдалану тапсырыс берушілерге сыртқы әзірлеушілер ынтымақтастықты тоқтатқан немесе жобадан кеткен жағдайда, құралдар мен жүйелердің әрі қарай дамуына әсер етпейтіндігіне сенімділік береді. Ашық деректердің басты артықшылықтарының бірі — әзірлеушілердің деректерге қол жеткізу үшін нақты платформаларға тәуелділігін жою қабілеті.

Егер компания толықтай меншікті шешімдерден бас тарта алмаса, реверс инжинириング әдістерін пайдалану мүмкін компромисс болып табылады. Бұл заңды және техникалық тұрғыдан негізделген әдістер жабық форматтарды қолжетімді, құрылымдалған және интеграциялауға жарамды форматтарға түрлендіруге мүмкіндік береді. Бұл әсіресе ескірген жүйелермен байланыс орнату немесе ақпаратты бір бағдарламалық ландшафттан екіншісіне көшіру қажет болған жағдайда маңызды.

Ашық форматтарға көшу және құрылыш саласында реверс инжинириング (меншікті жүйелерді заңды түрде бұзу) қолданудың жарқын мысалдарының бірі — автоматтандырылған жобалау жүйелерінде (CAD) кеңінен қолданылатын DWG форматының ашылуы үшін күрес тарихы. 1998 жылы бағдарламалық қамтамасыз етудің бір вендорының монополиясына жауап ретінде, басқа 15 CAD вендоры «Open DWG» атты жаңа альянс құрды, оның мақсаты әзірлеушілерге DWG форматымен (суреттерді беру стандарттары) жұмыс істеу үшін тегін және тәуелсіз құралдар ұсыну болды, меншікті бағдарламалық қамтамасыз етуді немесе жабық API-ді пайдаланбай. Бұл оқиға, 1980-ші жылдардың соңынан бастап бүгінгі күнге дейін танымал CAD шешімінің жабық форматына еркін қол жеткізуге мүмкіндік беріп, бәсекелестікті дамытуға ықпал еткен, ондаған мың компания үшін шешуші сәт болды. Бүгінгі таңда 1996 жылы алғаш рет құрылған «Open DWG» SDK, DWG форматында импорттау, өңдеу және экспорттау мүмкіндігі бар барлық шешімдерде, DWG форматты әзірлеушінің ресми қосымшасынан тыс, қолданылады.

Үқсас трансформациялар басқа технологиялық алыптарда да мәжбүрлі түрде орын алада. Microsoft, бұрын меншікті тәсілдің символы,.NET Framework-тің бастапқы кодын ашып, Azure бұлттық қызметінің инфрақұрылғында Linux-ты пайдалана бастады және Open Source қауымдастығында позицияларын ұйғайту үшін GitHub-ты сатып алды. Meta (бұрынғы Facebook) ИИ агенттерін әзірлеу саласындағы инновациялар мен ынтымақтастықты қолдау үшін Llama сериялы ашық бастапқы кодты ИИ модельдерін шығарды. Генеральды директор Марк Цукерберг ашық бастапқы кодты платформалардың алдағы онжылдықта технологиялық прогрестің көшбасшысы болатынын болжайды.

Open Source — бұл бағдарламалық қамтамасыз етуді әзірлеу және тарату моделі, мұнда бастапқы код еркін пайдалану, зерттеу, модификациялау және тарату үшін ашық.

Ашық деректер мен ашық кодты шешімдер тек тренд емес, цифрлық тұрақтылықтың негізіне айналуда. Олар компанияларға икемділік, өзгерістерге тәзімділік, өз шешімдері бойынша бақылау және цифрлық процестерді жеткізу шілердің саясатына тәуелді болмай масштабтау мүмкіндігін береді. Сонымен қатар, олар бизнестің XXI ғасырдың ең құнды ресурсы — өз деректері бойынша бақылауды қайтара алады.

Парадигманың өзгеруі: Ашық Код бағдарламалық қамтамасыз етуші вендорлардың үстемдік дәуірінің аяқталуы

Құрылыш индустриясын монетизациялаудың дәстүрлі тәсілімен жүзеге асырылмайтын өзгеріс күтіп тұр. Деректерді пайдалану, деректерге негізделген тәсіл және ашық кодты құралдарды қолдану бағдарламалық қамтамасыз ету нарығының алыптарын ұстап тұрған ойын ережелерін қайта қарауға әкеледі.

Алдыңғы технологиялық трансформациялардан айырмашылығы, бұл өтпелі кезең вендорлар тараپынан белсенді түрде ілгерілетілмейді. Парадигманың өзгеруі олардың лицензиялау, жазылымдар және консалтингке негізделген дәстүрлі бизнес модельдеріне қауіп тәндіреді. Жаңа шындық «қораптағы» дайын өнім немесе ақылды жазылымды талап етпейді – ол процестер мен ойлауды қайта құруды қажет етеді.

Ашық технологияларға негізделген деректерге бағытталған шешімдерді басқару және дамыту үшін компанияларға ішкі процестерді қайта қарау қажет болады. Әр түрлі бөлімдердің мамандары тек өзара әрекеттесіп қана қоймай, бірлескен жұмысқа деген көзқарастарын қайта қарауы тиіс.

Жаңа парадигма ашық деректер мен ашық кодты шешімдерді пайдалануды білдіреді, мұнда бағдарламалық кодты жасауда программистердің орнына жасанды интеллект пен үлкен тілдік модельдер (LLM) негізіндегі құралдар маңызды рөл атқарады. 2024 жылдың ортасына қарай Google-да жаңа кодтың 25%-дан астамы ИИ көмегімен жасалатын болады. Болашақта LLM-мен кодтау барлық жұмыстың 80%-ын 20%-ға жуық уақыт ішінде орындауға мүмкіндік береді.

McKinsey 2020 жылғы зерттеуге сәйкес, аналитика саласында CPU-ның орнына GPU-лар белсенді түрде қолданылуда – олардың жоғары өнімділігі мен заманауи ашық кодты құралдармен қолдауы арқасында. Бұл компанияларға деректерді өңдеуді қыимбат бағдарламалық қамтамасыз етуге немесе тапшылыққа ұшыраған мамандарды жалдауға үлкен инвестицияларсыз жеделдетуге мүмкіндік береді.

McKinsey, PwC және Deloitte сияқты жетекші консалтингтік компаниялар әртүрлі салаларда ашық стандарттар мен ашық кодты қосымшалардың өсіп келе жатқан маңыздылығын атап өтеді.

PwC Open Source Monitor 2019 жылғы есебіне сәйкес, 100-ден астам қызметкері бар компаниялардың 69%-ы ашық кодты шешімдерді саналы түрде пайдаланады. Әсіресе, OSS ірі компанияларда белсенді қолданылады: 200-499 қызметкері бар компаниялардың 71%-ы, 500-1999 қызметкері бар компаниялардың 78%-ы және 2000-нан астам қызметкері бар компаниялардың 86%-ы. 2023 жылғы Synopsys OSSRA есебіне сәйкес, талданған код базаларының 96%-ы ашық кодты компоненттерді қамтыды.

Болашақта әзірлеушінің рөлі қолмен код жазудан гөрі деректер модельдерін, ағын архитектурасын жобалау және қажетті есептеулерді сұрау бойынша жасайтын ИИ-агенттерді басқаруға айналады. Пайдаланушы интерфейстері минималистік болады, ал өзара әрекеттесу диалогтық форматта өтеді. Дәстүрлі бағдарламалу жоғары деңгейлі жобалау мен цифрлық шешімдерді оркестрациялауға орын береді (Рис. 3.214). Қазіргі заманғы трендтер – low-code платформалар (Рис. 7.46) және LLM қолдауымен экожүйелер (Рис. 7.44) – IT-жүйелерді әзірлеу мен қолдауға кететін шығындарды едәуір азайтуға мүмкіндік береді.–



Рис. 3.214 Егер бұғінгі күні қосымшалар бағдарламашылармен қолмен жасалса, болашақта кодтың едәуір бөлігі ИИ және LLM негізіндегі шешімдермен генерацияланады.

Бұл өтпелі кезең бұрынғы кезеңдерге ұқсамайды, ірі бағдарламалық қамтамасыз ету жеткізушилері, ең алдымен, оның катализаторлары болмайды.

Гарвард бизнес мектебінің 2024 жылғы «Ашық кодты бағдарламалық қамтамасыз етудің құндылығы» зерттеуіне сәйкес [40], ашық кодты бағдарламалық қамтамасыз етудің жалпы құны екі көзқарас бойынша бағаланады. Бір жағынан, барлық бар Open Source шешімдерін нөлден жасау үшін қанша қаражат қажет болар еді, бұл сома шама мен 4,15 миллиард долларды құрайды. Екінші жағынан, әрбір компанияның ашық кодты шешімдердің аналогтарын өз бетінше әзірлеуі (бұл кең таралған процесс) және бұрыннан бар құралдарға қол жеткізбеген жағдайда, бизнестің жалпы шығындары 8,8 триллион долларға жетеді – бұл сұраныстың құны.

Ешбір ірі бағдарламалық қамтамасыз ету жеткізушиі 8,8 триллион долларлық әлеуетті бағдарламалық қамтамасыз ету нарығын 4,15 миллиард долларға дейін қысқартуға мүдделі емес екенін түсіну қыын емес. Бұл сұраныс көлемінің 2000 есе азаюын білдіреді. Мұндай трансформация жеткізушилер үшін пайдасыз, өйткені олардың бизнес модельдері клиенттердің жабық шешімдерге тәуелділігін ұзақ мерзімді қолдауға негізделген. Сондықтан, «кез келген уақытта» ыңғайлы және ашық шешім ұсынатын компанияларды күтетін үйімдар көнілсіз болуы мүмкін – мұндай жеткізушилер пайда болмайды.

Ашық цифрлық архитектураға көшу жұмыс орындарының немесе табыстың қысқаруын білдірмейді. Керісінше, бұл дәстүрлі лицензиялар мен қораптық бағдарламалық қамтамасыз ету нарығын уақыт өте келе ығыстыра алатын икемді және бейімделгіш бизнес модельдеріне жағдай жасайды.

Лицензияларды сатудың орнына – қызметтер, жабық форматтардың орнына – ашық платформалар, жеткізушиге тәуелділіктің орнына – тәуелсіздік және нақты қажеттіліктерге сәйкес шешімдер құру мүмкіндігі. Бұрын тек құралдарды пайдаланғандар, енді олардың соавторлары бола алады. Ал деректермен, модельдермен, сценарийлермен және логикамен жұмыс істей алатындар жаңа цифрлық экономика орталығында болады. Бұл өзгерістер мен ашық деректер айналасында қалыптасып жатқан жаңа рөлдер, бизнес модельдер мен ынтымақтастық форматтары туралы кітаптың соңғы, оныншы бөлімінде толығырақ сөйлесеміз.

Ашық деректер мен ашық кодқа негізделген шешімдер компанияларға ескірген API-мен құресуге және жабық жүйелерді интеграциялауға емес, бизнес-процестердің тиімділігіне назар аударуға мүмкіндік береді. Ашық архитектураға саналы түрде көшу өнімділікті айтартылған арттыруға және вендорларға тәуелділікті азайтуға мүмкіндік береді.

Жаңа шындыққа көшу – бұл бағдарламалық қамтамасыз етуді әзірлеу тәсілдерін өзгерту ғана емес, сонымен қатар деректермен жұмыс істеу принципін қайта қарау. Бұл трансформацияның орталығында код емес, ақпарат: оның құрылымы, қолжетімділігі және интерпретациялануы. Дәл осы жерде ашық және құрылымдалған деректер алдыңғы қатарға шығып, жаңа цифрлық архитектураның ажырамас бөлігіне айналады.

Структурленген ашық деректер: цифрлық трансформацияның негізі

Егер өткен онжылдықтарда бизнестің тұрақтылығы көбінесе бағдарламалық шешімдерді таңдау мен нақты вендорларға тәуелділікпен анықталса, қазіргі цифрлық экономика жағдайында деректердің сапасы мен олармен тиімді жұмыс істеу қабілеті негізгі факторға айналады. Ашық базастапқы код жаңа технологиялық парадигманың маңызды бөлігі, бірақ оның әлеуеті тек түсінікті, үйімдасқан және машинамен оқылатын деректер болғанда ғана толық ашылады. Деректердің барлық түрлері арасында құрылымдалған ашық деректер тұрақты цифрлық трансформацияның негізі болып табылады.

Құрылымдалған ашық деректердің басты артықшылығы – бірмәнді интерпретация және автоматтандырылған өңдеу мүмкіндігі. Бұл жеке операциялар деңгейінде де, бүкіл ұйым ауқымында да тиімділікті айтартылған арттыруға мүмкіндік береді.

Deloitte компаниясының "Корпоративтік трансформациялардағы деректерді беру процесі" [82] есебіне сәйкес, құрылымдалған деректерді беру үшін IT-мен жұмыс істеу шешуші маңызға ие. 2024 жылғы Ұлыбритания үкіметінің "Мемлекеттік жобаларды жеткізуде деректерді талдау және AI" [83] есебіне сәйкес, әртүрлі жобалар мен ұйымдар арасындағы деректер алмасудағы кедергілерді жою жобаларды басқаруда тиімділікті арттырудың кілтті факторы болып табылады. Құжатта деректер форматтарын стандарттау және ашық деректер принциптерін енгізу ақпаратты қайталауды болдырмауға, уақытты жоғалтуды минимизациялауға және болжамдардың дәлдігін арттыруға мүмкіндік беретіні атап өтілген.

Дәстүрлі түрде жоғары фрагментация мен форматтардың әртүрлілігі басым құрылыш саласында құрылымдау-унификациялау процесі және құрылымдалған ашық деректер келісілген және басқарылатын процестерді қалыптастыруда шешуші рөл атқарады (Рис. 4.114). Олар жобаның

қатысушыларына жабық платформалардың, деректер модельдерінің және форматтарының үйлесімсіздігімен байланысты техникалық мәселелерді шешуге емес, өнімділікті арттыруға назар аударуға мүмкіндік береді.

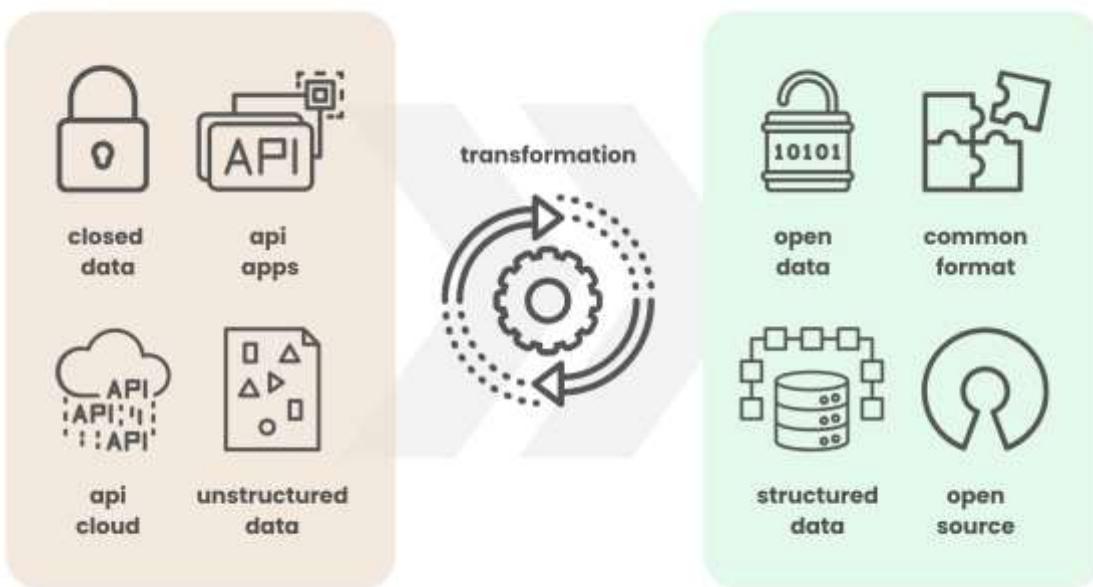


Рис. 3.215 Ашық құрылымдалған деректер бағдарламалық шешімдер мен платформаларға тәуелділікті азайтады және инновацияларды жеделдетеді.

Біз одан әрі кітапта егжей-тегжейлі қарастыратын заманауи технологиялар мен құралдар ақпаратты жинап қана қоймай, автоматты түрде тазартуға: қайталауды жою, қателерді түзету, мәндерді нормализациялауға мүмкіндік береді. Бұл аналитиктер мен инженерлердің шашыраңқы құжаттармен емес, талдауға, автоматтандыруға және шешім қабылдауға жарамды үйімдасқан білім базасымен жұмыс істейтінін білдіреді.

Мүмкіндігінше қарапайым жасаңыз, бірақ одан да қарапайым емес.

– Альберт Эйнштейн, теориялық физик (цитатаның принадлежности даулы [84]).

Бұғынгі таңда деректермен жұмыс істеуге арналған пайдаланушы интерфейстерінің көпшілігін автоматты түрде жасауға болады — әр бизнес жағдайы үшін кодты қолмен жазудың қажеті жоқ. Бұл үшін деректердің құрылымын, моделін және логикасын қосымша нұсқауларсыз түсінетін инфрақұрылымдық қабат қажет (Рис. 4.115). Деректердің құрылымдалған болуы осынданай тәсілді мүмкін етеді: формалар, кестелер, сұзгілер және көріністер минималды бағдарламалау шығындарамын автоматты түрде генерациялануы мүмкін.

Пайдаланушы үшін критикалық маңызды интерфейстер әлі де қолмен өңдеуді талап етуі мүмкін. Бірақ көп жағдайда — бұл 50-ден 90 пайызға дейінгі жұмыс сценарийлері — арнайы қосымшаларды пайдаланбай-ақ автоматты генерациялау жеткілікті (Рис. 3.216), бұл әзірлеу және қолдау шығындарын едәуір азайтады, қателер санын азайтады және цифрлық шешімдерді жүзеге асыруды жеделдетеді.

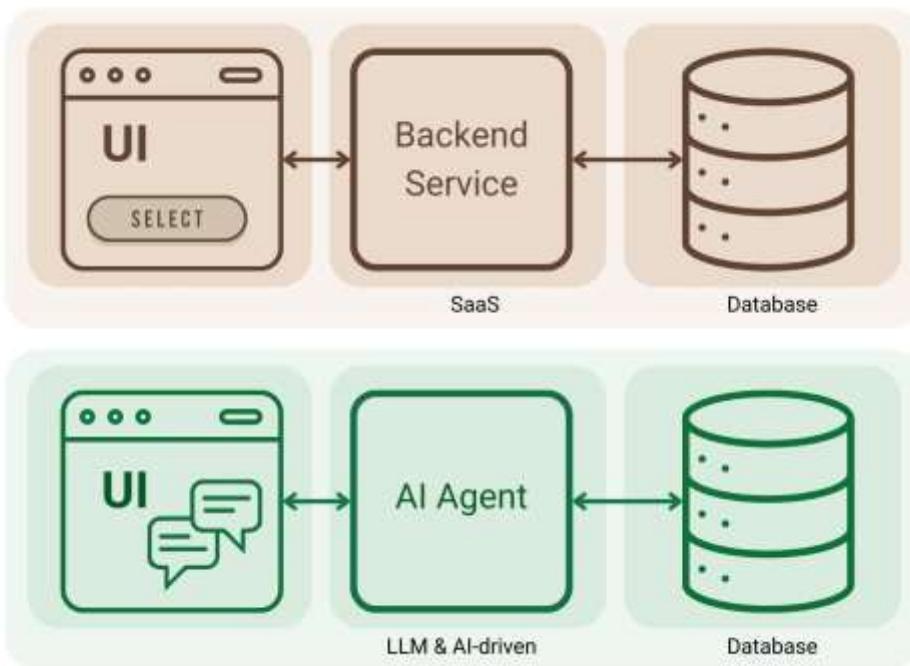
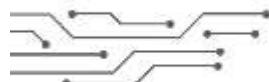


Рис. 3.216 Деректермен жұмыс істеудің архитектуралық модельдері: дәстүрлі қосымша архитектурасы және LLM-ге негізделген AI-бағытталған модель.

Жеке қосымшаларға негізделген архитектурадан тілдік модельдерге (LLM) сүйенетін интеллектуалды басқарылатын жүйелерге көшу – цифрлық эволюцияның келесі қадамы. Мұндай архитектурада құрылымдалған деректер тек сақтау объектісі емес, сонымен қатар контекст негізінде әрекеттерді талдайтын, интерпретациялайтын және ұсынатын ИИ құралдарымен өзара әрекеттесудің негізі болып табылады.

Келесі тарауларда біз ашық құрылымдалған деректерге негізделген архитектуралық модельдерге қарастырамыз, сондай-ақ тілдік модельдердің деректерді автоматты интерпретациялау, валидациялау және өңдеу үшін қалай қолданылатынын көрсетеміз. Бұл практикалық жағдайлар жаңа цифрлық логиканың қалай жұмыс істейтінін және трансформацияға дайын компанияларға қандай артықшылықтар беретінін жақсы түсінуге көмектеседі.



ТАРАУ 3.3. LLM ЖӘНЕ ОЛАРДЫҢ ДЕРЕКТЕРДІ ӨНДЕУ ЖӘНЕ БИЗНЕС-ПРОЦЕСТЕРДЕГІ РӨЛІ

**LLM чаттары: ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok
деректерді өңдеу процестерін автоматтандыру үшін**

Үлкен тілдік модельдердің (LLM) пайда болуы құрылымдалған ашық деректер мен Open Source философиясына қарай қозғалыстың табиғи жалғасы болды. Деректер ұйымдастырылған, қолжетімді және машинамен оқылатын болғанда, келесі қадам – күрделі код жазудың немесе арнайы техникалық білімді меңгерудің қажеті жок, осы ақпаратпен өзара әрекеттесуге қабілетті құралды жасау.

LLM – ашықтықтың тікелей өнімі: үлкен ашық деректер жиынтықтары, жарияланымдар және Open Source қозғалысы. Ашық ғылыми мақалалар, жалпыға қолжетімді мәтіндік деректер және бірлескен әзірлеу мәдениеті болмаса, ChatGPT немесе басқа LLM-дер болмас еді. LLM – бұл белгілі бір мағынада адамзаттың жинақталған цифрлық білімінің «дистилляты», ашықтық принциптері арқасында жиналған және оқытылған.

Заманауи үлкен тілдік модельдер (LLM – Large Language Models), мысалы, ChatGPT® (OpenAI), LLaMa™ (Meta AI), Mistral DeepSeek™, Grok™ (xAI), Claude™ (Anthropic), QWEN™ пайдаланушыларға деректерге табиғи тілде сұрау формулоревкалауға мүмкіндік береді. Бұл ақпаратпен жұмыс істеуді тек әзірлеушілер үшін ғана емес, сонымен қатар аналитиктер, инженерлер, жобалаушылар, менеджерлер және бұрын бағдарламалаудан алыс болған басқа мамандар үшін де қолжетімді етеді.

LLM (Large Language Model) – бұл интернеттен жиналған үлкен деректер негізінде мәтінді түсінуге және генерациялауға үйретілген жасанды интеллект. Ол контекстті талдай алады, сұрақтарға жауап бере алады, диалог жүргізе алады, мәтін жаза алады және бағдарламалық кодты генерациялай алады.

Егер бұрын визуализация, деректерді өңдеу немесе талдау арнайы бағдарламалау тілдерін: Python, SQL, R немесе Scala білуді, сондай-ақ Pandas, Polars немесе DuckDB сияқты кітапханалармен жұмыс істей білуді талап етсе, 2023 жылдан бастап жағдай түбегейлі өзгерді. Енді пайдаланушы не алғысы келетінін жай ғана сипаттай алады – және модель өзі кодты генерациялап, оны орындаپ, кесте немесе график шығарып, нәтижені түсіндіреді. Дауірлер бойы технологиялардың дамуы күрделену жолымен емес, түбегейлі жеңілдету және қолжетімділік жолымен жүріп жатыр.

Бұл принцип – «деректерді сөздермен (промпттармен) өңдеу» – ақпаратпен жұмыс істеудің эволюциясында жаңа кезеңді білдіреді, фактілі түрде шешімдер жасауды одан да жоғары абстракция деңгейіне көтерді. Уақытында пайдаланушылар интернеттің техникалық негіздерін түсінбей-ақ онлайн-дүкендер ашып немесе WordPress, Joomla және басқа модульдік ашық кодты жүйелер арқылы сайттар құра алатын болғаны сияқты (автор 2005 жылдан бері осындай жүйелермен, соның ішінде білім беру және инженерлік онлайн-платформалар саласында жұмыс істейді) – бұл, өз кезегінде, цифрлық контент пен интернет-бизнестің қарқынды өсуіне әкелді – бүгінгі күні инженерлер, аналитиктер және менеджерлер бағдарламалау тілдерін білмей-ақ жұмыс процестерін автоматтандыра алады. Бұған қуатты LLM-дер – тегін және ашық кодты, мысалы, LLaMA, Mistral, Qwen,

DeepSeek және басқалары – заманауи технологияларды ең кең аудиторияға қолжетімді етуде көмектеседі.

Үлкен тілдік модельдер LLM: бұл қалай жұмыс істейді

Үлкен тілдік модельдер (ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok) – интернет, кітаптар, мақалалар және басқа көздерден алынған үлкен мәтін деректер массивтерінде оқытылған нейрондық желілер. Олардың негізгі міндеті – адам сөйлеуінің контекстін түсіну және мағыналы жауаптар генерациялау.

Заманауи LLM-дердің негізінде 2017 жылы Google зерттеушілері ұсынған трансформер архитектуры жатыр. Бұл архитектураның негізгі компоненті – назар механизмі (attention), ол модельге мәтіндегі сөздер арасындағы байланыстарды олардың позициясынан тәуелсіз ескеру мүмкіндігін береді.

LLM-дерді оқыту процесі адамның тіл меңгеруіне алыстап ұқсас – тек миллиондаған есе ауқымды. Модель миллиардтаған сөздер мен сөз тіркестерінің қолданылу мысалдарын талдай отырып, тіл құрылымындағы және мағыналық ауысулар логикасындағы заңдылықтарды анықтайды. Осы кезде барлық мәтін токендерге – минималды мағыналық бірліктерге (сөздер немесе олардың бөліктері) бөлінеді, олар кейін көп өлшемді кеңістікте векторларға түрлендіріледі. Бұл векторлық ұсыныстар машинаға ұғымдар арасындағы жасырын байланыстарды «түсінуге», тек символдар тізбегі ретінде мәтінмен жұмыс іstemей, мүмкіндік береді.-

Үлкен тілдік модельдер – тек мәтін генерациялау құралдары емес. Олар мағынаны тану, ұғымдар арасындағы байланыстарды табу және деректермен жұмыс істей қабілетіне ие, тіпті олар әртүрлі форматтарда ұсынылса да. Негізгі – ақпараттың түсінікті модельдерге бөлініп, LLM-мен жұмыс істей алатын токендер түрінде ұсынылуы.

Сол тәсілді құрылыс жобаларына да қолдануға болады. Жобаны ерекше мәтін ретінде елестетсек, онда әрбір ғимарат, элемент немесе құрылым – бұл токен, біз осындағы ақпаратты өндеуді ұқсас түрде бастай аламыз. Құрылыс жобаларын кітаптармен салыстыруға болады, олар категорияларға, тарауларға және минималды токендерден – құрылыс жобасының элементтерінен тұратын абзацтар топтарына бөлінеді (Рис. 3.31). Деректер модельдерін құрылымдық форматқа аударған кезде, біз құрылымдық деректерді векторлық базаларға (Рис. 8.22) да аудара аламыз, олар мағыналық оқыту және LLM сияқты технологиялар үшін идеалды дереккөз болып табылады.-

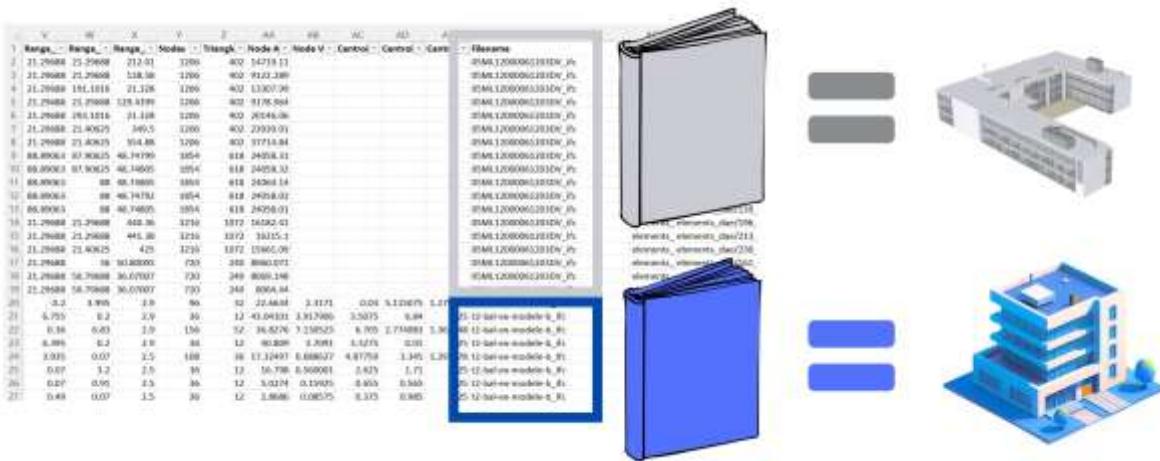


Рис. 3.31 Құрылыс жобасының элементі – мәтіндегі токен сияқты: бүкіл жобаның категориялары (бөлімдері) мен топтарын құрайтын минималды бірлік.

Егер құрылыс жобасы цифрландырылған және оның элементтері токендер немесе векторлар түрінде ұсынылса, онда оларға қатал формалды сұраулар арқылы емес, табиғи тілде жүгінуге мүмкіндік пайда болады. Мұнда LLM-нің негізгі артықшылықтарының бірі – сұраудың мағынасын түсіну және оны сәйкес деректермен байланыстыру қабілеті көрінеді.

Инженер SQL сұрауын немесе Python кодын жазу міндетті емес – ол LLM-нің жұмысын және деректер құрылымын түсініп, тапсырманы әдettегі түрде формулировка жасай алады: "B30 класынан жоғары темірбетон конструкцияларын тауып, олардың жалпы көлемін есепте." Модель сұраудың мағынасын таниды, оны машина оқытын формаға айналдырады, деректерді табады (топтастырады және түрлендіреді) және қорытынды нәтижені қайтарады.

Құжаттар, кестелер, жобалардың модельдері векторлық көріністерге (эмбеддинг) айналдырылады және базада сақталады. Пайдаланушы сұрақ қойған кезде, сұрау да векторға айналады, ал жүйе мағынасы бойынша ең жақын деректерді табады. Бұл LLM-ге тек өздерінің оқытылған білімдеріне ғана емес, сонымен қатар, модельдің оқытуы аяқталғаннан кейін пайда болған өзекті корпоративтік деректерге де сүйенуге мүмкіндік береді.

Құрылыс саласындағы LLM-нің ең маңызды артықшылықтарының бірі – бағдарламалық кодты генерациялау қабілеті. Техникалық тапсырманы программистке беру орнына, мамандар тапсырманы табиғи тілде сипаттай алады, ал модель қажетті кодты жасайды, оны автоматтандыру процестерін жасау үшін (чатаудан көшіріп) пайдалануға болады. LLM модельдері терең бағдарламалау білімдері жоқ мамандарға компанияның бизнес-процессерін автоматтандыру және жақсартуға үлес қосуға мүмкіндік береді.



Рис. 3.32 LLM пайдаланушыларға код жазуға және бағдарламалу дағдыларын менгермей-ақ нәтижелер алуға мүмкіндік береді.

2024 жылы Wakefield Research жүргізген және SAP-тың демеушілігімен өткізілген зерттеуге [36] сәйкес, жылдық табысы 1 миллиард доллардан асатын компаниялардың 300 жоғары басшысынан тұратын респонденттер: 52% жоғары басшылар деректерді талдау және шешім қабылдау үшін ұсыныстар беру мәселелерінде ИИ-ге сенеді. 48% ИИ-ді бұрын ескерілмеген тәуекелдерді анықтау үшін, ал 47% – баламалы жоспарларды ұсыну үшін пайдаланады. Сонымен қатар, 40% ИИ-ді жаңа өнімдерді әзірлеу, бюджеттік жоспарлау және маркетингтік зерттеулер жүргізу үшін қолданады. Зерттеу ИИ-дің жеке өмірге оң әсерін де көрсетті: респонденттердің 39% жұмыс пен жеке өмір арасындағы төптендіктің жақсарғанын, 38% психикалық денсаулықтың жақсарғанын, ал 31% стресстің деңгейінің төмендегенін атап етті.

Дегенмен, LLM-нің барлық қуатына қарамастан, оларды саналы түрде пайдалану маңызды құрал болып табылады. Әрбір технология сияқты, оларда шектеулер бар. Ең танымал мәселелердің бірі - «галлюцинациялар» деп аталатын жағдайлар, яғни модель шынайы, бірақ факті бойынша дұрыс емес жауапты сенімді түрде береді. Сондықтан модельдің қалай жұмыс істейтінін түсіну өте маңызды: қандай деректер мен деректер модельдерін қателіксіз интерпретациялай алатыны, сұрауларды қалай интерпретациялайтыны және ақпаратты қайдан алатыны. Сондай-ақ, LLM-нің білімдері оқыту қүнімен шектелетінін ескеру қажет, және сыртқы деректермен байланыссыз модель қазіргі нормаларды, стандарттарды, бағаларды немесе технологияларды ескермей қалуы мүмкін.

Бұл мәселелердің шешімі векторлық базаларды тұрақты жаңарту, өзекті көздерге қосылу және тек сұрақтарға жауап беріп қана қоймай, деректерді оқыту үшін проактивті пайдаланатын, тапсырма-ларды басқаруға, тәуекелдерді анықтауға, оңтайландыру нұсқаларын ұсынуға және жобаның орындалуын бақылауға арналған автономды AI-агенттерді әзірлеу болып табылады.

Құрылыш саласындағы LLM-интерфейстерге көшу - бұл тек технологиялық жаңалық емес. Бұл парадигмалық өзгеріс, адамдар мен деректер арасындағы кедергілерді жою. Бұл ақпаратпен өзара әрекеттесуді біз бір-бірімізben сөйлескендей оқай жасау мүмкіндігі - және дәл, тексерілген және әрекет етуге дайын нәтижелер алу.

Мұндай құралдарды бірінші болып пайдаланатын компаниялар айтарлықтай бәсекелестік

артықшылықта ие болады. Бұл жұмыс жылдамдығын арттыру, шығындарды азайту және деректерді жылдам талдау арқылы күрделі сұрақтарға жылдам жауп табу мүмкіндігі арқасында жобалық шешімдердің сапасын арттыру. Бірақ қауіпсіздік мәселелерін де ескеру қажет. Бұлттық LLM-сервистерін пайдалану деректердің ағып кету қаупімен байланысты болуы мүмкін. Сондықтан үйімдар жиі өздерінің инфрақұрылымында LLM-құралдарын локалды түрде, ақпаратты толық қорғау және бақылау арқылы орналастыруға мүмкіндік беретін баламалы шешімдерді іздейді.

Компанияның сезімтал деректері үшін жергілікті LLM пайдалану

2022 жылды алғашқы чат-LLM-нің пайда болуы жасанды интеллектінің дамуының жаңа кезеңін білдіреді. Дегенмен, осы модельдердің кеңінен таралуымен бірге, компанияға қатысты деректер мен сұрауларды бұлтқа беру қаншалықты қауіпсіз деген заңды сұрақ туындағы. Көптеген бұлттық тілдік модельдер сәйлесу тарихын және жүктелген құжаттарды өз серверлерінде сақтады, ал конфиденциалды ақпаратпен жұмыс істейтін компаниялар үшін бұл ИИ-ді енгізуге кедергі болды.

Бұл мәселенің ең тұрақты және логикалық шешімдерінің бірі Open Source LLM-ді корпоративтік IT-инфрақұрылымында локалды түрде орналастыру болды. Бұлттық қызметтерден айырмашылығы, локалды модельдер интернетке қосылуынан жұмыс істейді, деректерді сыртқы серверлерге жібермейді және компанияларға ақпаратты толық бақылау мүмкіндігін береді.

Бұгінгі таңда ең жақсы ашық модель [Open Source LLM] өнімділігі бойынша жабық модельдермен [мысалы, ChatGPT, Claude] салыстырмалы, бірақ шамамен бір жылға артта қалуда.

– Бен Коттье, Epoch AI коммерциялық емес зерттеу үйімінің жетекші зерттеушісі, 2024.

Ірі технологиялық компаниялар өздерінің LLM-дерін жергілікті пайдалану үшін қолжетімді ете бағытады. Meta-ның ашық LLaMA сериясы мен Қытайдағы DeepSeek жобасының қарқынды дамуы ашық архитектураға көшу мысалдарына айналды. Олардың қатарында Mistral мен Falcon да бар, олар да меншікті платформалардың шектеулерінен босатылған қуатты модельдерді шығарды. Бұл бастамалар тек жаһандық ИИ-дің дамуын жеделдетіп қана қоймай, деректердің құпиялығы маңызды компанияларға тәуелсіздік, икемділік және қауіпсіздік нормаларын сақтау мәселелерінде нақты баламаларды ұсынды.

Корпоративтік ортада, әсіресе құрылыш саласында, деректерді қорғау тек ыңғайлылық мәселесі емес, сонымен қатар реттеуші талаптарға сәйкестік мәселесі. Тендерлік құжаттама, сметалар, сыйбалар және құпия хат алмасу бойынша жұмыс қатаң бақылауды талап етеді. Мұнда жергілікті LLM компанияның периметрі ішінде деректердің сақталуына қажетті сенімділік береді.

	Cloud LLMs (OpenAI, Claude)	Local LLMs (DeepSeek, LLaMA)
Data Control	Data is transmitted to third parties	Data remains within the company's network
License	Proprietary, paid	Open-source (Apache 2.0, MIT)
Infrastructure	Requires internet	Operates in an isolated environment
Customization	Limited	Full adaptation to company needs
Cost	Pay-per-token/request	One-time hardware investment + maintenance costs
Scalability	Easily scalable with cloud resources	Scaling requires additional local hardware
Security & Compliance	Risk of data leaks, may not meet strict regulations (GDPR, HIPAA)	Full compliance with internal security policies
Performance & Latency	Faster inference due to cloud infrastructure	Dependent on local hardware, may have higher latency
Integration	API-based integration, requires internet access	Can be tightly integrated with on-premise systems
Updates & Maintenance	Automatically updated by provider	Requires manual updates and model retraining
Energy Consumption	Energy cost is covered by provider	High power consumption for inference and training
Offline Availability	Not available without an internet connection	Works completely offline
Inference Cost	Pay-per-use model (cost scales with usage)	Fixed cost after initial investment

Рис. 3.33 Жергілікті модельдер толық бақылау мен қауіпсіздікті қамтамасыз етеді, ал бұлттық шешімдер ыңғайлы интеграция мен автоматты жаңартуларды ұсынады.

Жергілікті Open Source LLM-нің негізгі артықшылықтары:

- Деректерге толық бақылау. Барлық ақпарат компания ішінде қалады, бұл рүқсатсыз қол жеткізуді және деректердің ағып кетуін болдырмайды.
- Автономды жұмыс. Интернет байланысина тәуелділік жойылады, бұл оқшауланған IT-инфрақұрылымдарда жұмыс істеу үшін ерекше маңызды. Бұл сонымен қатар санкциялар немесе бұлттық қызметтердің блокталуы жағдайында үздіксіз жұмысты қамтамасыз етеді.
- Қолдану икемділігі. Модель мәтіндерді генерациялау, деректерді талдау, бағдарламалық код жазу, жобалау мен бизнес-процесстерді басқару үшін пайдаланылуы мүмкін.
- Корпоративтік міндеттерге бейімделу. LLM-ді ішкі құжаттарда оқытуға болады, бұл компанияның жұмыс ерекшеліктері мен салалық ерекшеліктерін ескеруге мүмкіндік

береді. Жергілікті LLM-ді CRM, ERP немесе BI платформаларына қосуға болады, бұл клиенттік сұрауларды талдауды, есептерді жасауды немесе тіпті трендтерді болжауды автоматтандыруға мүмкіндік береді.

DeepSeek-R1-7B тегін және ашық моделін серверде орналастыру, бүкіл команда пайдаланушыларына қолжетімді ету, айна \$1000 құны, ChatGPT немесе Claude сияқты бұлттық API-ларға жыл сайынғы төлемдерден арзанға түсі мүмкін және компанияларға деректерді толық бақылауға, интернетке жіберуді болдырмауға және GDPR сияқты реттеуші талаптарға сәйкестікті қамтамасыз етуге көмектеседі.

Басқа салаларда жергілікті LLM автоматтандыруға көзқарасты өзгерту үстінде. Қолдау қызметтерінде олар клиенттердің жиі қойылатын сұрақтарына жауап беріп, операторларға жүктемені азайтады. HR бөлімдерінде – резюме талдап, сәйкес кандидаттарды таңдайды. Электрондық коммерцияда – пайдаланушы деректерін ашпай, жеке ұсыныстар қалыптастырады.

Құрылыш саласында да осындағы әсер күтілуде. LLM-ді жобалық деректермен және нормативтермен интеграциялау құжаттаманы дайындауды жеделдетуге, сметаларды автоматтандыруға және шығындарды болжай анализін жүргізуға мүмкіндік береді. LLM-ді құрылымдалған кестелер мен датафреймдермен бірге пайдалану әсіресе перспективалы бағыт болып табылады.

Компанияда ИИ-ға толық бақылау және өз LLM-ді іске қосу жолдары

Заманауи құралдар компанияларға үлкен тілдік модельді (LLM) жергілікті түрде бірнеше сағаттың ішінде іске қосуға мүмкіндік береді. Бұл деректер мен инфрақұрылымға толық бақылау жасауға, сыртқы бұлттық қызметтерге тәуелділікті жоюға және ақпараттың 漏 у қаупін азайтуға мүмкіндік береді. Мұндай шешім әсіресе сезімтал жобалық құжаттама немесе конфиденциалды коммерциялық деректермен жұмыс істейтін үйімдар үшін өзекті.

Міндеттер мен ресурстарға байланысты әртүрлі іске қосу сценарийлері қолжетімді – дайын шешімдерден "қораптан шыққан" нұсқаларға дейін, икемді және масштабталатын архитектураларға дейін. Ең қарапайым құралдардың бірі – Ollama, ол тілдік модельдерді терең техникалық білімсіз бір басу арқылы іске қосуға мүмкіндік береді. Ollama-мен жылдам бастау:

- Ресми сайттан (ollama.com) операциялық жүйеңізге (Windows / Linux / macOS) арналған дистрибутивті жүктеп алыңыз.
- Модельді командалық жол арқылы орнатыңыз. Мысалы, Mistral модельі үшін:

```
ollama run mistral
```

- Модель іске қосылғаннан кейін, ол жұмыс істеуге дайын – сіз терминал арқылы мәтіндік сұраулар жібере аласыз немесе оны басқа құралдарға интеграциялай аласыз. Модельді іске қосу және сұрау орындау:

ollama run mistral "100 мм көндіктегі гипсокартон қабырғасын орнату үшін барлық ресурстармен есептеу қалай жасауға болады?"

Дәстүрлі визуалды ортада жұмыс істегенді ұнататындар үшін LM Studio бар – ChatGPT-ге үқсас интерфейсі бар тегін қосымша:

- LM Studio-ны ресми сайттан (lmstudio.ai) дистрибутивті жүктеп алып орнатыңыз.
- Встроенный каталог арқылы модельді (мысалы, Falcon немесе GPT-Neo-X) таңдаңыз және жүктеп алыңыз.
- Модельмен ChatGPT-ге үқсас интуитивті интерфейс арқылы жұмыс істеңіз, бірақ толықтай жергілікті.

	Developer	Parameters	GPU Requirements (GB)	Features	Best For
Mistral 7B	Mistral AI	7	8 (FP16)	Fast, supports multimodal tasks (text + images), fully open-source code	Lightweight tasks, mobile devices, laptops
LLaMA 2	Meta	7–70	16–48 (FP16)	High text generation accuracy, adaptable for technical tasks, CC-BY-SA license	Complex analytical and technical tasks
Baichuan 7B/13B	Baichuan Intelligence	7–13	8–16 (FP16)	Fast and efficient, great for large data processing, fully open-source code	Data processing, automating routine tasks
Falcon 7B/40B	Technology Innovation Institute (TII)	7–40	8–32 (FP16)	Open-source, high performance, optimized for fast work	Workloads with limited computational resources
DeepSeek-V3	DeepSeek	671	1543 (FP16) / 386 (4-bit)	Multilingual, 128K token context window, balanced speed and accuracy	Large enterprises, SaaS platforms, multitasking scenarios
DeepSeek-R1-7B	DeepSeek	7	18 (FP16) / 4.5 (4-bit)	Retains 92% of R1 capabilities in MATH-500, local deployment support	Budget solutions, IoT devices, edge computing

Рис. 3.34 Танымаған жергілікті open source LLM-модельдерінің салыстырмасы.

Модельді таңдау жылдамдық, дәлдік және қолжетімді аппараттық мүмкіндіктерге байланысты (Рис. 3.34). Mistral 7B және Baichuan 7B сияқты шағын модельдер жеңіл тапсырмалар мен мобиЛЬДІ құрылғылар үшін жарамды, ал DeepSeek-V3 сияқты қуатты модельдер елеулі есептеу ресурстарын талап етеді, бірақ жоғары өнімділік пен көптеген тілдерді қолдауды қамтамасыз етеді. Алдағы жылдары LLM нарығы қарқынды дамитын болады – біз жеңіл және мамандандырылған модельдердің көбеюін көретін боламыз. Барлық адамзат мазмұнын қамтитын әмбебап LLM-дердің ор-

нына, тар профильді сараптамаға арналған модельдер пайда болады. Мысалы, инженерлік калькуляциялар, құрылымдың сметалары немесе CAD форматындағы деректермен жұмыс істеуге арналған модельдердің пайда болуын күтүге болады. Мұндай мамандандырылған модельдер кәсіби ортада жоғары сенімділік пен тақырыптық тереңдік маңызды болғандықтан, жылдамырақ, дәлірек және қауіпсіз болады.

Жергілікті LLM іске қосылғаннан кейін, оны компанияның нақты міндеттеріне бейімдеуге болады. Бұл үшін модельді ішкі құжаттар, техникалық нұсқаулықтар, келісім-шарт шаблондары немесе жобалық құжаттама бойынша қосымша оқыту (fine-tuning) техникасы қолданылады.

RAG: Корпоративтік деректерге қол жеткізетін интеллектуалды LLM-ассистенттер

Келесі кезең LLM-ді бизнесте қолданудың эволюциясы – бұл модельдерді нақты корпоративтік деректермен интеграциялау. Бұл тәсіл RAG (Retrieval-Augmented Generation) деп аталады – ақпаратты алу арқылы генерациялау. Бұл архитектурада тілдік модель тек диалогтық интерфейс емес, құжаттар, сызбалар, деректер базалары бойынша бағдарлай алатын толыққанды интеллектуалды ассистентке айналады және нақты, контекстуалды жауаптар береді.

RAG-тың негізгі артықшылығы – компанияның ішкі деректерін модельді қайта оқыту-дың қажеті жоқ, сонымен қатар ақпаратпен жұмыс істеуде жоғары дәлдік пен икемділікті сақтай отырып пайдалану мүмкіндігі.

RAG технологиясы екі негізгі компонентті біріктіреді:

- Ақпаратты алу (Retrieval): модель деректер қоймаларына – құжаттарға, кестелерге, PDF-файлдарға, сызбаларға қосылып, пайдалануышының сұрауы бойынша релевантты ақпаратты алады.
- Жауапты генерациялау (Augmented Generation): алынған деректер негізінде модель контексті мен сұраудың ерекшелігін ескере отырып, дәл, негізделген жауап формалдайды.

RAG қолдайтын LLM-ді іске қосу үшін бірнеше қадамды орындау қажет:

- Деректерді дайындау: қажетті құжаттарды, сызбаларды, спецификацияларды, кестелерді жинаңыз. Олар PDF-тен Excel-ге дейін әртүрлі форматтар мен құрылымдарда болуы мүмкін.
- Индексация және векторизация: LlamaIndex немесе LangChain сияқты құралдардың көмегімен деректер векторлық көріністерге түрлендіріледі, бұл мәтін фрагменттері арасындағы мағыналық байланыстарды табуға мүмкіндік береді (векторлық деректер базалары мен үлкен массивтерді векторлық көрініске аудару, соның ішінде CAD жобалары, 8 бөлімде толығырақ).
- Ассистентке сұрау: деректер жүктелгеннен кейін модельге сұрақтар қоюға болады, және ол жауаптарды корпоративтік базада іздейді, интернеттен жиналған жалпы білімдерден емес.

Мысалы, компанияда constructionsite_docs папкасы бар, онда келісімшарттар, нұсқаулықтар, сметалар мен кестелер сақталады. Python-скриптінің көмегімен (Рис. 3.35) осы папканы сканерлеп, векторлық индексация құруға болады: әр құжат мәтіннің мағыналық мазмұнын көрсететін векторлар жиынтығына түрлендіріледі. Бұл құжаттарды модельдің тиімді бағдарлай алатын және терминдер мен фразалар арасындағы байланыстарды таба алатын «мағыналар картасына» айналдырады.-

Мысалы, модель «қайтару» және «рекламация» сөздерінің жиі келісімшарттың материалдарды құрылыс алаңына жеткізуге қатысты бөлімінде кездесетін «есте сақтайды». Содан кейін, егер сұрақ қойылса – мысалы, «Тауарды қайтару мерзімі қандай?» (Рис. 3.35 - 11 жол коды) – LLM ішкі құжаттарды талдайды және нақты ақпаратты табады, барлық корпоративтік файлдардың мазмұнын оқып, түсінетін интеллектуалды көмекші ретінде әрекет етеді.



```

1 from llama_index import SimpleDirectoryReader, VectorStoreIndex
2
3 # Load documents from the folder
4 documents = SimpleDirectoryReader("constructionsite_docs").load_data()
5
6 # Creating a vector index for semantic search
7 index = VectorStoreIndex.from_documents(documents)
8
9 # Integration with LLM (e.g. Llama 3)
10 query_engine = index.as_query_engine()
11 response = query_engine.query("What are the return terms in the contracts?")
12 print(response)

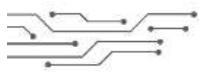
```

Рис. 3.35 LM файлдар папкасын оқиды – адам папканы ашып, қажетті құжатты іздегендей.

Кодты Python орнатылған кез келген компьютерде іске қосуға болады. Кодты іске қосу үшін Python мен IDE-ді қолдану туралы келесі бөлімде сөйлесеміз.

LLM-ді жергілікті орналастыру – бұл тек тренд емес, қауіпсіздік пен икемділікті бағалайтын компаниялар үшін стратегиялық шешім. Алайда, LLM-ді орналастыру, компанияның жергілікті компьютерлерінде немесе онлайн шешімдерді пайдалану кезінде – бұл тек алғашқы қадам. LLM мүмкіндіктерін нақты тапсырмаларда қолдану үшін, компаниялар тек чаттарда жауап алуға ғана емес, сонымен қатар LLM-ді қолдану контекстінен тыс іске қосуға болатын код түрінде жасалған логиканы сақтауға мүмкіндік беретін құралдарды пайдалану керек. Бұл шешімдерді масштабтау үшін маңызды - дұрыс ұйымдастырылған процестер ИИ-дің жетістіктерін бірден бірнеше жобаларда немесе тіпті бүкіл компанияда қолдануға мүмкіндік береді.

Осы контексте сәйкес бағдарламалу ортасын (IDE) таңдау маңызды рөл атқарады. Заманауи бағдарламалу құралдары LLM негізінде шешімдер әзірлеуге ғана емес, сонымен қатар оларды бар бизнес-процессерге интеграциялауға мүмкіндік береді, автоматтандырылған ETL-Pipeline-ге айналдырады.



ТАРАУ 3.4. LLM ҚОЛДАУЫ БАР IDE ЖӘНЕ БАҒДАРЛАМАЛАУДАҒЫ БОЛАШАҚ ӨЗГЕРИСТЕР

IDE таңдау: LLM эксперименттерінен бизнес шешімдеріне

Автоматтандыру, деректерді талдау және жасанды интеллект әлеміне – әсіреле үлкен тілдік модельдермен (LLM) жұмыс істегендегі – сәйкес интеграцияланған даму ортасын (IDE) таңдау өте маңызды. Ол сіздің негізгі жұмыс құралыңыз болады: LLM-мен жасалған кодты іске қосатын орын, жергілікті компьютерде де, корпоративтік желі ішінде де. IDE таңдау тек жұмыс ыңғайлышына ғана емес, сонымен қатар эксперименттік LLM сұрауларынан нақты бизнес-процестерге енгізілген толыққанды шешімдерге қаншалықты тез өтетініңізге де әсер етеді.

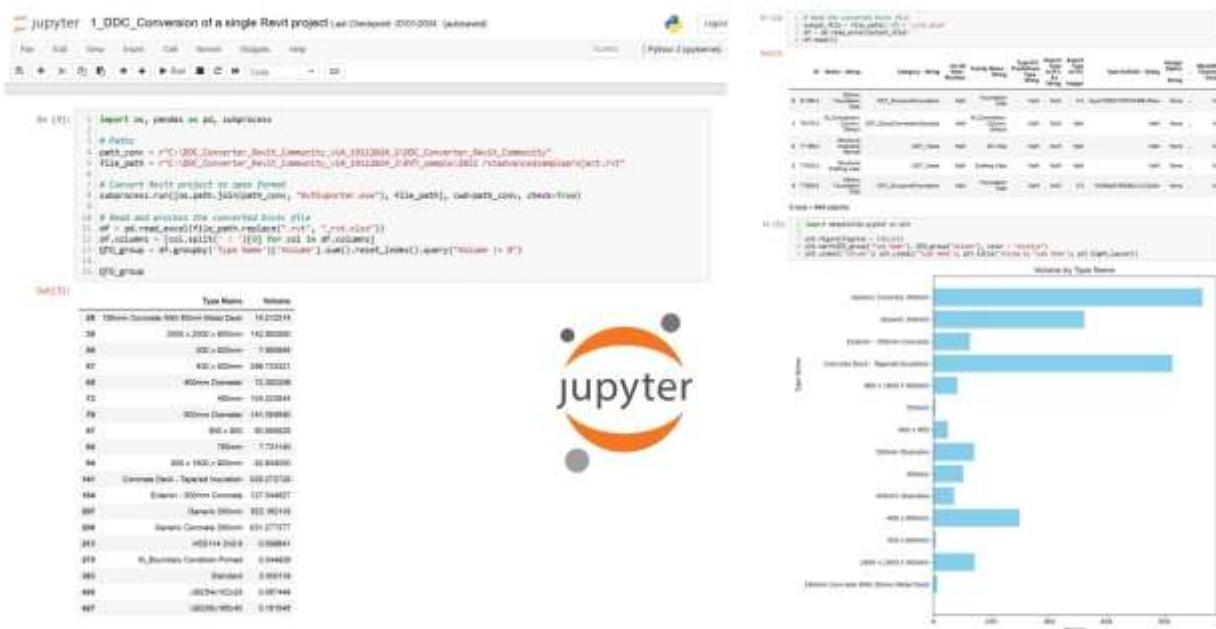
IDE (Интеграцияланған Даму Ортасы) – бұл автоматтандыру процестері мен деректерді өндөу үшін компьютеріңіздегі әмбебап құрылыш комбинаты. Пила, балға, бұрғы және басқа құралдарды бөлек сақтаудың орнына, сізде бәрін – кесу, бекіту, бұрғылау және тіпті материалдардың сапасын тексеру – жасай алатын бір құрылғы бар. Бағдарламашылар үшін IDE – код жазуға (құрылыш аналогиясында – сызыбаларды жасау), оның жұмысын тестілеуге (ғимарат моделін жинау), қателерді табуға (құрылымдардың беріктігін бақылау) және дайын жобаны іске қосуға (үйді пайдалануға беру) арналған біртұтас кеңістік.

Танымал IDE-лерге шолу:

- PyCharm® (JetBrains) – Python үшін қуатты қасіби IDE. Ол көптеген кіріктірілген функциялары арқасында маңызды жобалар үшін тамаша. Алайда, интерактивті Jupyter-файлдарын (IPYNB) базалық қолдау тек акылы нұсқада қолжетімді, ал жаңадан бастаушылар үшін интерфейс артық жүктелген болып көрінуі мүмкін.

IPYNB кеңейтімі бар файл – Jupyter® Notebook интерактивті ноутбуектерінің форматы, онда код, визуализациялар және түсіндірмелер бір құжатта біріктірілген. Мұндан формат есептер, аналитика және оқу сценарийлерін құру үшін өте қолайлы.-

- VS Code® (Microsoft) – жылдам, икемді және баптауға болатын құрал, IPYNB-ді тегін қолдаумен және көптеген плагиндермен. Ол жаңадан бастаушылар мен қасіби мамандарға арналған. GitHub Copilot және тілдік модельдермен жұмыс істеу үшін плагиндерді интеграциялауға мүмкіндік береді, бұл оны ИИ және деректер ғылыми саласындағы жобалар үшін тамаша таңдау етеді.
- Jupyter Notebook – эксперименттер мен оқыту үшін классикалық және танымал таңдау. Код жазуға, түсініктемелер қосуға, нәтижелерді бір интерфейсте визуализациялауға мүмкіндік береді (Сур. 3.41). Гипотезаларды жылдам тексеру, LLM-мен жұмыс істеу және деректерді талдау бойынша қайталанатын қадамдарды жасау үшін тамаша. Тәуелділіктер мен кітапханаларды басқару үшін Anaconda Navigator-ды пайдалану ұсынылады – Python ортасын басқару үшін визуалды интерфейс.



Сүр. 3.41 Jupyter Notebook – Pipeline процестерін құру үшін ең ыңғайлы және танымал құралдардың бірі.

■ Google Collab™ (және Kaggle платформасы (Сүр. 9.25)) – Jupyter-дің бұлттық баламасы, GPU/TPU-ға тегін қолжетімділік үсынады. Бұл орнатусыз, тікелей браузерден жұмыс істөу мүмкіндігімен бастау үшін тамаша шешімі. Google Drive-пен интеграцияны және жақында – Google-дың Gemini (LLM) жүйесімен интеграцияны қолдайды.-

	PyCharm	VS Code	Jupyter Notebook	Google Colab
Complexity	High	Medium	Low	Low
.ipynb support	Paid	Free	Built-in	Built-in
Copilots	Yes	Yes	Yes	Yes
Computing resources	Local	Local	Local	Cloud
For whom	Professionals	Universal	Beginners	Experimenters

Сүр. 3.42 IDE салыстырыруы: Jupyter Notebook – Pipeline процестерін құру үшін ең ыңғайлы және қарапайым құралдардың бірі.

IDE таңдау сіздің міндеттеріңізге байланысты. Егер сіз ИИ-мен жұмыс істеуді жылдам бастауғыңыз келсе, Jupyter Notebook немесе Google Collab-ты сынап көріңіз. Маңызды жобалар үшін PyCharm немесе VS Code-ты пайдалану жақсы. Негізгі – бастау. Заманауи құралдар эксперименттерді жұмыс шешімдеріне жылдам айналдыруға мүмкіндік береді.

Барлық сипатталған IDE-лер деректерді өңдеу пайплайндарын құруға мүмкіндік береді – яғни, әрқайсысы өз кезеңі үшін жауап беретін код блоктарының модульдерінен тұратын тізбектер (оларды LLM генерациялауы мүмкін).

- аналитикалық сценарийлер,
- құжаттардан ақпарат алу тізбектері,
- RAG негізінде автоматты жауаптар,
- есептілік және визуализацияларды генерациялау.

Модульдік құрылымның арқасында әр қадамды жеке блок ретінде көрсетуге болады: деректерді жүктеу → сұзу → талдау → визуализация → нәтижелерді экспорттау. Бұл блоктарды қайта пайдалану, бейімдеу және жаңа тізбектерге жинау, деректер үшін конструктор сияқты.

Инженерлер, басшылар және аналитиктер үшін бұл код түрінде шешім қабылдау логикасын құжаттауға мүмкіндік береді, ол LLM көмегімен генерациялануы мүмкін. Мұндай тәсіл рутиналық тапсырмаларды жеделдетуге, стандартты операцияларды автоматтандыруға және әр қадамның команда мүшелері үшін айқын және түсінікті болатын қайталанатын процестерді қалыптастыруға көмектеседі.

Автоматтандырылған ETL Pipelines (Сур. 7.23), Apache Airflow (Сур. 7.44), Apache NiFi (Сур. 7.45) және n8n (Сур. 7.46) құралдары автоматтандыру процестерінде логика блоктарын құру үшін 7 және 8 бөлімдерде талқыланады.—

LLM қолдауымен IDE және бағдарламалаудағы болашақ өзгерістер

Жасанды интеллектінің әзірлеу процестеріне интеграциялау бағдарламалаудың ландшафттың өзгерітеді. Заманауи орта тек синтаксистік бөлектеу бар мәтіндік редакторлар емес – олар жобаның логикасын түсінетін, кодты толықтыратын және тіпті кодтың белгілі бір фрагменті қалай жұмыс істейтін түсіндіретін интеллектуалды ассистенттерге айналуда. Нарықта ИИ көмегімен дәстүрлі әзірлеудің кеңейтетін өнімдер пайда болуда:

- GitHub Copilot (VS Code, PyCharm-да интеграцияланады): түсініктемелер немесе жартылай сипаттамалар негізінде кодты генерациялайтын ИИ-ассистент, мәтіндік нұсқауларды дайын шешімдерге айналдырады.
- Cursor (ИИ-ядросы бар VS Code форкы): кодты толықтырумен қатар, жобаға сұрақ қоюға, тәуелділіктерді іздеуге және код базасында оқуға мүмкіндік береді.
- JetBrains AI Assistant: JetBrains IDE (оның ішінде PyCharm) үшін құрделі кодты түсіндіру, оңтайландыру және тесттер жасау функциясымен плагин.
- Amazon CodeWhisperer: Amazon-ның AWS қызметтерін қолдауға және қауіпсіздікке баса назар аударатын Copilot аналогы.

Бағдарламалау алдағы жылдары түбектердің өзгерістерге үшірайды. Негізгі назар рутиналық код жазудан деректер модельін және архитектурасын жобалауға аудысады – әзірлеушілер жүйелерді жобалаумен көбірек айналысады, ал ИИ стандартты тапсырмаларды: код, тест, құжаттама және негізгі функцияларды генерациялауды өзіне алады. Бағдарламалаудың болашағы – адамның және ИИ-дің ынтымақтастығы, мұнда машиналар техникалық рутинаны өзіне алады, ал адамдар шығармашылыққа назар аударады.

Табиғи тілде бағдарламалау күнделікті өмірге айналады. IDE-дің персонализациясы жаңа деңгейге көтеріледі – әзірлеу ортасы пайдаланушының жұмыс стиліне бейімделуді үйренеді, ал компаниялар паттерндерді болжай отырып, контекстік шешімдерді ұсынып, алдыңғы жобалар негізінде үйренеді.

Бұл әзірлеушінің рөлін жоймайды, бірақ оны түбегейлі трансформациялайды: код жазудан – білім, сапа және процестерді басқаруға. Мұндай эволюция бизнес-аналитика саласына да әсер етеді, мұнда есептер, визуализациялар және шешім қабылдауды қолдау үшін қосымшаларды жасау көбінесе ИИ және LLM, чат және интерфейс агенттері арқылы код пен логиканы генерациялау арқылы жүзеге асырылады.

Компания LLM-чаттарын орнатып, тиісті әзірлеу ортасын таңдағаннан кейін, келесі маңызды кезең – деректерді ұйымдастыру. Бұл процесс әртүрлі көздерден ақпаратты алу, оны тазалау, құрылымдық түрге түрлендіру және корпоративтік жүйелерге интеграциялауды қамтиды.

Заманауи Data-Centric деректерді басқару тәсілінде негізгі мақсат – деректерді көптеген құралдар мен қосымшалармен үйлесімді бірегей әмбебап формаға келтіру. Құрылымдау және құрылымдық деректермен жұмыс істеу үшін арнайы кітапханалар қажет. Python үшін ең қуатты, икемді және танымал кітапханалардың бірі – Pandas. Ол кестелік деректерді ыңғайлы өндөуге мүмкіндік береді: сұзу, топтау, тазалау, толықтыру, агрегацияларды орындау және есептер құру.

Python Pandas: деректермен жұмыс істеудің алмастырылмайтын құралы

Деректерді талдау және автоматтандыру әлемінде Pandas ерекше орын алады. Бұл Python бағдарламалау тілінің ең танымал және кеңінен қолданылатын кітапханаларының бірі, құрылымдық деректермен жұмыс істеуге арналған.

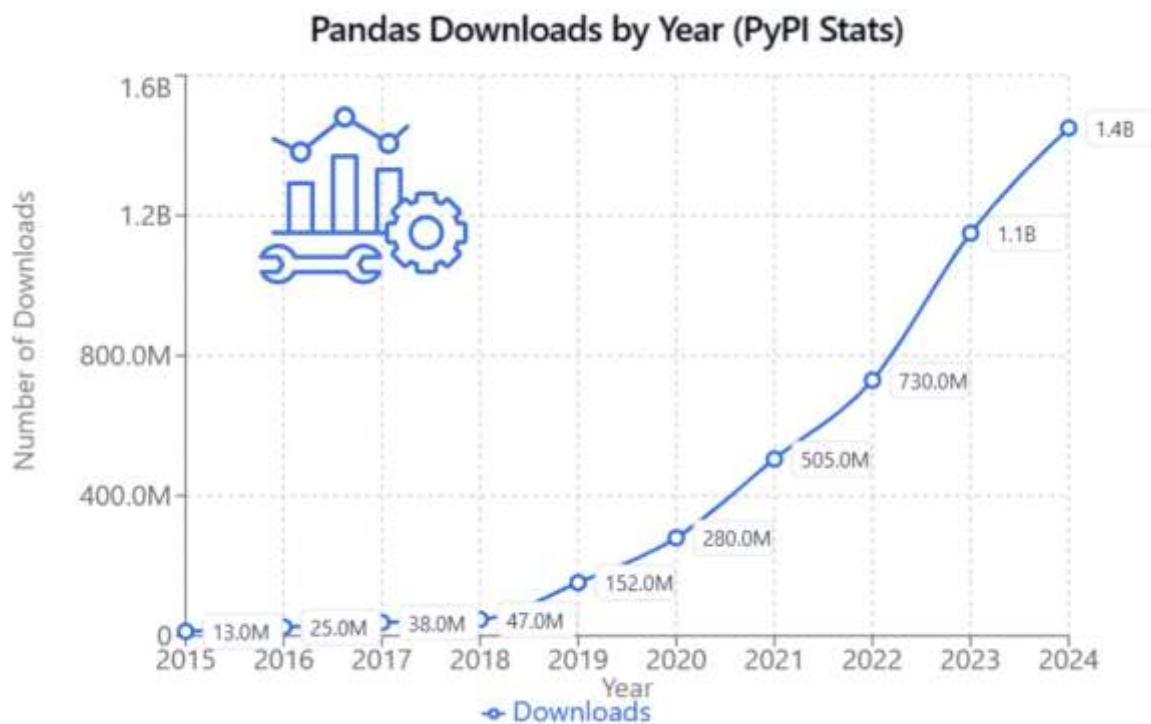
Кітапхана – бұл дайын құралдар жиынтығы: функциялар, модульдер, кластар. Құрылым алаңында әрдайым балға немесе деңгей ойлап табудың қажеті жоқ, дәл сол сияқты бағдарламалауда кітапханалар негізгі функциялар мен шешімдерді қайтадан ойлап табпай-ақ тапсырмаларды жылдам шешүге мүмкіндік береді.

Pandas – ашық кодты Python кітапханасы, жоғары өнімді және интуитивті деректер құрылымдарын, атап айтқанда, кестелермен жұмыс істеуге арналған әмбебап формат – DataFrame-ді ұсынады. Pandas – деректермен жұмыс істейтін аналитиктер, инженерлер және әзірлеушілер үшін швейцариялық пышақ.

Python – бұл жоғары деңгейлі бағдарламалау тілі, қарапайым синтаксисімен, аналитика, автоматтандыру, машиналық оқыту және веб-дамыту салаларында белсенді қолданылады. Оның танымалдылығы кодтың оқылу жеңілдігімен, кроссплатформалылығымен және бай кітапхана экожүйесімен түсіндіріледі. Бұғынгі күні Python үшін 137 000-нан астам ашық кодты пакет жасалған, және бұл сан күн сайын өсіп келеді. Әрбір осындағы кітапхана – бұл дайын функциялардың бір түрі: ең қарапайым математикалық операциялардан бастап, күрделі кескіндерді өңдеу, үлкен деректерді талдау, нейрондық желілермен жұмыс істеу және сыртқы қызметтермен интеграциялау құралдана дейін.

Басқаша айтқанда, сізде жүздеген мың дайын бағдарламалық шешімдерге – кітапханалар мен құралдарға тегін және ашық қолжетімділік бар деп елестетіңіз, оларды тікелей бизнес-процестеріңізге енгізе аласыз. Бұл автоматтандыру, талдау, визуализация, интеграция және тағы басқалар үшін арналған қосымшалардың үлкен каталогы сияқты – және бұның бәрі Python-ды орнатқаннан кейін бірден қолжетімді.

Pandas – Python экожүйесіндегі ең танымал пакеттердің бірі. 2022 жылы Pandas кітапханасының орташа жүктелу саны күніне 4 миллионды құрады, ал 2025 жылдың басына қарай бұл көрсеткіш 12 миллион жүктелуге дейін өсті, бұл оның деректер аналитикасы мен LLM чаттарындағы кеңінен қолданылуын көрсетеді. -



Pandas – ең көп жүктелетін кітапханалардың бірі. 2024 жылы оның жылдық жүктелу саны 1,4 миллиардтан асты.

Pandas кітапханасындағы сұраныс тілі функционалдығы жағынан SQL сұраныс тіліне үксас, біз "Реляциялық деректер базалары және SQL сұраныс тілі" бөлімінде қарастырғанымыздай.

Аналитика және құрылымдалған деректерді басқару әлемінде Pandas өзінің қаралайымдылығы, жылдамдығы және қуаттылығымен ерекшеленеді, пайдаланушыларға ақпаратты тиімді талдау және өңдеу үшін кең ауқымды құралдар ұсынады.

SQL және Pandas екі құралы да деректермен жұмыс істеу үшін қуатты мүмкіндіктерді ұсынады, әсіресе дәстүрлі Excel-мен салыстырғанда. Олар таңдау, сұзу сияқты операцияларды қолдайды,

тек SQL реляциялық деректер базаларымен жұмыс істеуге оңтайландырылған, ал Pandas деректерді оперативті жадта (RAM) өндейді, бұл оны кез келген компьютерде іске қосуға мүмкіндік береді, деректер базаларын құру және бөлек инфрақұрылымды орналастыру қажеттілігінсіз.-

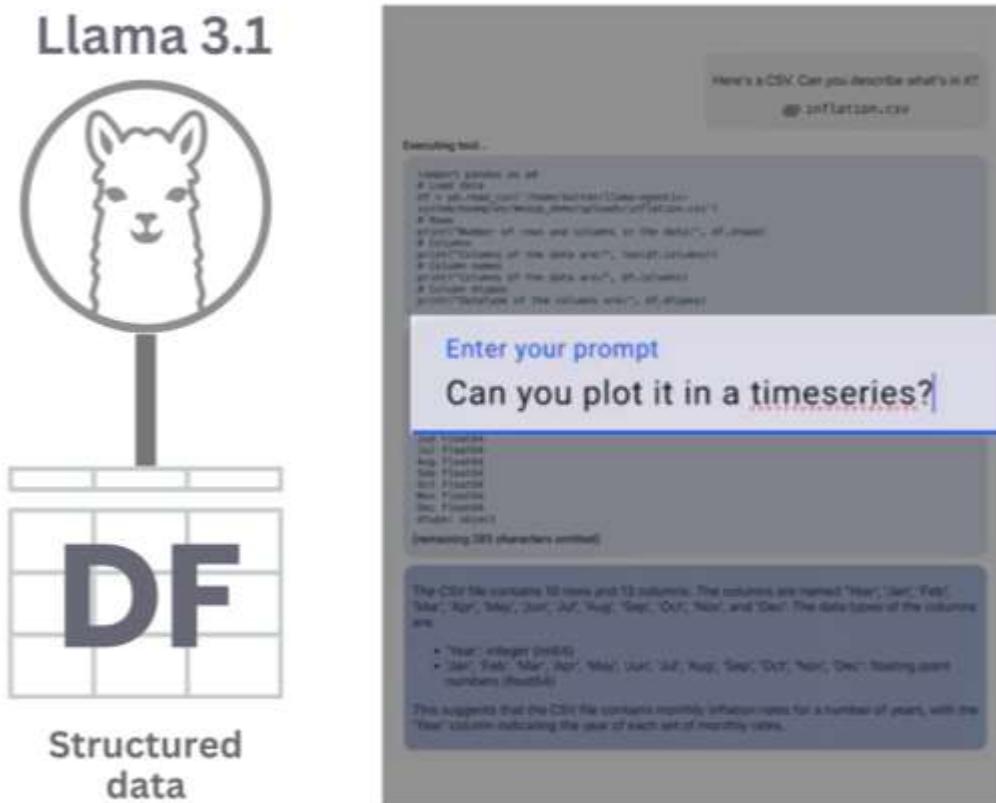


Pandas, SQL-дан айырмашылығы, деректердің әртүрлі форматтарымен жұмыс істеуде икемділікке ие, тек деректер базаларымен шектелмейді.

Pandas көбінесе ғылыми зерттеулерде, процестерді автоматтандыруда, конвейерлер (соның ішінде ETL) құруда және Python-да деректермен манипуляциялауда қолданылады, ал SQL деректер базаларын басқарудың стандартты құралы болып табылады және көбінесе корпоративтік ортада үлкен деректермен жұмыс істеу үшін пайдаланылады.

Python бағдарламалау тілінің Pandas кітапханасы тек кестелерді оқу және жазу сияқты негізгі операцияларды ғана емес, сонымен қатар деректерді біріктіру, деректерді топтастыру және күрделі аналитикалық есептеулер жүргізу сияқты күрделі тапсырмаларды орындауға мүмкіндік береді.

Бұғынгі таңда Pandas кітапханасы тек академиялық зерттеулер мен бизнес-аналитикада ғана емес, сонымен қатар LLM-модельдерімен бірге де қолданылады. Мысалы, Meta® (Facebook™) компаниясы 2024 жылы LLaMa 3.1 ашық кодты моделін жариялағанда, құрылымдалған деректермен жұмыс істеуге ерекше назар аударды, бұл өз релизінде құрылымдалған деректер фреймдерін (Сурет 3.45) CSV форматында өндеуді және Pandas кітапханасымен чатта интеграциялауды басты және алғашқы кейстердің бірі ретінде көрсетті.-



Сурет 3.45 Meta командасының LlaMa 3.1-де 2024 жылы таныстырылған алғашқы және негізгі кейстерінің бірі Pandas пайдалану арқылы қосымшаларды құру болды.

Pandas – генеративті ИИ үшін деректерді өндеп, дайындастырып миллиондаған деректер ғалымдарының қажетті құралы. Кодта нәлдік өзгерістермен Pandas-тың жылдамдатылуы үлкен қадам болады. Деректер ғалымдары деректерді минуттарда, ал сағаттарда емес, өндей алады және генеративті ИИ модельдерін оқыту үшін бірнеше есе көп деректер ала алады [88]. – Дженсен Хуанг, NVIDIA негізін қалаушы және бас директоры

Pandas пайдалану арқылы Excel мүмкіндітерінен әлдеқайда асып түсетін деректер жиынтықтарын басқаруға және талдауға болады. Excel, әдетте, 1 миллион жол деректерді өндей алса, Pandas он миллиондаған жолдарды (Сурет 9.12, Сурет 9.110) еш қызындықсыз өндей алады [89]. Бұл мүмкіндік пайдаланушыларға үлкен деректер жиынтықтарында күрделі деректерді талдау және визуализация жүргізуге мүмкіндік береді, терең түсінік алуға және деректер негізінде шешім қабылдауды жеңілдетеді. Сонымен қатар, Pandas-тың мықты қауымдастырыбы бар [90]: әлем бойынша жүздеген миллион әзірлеушілер мен аналитиктер (Kaggle.com, Google Collab, Microsoft® Azure™ Notebooks, Amazon SageMaker) күн сайын онлайн немесе оффлайн режимінде оны пайдаланып, кез келген бизнес міндеттері үшін дайын шешімдердің үлкен санын ұсынады.--

Python-дағы көптеген аналитикалық процестер Pandas кітапханасы ұсынатын DataFrame деп ататын құрылымдалған деректер формасына негізделген. Бұл – кестелік деректерді ұйымдастыру, талдау және визуализациялау үшін қуатты және икемді құрал.

DataFrame: кестелік деректердің әмбебап форматы

DataFrame – Pandas кітапханасындағы орталық құрылым, ол екі өлшемді кестені (Сурет 3.46) білдіреді, мұнда жолдар жеке объектілерге немесе жазбаларға, ал бағандар олардың сипаттамалығына, параметрлеріне немесе категорияларына сәйкес келеді. Мұндай құрылым визуалды түрде Excel кестелеріне ұқсас, бірақ икемділік, масштабталу және функционалдылық бойынша олардан айтарлықтай асып түседі.

DataFrame – компьютердің жедел жадында сақталатын кестелік деректердің көрсету және өңдеу тәсілі.

DataFrame – компьютердің жедел жадында сақталатын кестелік деректердің көрсету және өңдеу тәсілі. Кестеде жолдар, мысалы, құрылымың жобасының элементтерін, ал бағандар – олардың қасиеттерін: категориялар, габариттер, координаттар, құны, мерзімдері және т.б. көрсетуі мүмкін. Мұндай кестеде бір жоба туралы ақпарат (Сурет 4.113) немесе мындаған әртүрлі жобалардан миллиондаған объектілер туралы деректер (Сурет 9.110) болуы мүмкін. Pandas-тың векторизацияланған операцияларының арқасында, мұндай ақпарат көлемдерін жоғары жылдамдықпен оңай сүзуге, топтастыруға және агрегаттауға болады.-

The diagram illustrates the DataFrame structure. It features a table with columns: ID, Name, Category, Family Name, Height, BoundingBoxMin_X, BoundingBoxMin_Y, BoundingBoxMin_Z, and Level. A crane icon is positioned on the left, and a 'STRUCTURED DATA' logo with a grid icon is on the right. Annotations explain the structure: 'Index label' points to the first column (ID); 'Columns axis = 1' points to the second column (Name); 'Index axis = 0' points to the first row; 'Missing value' points to a blank cell in the Height column; 'Data' points to the main body of the table; and 'Column names' points to the header row.

ID	Name	Category	Family Name	Height	BoundingBoxMin_X	BoundingBoxMin_Y	BoundingBoxMin_Z	Level
431144	Single-Flush	OST_Doors	Single-Flush	6.88976378	20.1503	-10.438	9.84252	Level 1
431198	Single-Flush	OST_Doors		6.88976378	13.2281	-1.1207	9.84252	Level 2
457479	Single Window	OST_Windows	Single Window	8.858267717	-11.434	-11.985	9.80971	Level 2
485432	Single Window	OST_Windows	Single Window	8.858267717	-11.434	4.25986	9.80971	
490150	Single-Flush	OST_Doors	Single-Flush	6.88976378	-1.5748	-2.9565	-1E-16	Level 1
493697	Basic Wall	OST_Walls	Basic Wall		-38.15	20.1656	-4.9213	Level 2
497540	Basic Wall	OST_Walls	Basic Wall		-4.5212	0.0708	9.84252	Level 1

Сурет 3.46 Құрылымы жобасы DataFrame түрінде – элементтері жолдарда және атрибуттары бағандарда орналасқан екі өлшемді кесте.

Nvidia бағалауы бойынша, бүгінгі таңда барлық есептеу ресурстарының 30%-ға жуығы құрылымдық деректерді, яғни датафреймдерді өңдеуге жұмсалуда, және бұл үлес өсуді жалғастыруда.

Деректерді өңдеу – бұл, бәлкім, әлемдегі барлық компаниялардың есептесулерінің үштен бірін құрайтын іс. Деректерді өңдеу және көптеген компаниялардың деректері DataFrame форматында, яғни кестелік форматта орналасқан.

– Дженсен Хуанг, Nvidia Бас директоры

Pandas-тағы DataFrame-нің кейбір негізгі ерекшеліктерін атап өтейік:

- **Бағандар:** DataFrame-де деректер бағандарда үйымдастырылған, әрбір бағанның өзіндік атауы бар. Баған-атрибуттар әртүрлі типтегі деректерді қамтуы мүмкін, бұл деректер базасындағы немесе кестелердегі бағандарға үқсас.
- **Pandas Series** – бұл Pandas-тағы бір өлшемді деректер құрылымы, тізімге немесе кестедегі бағанға үқсас, мұнда әрбір мәнге өзінің индексі сәйкес келеді.

Pandas Series-те 400-ден астам атрибуттар мен әдістер бар, бұл деректермен жұмыс істеуді өте икемді етеді. Тікелей 400-ден астам қолжетімді функциялардың бірін бағанға қолдануға, математикалық операцияларды орындауға, деректерді сұзуге, мәндерді ауыстыруға, даталармен, жолдармен және тағы басқалармен жұмыс істеуге болады. Сонымен қатар, Series векторизацияланған операцияларды қолдайды, бұл үлкен деректер жиынтықтарын циклдік есептеулермен салыстырғанда өңдеуді едәуір жылдамдатады. Мысалы, барлық мәндерді санға көбейту, жоғалған деректерді ауыстыру немесе күрделі трансформацияларды жазбаша күрделі циклдарсыз қолдану оңай.

- **Жолдар:** DataFrame-де жолдар уникалды мәндермен индекстелуі мүмкін. Бұл индекс белгілі бір жолдардағы деректерді жылдам өзгертуге және түзетуге мүмкіндік береді.
- **Индекс:** DataFrame-ді жасаған кезде Pandas әр жолға 0-ден N-1-ге дейін (мұндағы N – DataFrame-дегі барлық жолдардың саны) индекс тағайындейды. Дегенмен, индекті өзгертуге болады, оған ерекше белгілер, мысалы, даталар немесе уникалды сипаттамалар енгізуге болады.
- **DataFrame-дегі жолдарды индекстелу** әр жолға уникалды атау немесе белгі, яғни DataFrame индексі беріледі.
- **Деректер типтері:** DataFrame әртүрлі деректер типтерін, соның ішінде: `int`, `float`, `bool`, `datetime64` және мәтіндік деректер үшін `object` қолдайды. Әрбір DataFrame бағаны өзінің деректер типіне ие, бұл оның мазмұнына қандай операцияларды орындауға болатынын анықтайды.
- **Деректермен операциялар:** DataFrame деректерді өңдеу үшін кең ауқымды операцияларды, соның ішінде агрегаттау (`groupby`), біріктіру (`merge` және `join`), конкатенация (`concat`), бөлу-қолдану-біріктіру және көптеген басқа деректерді түрлендіру әдістерін қолдайды.
- **Өлшемдерді манипуляциялау:** DataFrame бағандар мен жолдарды қосуға және жоюға мүмкіндік береді, бұл оны деректерді талдау қажеттіліктеріне сәйкес өзгертуге болатын динамикалық құрылым етеді.
- **Деректерді визуализациялау:** кірістірілген визуализация әдістерін пайдалану немесе Matplotlib немесе Seaborn сияқты танымал деректер визуализациясы кітапханаларымен өзара

әрекеттесу арқылы DataFrame-ді графиктер мен диаграммаларға оңай түрлендіруге болады, бұл деректерді графикалық түрде көрсетуге мүмкіндік береді.

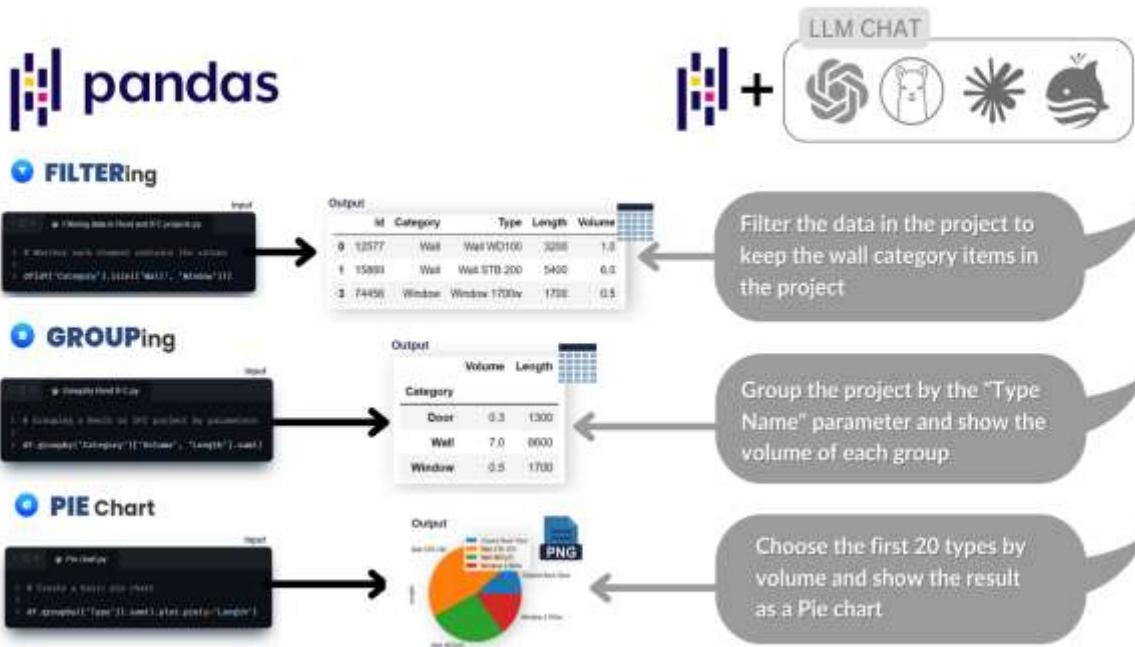
- Деректерді енгізу және шығару: Pandas әртүрлі файл форматтарына, мысалы, CSV, Excel, JSON, HTML және SQL форматтарына деректерді оқу, импорттау және экспорттау функцияларын ұсынады, бұл DataFrame-ді деректерді жинау және тарату үшін орталық торапқа айналдырады.

CSV және XLSX форматтарынан айырмашылығы, Pandas DataFrame деректермен жұмыс істеуде жоғары икемділік пен өнімділікті қамтамасыз етеді: ол оперативті жадта үлкен көлемдегі ақпаратты өңдеуге мүмкіндік береді, кеңейтілген деректер типтерін (соның ішінде даталар, логикалық мәндер және уақыттық қатарлар) қолдайды, сондай-ақ деректерді сұзу, агрегациялау, біріктіру және визуализациялау үшін кең мүмкіндіктер ұсынады. CSV деректер типтері мен құрылымы туралы ақпаратты сақтамаса, ал XLSX жиі форматтаумен артық жүктеліп, төмен масштабталушылыққа ие болса, DataFrame жылдам аналитика, процестерді автоматтандыру және ИИ модельдерімен интеграция үшін оңтайлы таңдау болып табылады. Келесі тарауларда біз осы деректер аспектілерін егжей-тегжейлі қарастырамыз, сондай-ақ кітаптың 8 бөлімінде Parquet, Apache Orc, JSON, Feather, HDF5 және деректер қоймалары сияқты ұқсас форматтарды теренірек зерттейміз.--

	XLSX	CSV	Pandas DataFrame
	Storage	Tabular	Tabular
	Usage	Office tasks, data presentation	Simple data exchange
	Compression	Built-in	None
	Performance	Low	Medium
	Complexity	High (formatting, styles)	Low
	Data Type Support	Limited	Very limited
	Scalability	Low	Extended
			Medium (memory limited)

DataFrame - жоғары өнімділік пен кеңейтілген деректер типтерін қолдаумен деректерді манипуляциялаудың оңтайлы таңдауы.

Икемділігі, қуаты және пайдаланудың қарапайымдылығы арқасында Pandas кітапханасы мен DataFrame форматы Python-да деректерді талдау саласында де-факто стандартқа айналды. Олар қарапайым есептерді құрудан бастап, LLM модельдерімен байланысты құрделі аналитикалық пайплиндарды құруға дейін тамаша жарамды.



LLM Pandas-пен өзара әрекеттесуді жеңілдетеді: кодтың орнына тек мәтіндік сұрау жеткілікті.

Бұғынгі таңда Pandas LLM негізіндегі чаттарда - ChatGPT, LLaMa, DeepSeek, QWEN және басқаларында белсенді қолданылады. Көптеген жағдайларда, модель кестелермен жұмыс, деректерді тексеру немесе аналитикамен байланысты сұрау алған кезде, кодты дәл Pandas кітапханасын пайдалана отырып генерациялайды. Бұл DataFrame-ді ИИ-мен диалогтарда деректерді ұсынудың табиғи «тіліне» айналдырады.-

Pandas сияқты заманауи деректерді өңдеу технологиялары аналитиканы, автоматтандыруды және деректерді бизнес-процестерге интеграциялауды айтарлықтай жеңілдетеді. Олар нәтижені жылдам алуға, мамандарға жүктемені азайтуға және операциялардың қайталанбалығын қамтамасыз етуге мүмкіндік береді.

Келесі қадамдар: тұрақты деректер каркасын құру

Осы бөлімде біз құрылыш саласында қолданылатын негізгі деректер типтерін қарастырдық, олардың сақтау форматтарымен таныстық және заманауи құралдардың, соның ішінде LLM және IDE, ақпаратты өңдеудегі рөлін талдадық. Деректерді тиімді басқару - негізделген шешімдер қабылдау мен бизнес-процестерді автоматтандыру үшін негіз болып табылатынын түсіндік. Деректерін құрылымдауға және жүйелендіруге қабілетті ұйымдар деректерді өңдеу және трансформациялау кезеңдерінде елеулі бәсекелестік артықшылықта ие болады.

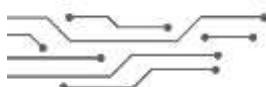
Осы бөлімнің қорытындысында, қарастырылған тәсілдердің күнделікті міндеттерінде қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

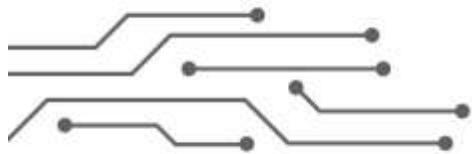
- Процестеріңіздегі деректерді аудит жүргізіңіз.
- Жобаларыңызда қолданылатын барлық деректер типтерінің тізімін жасаңыз.

- Сіздің бизнес-процесстеріңіз үшін ең маңызды деректер типтері мен модельдерін анықтаңыз.
- Ақпарат жиі құрылымдалмаған, әлсіз құрылымдалған немесе қолжетімсіз болатын проблемалық аймақтарды анықтаңыз.
- Деректерді басқару стратегиясын қалыптастыруды бастаңыз.
- Әртүрлі деректер типтерімен жұмыс істеу үшін саясат пен стандарттар туралы сұрақтарды көтеріңіз.
- Құрылымдалмаған деректерді құрылымдалғанға айналдыру арқылы қандай жұмыс процестеріңізді жақсартуға болатынын талдаңыз.
- Қауіпсіздік пен құпиялышты ескере отырып, деректерге сақтау және қол жеткізу регламентін жасаңыз.
- Деректермен жұмыс істеу үшін негізгі құралдарды орнатып, меңгеріңіз
- Сіздің міндеттеріңізге сәйкес келетін IDE таңдаңыз (мысалы, VS Code немесе Jupyter Notebook орнатыңыз)
- Жеке деректеріңізді құпия түрде өңдеу үшін жергілікті LLM орнатуға тырысыңыз
- XLSX кестелік деректерді өңдеу үшін Pandas кітапханасымен тәжірибе жасауды бастаңыз
- LLM-ге кестелік құралдарда немесе деректер базаларында өндейтін типтік тапсырмаларды сипаттап, Pandas көмегімен жұмысты автоматтандыруды сұраңыз

Мұндай қадамдарды қолдану сізге деректермен жұмыс істеу тәсілін біртіндеп трансформациялауға мүмкіндік береді, шашыраңқы, құрылымдалмаған ақпарат массивтерінен деректердің қолжетімді және түсінікті активке айналатын біртұтас экожүйеге өтуді қамтамасыз етеді. Кішкентайдан бастаңыз – Pandas-те алғашқы DataFrame жасаңыз, жергілікті LLM-ді іске қосыңыз, Python көмегімен алғашқы рутиналық тапсырманы автоматтандырыңыз (мысалы, Excel кестелерімен жұмыс істеу).

Кітаптың төртінші бөлімі деректердің сапасы, олардың ұйымдастырылуы, құрылымдалуы және модельдену мәселелеріне арналады. Біз PDF және мәтіндерден бастап, суреттер мен CAD-модельдеріне дейін шашыраңқы ақпарат көздерін құрылымдалған массивтерге айналдыруға мүмкіндік беретін әдістемелерге назар аударамыз. Сондай-ақ, деректерге қойылатын талаптардың қалай формализацияланатынын, құрылыш жобаларында концептуалдық және логикалық модельдердің қалай құрылатынын және осы процесте заманауи тілдік модельдердің (LLM) қалай көмектесетінін зерттейміз.





IV БӨЛІМ

ДЕРЕКТЕР САПАСЫ: ҰЙЫМДАСТЫРУ, СТРУКТУРЛАУ, МОДЕЛДЕУ

Төртінші бөлім шашыраңқы ақпаратты жоғары сапалы құрылымдалған деректер массивтеріне айналдыруды қамтамасыз ететін әдістемелер мен технологияларға назар аударады. Деректерге қойылатын талаптарды қалыптастыру және құжаттау процестері, құрылыш жобаларында тиімді ақпараттық архитектураның негізі ретінде егжей-тегжейлі қарастырылады. Әртүрлі көздерден (PDF құжаттары, суреттер, мәтіндік файлдар, CAD модельдері) құрылымдалған ақпаратты алу үшін практикалық әдістер ұсынылады. Деректерді автоматты түрде валидациялау және верификациялау үшін регулярлы өрнектер (RegEx) және басқа құралдарды қолдану талданады. Деректерді концептуалдық, логикалық және физикалық деңгейлерде моделдеу процесі құрылыш саласының ерекшеліктерін ескере отырып, кезең-кезеңімен сипатталады. Ақпаратты құрылымдау және тексеру процестерін автоматтандыру үшін тілдік модельдердің (LLM) нақты мысалдары көрсетіледі. Анализ нәтижелерін визуализациялаудың тиімді тәсілдері ұсынылады, бұл құрылыш жобаларын басқарудың барлық деңгейлері үшін аналитикалық ақпараттың қолжетімділігін арттырады.

ТАРАУ 4.1.

ДЕРЕКТЕРДІ СТРУКТУРЛАНГАН ФОРМАГА АЙНАЛДЫРУ

Data-driven экономика дәуірінде деректер кедегі емес, шешім қабылдаудың негізі болып табылады. Ақпаратты әр жаңа жүйе мен оның форматтарына үнемі бейімдеудің орнына, компаниялар жиі барлық процестер үшін универсалды шындық көзі қызметтің атқаратын біртұтас құрылымдалған деректер моделін қалыптастыруға үмтүлады. Заманауи ақпараттық жүйелер форматтар мен интерфейстердің айналасында емес, деректердің мағынасының айналасында жобаланады – өйткені құрылым өзгеруі мүмкін, ал ақпараттың мәні әлдеқайда ұзақ уақыт бойы өзгеріссіз қалады.

Деректермен тиімді жұмыс істеудің кілті олардың шексіз конвертациясы мен трансформациясында емес, бастапқыда дұрыс үйімдастыруды: жобаның өмірлік циклі кезеңдерінде ашықтықты, автоматтандыруды және интеграцияны қамтамасыз ететін универсалды құрылымды құруды.

Дәстүрлі тәсіл әр жаңа платформаны енгізгенде қолмен түзетуді талап етеді: деректерді көшіру, атрибут атауларын өзгерту, форматтарды реттеу. Бұл қадамдар деректердің сапасын жақсартпайды, тек мәселелерді жасырады, шексіз түрлендірулердің түйік циклын тудырады. Нәтижесінде компаниялар нақты бағдарламалық шешімдерге тәуелді болып, цифрлық трансформация баурайтады.

Келесі тарауларда деректерді дұрыс құрылымдауды, содан кейін әмбебап модельдер жасауды, платформаларға тәуелділікті минимизациялауды және басты назарды деректерді стратегиялық ресурс ретінде қарастыруға бағыттаймыз, оның айналасында тұрақты процестер құрылады.

Құжаттарды, PDF, суреттер мен мәтіндерді құрылымдық форматтарға айналдыруды үйрену

Құрылыш жобаларында ақпараттың басым көшілігі құрылымдалмаған түрде бар: техникалық құжаттар, орындалған жұмыстар актілері, сыйбалар, спецификациялар, графіктер, хаттамалар. Олардың әртүрлілігі – форматтары мен мазмұны бойынша – интеграция мен автоматизацияны қынданатады.

Құрылымдалмаған немесе жартылай құрылымдалған форматтарға түрлендіру процесі кіріс деректерінің түріне және өңдеу нәтижелеріне байланысты өзгеруі мүмкін.

Деректерді құрылымдалмағаннан құрылымдалған формаға түрлендіру – бұл өнер мен ғылым. Бұл процесс кіріс деректерінің түріне және талдау мақсаттарына байланысты өзгеріп, инженердің (Рис. 3.25) деректерді өңдеу және аналитика бойынша жұмысының едәуір бөлігін алады, таза, реттелген деректер жиынтығын алу мақсатында.



Рис. 4.11 Құрылымдалмаған сканерленген құжатты құрылымдалған кестелік форматқа түрлендіру.

Құжаттарды, PDF, суреттер мен мәтіндерді құрылымдалған форматқа (Рис. 4.11) айналдыру – бұл келесі кезеңдерді қамтитын кезең-кезеңмен жүзеге асырылатын процесс:

- Деректерді шығару (Extract): осы кезеңде құрылымдалмаған деректерді қамтитын бастапқы құжат немесе сурет жүктеледі. Мысалы, бұл PDF құжаты, фотосурет, сызба немесе схема болуы мүмкін.
- Деректерді түрлендіру (Transform): одан әрі құрылымдалмаған деректерді құрылымдалған форматқа түрлендіру кезеңі. Мысалы, бұл суреттерден мәтінді оптикалық тану (OCR) немесе басқа өндеу әдістері арқылы тану және интерпретациялауды қамтуы мүмкін.
- Деректерді жүктеу және сақтау (Load): соңғы кезең өнделген деректерді CSV, XLSX, XML, JSON сияқты әртүрлі форматтарда сақтау, мұнда форматты таңдау нақты талаптар мен қалауларға байланысты.

Бұл процесс ETL (Extract, Transform, Load) деп аталады, автоматтандырылған деректерді өндеуде маңызды рөл атқарады, оған біз "ETL және Pipeline: Extract, Transform, Load" тарауында толығырақ тоқталамыз. Әрі қарай, әртүрлі форматтағы құжаттардың құрылымдалған деректерге қалай түрленетінін мысалдармен қарастырамыз.

PDF құжатын кестеге айналдыру мысалы

Құрылыш жобаларындағы ең жиі кездесетін тапсырмалардың бірі – PDF форматындағы техникалық тапсырмаларды өндеу. Құрылымдалмаған деректерден құрылымдалған форматқа көшу процесін көрсету үшін практикалық мысал ретінде PDF құжатынан кестені шығарып, оны CSV немесе Excel форматына түрлендіруді қарастырамыз (Рис. 4.12). -

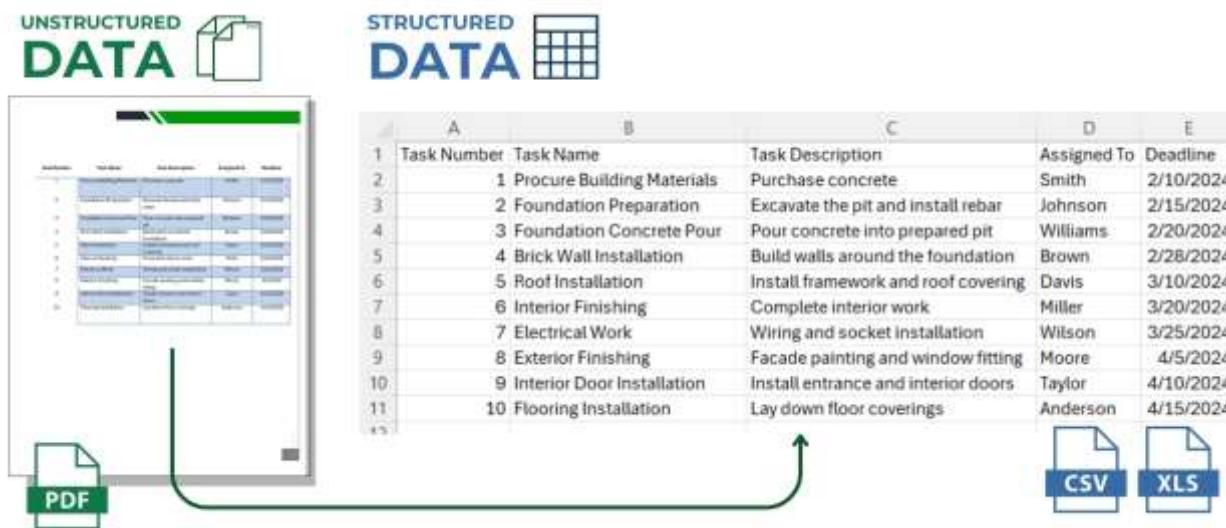


Рис. 4.12 PDF-ден айырмашылығы, CSV және XLSX форматтары кеңінен таралған және әртүрлі деректерді басқару жүйелеріне оңай интеграцияланады.

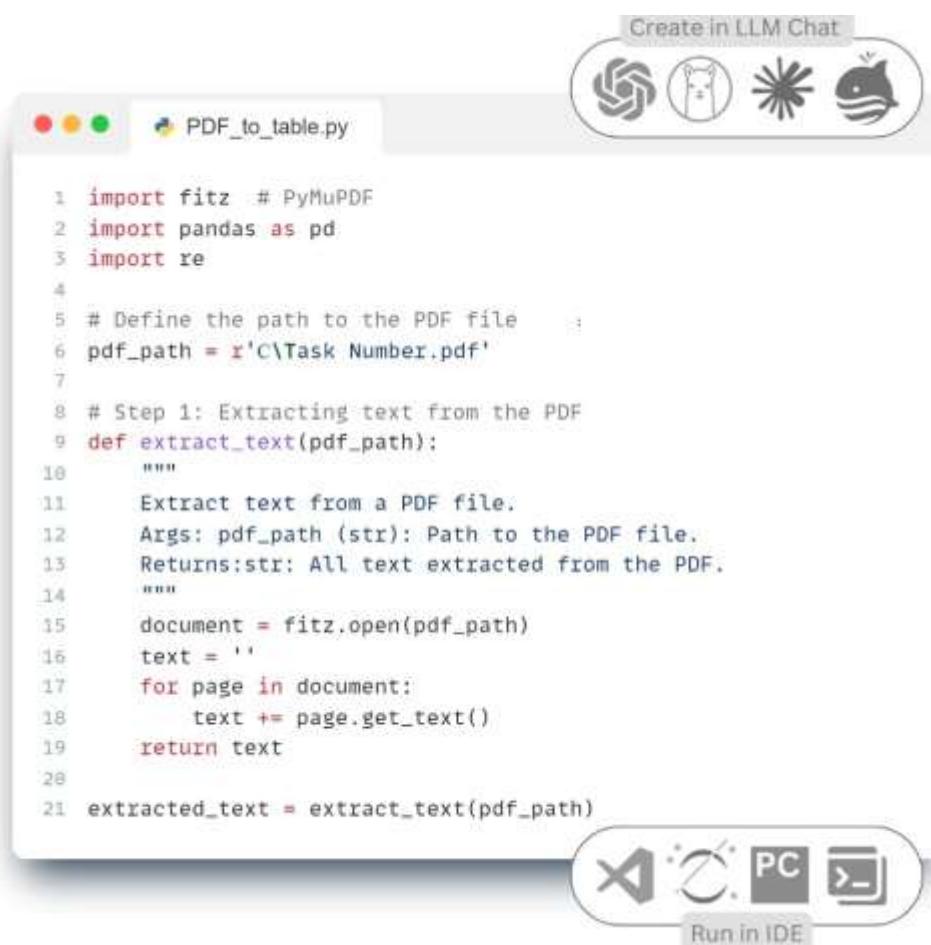
LLM тілдік модельдері, мысалы, ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN, деректермен жұмыс істейтін мамандардың жұмысын айтарлықтай жеңілдетеді, бағдарламалау тілдерін терең меңгеру қажеттілігін азайтады және көптеген тапсырмаларды мәтіндік сұраулар арқылы шешуге мүмкіндік береді.

Сондықтан, интернетте шешімдерді іздеуге (әдетте, бұл StackOverflow сайты немесе тақырыптық форумдар мен чаттар) немесе деректерді өңдеу мамандарына жүгінуге уақыт жұмсаудың орнына, біз қазіргі заманғы онлайн немесе локал LLM мүмкіндіктерін пайдалана аламыз. Сұрау жіберу жеткілікті, және модель PDF құжатын кестелік форматқа түрлендіруге арналған дайын кодты ұсынады.

- ❷ Келесі мәтіндік сұрауды кез келген LLM-модельге (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе басқа) жіберіңіз:

PDF-файлының мәтінін шығаратын кодты жазыңыз, ол кестені қамтиды. Код файлдың жолын аргумент ретінде қабылдауы және алынған кестені DataFrame түрінде қайтаруы тиіс.

- LLM-моделінің жауабы көбінесе Python тілінде код түрінде ұсынылады, себебі бұл тіл деректерді өңдеу, автоматтандыру және әртүрлі файл форматтарымен жұмыс істеу үшін қеңінен қолданылады:



The screenshot shows a window titled "PDF_to_table.py". The code is as follows:

```

1 import fitz # PyMuPDF
2 import pandas as pd
3 import re
4
5 # Define the path to the PDF file
6 pdf_path = r'C\Task Number.pdf'
7
8 # Step 1: Extracting text from the PDF
9 def extract_text(pdf_path):
10     """
11     Extract text from a PDF file.
12     Args: pdf_path (str): Path to the PDF file.
13     Returns: str: All text extracted from the PDF.
14     """
15     document = fitz.open(pdf_path)
16     text = ''
17     for page in document:
18         text += page.get_text()
19     return text
20
21 extracted_text = extract_text(pdf_path)

```

At the top right, there are icons for "Create in LLM Chat" and four AI models represented by icons: a brain, a cat, a star, and a whale. At the bottom right, there are icons for "Run in IDE" and "PC".

Рис. 4.13 LLM-нің Python кодындағы жауабы және PDF-файлының мәтінін шығаратын кітапханалар мен пакеттер (Pandas, Fitz).

Бұл кодты жоғарыда аталған танымал IDE-лердің бірінде оффлайн режимде іске қосуға болады: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагинімен, Thonny, Wing IDE, IntelliJ IDEA Python плагинімен, JupyterLab немесе танымал онлайн құралдар: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

- "Преобразование" кезеңінде, біз деректермен жұмыс істеу үшін қажетті Pandas кітапханасын қолданамыз (оның туралы «Python Pandas: деректермен жұмыс істеудің алмастырылмайтын құралы» тарауында егжей-тегжейлі айтылған), алынған мәтінді DataFrame-ге оқып, DataFrame-ди CSV немесе XLXS форматындағы кестелік файлға сақтаймыз:

PDF-файлының нәтижелік кестесін DataFrame-ге түрлендіретін код қажет. Сондай-ақ, DataFrame-ди CSV-файлға сақтау үшін код қосыңыз.

LLM-нің жауабы:

The screenshot shows a code editor window titled "PDF_to_table.py". The code is as follows:

```

1 # Step 2: Converting the extracted text to a DataFrame
2 def text_to_df(text):
3     """
4         Convert extracted text to a pandas DataFrame.
5         Args: text (str): Extracted text from a PDF file.
6         Returns: pandas.DataFrame: DataFrame representation of the text.
7     """
8     lines = text.split('\n')
9     data = []
10    for line in lines:
11        if line.strip(): # Check if line is not empty
12            data.append(re.split(r'\s{2,}', line))
13
14    # Determine the number of columns based on the longest row
15    num_columns = max(len(row) for row in data)
16
17    # Create DataFrame with dynamic columns
18    df = pd.DataFrame(data, columns=[f'Column{i+1}' for i in range(num_columns)])
19    return df
20
21 df = text_to_df(extracted_text)
22
23 # Save the DataFrame to a CSV file
24 csv_path = 'C:\extracted_table.csv'
25 df.to_csv(csv_path, index=False)

```

At the top of the window, there are icons for "Create in LLM Chat" and four AI models represented by icons: a spiral, a brain, a sunburst, and a whale. At the bottom right, there are icons for "Run in IDE" and "PC".

Рис. 4.14 PDF-ден алғынған кестені DataFrame-ге түрлендіру және кестені CSV-файлға сақтау.

Кодты орындау кезінде (Рис. 4.13, Рис. 4.14) қате туындаса – мысалы, кітапханалардың жоқтығы немесе файлдың дұрыс емес жолы – қате мәтінін бастапқы кодпен бірге көшіріп, LLM-модельге қайта жіберуге болады. Модель қате туралы хабарламаны талдайды, мәселені түсіндіреді және түзетулер немесе қосымша қадамдар ұсынады.

Осылайша, ИИ LLM-мен өзара әрекеттесу толық циклға айналады: сұрау → жауап → тест → кері байланыс → түзету – терең техникалық білімді қажет етпей.

Біз LLM чатында қарапайым мәтіндік сұрау мен кез келген IDE-де локалды іске қосуға болатын он шақты Python жолын пайдалана отырып, PDF құжатын CSV кестелік форматына түрлендірдік, бұл PDF құжатынан айырмашылығы, машинамен оңай оқылады және деректерді басқару жүйелеріне жылдам интеграцияланады.

Біз бұл кодты (Рис. 4.13, Рис. 4.14) LLM чатынан көшіріп, сервердегі ондаған және мыңдаған жаңа PDF құжаттарына қолдана аламыз, осылайша құрылымдалмаған құжаттар ағынын құрылымдық кестелік формат CSV-ге түрлендіру процесін автоматтандырамыз.-

PDF құжаттары әрдайым мәтінді қамтымайды, көбінесе бұл сканерленген құжаттар, оларды суреттер ретінде өңдеу қажет. Суреттер өз табиғаты бойынша құрылымдалмағанымен, тану кітапханаларын әзірлеу және қолдану олардың мазмұнын шығарып, өңдеуге және талдауға мүмкіндік береді, бұл бізге осы деректерді бизнес-процестерде толық пайдалануға мүмкіндік береді.

JPEG, PNG суреттерін құрылымдық формага айналдыру

Суреттер – құрылымдалмаған деректердің ең кең таралған түрлерінің бірі. Құрылыс және басқа да салаларда ақпараттың үлкен көлемі сканерленген құжаттар, сызбалар, фотосуреттер және сызбалар түрінде сақталады. Мұндай деректер құнды ақпаратты қамтиды, бірақ оларды тікелей Excel кестесі немесе деректер базасы сияқты өңдеу мүмкін емес. Суреттерде күрделі ақпарат көп, себебі олардың мазмұны, түстері, текстуралары әртүрлі, ал пайдалы ақпаратты шығару үшін арнайы өңдеу қажет.

Суреттерді деректер көзі ретінде пайдаланудың қыындығы құрылымның болмауында. Суреттер мағынаны тікелей, сандық бағалауға оңай болатын тәсілмен бермейді, оны компьютер бірден түсініп немесе өңдей алмайды, мысалы, Excel электрондық кестесі немесе деректер базасы. Неструктурланған сурет деректерін құрылымдалған формага айналдыру үшін, олардың визуалды ақпаратын интерпретациялай алатын арнайы кітапханаларды пайдалану қажет.-

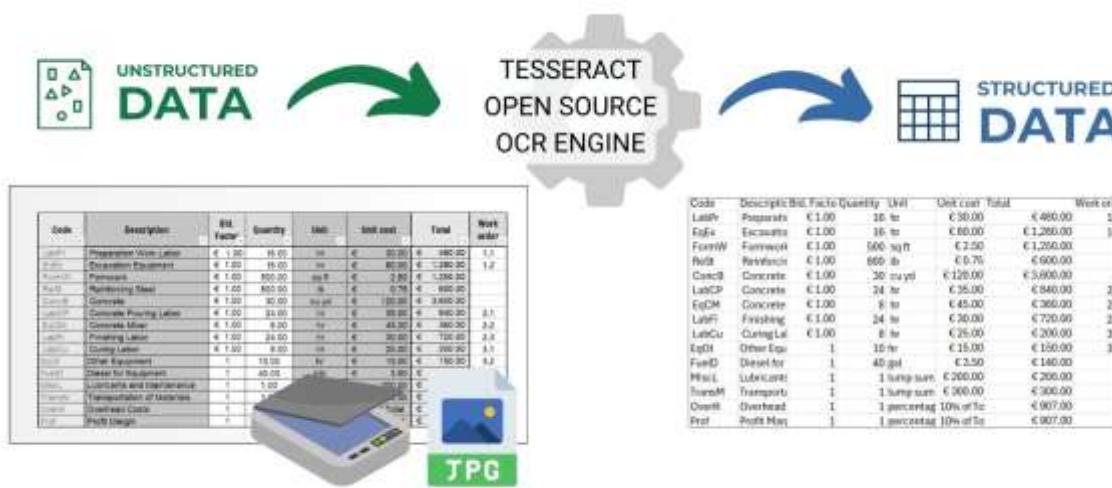


Рис. 4.15 Сканерленген құжаттар мен суреттерді құрылымдалған форматтарға айналдыру арнайы OCR құралдарының көмегімен мүмкін.

Суреттерден мәтінді шығару үшін OCR (Optical Character Recognition) – оптикалық символдарды

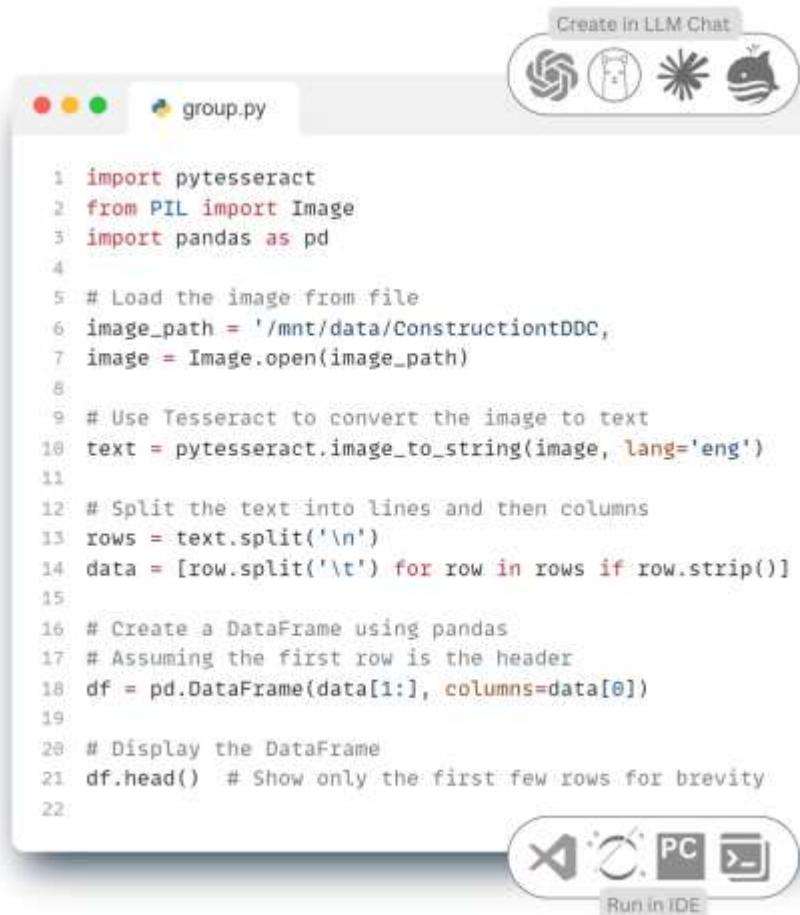
тану технологиясы қолданылады. Бұл технология сканерленген құжаттарда, фотосуреттерде және PDF файлдарында әріптер мен сандарды тануға мүмкіндік береді, оларды өңделетін және машина-мен оқылатын мәтінге айналдырады. OCR технологиялары құжат айналымын автоматтандыруды ұзақ уақыт бойы қолданылып келеді және бүгінгі күні олар кез келген бизнес-процестер мен Python қосымшаларына оңай интеграцияланады. Ең танымал OCR құралдарының бірі – Tesseract, ашық кодты, бастапқыда HP™ әзірлеген және қазір Google™ қолдайтын. Ол 100-ден астам тілдерді қолдайды және танудың жоғары дәлдігімен ерекшеленеді.

LLM чатынан сканерленген немесе фотосуретке түсірілген кестеден құрылымдалған түрде деректер алу үшін код мысалын жазуды сұрайық.

- ❷ LLM чатында (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе кез келген басқа) мәтіндік сұраныс жіберіңіз:

JPEG суретіндегі кестені DataFrame кестесіне айналдыруға мүмкіндік беретін код жазыңыз.

- LLM-нің жауабы көбінесе суреттердегі мәтінді тану үшін Pytesseract кітапханасын пайдалануды ұсынады:



```

1 import pytesseract
2 from PIL import Image
3 import pandas as pd
4
5 # Load the image from file
6 image_path = '/mnt/data/ConstructionDDC,
7 image = Image.open(image_path)
8
9 # Use Tesseract to convert the image to text
10 text = pytesseract.image_to_string(image, lang='eng')
11
12 # Split the text into lines and then columns
13 rows = text.split('\n')
14 data = [row.split('\t') for row in rows if row.strip()]
15
16 # Create a DataFrame using pandas
17 # Assuming the first row is the header
18 df = pd.DataFrame(data[1:], columns=data[0])
19
20 # Display the DataFrame
21 df.head() # Show only the first few rows for brevity
22

```

Рис. 4.16 Сурет немесе фотосурет кестесінен алынған мәтінді құрылымдалған кестелік көрініске айналдыру.

Бұл мысалда – LLM-ден алынған код (Рис. 4.16) суретті мәтінге айналдыру үшін OCR (оптикалық символдарды тану) көмегімен pytesseract (Python үшін Tesseract) кітапханасын және осы мәтінді құрылымдалған формаға, яғни DataFrame-ге айналдыру үшін Pandas кітапханасын пайдаланады.

Айналдыру процесі әдетте суреттің сапасын жақсарту үшін алдын ала өңдеуді қамтиды, содан кейін әртүрлі алгоритмдер бейнелерді анықтау, ерекшеліктерді шығару немесе объектілерді тану үшін қолданылады. Нәтижесінде құрылымдалмаған визуалды ақпарат құрылымдалған деректерге айналады.

PDF және суреттер – құрылымдалмаған ақпараттың негізгі көздері болса да, шын мәнінде көлемі бойынша чемпион – хаттарда, чаттарда, кездесулерде, мессенджерлерде жасалатын мәтін. Бұл дөректер тек көп емес – олар шашыранқы, формализденбеген және өте нашар құрылымдалған.

Мәтіндік деректерді құрылымдық формаға айналдыру

PDF құжаттарындағы кестелер (Рис. 4.12) және сканерленген кестелік формалардың нұсқалары (Рис. 4.15) қатар, жобалық құжаттамадағы ақпараттың едәүір бөлігі мәтін түрінде ұсынылған. Бұл мәтіндік құжаттардағы байланысты сөйлемдер немесе сызбалар мен схемалар бойынша шашыраңқы жазбалар болуы мүмкін. Қазіргі деректерді өңдеу жағдайында жиі кездесетін міндеттердің бірі — мұндан мәтінді талдау, визуализация және шешім қабылдау үшін жарамды құрылымдалған форматқа түрлендіру.-

Бұл процестің орталық элементі таксономия — ақпаратты категориялар мен подкатегориялар бойынша үйімдастыруға мүмкіндік беретін класификация жүйесі.

Таксономия — обьектілерді топтастыру және үйімдастыру үшін қолданылатын иерархиялық класификация құрылымы. Мәтінді өңдеу контекстінде ол элементтерді мағыналық категорияларға жүйелі түрде бөлу үшін негіз болып табылады, бұл талдауды жеңілдетуге және деректерді өңдеу сапасын арттыруға мүмкіндік береді.

Таксономияны құру кезеңдерге бөлінеді: сущностерді шығару, оларды категориялау және контекстпен байланыстыру. Мәтіндік деректерден ақпаратты шығару процесін модельдеу үшін, PDF құжаттарынан деректерді құрылымдау кезінде қолданған қадамдарға ұқсас келесі қадамдарды орындау қажет:

- Деректерді шығару (Extract): мәтіндік деректерді талдап, жобаның кестесіндегі кешігулер мен өзгерістер туралы ақпаратты шығару қажет.
- Категориялау және класификация (Transform): алынған ақпаратты категорияларға, мысалы, кешігулер мен кестедегі өзгерістердің себептеріне бөлу.
- Интеграция (Load): соңында құрылымдалған деректерді сыртқы деректер басқару жүйелеріне интеграциялауға дайындау.

Мысалды қарастырайық: жобаның менеджері мен инженер арасында жұмыс кестесіндегі мәселелер талқыланатын диалог бар. Біздің мақсатымыз — негізгі элементтерді (кешигү сеңептері, мерзімдерді түзету) шығарып, оларды құрылымдалған түрде ұсыну (Рис. 4.17).

Күтілетін кілт сөздер негізінде шығару жүргіземіз, деректерді шығару имитациясы үшін DataFrame құрамыз, содан кейін трансформациядан кейін жаңа DataFrame кестесін жасаймыз, ол күн, оқиға (мысалы, кешігу себебі) және әрекет (мысалы, кестені өзгерту) үшін бағандарды қамтиды.

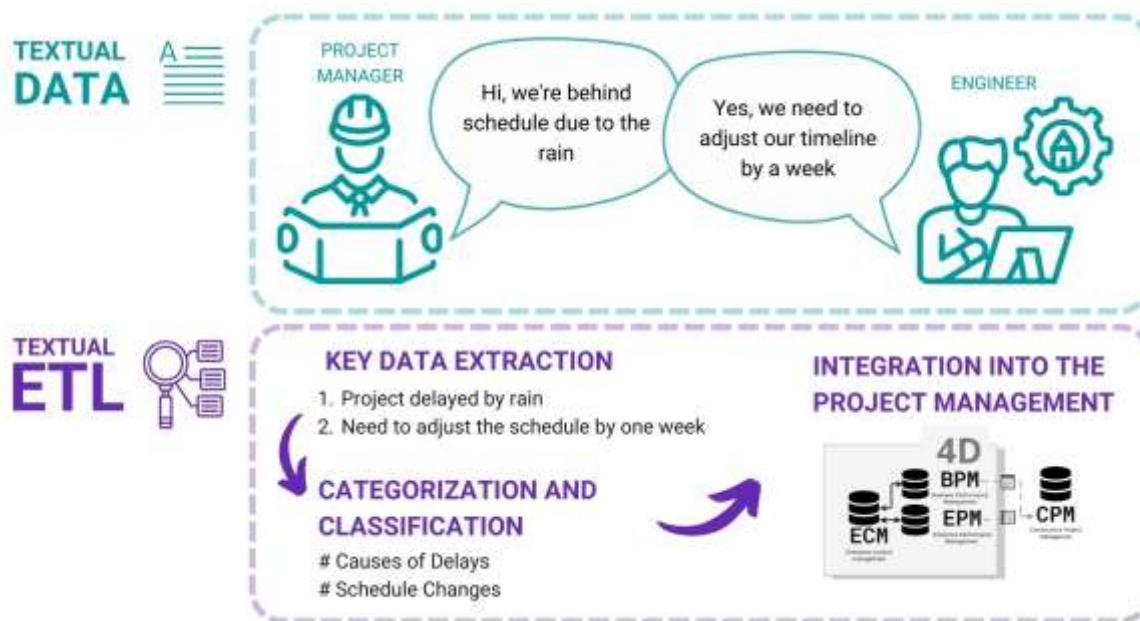


Рис. 4.17 Уақытты түзету қажеттілігі туралы мәтіннен негізгі ақпаратты бөліп алу және жобаны басқару жүйесіне өзгерістерді интеграциялау.

Мәтіндік сұраныс арқылы тапсырманы шешу үшін кодты келтірейік, алдыңғы мысалдардағыдай.

❷ Мәтіндік сұранысты кез келген LLM чатында жіберіңіз:

Менде менеджердің: "Сәлеметсіз бе, біз жаңбырдан кестеден кешігеміз" және инженердің: "Иә, мерзімдерді бір аптаға түзетуіміз керек" деген әңгімесі бар. Маған болашақта осындай мәтіндік диалогтарды талдайтын сценарий қажет, кешігулердің себептерін және мерзімдерді түзетуді шығарып, содан кейін осы деректерден DataFrame жасаңыз. Содан кейін DataFrame CSV-файлға сақталуы тиіс.

- LLM-ден жауап әдette Python кодын қамтиды, онда регулярлы өрнектер (re - Regex) және Pandas (pd) кітапханасы пайдаланылады:

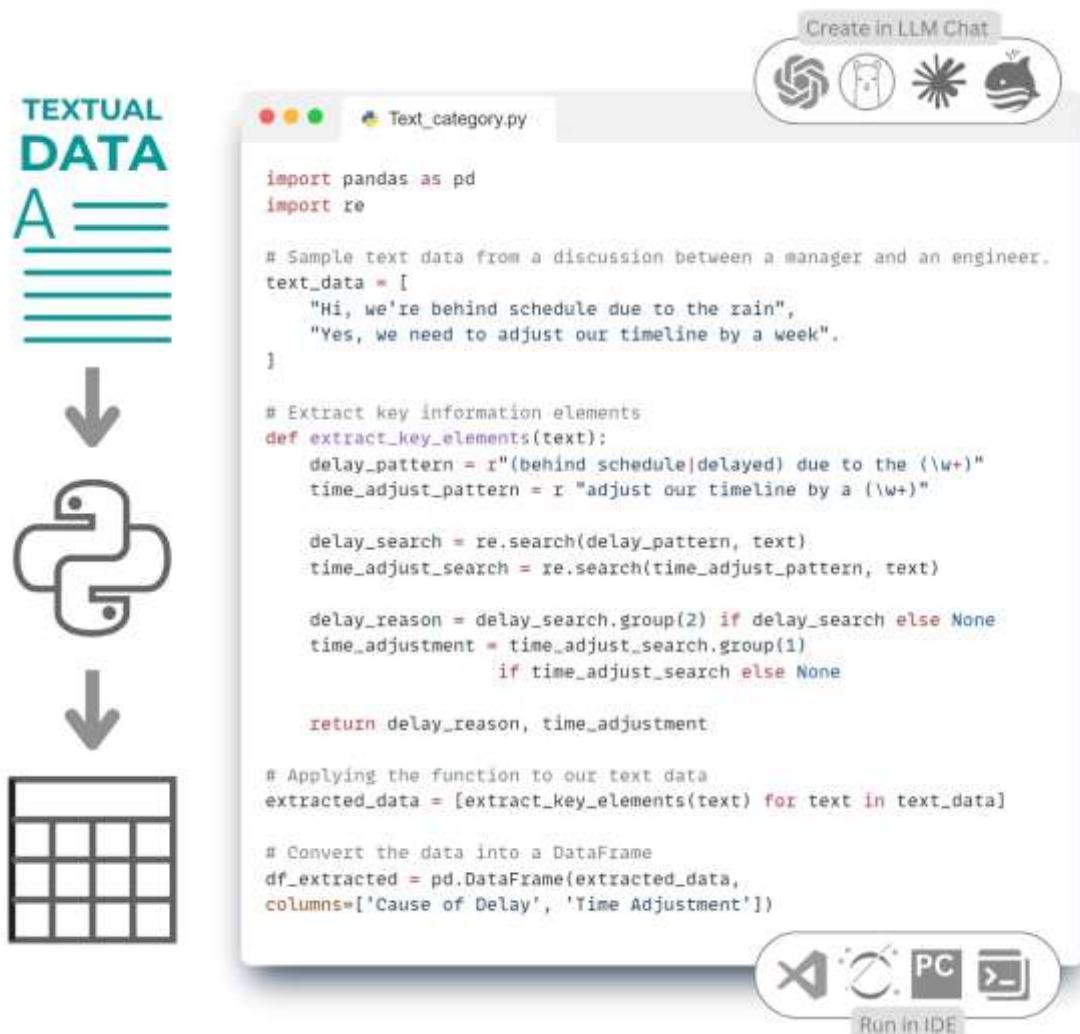


Рис. 4.18 Мәтіннен мерзімдерді түзету қажеттілігі туралы негізгі ақпаратты кесте түрінде бөліп көрсету.

Бұл мысалда (Рис. 4.17) жобаның менеджері мен инженер арасындағы хат алмасу қамтылған мәтіндік деректер талданады, болашақ жобаларды басқаруға әсер етуі мүмкін нақты ақпаратты анықтау және шығару үшін. Регулярлы өрнектердің көмегімен (регулярлы өрнектер туралы толығырақ «Құрылымдық талаптар және регулярлы өрнектер RegEx» бөлімінде талқылаймыз) жобаның кешігу себептері мен уақыт кестесін түзету қажеттіліктері паттерндер арқылы анықталады. Бұл мысалда жазылған функция жолдардан кешігу себептерін немесе уақыт түзетулерін шығараады, "из-за" сөзінен кейінгі сөзді кешігу себебі ретінде немесе "по" сөзінен кейінгі сөзді уақыт түзету ретінде негізге ала отырып.

Егер жолда ауа-райы себепті кешігу туралы айтылса, онда себеп ретінде "жауын" анықталады; егер

жолда белгілі бір мерзімге кестені түзету туралы айтылса, онда бұл мерзім уақыт түзету ретінде шығарылады (Рис. 4.19). Бұл сөздердің кез келгенінің жолда болмауы сәйкес атрибут-баған үшін "Жоқ" мәніне әкеледі.

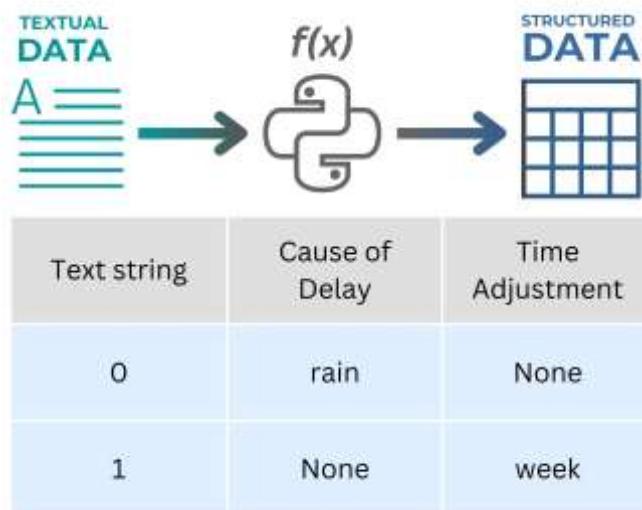


Рис. 4.19 Кодты орындағаннан кейін алынған DataFrame түріндегі жинақталған кесте кешігүлер мен қажетті уақыт түзетулері туралы ақпаратты қамтиды.

Мәтіннен (диалог, хат, құжат) жағдайларды құрылымдау және параметрлеу құрылым кешігүлерін жедел жоюға мүмкіндік береді: мысалы, жұмысшылардың жетіспеушілігі нашар ауа-райында жұмыс қарқынына асер етуі мүмкін, сондықтан компаниялар диалогтардан (Рис. 4.19) кешігу параметрлерін білгендіктен, қолайсыз болжам кезінде бригаданы қүшайте алады.

Құжаттар мен суреттерді құрылымдық форматқа түрлендіру салыстырмалы түрде қарапайым, ашық және тегін құралдардың көмегімен жүзеге асырылуы мүмкін, олар категориялау негізінде жұмыс істейді.

Элементтерді категориялау жобалық деректермен жұмыс істеудің маңызды бөлігі болып табылады, әсіресе CAD (BIM) бағдарламаларын қолдану контекстінде.

CAD (BIM) деректерін құрылымдық формага аудару

CAD (BIM) деректерін құрылымдау және категориялау - бұл қурделі міндет, себебі CAD (BIM) деректерларынан сақталатын деректер көбінесе жабық немесе қурделі параметрлік форматтарда ұсынылады, олар көбінесе геометриялық деректер элементтерін (жартылай құрылымдық) және метаақпарат элементтерін (жартылай құрылымдық немесе құрылымдық деректер) бір уақытта біріктіреді.

CAD (BIM) жүйелеріндегі нативті деректер форматтары, әдетте, қорғалған және тікелей пайдалануға қолжетімді емес, егер тек мамандандырылған бағдарламалық қамтамасыз ету немесе әзірлеушінің API-интерфейстерін қолданбаса (Рис. 4.110). Мұндай деректердің оқшаулануы «силос» дег аталатын жабық қоймаларды қалыптастырады, олар ақпараттың еркін алмасуын шектейді және компанияда цифрлық процестерді құруды тежейді.-

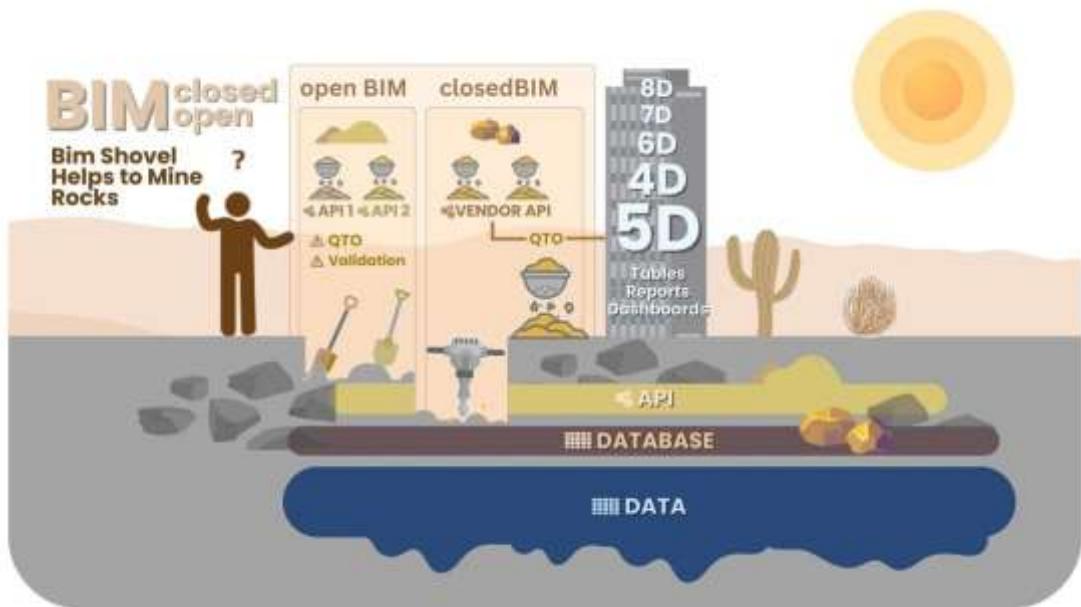


Рис. 4.110 CAD (BIM) мамандары нативті деректерге API байланыстары немесе вендор құралдары арқылы қол жеткізе алады.

Арнайы CAD (BIM) форматтарында жобаның элементтерінің сипаттамалары мен атрибуттары иерархиялық класификация жүйесінде жиналады, мұнда сәйкес қасиеттері бар объектілер деректер класификациясының бұтақтарының ең соңғы түйіндерінде орналасады (Рис. 4.111).-

Мұндай иерархиялардан деректерді алу екі жолмен жүзеге асырылады: біріншіден, әрбір түйінді қолмен басу арқылы, ағашты өндегендегі, таңдалған категориялар мен типтердің бұтақтарын балтамен кесу. Екінші нұсқа - бағдарламалық интерфейстерді (API) пайдалану - деректерді алу мен топтастырудың тиімді, автоматтандырылған тәсілін ұсынады, нәтижесінде оларды басқа жүйелерде пайдалану үшін құрылымдық кестеге айналдырады.

CAD (BIM) жобаларынан құрылымдық деректер кестелерін алу үшін Dynamo, pyRvt, Pandamo (Pandas + Dynamo), ACC сияқты әртүрлі құралдарды немесе IFC форматындағы IfcOpSh немесе IFCjs сияқты ашық кодты шешімдерді қолдануға болады.

Заманауи деректерді экспорттау және конвертациялау құралдары CAD модельдерінің мазмұнын өңдеу мен дайындауды жеңілдету үшін геометриялық ақпарат пен атрибутивтік деректер (Рис. 4.113) - құрылымдық элементтерінің қасиеттерін сипаттайтын метақпаратты екі негізгі компонентке бөлуге мүмкіндік береді (Рис. 3.116). Бұл екі деректер қабаты бір-бірімен уникалды идентификаторлар арқылы байланысты, олардың арқасында әрбір элементтің геометрия сипаттамасымен (параметрлер немесе полигондар арқылы) атрибуттарын: атавы, материалы, орындалу кезеңі, құны және т.б. дәл сәйкестендіруге болады. Мұндай тәсіл модельдің тұтастығын қамтамасыз етеді және деректерді визуализация (модельдің геометриялық деректері) және аналитикалық немесе басқару міндеттері (құрылымдық немесе әлсіз құрылымдық) үшін икемді пайдалануға мүмкіндік береді, екі деректер типімен бөлек немесе параллель жұмыс істей отырып.-

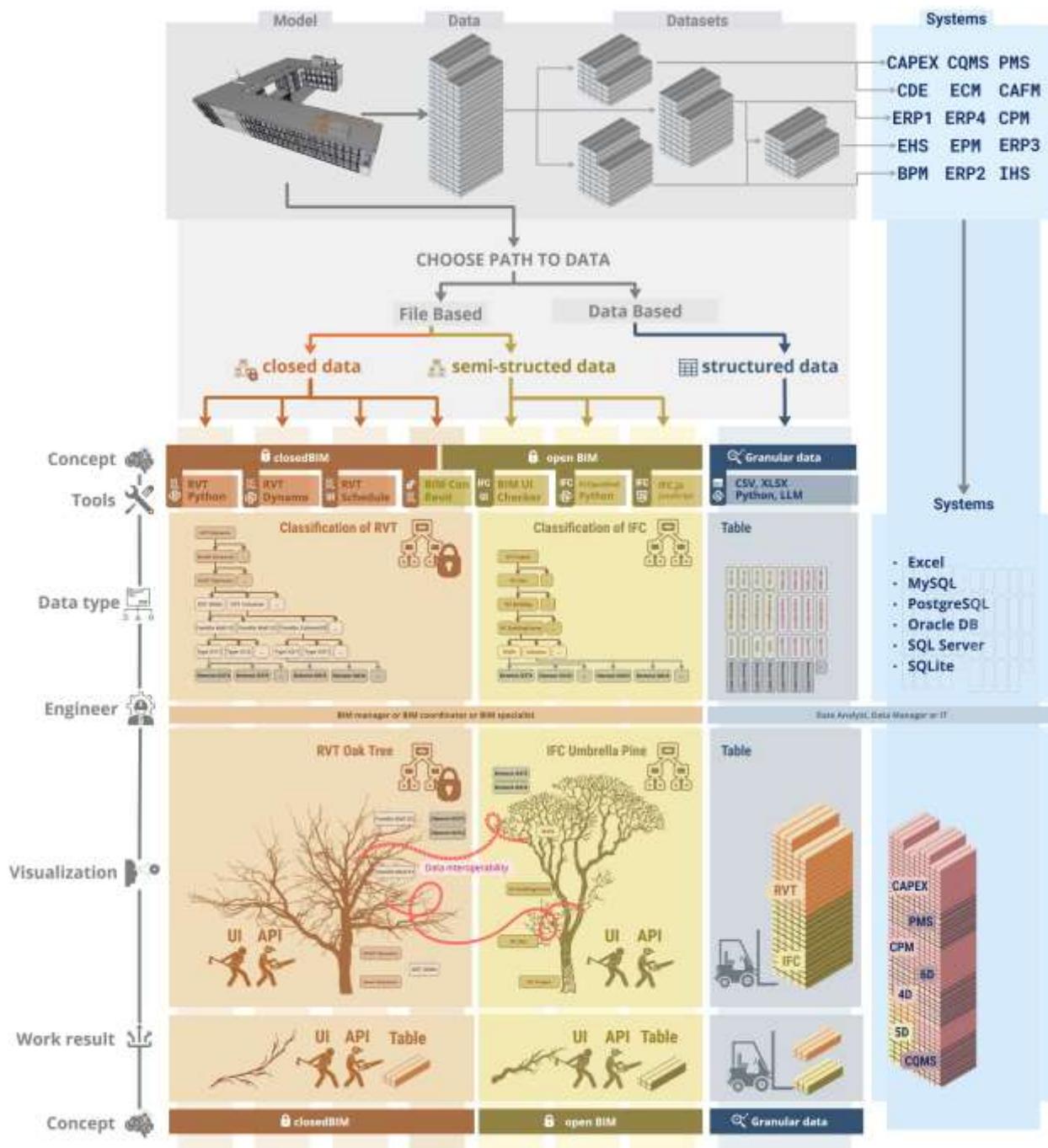


Рис. 4.111 CAD (BIM) деректер базасынан алынған ақпарат пайдаланушыға классификация ағаштары түрінде ұсынылады.

Инженерлік кері жобалау технологияларының дамуы мен CAD деректерін конвертациялау үшін бағдарламалық қамтамасыз ету жинақтары (SDK) пайда болуымен, CAD (BIM) жабық форматтарынан деректерді алу және конвертациялау әлдеқайда жөнделдеді. Енді жабық форматтардан деректерді заңды және қауіпсіз түрде универсал форматтарға, талдауға және басқа жүйелерде пайдалануға жарамды етіп конвертациялау мүмкіндігі бар. Алғашқы реверс-инжинириング құралдарының

(«Open DWG») пайда болуы мен CAD форматтары бойынша вендорлардың үстемдігі үшін күрес туралы біз «Құрылымдық деректер: цифрлық трансформацияның негізі» бөлімінде айттық.

Кепі жобалау құралдары жабық меншікті форматтардан деректерді заңды түрде алуға мүмкіндік береді, CAD (BIM) аралас форматындағы ақпаратты пайдаланушыға қажетті деректер мен форматтарға бөле отырып, олардың өнделуін және талдануын жеңілдеді.

Кепі жобалауды және CAD деректер базасынан ақпаратқа тікелей қолжетімділікті пайдалану ақпаратты қолжетімді етеді, ашық деректер мен ашық құралдарды пайдалануға, сондай-ақ стандартты құралдармен деректерді талдауға, есептер, визуализациялар құруға және басқа цифрлық жүйелермен интеграциялауға мүмкіндік береді (Рис. 4.112). -

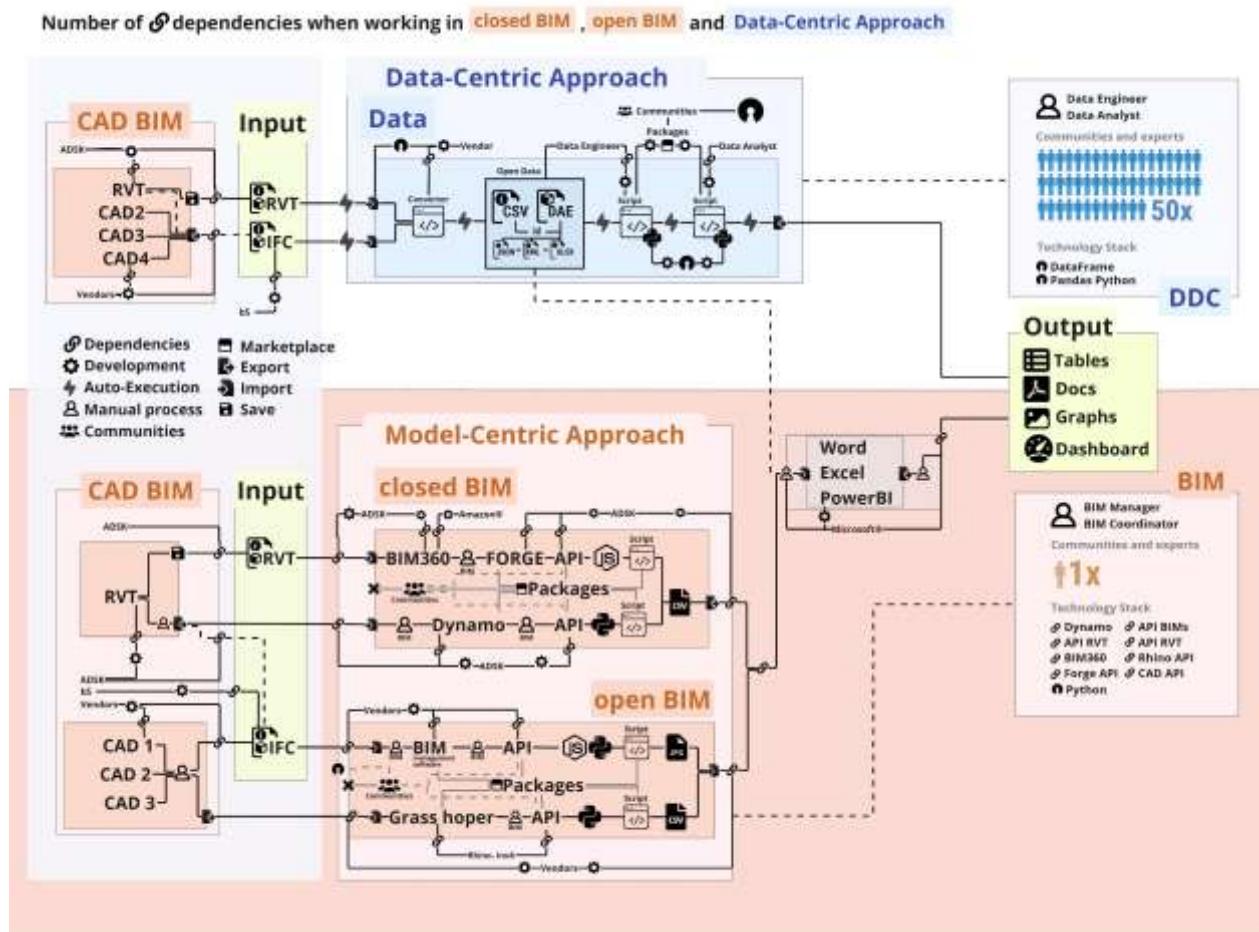


Рис. 4.112 CAD деректеріне тікелей қолжетімділік бағдарламалық платформаларға тәуелділікті минимизациялауға және деректерге негізделген тәсілге көшуге мүмкіндік береді.

1996 жылдан бастап DWG форматы, 2008 жылдан бастап DGN форматы және 2018 жылдан бастап RVT форматы жабық CAD деректер форматтарын ыңғайлы және тиімді түрде кез келген басқа фор-

маттарға, соның ішінде құрылымдық форматтарға, кері жобалау құралдары арқылы конвертациялауға мүмкіндік берді (Рис. 4.113). Бұгінде әлемдегі барлық ірі CAD (BIM) және ірі инженерлік компаниялар жабық CAD (BIM) форматтарынан деректерді алу үшін SDK-реверс-инжиниринг құралдағын пайдаланады.-



Рис. 4.113 Кері инжинириング құралдарын пайдалану CAD (BIM) бағдарламасының деректер базаларын кез келген ыңғайлы деректер моделіне түрлендіруге мүмкіндік береді.

Жабық, меншікті форматтардан ашық форматтарға деректерді түрлендіру және аралас CAD (BIM) форматтарын геометриялық және метаақпараттық атрибуттық деректерге бөлу, олармен жұмыс істеу процесін жеңілдетеді, оларды талдау, манипуляциялау және басқа жүйелермен интеграциялау үшін қолжетімді етеді (Рис. 4.114). -

CAD (BIM) деректерімен қазіргі жұмысымызда CAD форматтарынан ақпаратқа қол жеткізу үшін CAD (BIM) жеткізушілерінен рұқсат сұраудың қажеті жоқ.

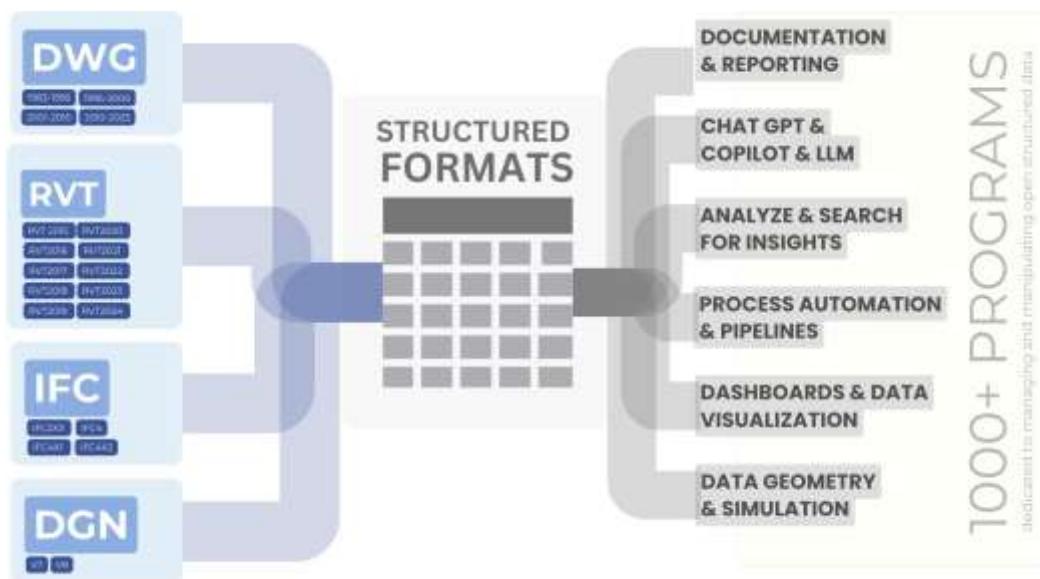


Рис. 4.114 Заманауи SDK құралдары меншікті CAD (BIM) деректер базалары форматтарынан заңды түрде деректерді түрлендіруге мүмкіндік береді.

Жобалау CAD деректерін өңдеудегі заманауи трендтер нарықтағы негізгі ойыншылар – CAD жеткізушилдері тараапынан қалыптасып, олар деректер әлеміндегі өз позицияларын нығайтуға және жаңа форматтар мен концепцияларды жасауға бағытталған.

CAD шешімдерінің вендорлары құрылымдық деректерге көшуде

2024 жылдан бастап жобалау және құрылыш саласында деректерді пайдалану мен өңдеуде елеулі технологиялық өзгерістер орын алуда. Жобалық деректерге еркін қолжетімділіктен, CAD жүйелерінің өндірушілері жаңа концепцияларды ілгерілетуге назар аударады. 2002 жылдың құрылған BIM және 2012 жылдың құрылған open BIM сияқты тәсілдер, CAD жеткізушилдері ұсынатын заманауи технологиялық шешімдерге орын беруде [93]:

- Ақпаратты тиімді басқаруға және деректер аналитикасына көшуге мүмкіндік беретін «гранулирленген» деректерді пайдалану.
- USD форматтарының пайда болуы және деректерді икемді үйымдастыру үшін Entity-component-system (ECS) тәсілінің енгізілуі.
- Деректерді өңдеуде, процестерді автоматтандыруда және деректер аналитикасында жаңанды интеллектінің белсенді пайдаланылуы.
- Интероперабельділіктің дамуы - әртүрлі бағдарламалар, жүйелер мен деректер базалары арасындағы жақсартылған өзара әрекеттестік.

Әрбір аспектіні «CAD және BIM: маркетинг, шындық және құрылыш саласындағы жобалық деректердің болашағы» атты кітаптың алтыншы бөлімінде толығырақ қарастырамыз. Осы тарауда біз тек өзгерістердің жалпы векторын қысқаша белгілейміз: бұғынгі таңда ірі CAD жеткізушилдері жобалық ақпаратты құрылымдаудың тәсілдерін қайта ойластыруға ұмтылуда. Негізгі өзгерістердің бірі – классикалық файлдық сақтау моделінен деректер архитектурасының аналитикаға бағытталған гранулирленген моделіне көшу, бұл модельдің жеке компоненттеріне тұрақты қолжетімділікті қамтамасыз етеді [93].

Бұғынгі күні индустрия көлемді, мамандандырылған және параметрлік форматтардан, геометриялық ядроларды пайдалануды талап ететін, әмбебап, машина оқытын және икемді шешімдерге көшу процесін бастап кешіруде.

Өзгерістердің бір драйвері USD (Universal Scene Description) форматы болды, ол бастапқыда компьютерлік графика индустриясында әзірленген, бірақ NVIDIA Omniverse (және Isaac Sim) платформасының дамуы арқасында инженерлік қосымшаларда да танымал болды. Параметрлік IFC-дан айырмашылығы, USD қарапайым құрылымды ұсынады, геометрия мен объектілердің қасиеттерін JSON форматында сипаттауға мүмкіндік береді, бұл ақпаратты өңдеуді жеңілдетеді және оны цифровық процестерге интеграциялауды жеделдетеді. Жаңа формат геометрияны (BREP-NURBS-тан басқа – кітаптың 6 бөлімінде толығырақ) MESH-полигоны түрінде сақтауға, ал объектілердің қасиеттерін JSON форматында сақтауға мүмкіндік береді, бұл автоматтандырылған процестер мен бұлтты экожүйелерде жұмыс істеуді жеңілдетеді.-

Кейбір CAD және ERP жеткізушилери ұқсас форматтарды (мысалы, NWD, SVF, CP2, CPIXML) пайдалана бастады, алайда олардың көпшілігі жабық болып қалады және сыртқы пайдалануға

қолжетімді емес, бұл интеграция мен деректердің қайта пайдаланудың мүмкіндіктерін шектейді. Осы контексте USD DXF-тің бұрынғы рөлін атқаруы мүмкін – DWG сияқты меншікті форматтарға ашық балама.

General Information				Comparison / Notes
Year of format creation	1991	2016	IFC focuses on construction data, USD on 3D graphics	
Creator-developer	TU Munich	Pixar	IFC was founded in Germany, USD in America	
Prototypes and predecessors	IGES, STEP	PTEX, DAE, GLTF	IFC evolved from IGES/STEP, USD from PTEX/DAE/GLTF	
Initiator in Construction	ADSK	ADSK	ADSK initiated the adoption of both formats in construction	
Organizer of the Alliance	ADSK	ADSK	ADSK organized both alliances	
Name of the Alliance	b5 (IAI)	AOUSD	Different alliances for each format	
Year of Alliance Formation	1994	2023	The IFC alliance was formed in 1994, AOUSD for USD in 2023	
Promoting in the construction	ADSK and Co	ADSK and Co	ADSK and Co actively promotes both formats in b5 (IAI) since the introduction	

Purpose and Usage				Comparison / Notes
Purpose	Semantic description and interoperability	Data simplification, visualization unification	IFC for semantics and exchange; USD for simplification and visualization	
Goals and Objectives	Interoperability and semantics	Unification for visualization and data processing	IFC focuses on semantics; USD on visualization	
Use in Other Industries	Predominantly in construction	In film, games, VR/AR, and now in construction	USD is versatile and used in various fields	
Supported Data Types	Geometry, object attributes, metadata	Geometry, shaders, animation, light, and camera	USD supports a wider range of data types suitable for complex visualizations; IFC focuses on construction-specific data	

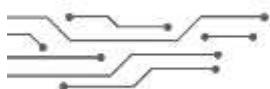
Рис. 4.115 USD форматы, CAD жеткізушілерінің интероперабельділікке және жобалық, деректердің геометриялық ядролардан тәуелсіздігіне деген сұраныстың қанағаттандыруға тырысуы ретінде.

Ірі әзірлеушілердің ашық және жеңілдетілген USD, GLTF, OBJ, XML (жабық NWD, CP2, SVF, SVF2, CPIXML) және үқсас форматтарға көшүи (Рис. 3.117) деректердің жеңілдету және олардың қолжетімділігін арттыруға деген жаһандық тенденцияны көрсетеді. Алдағы жылдары күрделі параметрлік стандарттар мен геометриялық ядроларға тәуелді форматтардан жеңіл және құрылымдалған шешімдерге көшу күтілуде. Мұндай көшү құрылымындағы саласының цифрандыруын жеделдетеді, процестерді автоматтандыруды жеңілдетеді және деректер алмасуды оңайлатады. -

CAD жеткізушілерінің жаңа ашық форматтарды ілгерілету жөніндегі стратегиялық жоспарларына қарамастан, құрылымдық саласының мамандары CAD (BIM) құралдарын пайдаланбай-ақ жабық CAD жүйелерінен деректерге толық қол жеткізе алады, бұл үшін кері инженерия құралдарын пайдалануарқылы.

Барлық осы тенденциялар ауыр, монолитті 3D модельдерден әмбебап, құрылымдалған деректерге және басқа салаларда бұрыннан танымал форматтарды пайдалануға көшуге әкеледі. Жобалық командалар CAD модельдерін тек визуалды объектілер немесе файлдар жиынтығы ретінде емес, білім мен ақпаратты қамтитын деректер базасы ретінде қабылдай бастаған кезде, жобалау мен басқару тәсілі түбекейлі өзгереді.

Командалар құжаттардан, мәтіндерден, сызбалардан және CAD модельдерінен құрылымдалған деректерді шығарып, деректер базаларына қол жеткізгеннен кейін, келесі маңызды қадам деректерді модельдеу және олардың сапасын қамтамасыз ету болады. Дәл осы кезең ақпаратты өңдеу мен трансформациялау жылдамдығына көп әсер етеді, ол ақырында нақты қолданбалы міндеттер бойынша шешімдер қабылдау үшін пайдаланылатын болады.



ТАРАУ 4.2. СЫНЫПТАУ ЖӘНЕ ИНТЕГРАЦИЯ: ҚҰРЫЛЫС ДЕРЕКТЕРІНІҢ БІРДЕЙ ТІЛІ

Шешім қабылдау жылдамдығы деректер сапасына байланысты

Қазіргі заманғы жобалық деректер архитектурасы түбегейлі өзгерістерге ұшырауда. Сала ауыр, оқшауланған модельдер мен жабық форматтардан аналитика, интеграция және процестерді автоматтандыруға бағытталған икемді, машина оқытын құрылымдарға көшуде. Алайда жаңа форматтарға көшу өздігінен тиімділікті қамтамасыз етпейді – назар деректердің сапасына ауысады.

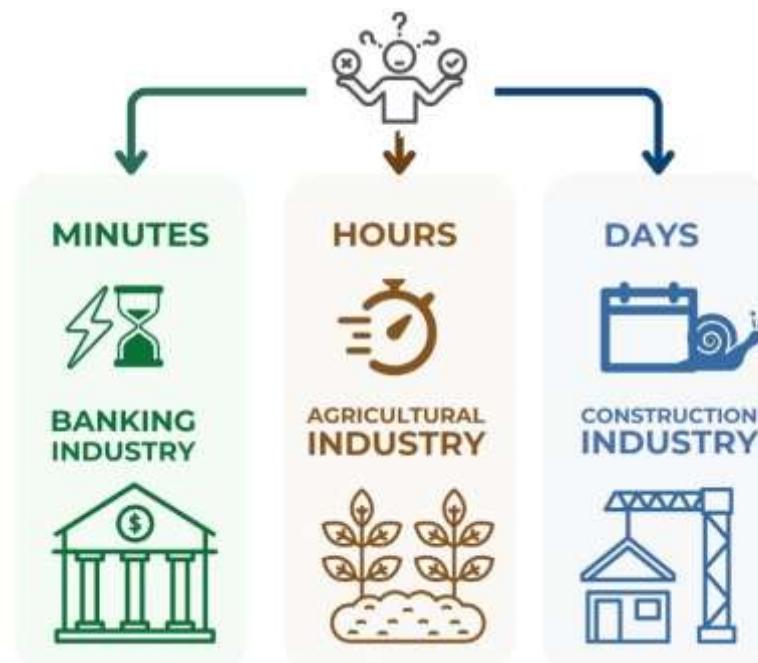
Бұл кітаптың беттерінде біз форматтар, жүйелер және процестер туралы көп айтамыз. Бірақ барлық осы күш-жігер бір негізгі элементсіз мағынасыз – сенімді деректер. Деректердің сапасы – цифрландырудың негізгі тасы, оған біз барлық келесі бөлімдерде қайта ораламыз.

Қазіргі заманғы құрылыш компаниялары – əсіресе ірі компаниялар – ондаған, тіпті мындаған әртүрлі жүйелер мен деректер базаларын пайдаланады. Бұл жүйелер тек жаңа ақпаратпен тұрақты түрде толтырылып қана қоймай, сонымен қатар тиімді өзара әрекеттесуі тиіс. Жаңа деректер, келіп түскен ақпаратты өңдеу нәтижесінде қалыптасқан, осы ортаға интеграцияланып, нақты бизнес міндеттерін шешуге қызмет етеді.

Егер бұрын нақты бизнес міндеттерін шешу жоғары басшылық – HiPPO деп аталатын – тәжірибе мен интуиция негізінде қабылданса, бүгінгі күні ақпарат көлемінің күрт өсуі жағдайында мұндай тәсіл даулы болып отыр. Оның орнына нақты уақыт режимінде деректермен жұмыс істейтін автоматтандырылған аналитика келеді.

"Дәстүрлі-қолмен" бизнес процестерін талқылау деңгейі басшылардан операциялық аналитикаға ауысады, ол бизнес сұраныстарына жылдам жауаптарды талап етеді.

Бухгалтерлер, прорабтар және сметашылар жобалар бойынша есептер мен жинақталған кестелерді бірнеше күн мен апта бойы қолмен қалыптастырған дәүірі артта қалуда. Бүгінгі күні шешім қабылдаудың жылдамдығы мен уақтылы болуы бәсекелестік артықшылықтың негізгі факторы болып табылады.



Құрылыш саласында есептеулер мен шешім қабылдауға бірнеше күн кетеді, ал басқа салаларда бұл сағаттар немесе минуттар ішінде жүзеге асырылады.

Құрылыш саласының цифрландыру деңгейі жоғары дамыған салалардан басты айырмашылығы – деректердің сапасы мен стандартталу деңгейінің төмендігі. Ақпаратты қалыптастыру, беру және өңдеу бойынша ескірген тәсілдер процестерді баяулатып, хаос тудырады. Біртұтас деректер сапасы стандарттарының болмауы кешенді автоматтандыруды енгізуге кедергі келтіреді.

Негізгі мәселелердің бірі бастапқы деректердің төмен сапасы, сондай-ақ олардың дайындау және тексеру бойынша формализденген процестердің болмауы болып табылады. Сенімді және келісілген деректерсіз жүйелер арасында тиімді интеграция мүмкін емес. Бұл жобаның өмірлік циклінің әр кезеңінде кідірістерге, қателіктерге және шығындардың артуына әкеледі.

Кітаптың келесі бөлімдерінде деректердің сапасын қалай арттыруға, процестерді стандарттауға және ақпаратты алу кезеңінен сапалы, тексерілген және келісілген деректерге дейінгі жолды қысқартуға ежей-тегжейлі тоқталамыз.

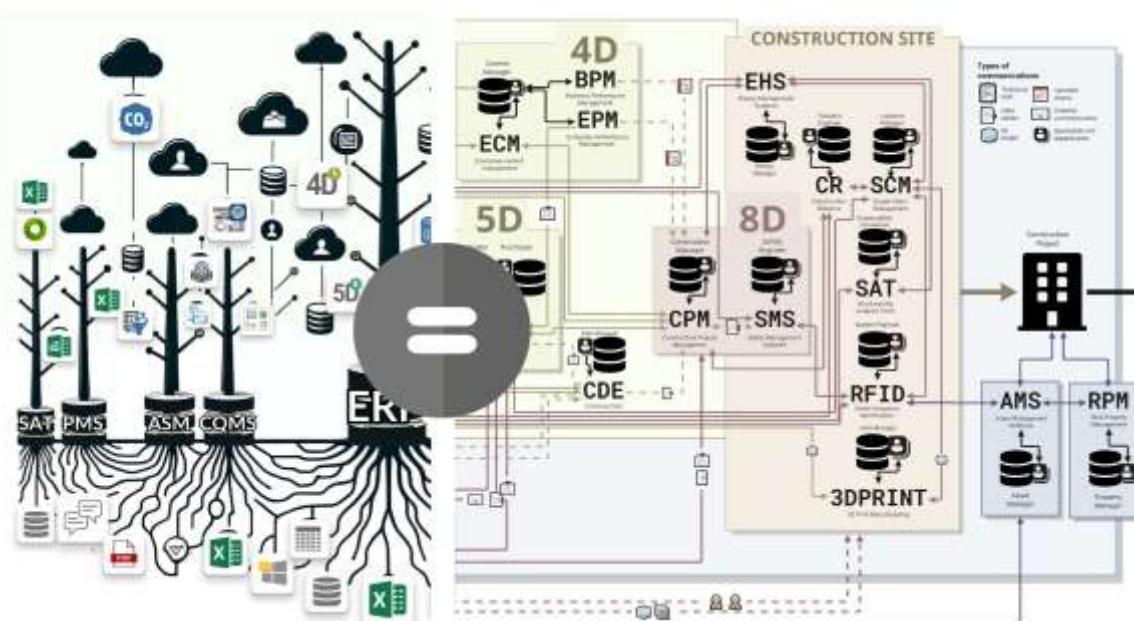
Деректерді стандарттау және интеграциялау

Деректермен тиімді жұмыс істеу нақты стандарттау стратегиясын талап етеді. Деректердің құрылымы мен сапасына нақты талаптар болғанда ғана олардың тексеруін автоматтандыруға, қолмен операциялардың санын азайтуға және жобаның барлық кезеңдерінде негізделген шешімдерді қабылдауды жеделдетуге болады.

Құрылыш компаниясының құнделікті практикасында жүздеген файлдарды өңдеу қажет: электрондық хаттар, PDF құжаттары, CAD жобалық файлдары, IoT датчиктерінен алынған деректер, оларды компанияның бизнес-процестеріне интеграциялау қажет.

Компания экожүйесінің орманы, деректер базалары мен құралдардан тұратын (Сурет 4.22), келіп түсетін әртүрлі форматтағы деректерден қоректік заттарды алууды үйренуі керек, компанияның қажетті нәтижелеріне қол жеткізу үшін.

Деректер ағынымен тиімді жұмыс істеу үшін бүтін бір менеджерлер армиясын жалдаудың қажеті жоқ, ен алдымен, деректерге қатаң талаптар мен стандарттарды әзірлеу және олардың автоматты тексеру, унификациялау және өндеге үшін тиісті құралдарды пайдалану қажет.



Сурет 4.22 Компания экожүйесінің сау өмір сүруін қамтамасыз ету үшін оның жүйелерін ресурстармен сапалы және уақытылы қамтамасыз ету қажет.

Деректерді тексеру және унификациялау (кейінгі автоматты интеграция үшін) процесін автоматтандыру үшін әрбір нақты жүйе үшін минималды қажетті деректер талаптарын сипаттаудан бастау керек. Бұл талаптар анықтайды:

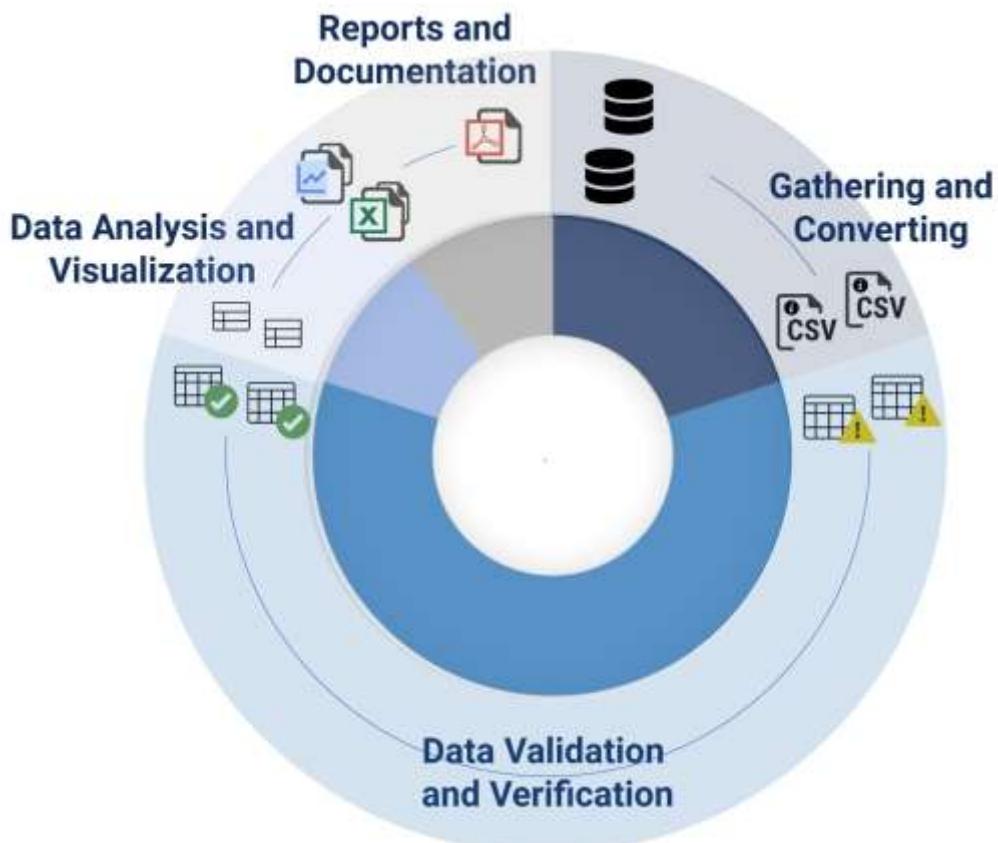
- Нені алу қажет?
- Қандай түрде (құрылым, формат)?
- Қандай атрибуттар міндетті?
- Дәлдік пен толықтық бойынша қандай рұқсаттар қабылданады?

Деректерге қойылатын талаптар алынған және өндөлетін ақпараттың сапасы, құрылымы мен толықтығының критерийлерін сипаттайты. Мысалы, PDF құжаттарындағы мәтіндер үшін сала стандарттарына сәйкес дәл рәсімдеу маңызды (Сурет 7.214 - Сурет 7.216). CAD модельдеріндегі объектілерде дұрыс атрибуттар (өлшемдер, кодтар, классификаторларға сілтемелер) болуы тиіс (Сурет 7.39, Сурет 7.310). Ал келісімшарттардың сканерлерінде нақты даталар мен соманы автоматты түрде шығару мүмкіндігі маңызды (Сурет 4.17 - Сурет 4.110).----

Деректерге қойылатын талаптарды формулировка жасау және олардың сәйкестігін автоматты түрде тексеру - ең көп енбек талап ететін, бірақ критикалық маңызды кезеңдердің бірі. Нақты осы процесс бизнес-процессерде көбінесе ең көп уақытты алады.

Кітаптың үшінші бөлімінде айтылғандай, бизнес-аналитика (BI) саласындағы мамандардың жұмыс уақытының 50%-дан 90%-ына дейінгі бөлігі талдауға емес, деректерді дайындауға жұмсалады (Сурет 3.25). Бұл процесс деректерді жинауды, тексеруді, валидациялауды, унификациялауды және құрылымдауды қамтиды.

2016 жылғы сауалнамаға сәйкес [95], деректерді өндөу мамандары әртүрлі салаларда өздерінің жұмыс уақытының көп бөлігін (шамамен 80%) ең аз ұнататын істеріне жұмсайтынын мәлімдеді (Сурет 4.23): бар деректер жиынтықтарын жинау және оларды ұйымдастыру (унификациялау, құрылымдау). Осылайша, шығармашылық тапсырмаларға, мысалы, жаңа инсайттар мен ашылуарға әкелетін заңдылықтар мен паттерндерді іздеуге 20%-дан аз уақыт қалады.



Сурет 4.23 Деректердің сапасын тексеру және қамтамасыз ету - басқа жүйелерге интеграциялау үшін деректерді дайындаудағы ең қыимбат, ұзақ және күрделі кезең.

Құрылыш компаниясында деректерді тиімді басқару кешенді тәсілді талап етеді, оған тапсырмаларды параметрлеу, деректердің сапасына қойылатын талаптарды формулировка жасау және олардың автоматтандырылған тексеру үшін тиісті құралдарды пайдалану кіреді.

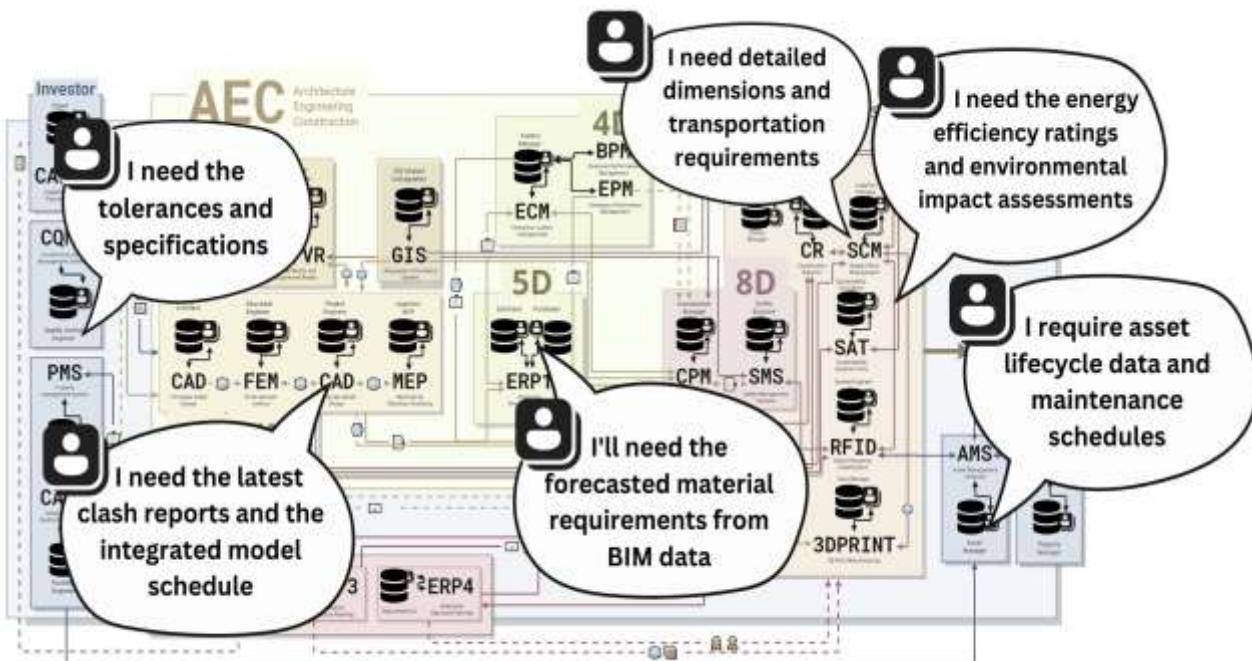
Цифрлық үйлесімділік талаптардан басталады

Цифрлық жүйелердің компаниялар ішінде санының артуымен, олардың арасында деректердің көлісімділігіне деген қажеттілік те артады. Әртүрлі IT жүйелеріне жауапты менеджерлер көбінесе ақпарат көлемінің және форматтарының әртүрлілігінің артуына ілесе алмайды. Мұндай жағдайда олар мамандардан деректерді басқа қосымшалар мен платформаларда пайдалануға жарамды формада жасауын сұрауға мәжбүр болады.

Бұл, өз кезегінде, инженерлер мен деректерді қалыптастырумен айналысатын қызметкерлерден көптеген талаптарға бейімделуді талап етеді, көбінесе деректердің қайда және қалай қолданылатыны туралы ашықтық пен нақты түсініксіз. Ақпаратпен жұмыс істеуде стандартталған тәсілдердің болмауы тиімділіктің жоғалуына және тексеру кезеңінде шығындардың артуына әкеледі, бұл көбінесе деректердің күрделілігі мен стандартталмағандығынан қолмен жүзеге асырылады.

Деректерді стандарттау мәселесі – бұл тек ыңғайлылық немесе автоматтандыру мәселеі емес. Бұл тікелей қаржылық шығындар. IBM-нің 2016 жылғы есебіне сәйкес, АҚШ-та төмен сапалы деректерден жыл сайынғы шығындар 3,1 триллион долларды құрайды. Сонымен қатар, MIT және басқа аналитикалық консалтингтік компаниялардың зерттеулері төмен сапалы деректердің құны компанияның табысының 15-25%-на дейін жетуі мүмкін екенін көрсетеді.

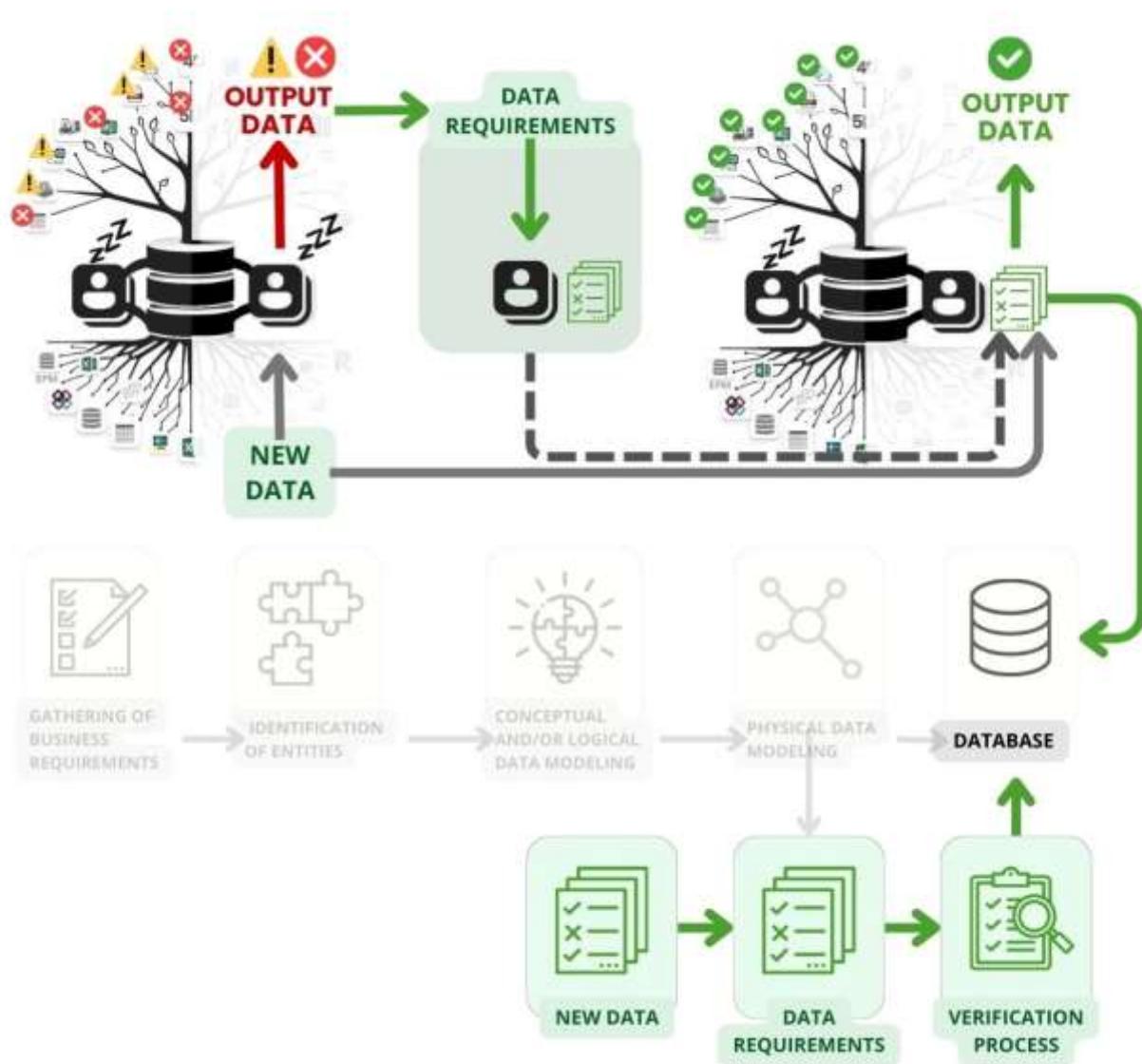
Мұндай жағдайда деректерге қойылатын нақты талаптар мен қандай параметрлер, қандай форматта және қандай деңгейде егжей-тегжейлі болуы тиіс екендігі туралы сипаттамалардың болуы критикалық маңызды болып табылады. Бұл талаптарды формализацияламай, жүйелер мен жобаның кезеңдері арасында деректердің сапасы мен үйлесімділігін қамтамасыз ету мүмкін емес.



Бизнес әртүрлі рөлдердің өзара әрекеттесуіне негізделген, олардың әрқайсысы бизнес міндеттерін орындау үшін маңызды параметрлер мен мәндерді талап етеді.

Деректерге қойылатын дұрыс талаптарды формулирлеу үшін деректер деңгейінде бизнес-процесстерін түсіну қажет. Құрылым жобалары типі, масштабы және қатысушылар саны бойынша ерекшеленеді, ал әрбір жүйе – моделдеу (CAD (BIM)), күнтізбелік жоспарлау (ERP 4D), құнды есептеу (ERP 5D) немесе логистика (SCM) – кіріс (енгізілетін элементтер) үшін өзіне тән параметрлерді талап етеді.

Осы қажеттіліктерге байланысты бизнес-менеджерлер жаңа деректер құрылымдарын белгіленген талаптарды ескере отырып жобалауы немесе бар кестелер мен деректер базаларын бейімдеуі тиіс. Бұл ретте жасалатын деректердің сапасы талаптардың қаншалықты дәл және дұрыс формулирленгеніне тікелей байланысты болады.



Деректердің сапасы деректерді пайдалану жағдайлары үшін жасалатын талаптардың сапасына байланысты.

Әрбір жүйе деректерге өзіне тән талаптарды қойып отырғандықтан, жалпы талаптарды формулируеудегі алғашқы қадам бизнес-процестеріне қатысатын барлық элементтерді категориялау болуы тиіс. Бұл объектілерді нақты жүйелерге немесе қолданбалы міндеттерге сәйкес келетін сыныптар мен сыныптар топтарына бөлу қажеттілігін білдіреді. Әрбір осындай топ үшін деректердің құрылымы, атрибуттары мен сапасына қатысты жеке талаптар әзірленеді.

Дегенмен, тәжірибеде осы тәсілді жүзеге асыруда маңызды қындықтармен бетпе-бет келу қажет: деректерді топтастырудың бірегей тілінің болмауы. Бөлшектелген классификациялар, идентификаторлардың қайталануы және форматтардың үйлесімсіздігі әрбір компания, әрбір бағдарламалық қамтамасыз ету және тіпті әрбір жоба өздерінің оқшауланған деректер модельдері мен сыныптарын қалыптастыруына әкеледі. Нәтижесінде, ақпаратты жүйелер арасында беру үшін қажетті деректер модельдері мен сыныптарына көптеген түрлендірuler, көбінесе қолмен орындалатын, қажет болатын цифрлық «Вавилон мұнарасы» пайда болады. Бұл кедергін жеңу тек әмбебап классификаторларға және стандартталған талаптар жиынтығына кешу арқылы мүмкін болады.

Құрылыш үшін бірдей тіл: цифрлық трансформациядағы классификаторлардың рөлі

Цифрландыру және тексеру мен өндеу процестерін автоматтандыру контекстінде элементтерді классификациялау жүйелері – объектілерді сипаттау мен параметрлеуде біртектілікті қамтамасыз ететін «цифрлық сөздіктер» ерекше рөл атқарады. Классификаторлар деректерді пайдалану мәні бойынша топтастыруға және деректерді әртүрлі жүйелер, басқару деңгейлері мен жобаның өмірлік циклы кезеңдері арасында интеграциялауға мүмкіндік беретін «бірегей тіл» қалыптастырады.

Классификаторлардың ең айқын әсері ғимараттың өмірлік циклы экономикасында байқалады, мұнда ең маңызды аспект ұзақ мерзімді эксплуатациялық шығындарды оңтайландыру болып табылады. Зерттеулер көрсеткендегі, эксплуатациялық шығындар ғимараттың жалпы меншік құнының 80%-на дейін құрайды, бұл құрылышқа жұмсалған бастапқы шығындардан үш есе көп (Рис. 4.26) [98]. Бұл, өз кезегінде, болашақ шығындар туралы шешімдер жобалау кезеңінде қалыптастынын білдіреді.

Сондықтан эксплуатация инженерлерінің талаптары (CAFM, AMS, PMS, RPM) жобалау кезеңінде деректерге қойылатын талаптарды қалыптастырудың бастапқы нұктесі болуы тиіс (Рис. 1.24). Бұл жүйелер жобаның соғы кезеңі ретінде емес, жобаның цифрлық экожүйесінің ажырамас бөлігі ретінде қарастырылуы керек – концепциядан бастап демонтажға дейін.-

Заманауи классификатор – бұл тек топтастыру үшін кодтар жүйесі емес. Бұл архитекторлар, инженерлер, сметашылар, логистер, эксплуатация қызметтері және IT жүйелері арасындағы өзара түсіністік механизмі. Автопилоттың жол объектілерін жоғары дәлдікпен анықтауы сияқты, цифрлық құрылыш жүйелері мен олардың пайдаланушылары әртүрлі жүйелер үшін жобаның бір және сол элементтің классификация арқылы біржақты интерпретациялауы тиіс.



Рис. 4.26 Эксплуатациялық және техникалық шығындар құрылым үш есе асып, ғимараттың өмірлік циклының барлық шығындарының 60–80%-ын құрайды (материалдар бойынша [99]).

Классификаторлардың даму деңгейі компанияның цифрландыру тереңдігімен және оның цифрлық жетілуімен тікелей корреляцияланады. Төмен цифрлық жетілу деңгейі бар үйлемдер деректердің фрагменттелуі, ақпараттық жүйелердің үйлесімсіздігі және, соның салдарынан, классификаторлардың үйлесімсіздігі мен тиімді еместігімен бетпе-бет келеді. Мұндай компанияларда бір және сол элемент әртүрлі жүйелерде жиі әртүрлі топтастыру идентификаторларына ие болуы мүмкін, бұл соңғы интеграцияны критикалық түрде қыннадады және процестерді автоматтандыруды мүмкін емес етеді.

Мысалы, бір және сол терезе жобада CAD моделінде, сметалық және эксплуатациялық жүйеде әртүрлі түрде белгіленуі мүмкін (Рис. 4.27) элементтердің көпасспекттілігінен, яғни процестің әртүрлі қатысушыларының элементтерді қабылдау ерекшеліктерінен. Сметашы үшін «Терезелер» категориясындағы элементте көлемдер мен құн маңызды, эксплуатациялық қызмет үшін – қолжетімділік пен жөндеуге жарамдылық, архитектор үшін – эстетикалық және функционалдық сипаттамалар. Нәтижесінде, бір және сол элемент әртүрлі параметрлерді талап етуі мүмкін.-

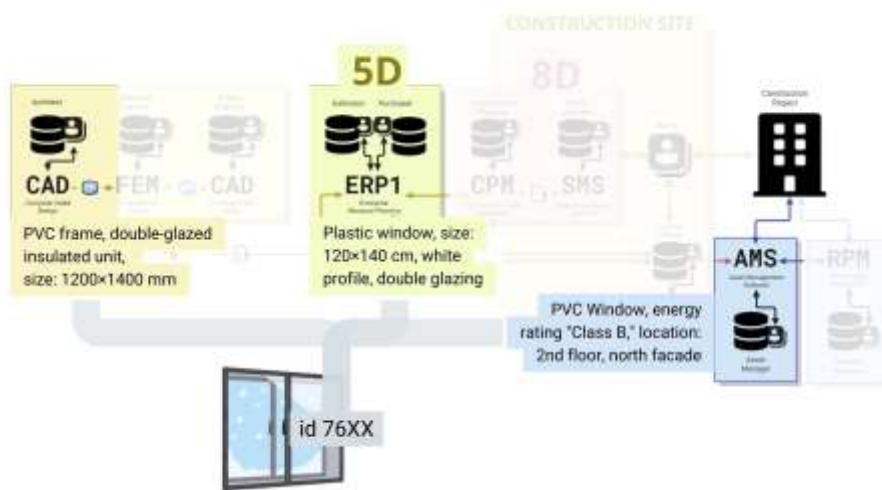


Рис. 4.27 Жүйелер арасында келісілмеген классификация кезінде элемент әр кезеңде басқа жүйеге өткенде атрибутивті ақпараттың бір бөлігін жоғалтады.

Құрылыш элементтерінің классификациясын біржакты анықтаудың қыындығынан, әртүрлі салалардың мамандары бір және сол элементке бір-бірімен үйлеспейтін кластарды жіне тағайындаиды. Бұл объекті туралы біртұтас түсініктің жоғалуына әкеледі, бұл әртүрлі классификация жүйелерін келісу және әртүрлі мамандармен анықталған типтер мен кластар арасындағы сәйкестікіт орнату үшін кейінгі қолмен араласуды талап етеді.

Мұндай келісілмегендік нәтижесінде, сатып алу бөлімі (ERP) өндірушіден құрылыш элементін сатып алған кезде алынған эксплуатациялық құжаттама, кебінесе, құрылыш алаңындағы (PMIS, SCM) осы элементтің классификациясына дұрыс байланыстырылмайды. Нәтижесінде, маңызды ақпараттың интеграцияланбау ықтималдығы жоғары, бұл инфрақұрылым мен активтерді басқару жүйелеріне (CAFM, AMS) елеулі проблемалар туғызады, сондай-ақ объектіні пайдалануға енгізу кезінде, сондай-ақ кейінгі қызмет көрсету (AMS, RPM) немесе осы элементті ауыстыру кезінде қыындықтар тудырады.

Жоғары цифрлық жетілуі бар компанияларда классификаторлар барлық ақпараттық ағындарды біріктіретін жүйке жүйесінің рөлін атқарады. Бір және сол элемент бірегей идентификатор алады, бұл оны CAD, ERP, AMS және CAFM жүйелері мен олардың классификаторлары арасында бүрмаланусыз немесе жоғалусыз беру мүмкіндігін қамтамасыз етеді.

Эффективті классификаторларды құру үшін деректердің қалай пайдаланылатынын түсіну қажет. Бір және сол инженер әртүрлі жобаларда элементті әртүрлі атаумен және классификациямен атауы мүмкін. Тек жылдар бойы пайдалану статистикасын жинау арқылы тұрақты классификация жүйесін қалыптастыруға болады. Бұл машиналық оқытуға көмектеседі: алгоритмдер мындаған жобаларды (Рис. 9.110) талдай отырып, машиналық оқыту арқылы ықтимал кластар мен параметрлерді анықтайты (Рис. 10.16). Автоматты классификация, деректер қөлемі үлкен болған жағдайда, қолмен классификациялау мүмкін болмаған кезде, әсіресе құнды. Автоматты классификация жүйелері элементтің минималды толтырылған параметрлері негізінде негізгі категорияларды ажыратады (книгадағы тоғызынышы және оныншы бөлімдерде толығырақ).--

Дамыған классификатор жүйелері одан әрі цифрандырудың катализаторына айналып, келесі негіздерді құрайды:

- Жобалардың құны мен іске асыру мерзімдерін автоматтандырылған бағалау.
- Потенциалды тәуекелдер мен коллизияларды предиктивті талдау.
- Сатып алу процестерін және логистикалық тізбектерді оңтайландыру.
- Ғимараттар мен құрылыштардың цифрлық егіздерін жасау.
- «Ақылды қала» және заттар интернетімен интеграция.

Трансформация үшін уақыт шектеулі – машиналық оқыту және компьютерлік көру технологияларының дамуы арқасында автоматты классификация мәселесі, ондаған жылдар бойы шешілмеген, алдағы жылдары шешілетін болады, және осыған бейімделмеген құрылыш және жобалуа компаниялары цифрлық платформалармен ығыстырылған такси парктерінің тағдырын қайталауы мүмкін.

Бағалау құны мен мерзімдерін автоматтандыру, сондай-ақ үлкен деректер мен машиналық оқыту

туралы толығырақ бесінші және тоғызынышы бөлімдерде баяндалады. Такси парктерінің тағдырын қайталау мәселелері және құрылыш саласының "Uberизациясы" он бірінші бөлімде егжей-тегжейлі қарастырылады.

Құрылыш саласының цифрлық трансформациясында классификаторлардың маңызды рөлін түсіну үшін олардың эволюция тарихына жүгінү қажет. Тарихи контекст қазіргі жағдайын анықтайдын классификацияға қатысты тәсілдердің қалай дамығанын түсінуге мүмкіндік береді.

Masterformat, OmniClass, Uniclass және CoClass: классификациялық жүйелердің эволюциясы

Тарихи түрғыдан құрылыш элементтері мен жұмыстарын классификациялаушылар үш үрпақта да-мыды, олардың әрқайсысы белгілі бір уақыт кезеңінде қолжетімді технологиялар мен саланың өзекті қажеттіліктерін көрсетеді (Рис. 4.28):

- Бірінші үрпақ (1950-жылдардың басы – 1980-жылдардың соңы) – қағаз анықтамалықтары, жергілікті қолданылатын иерархиялық классификаторлар (мысалы, Masterformat, SfB).
- Екінші үрпақ (1990-жылдардың соңы – 2010-жылдардың ортасы) – Excel және Access-те жүзеге асырылған кестелер мен құрылымдық деректер базалары (ASTM E 1557, Omni-Class, Uniclass 1997).
- Үшінші үрпақ (2010-жылдардан қазіргі уақытқа дейін) – цифрлық қызметтер мен API интерфейстері, CAD (BIM) интеграциясы, автоматтандыру (Uniclass 2015, CoClass).

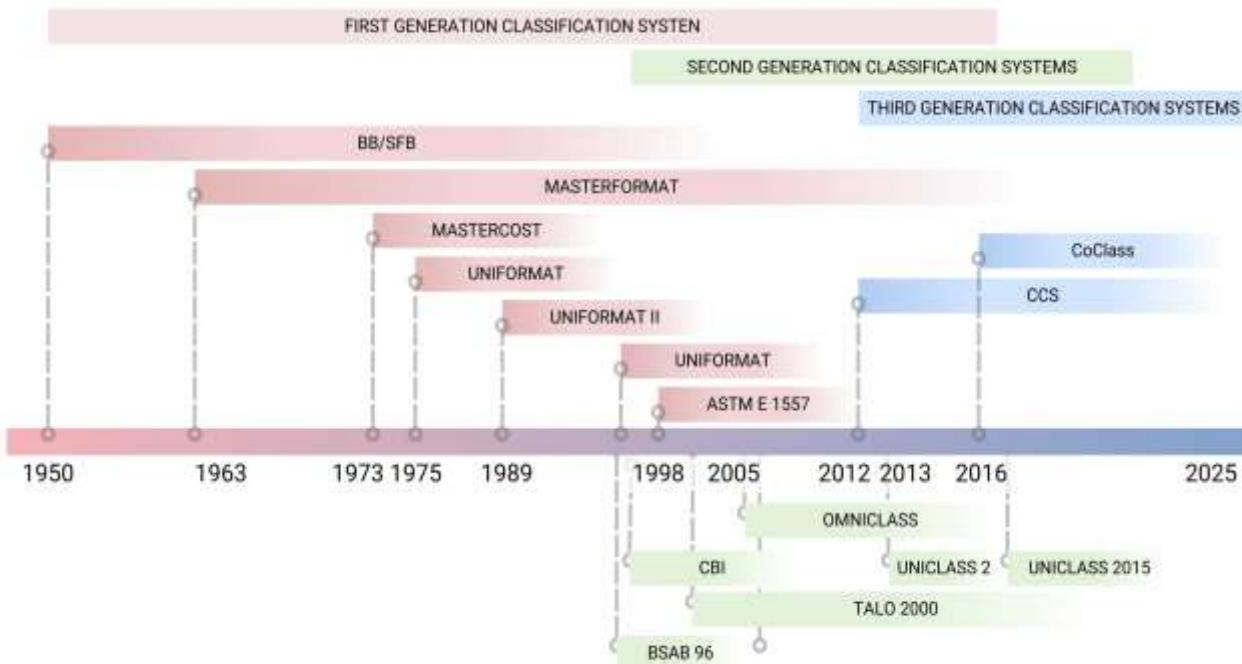


Рис. 4.28 Құрылыш саласының үш үрпақ классификаторлары.

Соңғы онжылдықтарда классификаторлардың иерархиялық күрделілігі азайып келеді (Рис. 4.29): егер алғашқы жүйелер, мысалы, OmniClass, 6887 классты сипаттау үшін 7 деңгейлі иерархияны қолданса, қазіргі шешімдер, мысалы, CoClass, 750 класспен 3 деңгеймен шектеледі. Бұл деректермен жұмыс істеуді жеңілдетеді, сонымен бірге қажетті егжей-тегжейлілікті сақтайды. Ұлыбританияда стандарт ретінде жиі қолданылатын Uniclass 2015 барлығы 4 деңгейде 7210 классты біркітреді, бұл оны CAD жобалары мен мемлекеттік сатып алулар үшін ыңғайлы етеді.

Classifier	Table / Objects	Number of classes	Nesting depth
OmniClass	Table 23 Products	6887	7 levels
Uniclass 2015	Pr — Products	7210	4 levels
CoClass, CCS	Components	750	3 levels

Рис. 4.29 Әр жаңа классификатор үрпағымен категориялау күрделілігі бірнеше есе азаяды.

Әртүрлі елдердің құрылыш сметалық бағалау жүйелерінде классификациялардың әртүрлілігіне байланысты, тіпті бетонды фундаменттік қабырға сияқты типтік элемент мүлдем әртүрлі сипатталармен сипатталуы мүмкін (Рис. 4.210). Бұл айырмашылықтар құрылыш практикасының үлттық ерекшеліктерін, қолданылатын өлшем жүйелерін, материалдарды классификациялау тәсілдерін, сондай-ақ әр елде қолданылатын нормативтік және техникалық талаптарды көрсетеді.



Рис. 4.210 Бір және сол элемент әртүрлі елдерде жобаларда әртүрлі сипаттамалар мен классификациялар арқылы қолданылады.

Бірдей элементтердің әртүрлі классификациялары халықаралық ынтымақтастықты қынданатады, халықаралық жобалар аясында жұмыстардың құны мен көлемін салыстыруды енбеккөр етеді, ал кейде – іс жүзінде мүмкін емес етеді. Қазіргі уақытта, жаһандық деңгейде бірегей универсалды классификатор жоқ – әрбір ел немесе өңір өздерінің жүйелерін әзірлейді, жергілікті нормаларға, тілге және бизнес-мәдениетке сүйене отырып.

- CCS (Дания): Нысанның өмірлік циклі бойында шығындарды классификациялау жүйесі (жобалау, құрылыш, пайдалану). Пайдалану және техникалық қызмет көрсету логикасына акцент жасалады, бірақ бюджеттер мен ресурстарды басқаруды да қамтиды.
- NS 3451 (Норвегия): Нысандарды функциялар, конструктивтік элементтер және өмірлік цикл кезеңдері бойынша классификациялады. Жобаларды басқару, құнды бағалау және ұзақ мерзімді жоспарлау үшін қолданылады.
- MasterFormat (АҚШ): Құрылыш спецификацияларын бөлімдер бойынша құрылымдау жүйесі (мысалы: бетон, электр монтажы, әрлеу). Функционалдық элементтерге емес, пәндер мен жұмыс түрлеріне назар аударады (UniFormat-тан айырмашылығы).
- Uniclass 2 (Ұлыбритания): Мемлекеттік сатып алулар мен BIM-жобаларда қолданылатын ең ежей-төгжелі класификаторлардың бірі. Нысандар, жұмыстар, материалдар және кеңістіктер туралы деректерді біртұтас жүйеге біріктіреді.
- OmniClass: Нысандар туралы ақпаратты басқару үшін халықаралық стандарт (АҚШ-та CSI әзірлеген): компоненттер кітапханаларынан электронды спецификацияларға дейін. Деректерді ұзақ мерзімді сақтау үшін қолайлы, CAD (BIM) және басқа цифрлық құралдармен үйлесімді.
- COBie: Құрылыш-эксплуатация ғимаратының ақпарат алмасу стандарты – жобалау, құрылыш және пайдалану кезеңдері арасында деректер алмасудың халықаралық стандарты. BS 1192-4:2014-ке енгізілген, «пайдалануға дайын BIM-моделі» концепциясының бір бөлігі. Ақпаратты беру (мысалы, жабдық спецификациялары, кепілдіктер, мердігерлердің байланыстары) бойынша акцент жасайды.

Құрылыш саласының глобализациясы, ықтимал, құрылыш элементтерінің классификация жүйелерінің біртіндеп унификациялануына әкеледі, бұл жергілікті ұлттық стандарттарға тәуелділікті едәуір төмендетеді. Бұл процесс интернет-коммуникациялардың эволюциясымен аналогия бойынша дамуы мүмкін, мұнда универсалды деректерді беру протоколдары ақырында бөлшектелген жергілікті форматтарды ығыстырып, жүйелердің жаһандық үйлесімділігін қамтамасыз етті.

Альтернативті даму жолы машиналық оқыту технологияларына негізделген автоматты классификация жүйелеріне тікелей көшу болуы мүмкін. Бұғынгі күні негізінен автономды көлік саласында әзірленіп жатқан бұл технологиялар CAD-жобалаудың үлкен деректер массивтеріне қолдану үшін едәуір әлеуетке ие.

Бұғынгі күні жағдай тек ұлттық кластеризациямен шектелмейді. Мемлекеттік деңгейде ескерілмеген көптеген ерекшеліктердің салдарынан, әрбір компания элементтер мен ресурстардың категорияларын унификациялау және стандарттау мәселелерімен өз бетінше айналысуға мәжбүр.

Әдетте, бұл процесс кішігірімнен басталады – объектілердің локалды кестелерінен немесе ішкі белгілер жүйесінен. Алайда стратегиялық мақсат – барлық элементтердің біртұтас сипаттамалық тілге көшу, ол тек компания ішінде ғана емес, сонымен қатар оның шегінен тыс, идеалында халықаралық немесе салалық класификаторлармен келісілген болуы тиіс. Мұндай тәсіл сыртқы серікtestермен, цифрлық жүйелермен интеграцияны жеңілдетеді және объектілердің өмірлік циклында біртұтас процестердің қалыптасуына ықпал етеді.

Автоматтандыру мен масштабталатын IT-жүйелеріне көшу алдында ұлттық деңгейдегі классифи-

каторларды пайдалану немесе элементтердің логикалық және бірмәнді идентификация құрылымын құру қажет. Әрбір объект – терезе, есік немесе инженерлік жүйе болсын – компанияның кез келген цифрлық жүйесінде қателіксіз танылатында сипатталуы тиіс. Бұл жазық сызбалардан цифрлық модельдерге көшу кезінде, жобалау кезеңін және ғимараттарды пайдалану кезеңін қамтитын, өте маңызды.

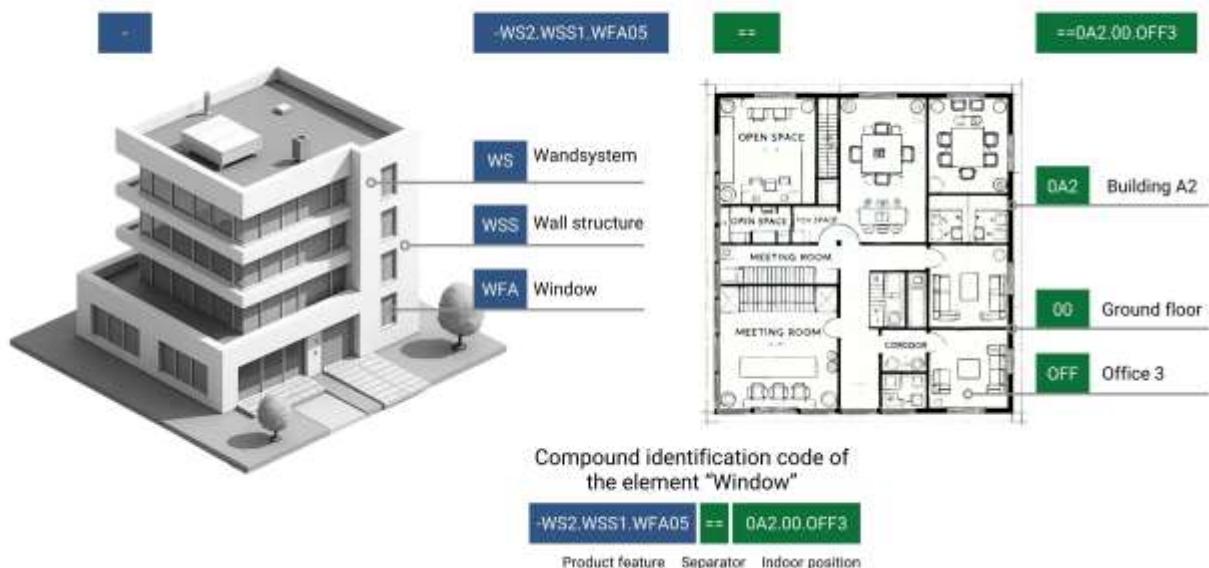


Рис. 4.211 Құрылыш элементінің терезе құрамды идентификаторына мысал, классификация мен ғимараттағы орнына негізделген.

Ішкі классификаторлардың бір мысалы – идентификацияның құрамды кодын әзірлеу. Мұндай код бірнеше ақпарат деңгейлерін біріктіреді: элементтің функционалдық мақсаты (мысалы, "Қабырғадағы терезе"), оның түрі, сондай-ақ нақты кеңістіктік байланыс – A2 ғимараты, 0 қабат, 3 бөлме. Мұндай көп деңгейлі құрылым цифрлық модельдер мен құжаттама бойынша біртұтас навигациялық жүйені құруға мүмкіндік береді, әсіресе деректерді тексеру және трансформациялау кезеңдерінде, мұнда элементтердің бірмәнді топтастырылуы қажет. Элементтің бірмәнді танылуы бөлімдер арасындағы келісімді қамтамасыз етеді және дублирлеу, қателіктер мен ақпарат жоғалту тәуекелдерін азайтады.

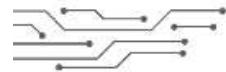
Жақсы құрылымдалған классификатор – бұл тек техникалық құжат емес, компанияның цифрлық экожүйесінің негізі:

- жүйелер арасында деректердің үйлесімділігін қамтамасыз етеді;
- ақпаратты іздеу және қайта өндеу шығындарын азайтады;
- ашықтық пен басқарушылықты арттырады;
- масштабтау мен автоматтандыру үшін негіз құрайды.

Объектілердің стандартталған сипаттамасы, ұлттық классификаторларды немесе өзіміздің құрамды идентификация кодтарын қолдану арқылы, біртұтас деректер, сенімді ақпарат алmasу және интеллектуалды қызметтерді енгізу – автоматтандырылған сатып алушардан цифрлық

егіздерге дейін – негіз болады.

Әртүрлі форматтағы деректерді құрылымдау кезеңі мен элементтерді тану және топтастыру үшін қолданылатын классификаторды таңдау аяқталғаннан кейін, келесі қадам – деректерді дұрыс мөдільдеу. Бұл процесс негізгі параметрлерді анықтауды, логикалық деректер құрылымын құруды



және элементтер арасындағы өзара байланыстарды сипаттауды қамтиды.

ТАРАУ 4.3.

ДЕРЕКТЕРДІ МОДЕЛДЕУ ЖӘНЕ ОЗЫҚ ТӘЖИРИБЕ ОРТАЛЫҒЫ

Деректерді моделдеу: концептуалдық, логикалық және физикалық модель

Деректерді (бұрын құрылымдалған және классификацияланған) тиімді басқару, сақтау және өндіру үшін ойластырылған құрылымсыз мүмкін емес. Ақпаратқа қолжетімділікті және келісімділікті қамтамасыз ету үшін компаниялар деректерді модельдеу әдістемесін пайдаланады – бұл кестелерді, деректер базаларын және олардың арасындағы байланыстарды бизнес талаптарына сәйкес жобалауға мүмкіндік беретін әдіс.

Деректерді модельдеу – бұл кез келген цифрлық экожүйенің негізі. Жүйелерді, талаптарды және деректерді модельдемей, деректерді жасайтын инженерлер мен мамандар өздерінің жасаған деректерінің қайда қолданылатынын білмейді және түсінбейді.

Фимарат салу сияқты, мұнда кірпіштерді тәсевуді жоспарсыз бастау мүмкін емес, деректерді сақтау жүйесін құру үшін қандай деректер қолданылатынын, олардың бір-бірімен қалай байланысатынын және кімдермен жұмыс істелетінін нақты түсіну қажет. Процестер мен талаптарды сипаттамай, инженерлер мен деректерді жасайтын мамандар деректердің қайда және қалай қолданылатынын көруден айрылады.

Деректер моделі бизнес пен IT арасындағы байланыстыруши буын қызметін атқарады. Ол талаптарды формализациялауға, ақпаратты құрылымдауға және мүдделі тараптар арасындағы коммуникацияны жеңілдетуге мүмкіндік береді. Осы тұрғыдан алғанда, деректерді модельдеу тапсырыс берушінің идеясы бойынша ғимараттың жоспарын әзірлейтін архитектордың жұмысымен үксас, содан кейін оны құрылышыларға – деректер базасының әкімшілері мен әзірлеушілеріне – жүзеге асыру үшін береді.

Осылайша, әрбір құрылым компаниясы элементтер мен ресурстарды құрылымдау мен категориялау (Рис. 4.211) жұмыстарымен қатар, деректер базаларын (kestelerді) "құру" өнерін меңгеруі және оларды бір-бірімен байланыстыруды үйренуі тиіс, бұл деректердің компаниядағы білімдерден тұратын мықты және берік қабырғасын құрастыру сияқты. Деректерді модельдеудегі негізгі үғымдар (Рис. 4.31) мыналарды қамтиды:–

- Субъектілер – бұл жинақталуы қажет деректер туралы объектілер. Жобалаудың бастапқы кезеңінде субъект жеке элемент (мысалы, "есік") болуы мүмкін, ал сметалық модельде – категориялар бойынша біріктілген элементтер тобы (мысалы, "ішкі есіктер").
- Атриуттар – субъектілердің сипаттамалары, маңызды мәліметтерді сипаттайтыды: өлшемдер, қасиеттер, жинақтау құны, логистика және басқа параметрлер.
- Байланыстар (қатынастар) – субъектілердің бір-бірімен қалай әрекеттесетінін көрсетеді. Олар "бір-біріне", "көп-біріне", "көп-көпке" типтерінің біріне жатуы мүмкін.
- ER-диаграммалар (Субъект-Қатынас диаграммалары) – субъектілерді, атриуттарды және олардың арасындағы байланыстарды көрсететін визуалды схемалар. ER-диаграммалары

концептуалдық, логикалық және физикалық болып табылады – әрқайсысы деректер құрылымын көрсету деңгейімен ерекшеленеді.

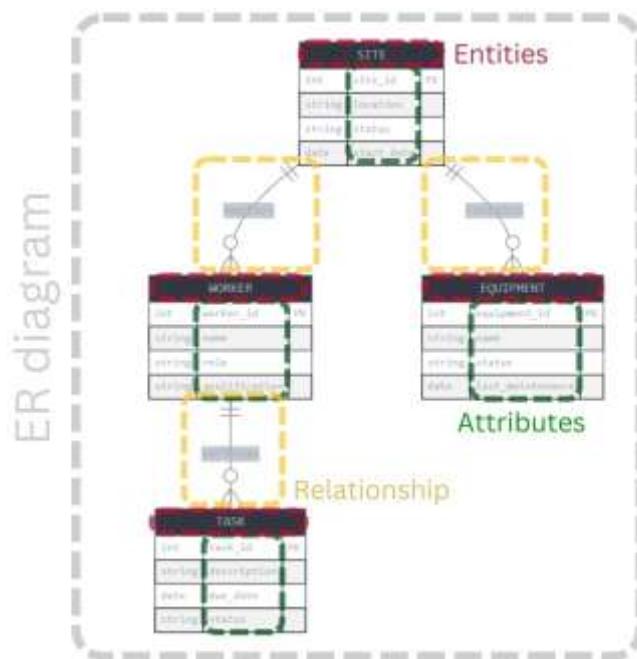


Рис. 4.31 ER-диаграмма концептуалдық деректер базасының құрылымын субъекттермен, атрибуттармен және байланыстармен көрсетеді.

Деректерді жобалау және олардың арасындағы өзара байланыстарды анықтау процесі дәстүрлі түрде үш негізгі модельге бөлінеді. Әрқайсысы белгілі бір функцияларды орындайды, деректер құрылымын көрсету деңгейі мен абстракция дәрежесі бойынша ерекшеленеді:

- Концептуалдық деректер моделі: бұл модель негізгі субъекттер мен олардың өзара байланыстарын сипаттайтын, атрибуттардың егжей-тегжейлеріне тереңдемейді. Эдетте, ол жоспарлаудың бастапқы кезеңдерінде қолданылады. Осы кезеңде біз деректер базалары мен жүйелердің сыйбаларын жасау арқылы әртүрлі бөлімдер мен мамандар арасындағы байланысты көрсетуіміз мүмкін.

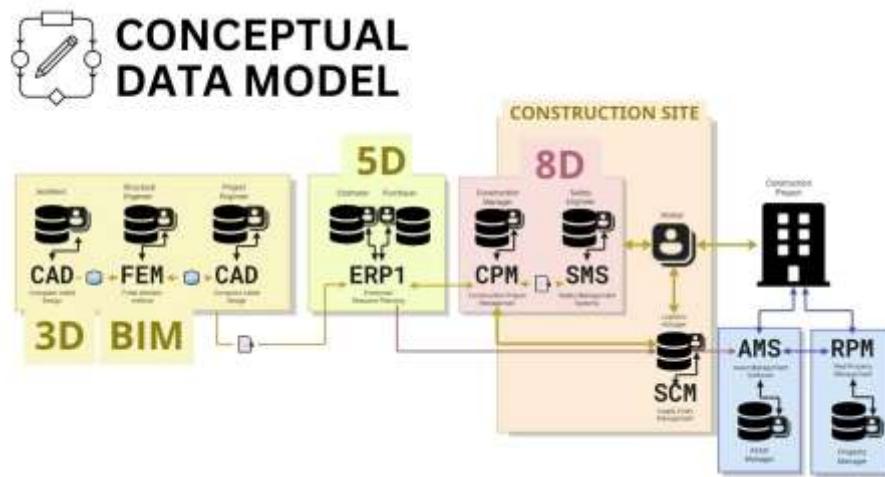


Рис. 4.32 Концептуалдық диаграмма жүйенің мазмұнын сипаттайтын: техникалық мәліметтерсіз жоғары деңгейдегі қатынастарды көрсетеді.

- Логикалық деректер моделі: концептуалдық модельге негізделе отырып, логикалық деректер моделі субъектілердің, атрибуттардың, кілттердің және қатынастардың егжей-тегжейлі сипаттамасын қамтиды, бизнес ақпаратын және ережелерді көрсетеді.

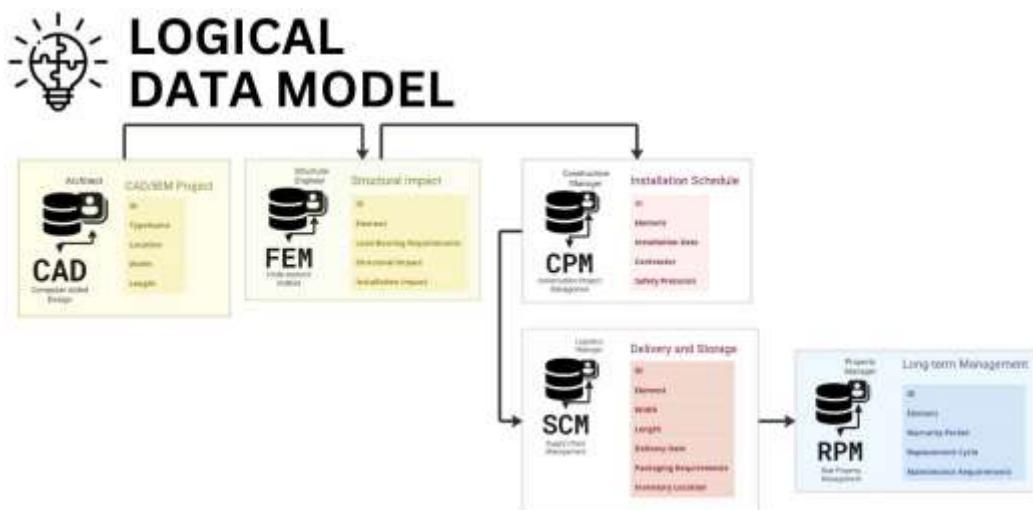


Рис. 4.33 Логикалық деректер моделі деректердің түрлерін, қатынастарды және кілттерді егжей-тегжейлі сипаттайды, бірақ жүйелік іске асырусыз.

- Физикалық деректер моделі: бұл модель деректер базасын іске асыру үшін қажетті құрылымдарды, соның ішінде кестелерді, бағандарды және қатынастарды сипаттайтын. Ол деректер базасының өнімділігіне, индекстеу стратегияларына және физикалық сақтауына назар аударады, деректер базаларын физикалық орналастыруды оңтайландыру үшін.

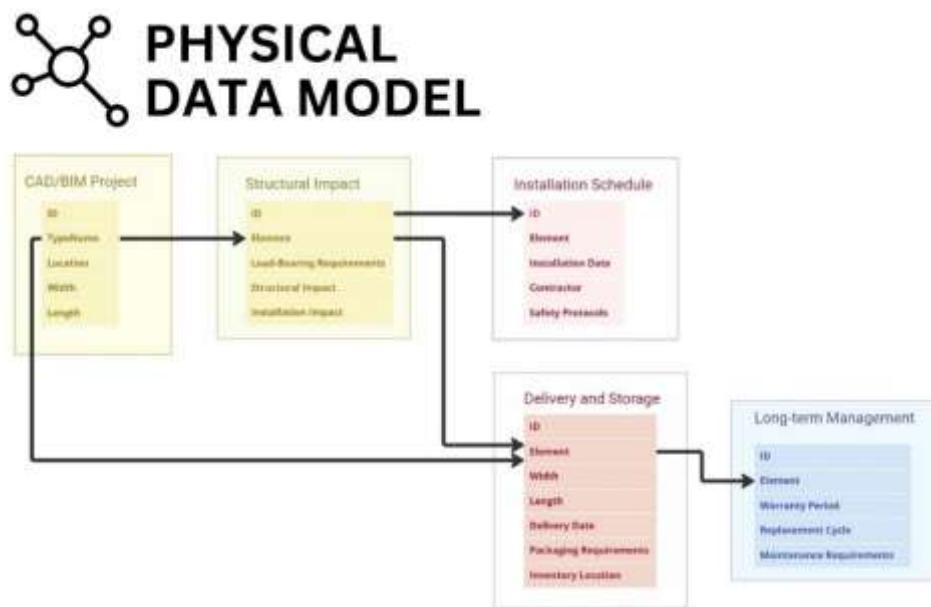


Рис. 4.34 Физикалық деректер моделі жүйенің қалай іске асырылатынын, соның ішінде кестелер мен деректер базасының нақты мәліметтерін анықтайды.

Деректер базаларын әзірлеу және кестелік қатынастарды жобалау кезінде абстракция деңгейлерін түсіну жүйенің тиімді архитектурасын құруда маңызды рөл атқарады.

Тиімді деректер моделдеу әдістемесі бизнес міндеттерін техникалық іске асырумен біріктіруге мүмкіндік береді, барлық процестер тізбегін айқын және басқарылатын етеді. Деректер моделдеу – бұл бір реттік міндет емес, кезең-кезеңімен жүзеге асырылатын процесс (Рис. 4.35): -

- Бизнес талаптарын жинау: негізгі міндеттер, мақсаттар және ақпарат ағындары анықталады. Бұл кезең сарапшылармен және пайдаланушылармен белсенді өзара әрекеттесуді қамтиды.
- Субъектілерді анықтау: болашақ жүйеде ескеру қажет негізгі обьектілер, санаттар және деректер түрлері бөлінеді.
- Концептуалдық және логикалық модельді әзірлеу: алдымен негізгі субъектілер мен олардың байланыстары фиксиленеді, содан кейін – атрибуттар, ережелер және егжеттегжейлі құрылым.
- Физикалық моделдеу: модельдің техникалық іске асырылуы жобаланады: кестелер, өрістер, байланыстар, шектеулер, индекстер.
- Деректер базасын құру: соңғы қадам – таңдалған СУБД-де физикалық модельді іске асыру, тестілеу жүргізу және пайдалануға дайындау.

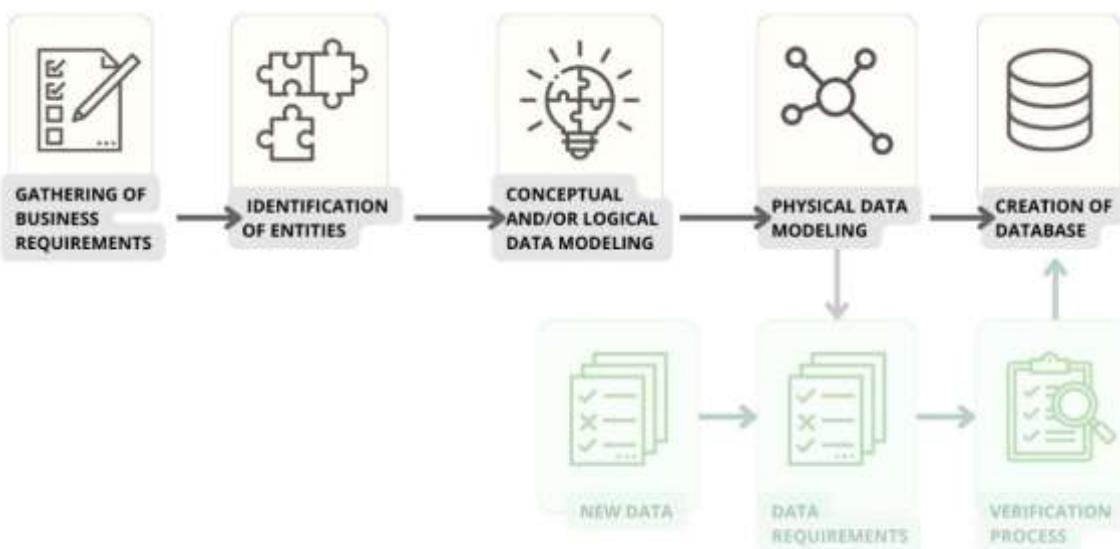


Рис. 4.35 Деректер базаларын және бизнес процестері үшін деректерді басқару құйеелерін құру талаптарды қалыптастыру мен деректер моделдеуінен басталады.

Деректер моделдеу процестерін дұрыс ұйымдастыру ақпарат ағындарының айқындығын қамтамасыз етеді, бұл құрылымдың жобасын немесе құрылымдың алаңын басқару сияқты курделі жобаларда ерекше маңызды. Концептуалдық модельден логикалық модельге, содан кейін физикалық модельге өту процестерді реттеуге қалай көмектесетінін қарастырайық.

Құрылымдық контекстінде практикалық деректер моделдеуі

Деректер моделдеу мысалы ретінде құрылымдық алаңын басқару міндеттін алайық және прорабтың талаптарын құрылымдалған логикалық модельге түрлендірейік. Құрылымдың негізгі қажеттіліктеріне сүйене отырып, біз: құрылымдық алаң (SITE), жұмысшылар (WORKER), жабдық (EQUIPMENT), тапсырмалар (TASK) және жабдықты пайдалану (EQUIPMENT_USAGE) үшін негізгі субъектілерді анықтаймыз. Әрбір субъект маңызды сипаттамаларды көрсететін атрибуттар жиынтығын қамтиды. Мысалы, TASK үшін бұл тапсырманың сипаттамасы, орындау мерзімі, статусы, приоритеті; WORKER үшін – аты, алаңдағы рөлі, ағымдағы жұмыспен қамтылуы және т. б.

Логикалық модельде осы объектілер арасындағы байланыстар орнатылады, олар нақты жұмыс процестерінде қалай өзара әрекеттесетінін көрсетеді (Рис. 4.36). Мысалы, алаң мен жұмысшылар арасындағы байланыс бір алаңда көптеген жұмысшылардың жұмыс істей алатынын көрсетеді, ал жұмысшылар мен тапсырмалар арасындағы байланыс бір жұмысшының бірнеше тапсырманы орындаған алатынын білдіреді.

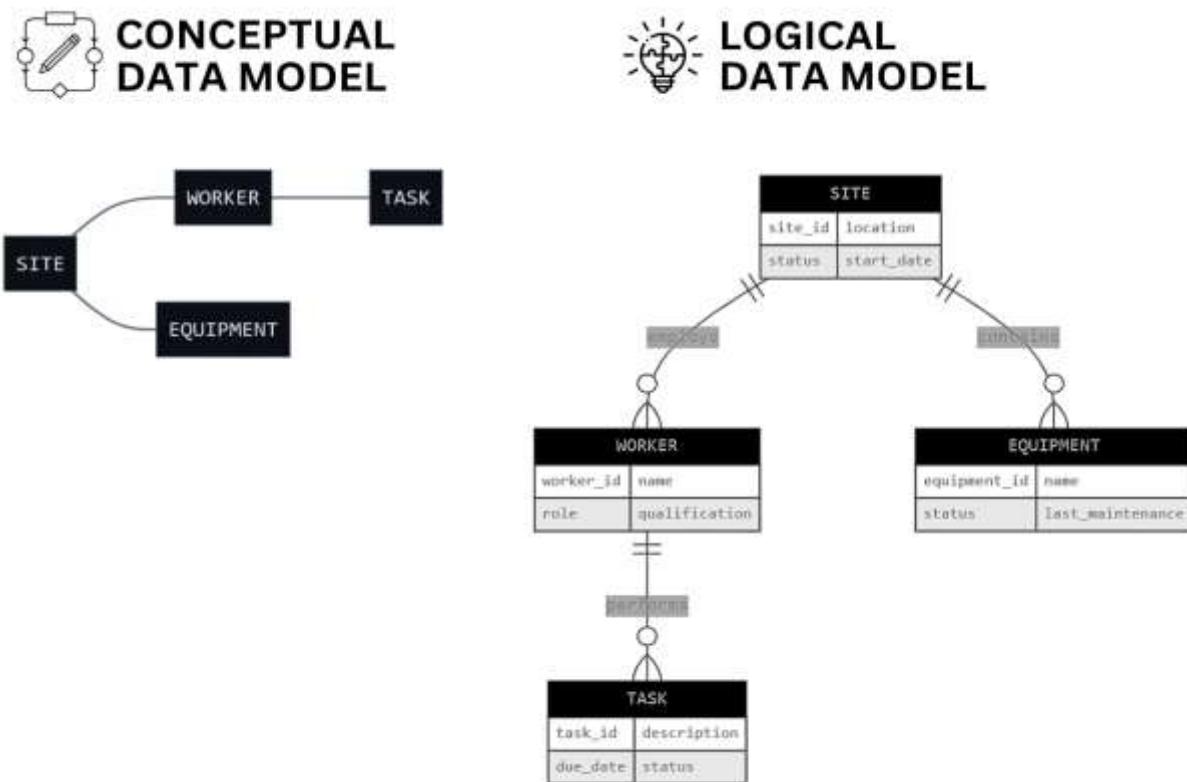


Рис. 4.36 Концептуалдық және логикалық деректер моделі, құрылыш алаңындағы процестерді сипаттау үшін прорабтың талаптарына негізделген.

Физикалық модельге көшкенде, іске асырудың техникалық мәліметтері қосылады: нақты деректер типтері (VARCHAR, INT, DATE), кестелер арасындағы байланыстар үшін бастапқы және сыртқы кілттер, сондай-ақ деректер базасының өнімділігін оңтайландыру үшін индекстер (Рис. 4.37).

Мысалы, статустар үшін мүмкін мәндері бар нақты типтерді анықтау қажет, ал іздеу өнімділігін артыру үшін `status` және `worker_id` сияқты кілттік өрістер бойынша индекстер қосу қажет. Бұл логикалық жүйе сипаттамасын деректер базасын құру және енгізу үшін дайын нақты жоспарға айналдырады.

PHYSICAL DATA MODEL

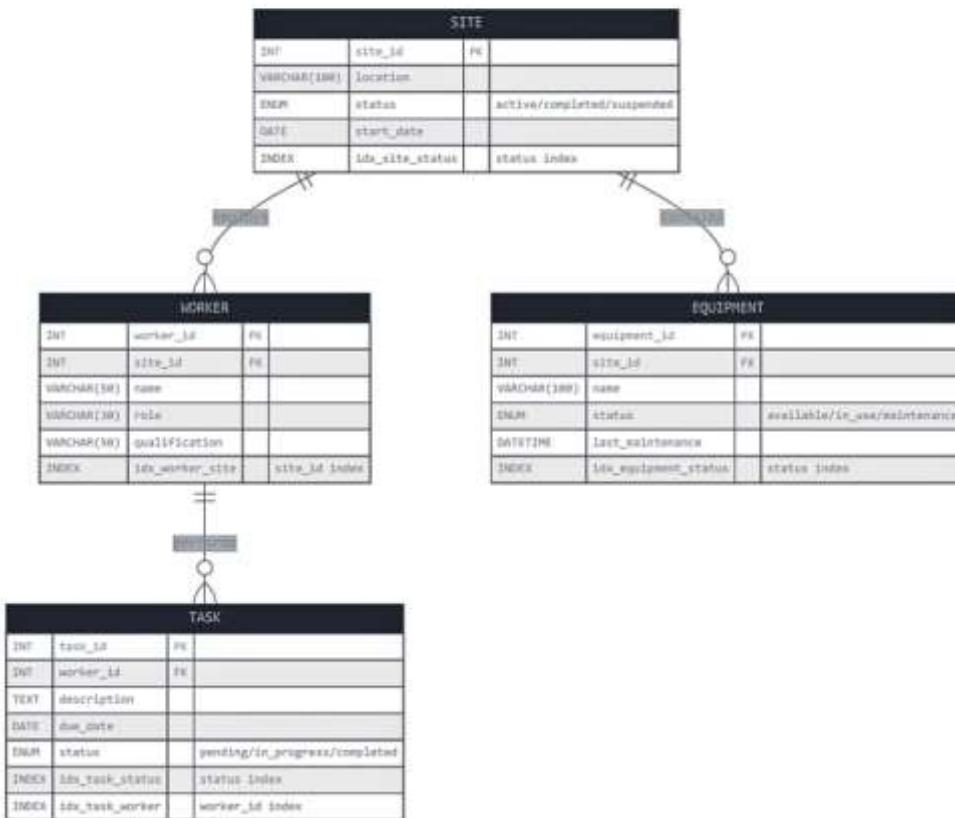


Рис. 4.37 Физикалық деректер моделі құрылыш алаңының объектілерін минималды қажетті параметрлер арқылы сипаттайты.

Физикалық модель логикалық модельден жиі ерекшеленеді. Орташа алғанда, модельдеуге жұмысалатын уақыттың үлестері былайша бөлінеді: концептуалдық модельге шамамен 50% (талаптарды жинау, процестерді талқылау, объектілерді анықтау), логикалық модельге 10% (атрибуттар мен байланыстарды нақтылау) және физикалық модельге 40% (иске асыру, тестілеу, СУБД-ге бейімдеу).

Мұндай баланс концептуалдық кезең деректер құрылымының негізін қалайды, ал логикалық модель тек байланыстар мен атрибуттарды нақтылайды. Физикалық модельге ең көп ресурстар қажет, себебі дәл осы кезеңде деректер нақты платформалар мен құралдарға енгізіледі.

LLM көмегімен деректер базасын құру

Деректер моделін және объектілерді параметрлер арқылы сипаттағаннан кейін, біз деректер базаларын құруға дайынбыз – ақпаратты сақтау орындары, мұнда біз құрылымдау кезеңінен кейін нақты процестер бойынша келетін ақпаратты сақтаймыз.

SQLite көмегімен Python бағдарламалау тілінің мысалында минималды кодпен қарапайым, бірақ функционалды деректер базасын құруды қарастырайық. Реляциялық деректер базалары "Структурированные реляционные базы данных и язык запросов SQL" бөлімінде егжей-тегжейлі қарастырылған.

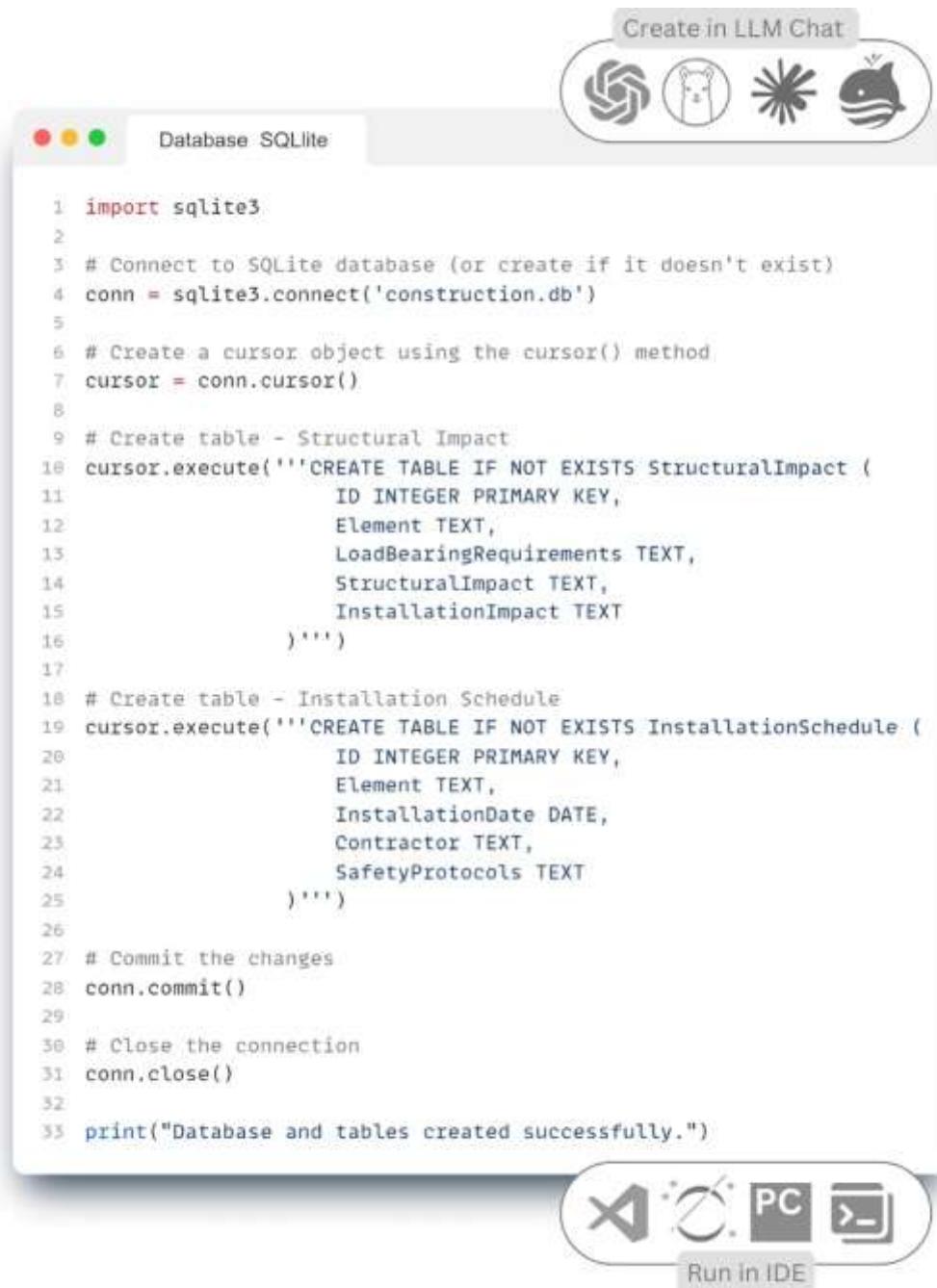
SQLite – жеңіл, open source, кіріктірілген, кроссплатформалық деректер базасы, негізгі SQL функцияларын қолдайды. SQLite CAD (BIM), ERP, EPM жүйелерінде кеңінен қолданылады және дерлік әр смартфонда бар.

LLM моделінен жоғарыда жасалған физикалық модель негізінде бізге шағын SQLite деректер базасын жасаудың сұраймызды (Рис. 4.34) FEM (құрылымдық әсер) және СРМ (монтаж кестесі) жүйелері үшін.-

- ❷ LLM чатында (ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе кез келген басқа) мәтіндік сұраныс жіберіңіз:

"Конструктивное воздействие" және "График установки" жүйелері үшін SQLite деректер базаларын жасаңыз, қосыншада көрсетілген суретке (Рис. 4.34) сілтеме жасаңыз.

■ LLM-нің жауабы:



The screenshot shows a window titled "Database SQLite". At the top right, there are four icons labeled "Create in LLM Chat": a neural network, a bear, a flower, and a whale. Below the title bar, there are standard Mac OS X window controls (red, yellow, green). The main area contains the following Python code:

```

1 import sqlite3
2
3 # Connect to SQLite database (or create if it doesn't exist)
4 conn = sqlite3.connect('construction.db')
5
6 # Create a cursor object using the cursor() method
7 cursor = conn.cursor()
8
9 # Create table - Structural Impact
10 cursor.execute('''CREATE TABLE IF NOT EXISTS StructuralImpact (
11         ID INTEGER PRIMARY KEY,
12         Element TEXT,
13         LoadBearingRequirements TEXT,
14         StructuralImpact TEXT,
15         InstallationImpact TEXT
16     )''' )
17
18 # Create table - Installation Schedule
19 cursor.execute('''CREATE TABLE IF NOT EXISTS InstallationSchedule (
20         ID INTEGER PRIMARY KEY,
21         Element TEXT,
22         InstallationDate DATE,
23         Contractor TEXT,
24         SafetyProtocols TEXT
25     )''' )
26
27 # Commit the changes
28 conn.commit()
29
30 # Close the connection
31 conn.close()
32
33 print("Database and tables created successfully.")

```

At the bottom right, there is a "Run in IDE" button with icons for VS Code, PyCharm, and a terminal.

Рис. 4.38 Генерацияланған код жүктелген суреттөн деректер базалары мен кестелерді құрады, енді біз деректерді жүктей аламыз.

Бұл деректер базасын құру коды Python ортасында SQLite қолдауымен орындалуы мүмкін, ол әдette Python стандартты ортасына енгізілген. Оны скрипттерді іске қосып, файлдар жасауға қабілетті кез келген жергілікті машинада іске қосуға болады. Иске қосылып, орындалғаннан кейін, бұл код қатты дискіде construction.db атты SQLite деректер базасы файлын жасайды (Excel файлағына ұқсас, жұмыс парақтарымен), ол "Структуралық әсер" және "Орнату кестесі" атты кестелерді қамтиды.

construction.db жаңа деректер базасында кестелерді SQL сұраулары немесе импорт арқылы деректерді қосуға болады, бұл кейін автоматтандырылған деректер өңдеуді бастауға мүмкіндік береді. Деректерді SQLite деректер базасына CSV файлдарынан, Excel электрондық кестелерінен немесе басқа деректер базалары мен қоймаларынан API арқылы экспорттау арқылы импорттауға болады.

Деректерді модельдеу және деректер базаларын тиімді басқару үшін компанияларға нақты анықталған стратегия қажет, сондай-ақ техникалық және бизнес командалары арасында үйлесімділік орнату қажет. Жеке жобалар мен көптеген дереккөздер жағдайында барлық деңгейлерде келісімділік, стандарттау және сапа бақылауын қамтамасыз ету қыын. Негізгі шешімдердің бірі компания ішінде Деректерді модельдеу бойынша Озық тәжірибе орталығын (Data Modeling Center of Excellence, CoE) құру болуы мүмкін.

Деректер моделдеуі бойынша Озық Тәжірибе Орталығы (CoE)

Деректер стратегиялық активтердің бірі болып табылатын жағдайда, компанияларға ақпаратты дұрыс жинап, сақтау жеткіліксіз – деректерді жүйелі түрде басқаруды үйрену маңызды. Деректерді классификациялау және модельдеу бойынша Озық тәжірибе орталығы (Center of Excellence, CoE) – бұл үйымдағы деректермен жұмыс істеудің келісімділігін, сапасын және тиімділігін қамтамасыз ететін құрылымдық бөлімше.

Озық тәжірибе орталығы (CoE) – компаниядағы цифровық трансформациялар үшін сараптамалық қолдау мен әдістемелік негіздің ядросы. Ол деректермен жұмыс істеу мәдениетін қалыптастырады және үйымдарға шешімдерді интуиция немесе жергілікті ақпарат негізінде емес, құрылымдалған, тексерілген және репрезентативті деректер негізінде қабылдауға мүмкіндік береді.

Деректер бойынша Озық тәжірибе орталығы әдette "екі пицца" принципі бойынша жұмыс істейтін кросс-функционалды командалардан құралады. Джек Бозос ұсынған бұл принцип команда көлемі екі пицзамен тамақтандыруға болатында болуы керек, яғни 6-10 адамнан аспауы тиіс. Мұндай тәсіл артық бюрократияны болдырмауға және жұмыстың икемділігін арттыруға көмектеседі. CoE командасы деректер аналитикасы мен машиналық оқудан бастап нақты бизнес салаларындағы сараптамаға дейін әртүрлі техникалық дағдылары бар қызметкерлерді қамтуы тиіс. Терен техникалық білімге ие деректерді өңдеу инженерлері процестерді оңтайландырумен және деректерді модельдеумен қатар, әріптерлеріне қолдау көрсетуі керек, рутиналық тапсырмаларды орындау уақытын қысқартады.-

Табиғатта экожүйенің тұрақтылығы биологиялық әртүрлілік арқылы қамтамасыз етілсе, цифрилік әлемде икемділік пен бейімделушілік деректермен жұмыс істеудің әртүрлі тәсілдері арқылы қол жеткізіледі. Алайда, бұл әртүрлілік бірегей ережелер мен түсініктеге негізделуі тиіс.

Озық тәжірибе орталығын (СоE) орман экожүйесінің "климаттық жағдайларымен" салыстыруға болады, олар қандай деректер түрлерінің гүлденетінін, ал қандай деректердің автоматты түрде жойылатынын анықтайды. Сапалы деректер үшін қолайлы "климат" құру арқылы СоE үздік тәжірибелер мен әдістемелердің табиғи сұрыпталуына ықпал етеді, олар кейіннен үйым стандарттарына айналады.

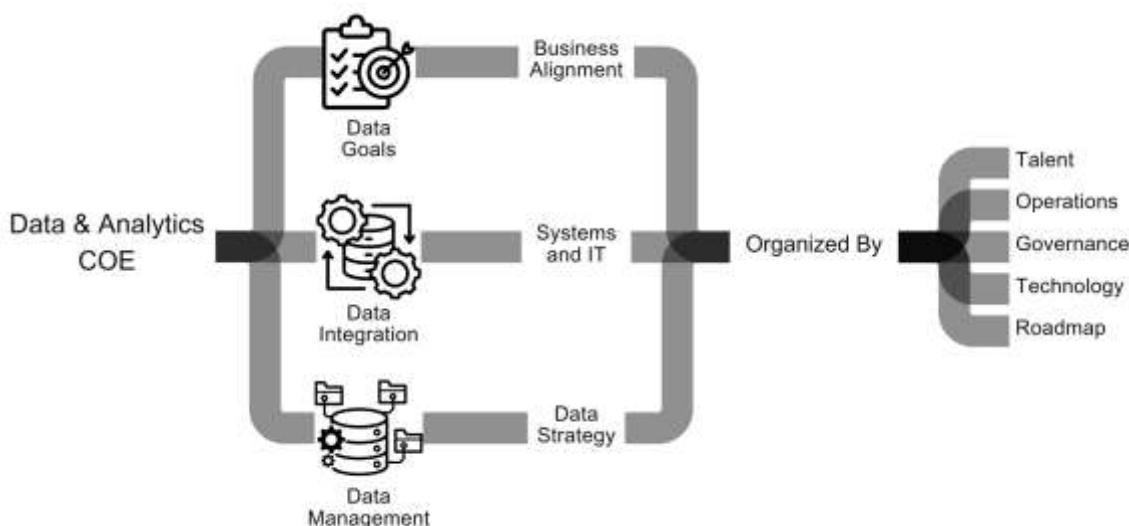


Рис. 4.39 Деректер мен аналитика бойынша Озық тәжірибе орталығы (СоE) деректердің басқарудың, интеграциялаудың және стратегияны әзірлеудің негізгі аспекттері бойынша саралтаманы біріктіреді.

Интеграциялық циклдарды жеделдету және жақсы нәтижеге қол жеткізу үшін СоE мүшелеріне шешім қабылдауда жеткілікті автономия беруі тиіс. Бұл әсіресе динамикалық ортада, сынақ және қателік әдісі, тұрақты кері байланыс және жиі шығарылымдар елеулі пайда әкелуі мүмкін. Дегенмен, бұл автономия тек жоғары басшылықтың нақты коммуникациялары мен қолдауы болған жағдайда тиімді. Стратегиялық көзқарас пен жоғарыдан үйлестірусіз, тіпті ең білікті команда да өз бастамаларын енгізуде кедергілерге тап болуы мүмкін.

Деректерді модельдеу тәсілі бір-екі жобамен шектелмей, ақпаратты басқару және бизнес-процестердің жалпы жүйесіне енгізуін қамтамасыз ету үшін СоE немесе компанияның жоғары басшылығы жауапты.

Озық тәжірибе орталығы (СоE) деректерді модельдеу және басқару (Data Governance) міндеттерінен басқа, деректер инфрақұрылымын енгізу және пайдалану бойынша біртұтас стандарттар мен тәсілдерді әзірлеуге жауапты. Сонымен қатар, ол үйымдағы процестерді үздіксіз жетілдіру, онтайландыру және деректерді тиімді пайдалану мәдениетін қалыптастырады (Рис. 4.310).

СоE ішіндегі деректерді және модельдерді басқаруға жүйелі көзқарас бірнеше негізгі блоктарға бөлінеді:

- Процестерді стандарттау және модельдердің өмірлік циклын басқару: СоE деректер модельдерін жасау мен басқаруды унификациялауға мүмкіндік беретін әдістемелерді әзірлеп, енгізеді. Бұл құрылымдық шаблондарды, сапаны бақылау әдістерін және деректердің барлық жұмыс кезеңдерінде үздіксіздігін қамтамасыз ететін нұсқаларды басқару жүйелерін қалыптастыруды қамтиды.
- Рөлдерді басқару және жауапкершілікті бөлу: СоE аясында деректерді модельдеу процесінде негізгі рөлдер анықталады. Жобаның әр қатысуышына нақты анықталған функциялар мен жауапкершілік аймақтары беріледі, бұл командалардың үйлесімді жұмысын қамтамасыз етеді және деректердің сәйкесіздік тәуекелдерін азайтады.
- Сапаны бақылау және аудит: құрылыш деректерін тиімді басқару олардың сапасын тұрақты мониторингтеуді талап етеді. Деректерді тексеру, қателерді анықтау, жетіспейтін атрибуттарды табу үшін автоматтандырылған механизмдер енгізіледі.
- Мета деректерді және ақпарат архитектурасын басқару: СоE біртұтас класификация және идентификаторлар, атау стандарттары мен объектілерді сипаттау жүйесін құруға жауапты, бұл жүйелер арасындағы интеграция үшін критикалық маңызды.

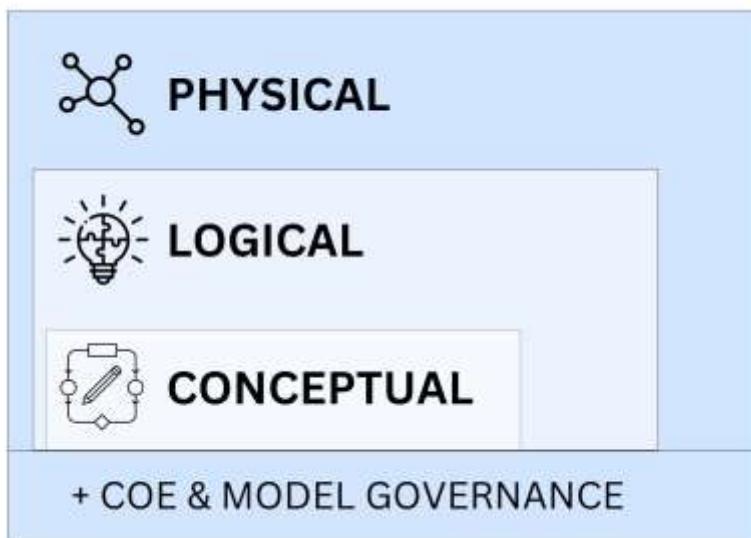


Рис. 4.310 Деректерді модельдеу және деректер сапасын басқару СоE-нің негізгі міндеттерінің бірі болып табылады.

Деректер бойынша Озық тәжірибе орталығы (СоE) тек сарапшылар тобы емес, жаңа data-driven мәдениетті қалыптастыратын және компанияның барлық деректермен жұмыс істеуіне біртұтас тәсілді қамтамасыз ететін жүйелі механизм. Модельдеу процестерін ақпаратты басқару, стандарттау, класификациялау және деректер сапасын бақылау жүйесіне тиімді интеграциялау арқылы СоE бизнеске өнімдерін және бизнес-процесстерін үздіксіз жетілдіруге, нарықтағы өзгерістерге жылдам жауап беруге және сенімді аналитика негізінде салмақты шешімдер қабылдауға көмектеседі.

Мұндай орталықтар заманауи DataOps принциптерімен үйлескенде ерекше тиімді болады – бұл тәсіл деректердің үздіксіз жеткізілуін, автоматтандыруды және сапа бақылауын қамтамасыз етеді.

DataOps туралы толығырақ біз сегізінші бөлімде, «Құрылым саласындағы деректермен жұмыс істеудің заманауи технологиялары» тарауында талқылаймыз.

Келесі тарауларда біз стратегиядан практикаға өтеміз — шартты түрде «деректерді өндеде орталығына» айналамыз: бірнеше мысалдар арқылы тапсырманың параметрленуі, талаптарды жинау және автоматты валидация процесі қалай жүзеге асатынын қарастырамыз.



ТАРАУ 4.4.

ТАЛАПТАРДЫ ЖҮЙЕЛЕУ ЖӘНЕ АҚПАРАТТЫ ВАЛИДАЦИЯЛАУ

Талаптарды жинау және талдау: коммуникацияларды құрылымдық деректерге айналдыру

Талаптарды жинау және басқару – деректер сапасын қамтамасыз етудің алғашқы қадамы. Цифрлық құралдардың дамуына қарамастан, талаптардың көпшілігі өлі күнге дейін құрылымдалмаған түрде қалыптасады: хаттар, кездесулердің хаттамалары, телефон қоңыраулары және ауызша талқылаулар арқылы. Мұндай коммуникация формасы автоматтандыруды, тексеруді және ақпаратты қайта пайдалануды қыындалатады. Бұл тарауда біз мәтіндік талаптарды формалды құрылымдарға қалай аударуға болатынын, бизнес міндеттерінің айқындығы мен жүйелілігін қамтамасыз етуді қарастырамыз.

Gartner компаниясының «Деректер сапасы: Дәл ақпарат алу үшін үздік тәжірибелер» зерттеуі деректер мен аналитика саласындағы сәтті бастамаларды жүзеге асыру үшін деректер сапасының критикалық маңыздылығын атап өтеді. Олар төмендеректер сапасының үйімдарға жыл сайын орташа 12,9 миллион доллардан кем емес шығынға әкеletінін және сенімді, жоғары сапалы деректердің деректермен басқарылатын компания құру үшін қажетті екенін атап көрсетеді.

Құрылымдалған талаптардың болмауы бірдей элемент (субъект) және оның параметрлері әртүрлі жүйелерде әртүрлі нұсқаларда сақталуына әкеледі. Бұл тек процестердің тиімділігін төмендетіп қана қоймай, уақытты жоғалтуға, ақпараттың дублиренеүіне және деректерді пайдалану алдында қайта тексеру қажеттілігіне әкеледі. Нәтижесінде, тіпті бір ғана жіберіп алған нәрсе – жоғалған параметр немесе дұрыс сипатталмаған элемент – шешім қабылдауды баяулатып, ресурстардың тиімсіз жұмысалуына себеп болуы мүмкін.

Гвоздь болмағанда, подкова жоғалды. Подкова болмағанда, ат жоғалды. Ат болмағанда, 骑士 жоғалды. 骑士 болмағанда, хабарлама жоғалды. Хабарлама болмағанда, шайқас жоғалды. Шайқас болмағанда, патшалық жоғалды. Және бәрі гвоздьтың болмауынан.

– Мақал

Деректерді толтыру және сақтау процесіне талаптарды талдау және жинау барлық мұдделі тараптарды анықтаудан басталады. Мақалда бір гвоздьтың жоғалуы критикалық салдардың тізбегіне әкелетіндей, бизнесте бір қатысушының жоғалуы, жіберілген талап немесе тіпті бір параметрдің жоғалуы жеке бизнес процесіне ғана емес, бүкіл жобаның және үйімнің экожүйесіне елеулі әсер

етуі мүмкін. Сондықтан, бірінші көзқарас бойынша маңызызы болып көрінетін элементтерді, параметрлерді және рөлдерді анықтау өте маңызды, бірақ кейін олар бизнестің тұрақтылығы үшін критикалық маңызды болуы мүмкін.

Компанияда клиент жаңа өтініш білдірген жобаны қарастырайық - "ғимараттың солтүстік жағында қосымша терезе қосу". "Клиенттің ағымдағы жобаға жаңа терезе қосу өтініші" атты шағын процесс архитектор, тапсырыс беруші, CAD (BIM) маманы, құрылыш менеджері, логистика менеджері, ERP-аналитик, сапа бақылаушы инженері, қауіпсіздік инженері, бақылау менеджері және жылжымайтын мүлік менеджерін қамтиды.

Тіпті шағын процесте ондаған түрлі мамандар қатыса алады. Әрбір процесс қатысуышы деректер деңгейінде байланысқан мамандардың талаптарын түсінуі тиіс.

Мәтіндік деңгейде (Рис. 4.41) клиент пен процесс тізбегіндегі мамандар арасындағы коммуникация келесідей жүзеге асырылады:-

- Тапсырыс беруші: "Біз солтүстік жағында қосымша терезе қосуға шешім қабылдадық, жағықтандыруды жақсарту үшін. Бұл іске асырылуы мүмкін бе?"
- Архитектор: "Әрине, мен жобаны қайта қарап, жаңа терезені қосып, жаңартылған CAD (BIM) жоспарларын жіберемін".
- CAD (BIM) маманы: "Жаңа жобаны алдым. Мен CAD (BIM) моделін қосымша тереземен жаңартып, FEM инженерімен келісімнен кейін жаңа терезенің дәл орналасуын және өлшемдерін ұсынамын".
- Құрылыш менеджері: "Жаңа жобаны алдық. Біз 4D орнату мерзімдерін түзетеміз және барлық тиісті субмердігерлерді хабардар етеміз".
- Объект инженері (CAFM): "Мен CAFM жүйесіне жаңа терезе туралы 6D деректерін енгіземін, болашақта объектіні басқару және техникалық қызмет көрсету үшін".
- Логистика менеджері: "Мен жаңа терезенің өлшемдері мен салмағын білуім керек, терезені объектіге жеткізу үшін".
- ERP-аналитик: "Мен жаңа терезенің көлемдері мен нақты түрін білуім керек, 5D бюджетімізді жаңарту үшін, жобаның жалпы сметасында жаңа терезенің құнын көрсету үшін".
- Сапа бақылаушы инженері: "Терезе спецификациялары дайын болғанда, мен олардың сапа және материал стандарттарына сәйкестігін тексеремін".
- Қауіпсіздік инженері: "Мен жаңа терезенің қауіпсіздік аспектілерін бағалаймын, 8D схемасы бойынша талаптарды және эвакуацияны сақтау мәселелеріне ерекше назар аударамын".
- Бақылау менеджері: "ERP-дан алынған нақты жұмыс көлеміне негізделе отырып, біз жаңа терезені орнату үшін 4D уақыт шкаlamызды жаңартамыз және жаңа деректерді жобаның контент басқару жүйесінде сақтаймыз".
- Жұмысшы (монтажник): "Орнату, жинау және жұмыс мерзімдері бойынша нұсқаулықтар қажет. Сонымен қатар, мен сақтауға тиіс арнайы қауіпсіздік ережелері енгізілді ме?"
- Жылжымайтын мүлік менеджері: "Орнатудан кейін мен ұзақ мерзімді басқару үшін кепілдік және қызмет көрсету туралы ақпаратты құжаттаймын".
- Активтер менеджері: "Жабдық инженері, активтерді бақылау және өмірлік циклды басқару үшін соңғы деректерді жіберіңіз".
- Клиент: "Күте тұрыңыз, мүмкін мен асығып жатырмын, және терезе қажет емес. Мүмкін, балкон жасау керек шығар".

Мұндай сценарийлер, жиі кездесетін, тіпті кішкентай өзгеріс көптеген жүйелер мен рөлдер арасында тізбекті реакцияны тудырады. Осы орайда, бастапқы кезеңде коммуникацияның басым бөлігі мәтіндік формада жүргізіледі: хаттар, чаттар, кездесулердің хаттамалары (Рис. 4.41).

Мәтіндік коммуникациялар жүйесінде құрылыштың жобасы үшін барлық деректер алмасу операциялары мен қабылданған шешімдердің заңды растау және тіркеу жүйесі өте маңызды. Бұл әрбір қабылданған шешім, нұсқаулық немесе өзгерісті заңды күшпен қамтамасыз ету және бақылау мүмкіндігін қамтамасыз ету үшін қажет, бұл болашақта «түсініспеушіліктер» туындау қаупін азайтады.

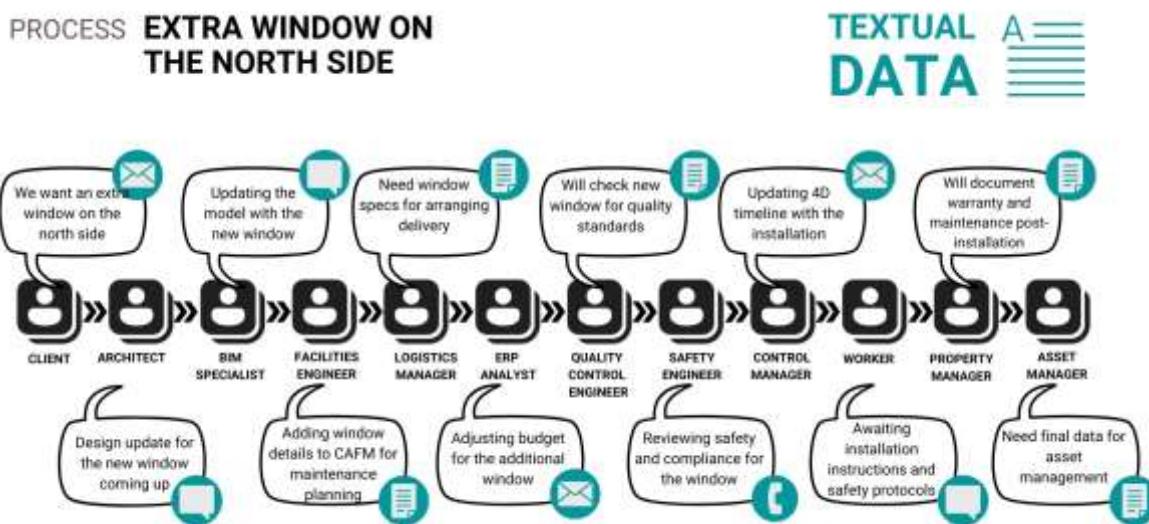


Рис. 4.41 Жобаның бастапқы кезеңдерінде клиент пен орындаушы арасындағы коммуникация жиі әртүрлі форматтағы мәтіндік деректерді қамтиды.

Құрылыштың жобасындағы тиісті жүйелерде заңды бақылау мен шешімдерді растаудың болмауы барлық қатысушылар үшін ауыр проблемаларға әкелуі мүмкін. Дұрыс құжаттық рәсімдеу мен растаусыз қабылданған әрбір шешім, бұйрық немесе өзгеріс дауларға (және сот талқылауларына) әкелуі мүмкін.

Мәтіндік коммуникациядағы барлық шешімдердің заңды бекітілуі тек көп мөлшерде қол қойылған құжаттарды талап етеді, олар барлық мәмілелерді тіркеуге міндетті басшылықтың мойнына түседі. Нәтижесінде, егер әрбір қатысушы әрбір әрекет үшін құжаттарға қол қоюға міндетті болса, жүйе икемділігін жоғалтып, бюрократиялық лабиринтке айналады. Транзакциялардың растауларының болмауы жобаның жүзеге асырылуын кешіктіріп қана қоймай, қаржылық шығындарға, сондай-ақ қатысушылар арасындағы қатынастардың нашарлауына, тіпті заңды мәселелерге әкелуі мүмкін.

Мәтіндік талқылаулардан басталатын транзакцияларды келісу және бекіту процесі, келесі кезеңдерде әртүрлі форматтағы құжаттармен алмасу форматына біртіндеп өтеді (Рис. 4.42), бұл тек мәтін арқылы жүзеге асырылған коммуникацияны айтарлықтай қынадатады. Айқын талаптардың болмауы, әртүрлі деректермен және көп мәтіндік талаптармен толтырылған мұндай процестерді автоматтандыруды іс жүзінде мүмкін емес етеді.

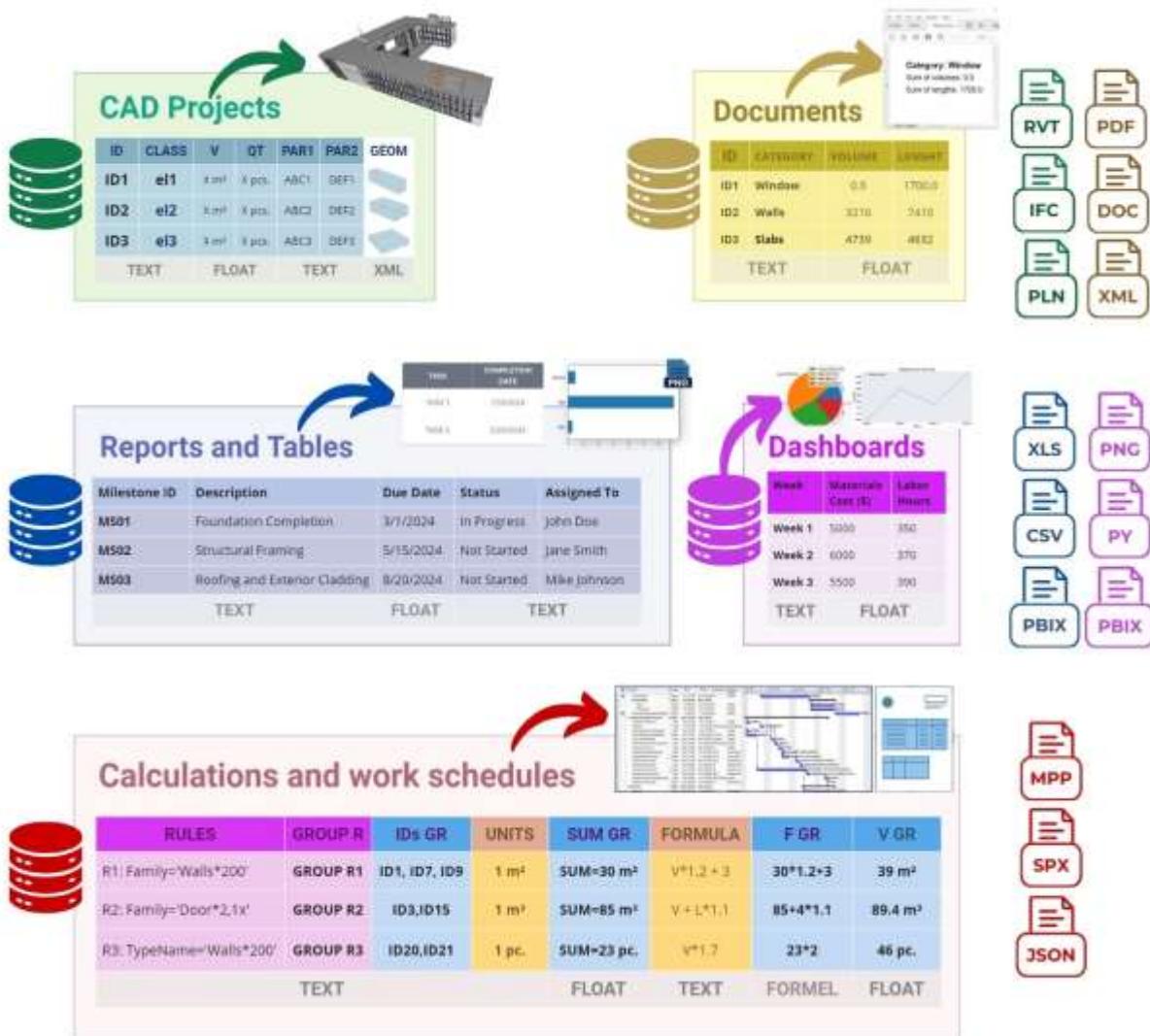


Рис. 4.42 Құрылыш компаниясының ландшафтындағы әрбір жүйе әртүрлі форматтағы зандау маңызы бар құжаттардың көзі болып табылады.

Мәтіндік коммуникациялар әрбір маманнан немесе толық хаттамамен танысады, немесе жобаның ағымдағы статусын түсіну үшін барлық кездесулерге тұрақты қатысады талап етеді.

Бұл шектеуді жену үшін мәтіндік коммуникациядан құрылымдық талаптар моделіне көшу қажет. Бұл жүйелі талдау, процесті визуализациялау және өзара әрекеттесулерді блок-схемалар мен де-ректер модельдері түрінде сипаттау арқылы ғана мүмкін болады (Рис. 4.43). Деректер моделде-үінде (Рис. 4.37) сияқты, біз контекстуалды-идеялық деңгейден концептуалды деңгейге өттік, қатысушылар қолданатын жүйелер мен құралдарды, сондай-ақ олардың арасындағы байланыстарды қосып.-

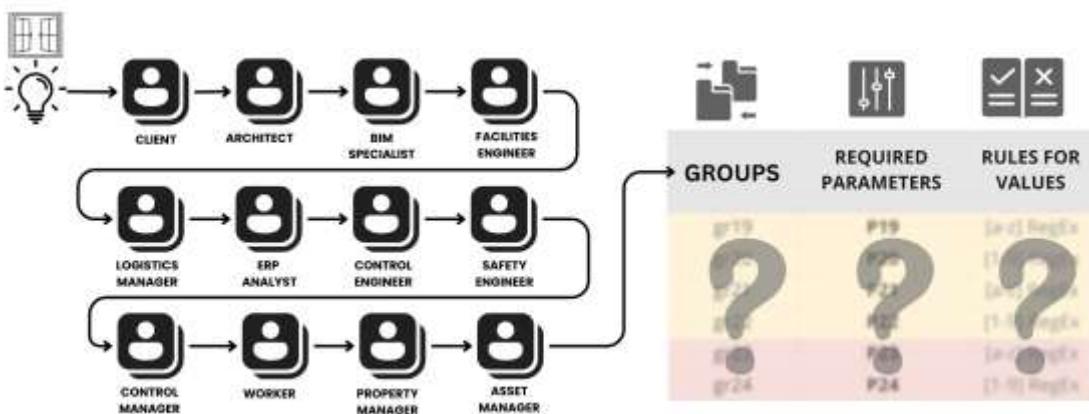


Рис. 4.43 Процесті валидациялауды басқару және автоматтандыруды үйрену үшін процестерді визуализациялау және талаптарды құрылымдау қажет.

Талаптар мен қатынастарды жүйелендірудің алғашқы қадамы - концептуалдық блок-схемалар арқылы барлық байланыстар мен қатынастарды визуализациялау. Концептуалдық деңгей барлық процесс қатысушыларына технологиялық тізбектің толық түсінігін жеңілдетіп қана қоймай, сонымен қатар әр кезеңдегі деректер мен талаптардың не үшін және кімге қажет екенін айқын көрсетеді.

Процесстердің блок-схемалары және концептуалдық схемалардың тиімділігі

Деректерді басқарудың дәстүрлі және заманауи тәсілдері арасындағы алшақтықты жою үшін компанияларға фрагментtelген мәтіндік сипаттамалардан құрылымдық процесс көрсетуге саналы түрде көшу қажет. Деректердің эволюциясы - балшық тақталардан цифрлық экожүйелерге дейін - жаңа ойлау құралдарын талап етеді. Осында құралдардың бірі - блок-схемалар арқылы концептуалдық моделдеу. Визуалдық схемаларды - блок-схемаларды, процесс диаграммаларын, өзара әрекет схемаларын жасау жобаның қатысушыларына олардың әрекеттері мен шешімдерінің бүкіл шешім қабылдау жүйесіне қалай әсер ететінін түсінуге мүмкіндік береді.

Егер процестер деректерді сақтау ғана емес, сонымен қатар оларды талдау немесе автоматтандыруды талап етсе, онда концептуалдық-визуалдық талаптар деңгейін құрумен айналысу қажет.

Біздің мысалымызда (Рис. 4.41) әрбір маман тек шағын командаға ғана емес, сонымен қатар бас менеджердің басқаруындағы оннан астам сарапшыны қамтитын үлкен бөлімге де кіруі мүмкін. Әр бөлім құжаттарды жасау, шешімдердің заңды мәртебесін тіркеу және процестерді басқару үшін қажетті кіріс ақпаратын тұрақты түрде толықтыратын арнайы қосымша деректер базасын пайдаланады (Рис. 1.24 мысалы ERP, CAD, MEP, CDE, ECM, CRM және т.б.).

Транзакция процесі 4000 жыл бұрынғы ежелгі менеджерлердің жұмысын еске салады, сол кезде шешімдерді занұды түрде растау үшін балшық тақталар мен папиустар қолданылған. Заманауи жүйелердің балшық және қағаз предшественниктерінен айырмашылығы, қазіргі әдістер мәтіндік ақпаратты цифрлық форматта түрлендіру процесін қосып, оны басқа жүйелер мен құралдарда автоматты түрде өндеге мүмкіндік береді.

Процесті концептуалдық блок-схемалар түрінде визуализациялау әр қадамды және әртүрлі рөлдер арасындағы өзара әрекеттестікті сипаттайға көмектеседі, күрделі жұмыс процесін түсінікті және қарапайым етеді.

Процестерді визуализациялау команда мүшелерінің барлығына процесс логикасының ашықтығы мен қолжетімділігін қамтамасыз етеді.

Жобада терезе қосу туралы коммуникациялық процесс мәтін, хабарламалар (Рис. 4.41) және блок-схема түрінде сипатталған, деректерді моделдеу туралы тарауда қарастырылған концептуалдық модельге үқсас (Рис. 4.44).

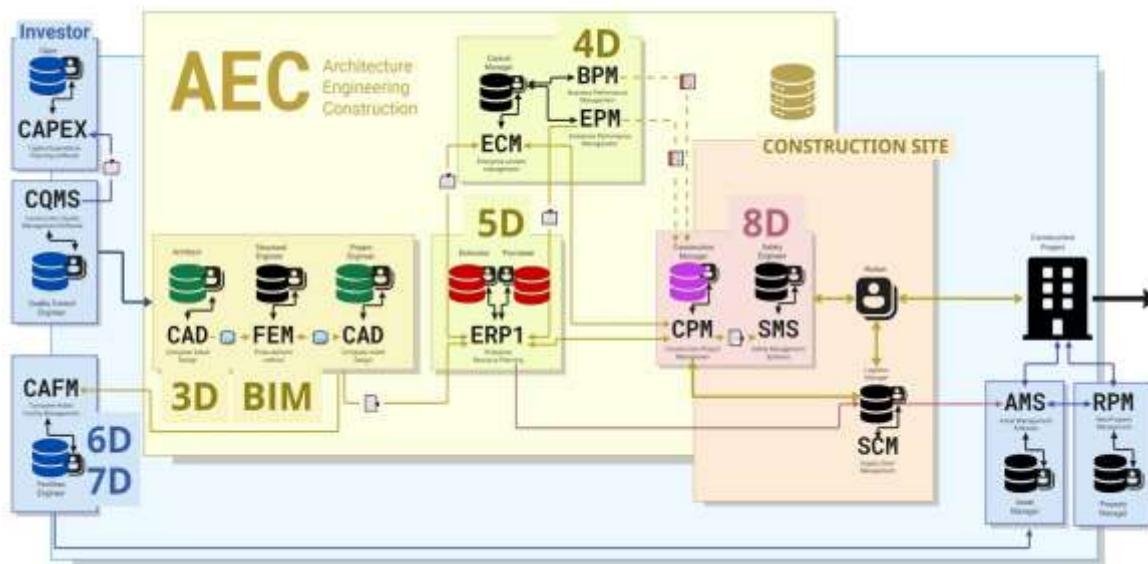


Рис. 4.44 Концептуалдық схемада жобаның қатысушылары деректер базасының пайдаланушылары ретінде көрсетілген, мұнда олардың сұраулары әртүрлі жүйелерді байланыстырады.

Концептуалдық схемалар маңызды қадам болып табылса да, көптеген компаниялар тек осы деңгеймен шектеледі, визуалдық схеманың процестерді түсіну үшін жеткілікті екеніне сенеді. Бұл басқарудың иллюзиясын тудырады: менеджерлерге мұндай блок-схема арқылы жалпы көріністі қабылдау, қатысушылар мен кезеңдер арасындағы байланыстарды көру оңайырақ. Алайда, мұндай схемалар әрбір қатысушыға қандай деректер қажет, олар қандай форматта берілуі тиіс және автоматизацияны жүзеге асыру үшін қандай параметрлер мен атрибуттар міндетті екендігі туралы нақты түсінік бермейді. Концептуалдық блок-схема маршрут картасына үқсас: ол кімнің

кіммен өзара әрекеттесетінін көрсетеді, бірақ осы өзара әрекеттесулерде не берілетінін ашпайды.

Процесс концептуалдық деңгейде блок-схема арқылы егжей-тегжейлі сипатталған болса да, оның тиімділігін қамтамасыз етпейді. Визуализация менеджерлердің жұмысын жіңінен көзінде дәректер базасын басқарушы инженерлер үшін концептуалдық көрініс жеткілікті түсінікті болмауы мүмкін және процесті параметрлер мен талаптар деңгейінде қалай жүзеге асыру керектігін нақты түсінуге мүмкіндік бермейді.

Деректердің күрделі әкожүйелеріне қарай жылжу барысында концептуалдық және визуалдық құралдарды бастапқы енгізу деректерді өңдеу процестерінің тиімді ғана емес, сонымен қатар үйімнің стратегиялық мақсаттарына сәйкес болуын қамтамасыз ету үшін критикалық маңызды болып табылады. Бұл процесті деректерге қойылатын талаптар деңгейіне толық аудару үшін, біз концептуалдық визуализацияны логикалық және физикалық деңгейдегі деректер, қажетті атрибуттар және олардың шекті мәндері деңгейіне түсіруіміз қажет.

Құрылымдық талаптар және RegEx түрақты өрнектері

Компанияларда жасалатын деректердің 80%-ға жуығы құрылымдалмаған немесе жартылай құрылымдалған форматтарда ұсынылған – мәтін, құжаттар, хаттар, PDF файлдары, әңгімелер. Мұндай деректерді талдау, тексеру, жүйелер арасында беру және автоматизацияда пайдалану қыын. -

Басқаруды, ашықтықты және автоматты валидацияны қамтамасыз ету үшін мәтіндік және жартылай құрылымдалған талаптарды нақты анықталған, құрылымдалған форматтарға аудару қажет. Структуризация процесі тек деректерге ғана емес (бұл кітаптың осы бөлігінің алғашқы тарауында егжей-тегжейлі қарастырылған), сонымен қатар жобаның қатысушылары әдетте жобаның өмірлік циклі бойында еркін мәтін түрінде формулировка жасайтын талаптарға да қатысты, олар бұл процестерді автоматтандыруға болатынын жиі ескермейді.

Біз құрылымдалмаған мәтіндік формадан құрылымдалған формаға деректерді трансформациялағанымыздай, талаптармен жұмыс барысында мәтіндік талаптарды "логикалық және физикалық деңгей" құрылымдық форматына трансформациялаймыз.

Терезе қосу мысалында (Рис. 4.41) келесі қадам деректерге қойылатын талаптарды кесте түрінде сипаттау болады. Біз жобаның қатысушылары пайдаланатын әрбір жүйе үшін ақпаратты құрылымдалап, негізгі атрибуттар мен олардың шекті мәндерін көрсетеміз.

Мысал ретінде, құрылыс сапасын басқару жүйесін (CQMS) қарастырайық, оны тапсырыс берушінің сапа бақылаушысы пайдаланады. Оның көмегімен ол жобаның жаңа элементі – "жаңа терезе" – белгіленген стандарттар мен талаптарға сәйкес келетінін тексереді.

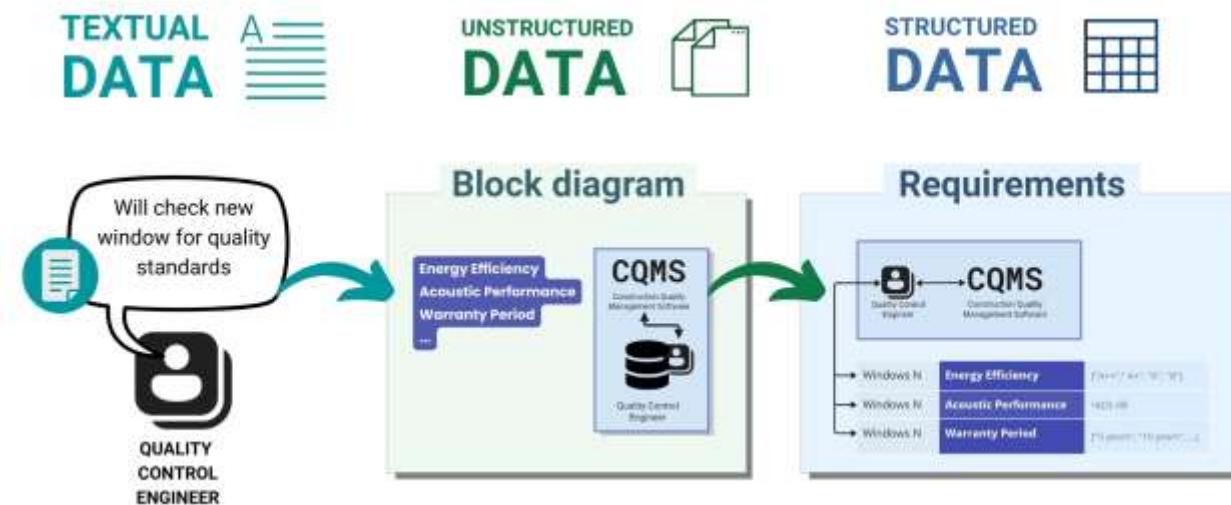


Рис. 4.45 Мәтіндік талаптарды кесте форматына түрлендіру, объектілердің атрибуттарын сипаттау арқылы басқа мамандар үшін түсінуді жеңілдетеді.

Мысал ретінде CQMS жүйесіндегі "терезе жүйелері" типіндегі объектілердің атрибуттарына қатысты кейбір маңызды талаптарды қарастырайық (Рис. 4.46): энергия тиімділігі, акустикалық сипаттамалар және кепілдік мерзімі. Эрбір категория жобалау және терезе жүйелерін орнату кезінде ес-керілуі тиіс стандарттар мен спецификацияларды қамтиды.-



Рис. 4.46 Сапа бақылаушы инженері "Терезе" типіндегі жаңа элементтердің энергия тиімділігі, дыбыс оқшаулау және кепілдік қызмет көрсету стандарттарына сәйкестігін тексеруі қажет.

Сапа бақылаушы инженерінің кесте түрінде белгілеген деректерге қойылатын талаптары, мысалы, келесі шекті мәндерге ие:

- Терезелердің энергия тиімділігі класы "A++" - ең жоғары тиімділікті білдіретін, "B" - минималды рұқсат етілген деңгей болып табылатын диапазонда болады, және бұл кластар рұқсат етілген мәндер тізімімен ["A++","A+","A","B"] ұсынылған.

- Терезелердің акустикалық оқшаулауы, децебелмен өлшенетін және олардың көшедегі шуды азайту қабілетін көрсететін, $\text{d}\{\cdot\}$ dB регулярлы өрнегімен анықталады.
- "Терезе типі" обьектісінің "Кепілдік мерзімі" атрибуты бес жылдан басталады, бұл мерзімді өнімді таңдау кезінде минималды рұқсат етілген мерзім ретінде белгілейді; сондай-ақ кепілдік мерзімінің мәндері, мысалы, ["5 жыл", "10 жыл" және т. б.] немесе логикалық шарт " >5 (жыл)" көрсетіледі.

Жиналған талаптарға сәйкес, белгіленген атрибуттар шеңберінде "Терезе" категориясы немесе класына жататын жаңа элементтер "B" класынан төмен, мысалы, "C" немесе "D" класындағы элементтер энергия тиімділігі бойынша тексеруден өтпейді. Сапа бақылаушы инженері қабылдайтын деректерде немесе құжаттарда терезелердің акустикалық оқшаулауы екі таңбалы санмен, "dB" постфиксімен белгіленуі тиіс, мысалы, "35 dB" немесе "40 dB", ал осы форматтан тыс мәндер, мысалы, "9 Д Б" немесе "100 децибел", қабылданбайды (RegEx жол паттерніне сәйкес келмейді). Кепілдік мерзімі кемінде "5 жыл" басталуы тиіс, ал "3 жыл" немесе "4 жыл" сияқты қысқа кепілдік мерзімі бар терезелер талаптарға сәйкес келмейді, оларды сапа инженері кесте форматында сипаттаған.

Талаптарға сәйкес атрибут-параметрлердің мәндерін тексеру үшін валидация процесінде рұқсат етілген мәндер тізімін ([A", "B", "C"], сөздіктерді ([A": "H1", "H2"; "B": "W1", "W2"]), логикалық операцияларды (мысалы, ">", "<", "<=", ">=", "==" сандық мәндер үшін) және регулярлы өрнектерді (стрингтік және мәтіндік мәндер үшін, мысалы, "Акустикалық өнімділік" атрибутында) қолданамыз. Регулярлы өрнектер - стрингтік мәндермен жұмыс істеуде өте маңызды құрал.

Регулярлы өрнектер (RegEx) программалау тілдерінде, соның ішінде Python (Re кітапханасы) арқылы, жолдарды іздеу және өзгерту үшін қолданылады. Regex - бұл жолдар әлеміндегі детектив, мәтінде текстік шаблондарды дәл анықтай алатын құрал.

Регулярлы өрнектерде әріптер тікелей алфавиттің сәйкес символдары арқылы сипатталады, ал сандар арнайы символ \d арқылы ұсынылуы мүмкін, ол 0-ден 9-ға дейінгі кез келген цифрага сәйкес келеді. Квадратты жақшалар әріптер немесе цифрлар диапазонын көрсету үшін қолданылады, мысалы, [a-z] латын алфавитінің кез келген кіші әрпіне немесе [0-9] \d-ге эквивалент. Сандық емес және әріптік емес символдар үшін \D және \W сәйкесінше пайдаланылады.

RegEx-тің танымал қолдану мысалдары (Рис. 4.47):-

- Электрондық пошта адресін тексеру: жолдың заңды электрондық пошта адресі болып табылатынын тексеру үшін «^a-zA-Z0-9._%+]+@[a-zA-Z0-9.]+\.[a-zA-Z]{2,}\$» шаблонын қолдануға болады.
- Күндізгі датаны шығару: «\b\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\b» шаблоны мәтіннен DD.MM.YYYY форматындағы датаны шығару үшін пайдаланылуы мүмкін.
- Телефон нөмірлерін тексеру: +49(000)000-0000 форматындағы телефон нөмірлерін тексеру үшін шаблон «\+\d{2}(\d{3})\d{3}-\d{4}» түрінде болады.
- Сапа бақылаушы инженердің талаптарын атрибуттар мен олардың шекті мәндері форматына ауыстыра отырып (Рис. 4.46), біз оларды бастапқы мәтіндік форматтан (сұхбаттар, хаттар және нормативтік құжаттар) үйымдасқан және құрылымдалған кестеге айналдырыңық, осылайша кез келген келіп түскен деректерді (мысалы, «Терезе» категориясының жаңа элементтері) автоматты түрде тексеру және талдау мүмкіндігін

қамтамасыз еттік. Талаптардың болуы тексеруден өтпеген деректерді автоматты түрде жоюға, ал тексерілген деректерді автоматты түрде өңдеу үшін жүйелерге жіберуге мүмкіндік береді.-

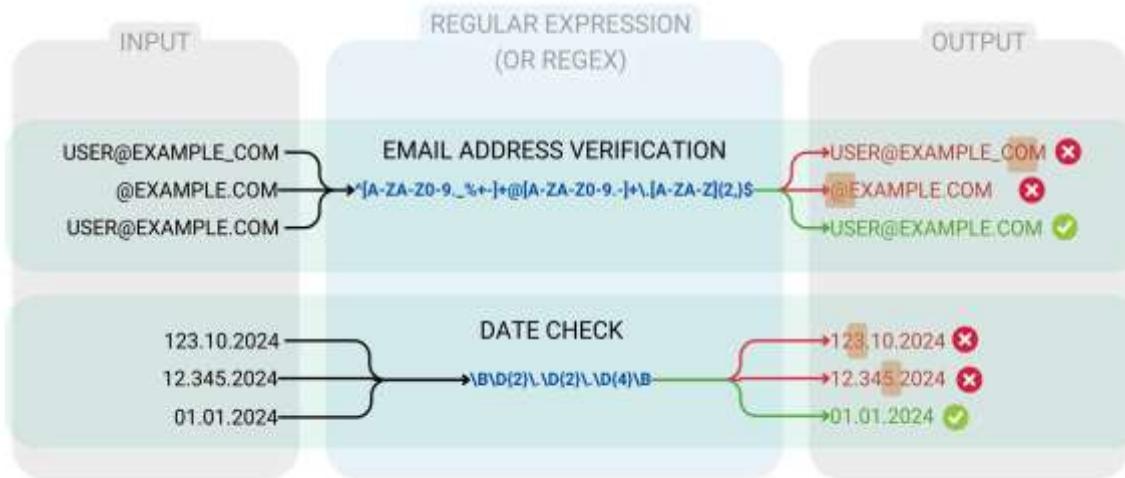


Рис. 4.47 Регулярлы өрнектерді пайдалану мәтіндік деректерді тексеру процесінің өтешшілдік маңызды куралы болып табылады.

Енді концептуалды деңгейден талаптармен жұмыс логикалық деңгейге ауысқанда, біз жаңа терезе орнату процесінің барлық мамандарының талаптарын (Рис. 4.44) атрибуттар форматында реттелген тізімге айналдырамыз және осы тізімдерді қажетті атрибуттармен біздің блок-схемамызыға қосамыз (Рис. 4.48). –

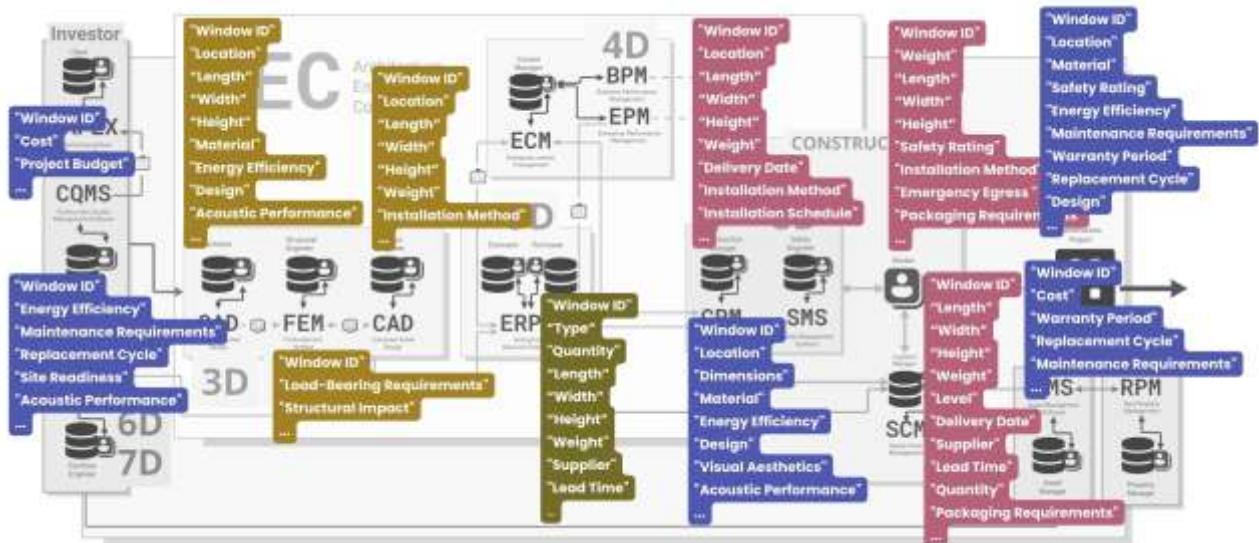


Рис. 4.48 Логикалық деңгейде әр маман өндійтін атрибуттар тиісті жүйелерге қосылады.

Барлық атрибуттарды бір жалпы процесс кестесінде қосу арқылы, біз бұрын мәтін және диалог түрінде ұсынылған ақпаратты концептуалды деңгейде (Рис. 4.41) физикалық деңгейдегі құрылым-далған және жуиеленген кесте формасына айналдырамыз (Рис. 4.49).-

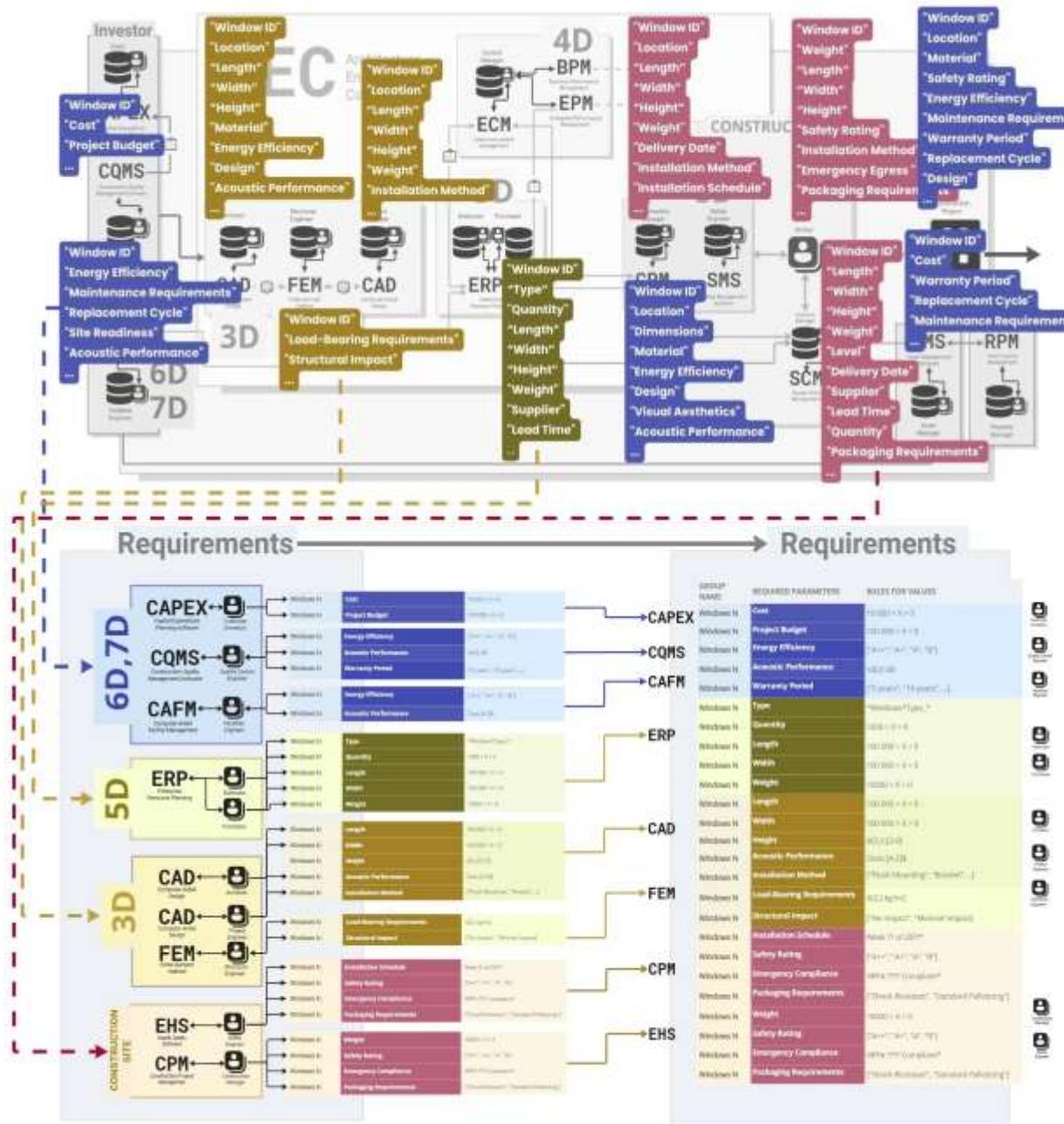


Рис. 4.49 Мамандардың құрылымдалмаған диалогын құрылымдалған кестелерге айналдыру физикалық деңгейде талаптарды түсінуге көмектеседі.

Енді деректерге қойылатын талаптарды нақты жүйелер үшін ақпарат жасайтын мамандарға жеткізу қажет. Мысалы, егер сіз CAD дерекқорында жұмыс істесеңіз, элементтерді модельдеуге кіріспес бұрын, осы деректерді пайдалану сценарийлеріне негізделген барлық қажетті параметрлерді жинау қажет. Әдетте, бұл пайдалану кезеңінен басталады, содан кейін құрылыш алаңы, логистика бөлімі, сметалық бөлім, конструктивтік есептеу бөлімі және т.б. жалғасады. Тек осы буындардың талаптарын ескергеннен кейін, жинақталған параметрлер негізінде деректерді жасауға кірісуге болады. Бұл болашақта деректерді тексеру мен беру процесін автоматтандыруға мүмкіндік береді.

Жаңа деректердің белгіленген талаптарға сәйкестігі жағдайында олар автоматты түрде компанияның деректер экожүйесіне интеграцияланады, тікелей сол пайдаланушылар мен жүйелерге бағытталады, олар үшін арналған. Деректердің атрибуттары мен олардың мәндерінің бар-жоғын және сәйкестігін тексеру ақпараттың қажетті сапа стандарттарына сәйкес келуін және компанияның сценарийлерінде қолдануға дайын болуын қамтамасыз етеді.

Деректерге қойылатын талаптар анықталды, енді тексеруді бастамас бұрын тексеруге жататын деректерді жасау, алу немесе жинау қажет, немесе деректер базасындағы ақпараттың ағымдағы күйін тіркеу керек, оны тексеру процесінде пайдалану үшін.

Тексеру процесі үшін деректерді жинау

Тексеруді бастамас бұрын, деректердің валидация процесіне қолайлы формада қолжетімді екеніне көз жеткізу маңызды. Бұл тек ақпараттың бар болуы емес, сонымен қатар оны дайындауды білдіреді: деректерді жинап, құрылымдаған, нашар құрылымдаған, мәтіндік және геометриялық форматтардан құрылымдаған түрге түрлендіру қажет. Бұл процесс алдыңғы тарауларда егжей-тегжейлі сипатталған, онда әртүрлі деректер түрлерін трансформациялау әдістері қарастырылған. Нәтижесінде барлық түрлендірулерден кейін кіріс деректері ашық құрылымдаған кесте түріне ие болады.-

Талаптар мен қажетті параметрлер мен шекті мәндері бар құрылымдаған кестелерге ие болғанда, біз деректерді тексеруге кіріс аламыз - бұл бір автоматтандырылған процесс (Pipeline) ретінде немесе әрбір кіріс құжатын кезең-кезеңімен тексеру форматы ретінде.

Тексеруді бастау үшін жаңа файл алу немесе деректердің ағымдағы күйін тіркеу қажет - ағымдағы және келіп жатқан мәліметтердің суретін жасау немесе экспорттау, немесе сыртқы немесе ішкі деректер базасына қосылуды орнату. Қарастырылып отырған мысалда, мұндай сурет CAD деректерін құрылымдаған форматқа автоматты түрде түрлендіру арқылы жасалады, мысалы, 2024 жылдың 29 наурызында, жұма күні, сағат 23:00:00-де, барлық жобалаушылар үйіне кеткеннен кейін.



CAD (BIM) деректер базасының суреті, жобаның ағымдағы нұсқасындағы «Терезе» класының жаңа элементі үшін атрибуттар туралы қазіргі ақпаратты көрсетеді.

"CAD (BIM) деректерін құрылымдаған формаға аудару" тарауында қарастырылған кері инжини-

ринг құралдарының арқасында, бұл ақпарат әртүрлі CAD (BIM) құралдарынан және редакторларынан жеке кестелерге (Рис. 4.411) ұйымдастырылуы немесе жобаның әртүрлі бөлімдерін біріктіретін бір жалпы кестеге біріктірілуі мүмкін. –

Мұндай кесте - деректер базасында терезелер мен есіктердің бірегей идентификаторлары (ID атрибуты), типтік атаулар (TypeName), өлшемдер (Width, Length), материалдар (Material), сондай-ақ энергетикалық және акустикалық тиімділік көрсеткіштері, сондай-ақ басқа сипаттамалар көрсетіледі. CAD (BIM) бағдарламасында толтырылған мұндай кесте жобалаушы инженер тарарапынан әртүрлі бөлімдер мен құжаттардан жиналышпайды, жобаның ақпараттық моделін қалыптастырады.



CAD жүйелерінен алынған құрылымдалған деректер элементтердің атрибуттарын белгілейтін бағандары бар екі өлшемді кесте түрінде болуы мүмкін.

Нақты CAD (BIM) жобалары ондаған немесе жүздеген мың элементтерді қамтиды (Рис. 9.110). CAD форматындағы элементтер автоматты түрде типтер мен категорияларға жіктеледі – терезелер мен есіктерден бастап плиталар, төбелер және қабырғаларға дейін. Бірегей идентификаторлар (мысалы, CAD шешімі автоматты түрде орнататын native ID) немесе тип атрибуттары (Type Name, Type, Family) бір және сол объектіні әртүрлі жүйелерде бақылауға мүмкіндік береді. Мысалы, ғимараттың солтүстік қабырғасындағы жаңа терезе барлық ұйымның тиісті жүйелерінде "W-NEW" идентификаторы арқылы анықталуы мүмкін.-

Субъектілердің атаулары мен идентификаторлары барлық жүйелерде бірдей болуы тиіс, алайда осы субъектілермен байланысты атрибуттар мен мәндер жиынтығы қолдану контекстіне байланысты айтартықтай өзгеруі мүмкін. Архитекторлар, құрылыш инженерлері, құрылыш, логистика және жылжымайтын мүлікті пайдалану мамандары бір және сол элементтерді әртүрлі қабылдайды. Әрқайсысы өз классификаторларына, стандарттарына және мақсаттарына сүйенеді: біреулер терезені тек эстетикалық түрғыдан қарап, оның формасы мен пропорцияларын бағаласа, басқалары инженерлік немесе эксплуатациялық түрғыдан қарап, жылу өткізгіштігін, монтаждау әдісін, массасын немесе техникалық қызмет көрсету талаптарын талдайды. Сондықтан деректерді модельдеу және элементтерді сипаттау кезінде олардың қолданылу көпқырлылығын ескеру және салалық ерекшеліктерді ескере отырып, деректердің келісімділігін қамтамасыз ету маңызды.

Компанияның процестерінде әр рөл үшін жобалау мен есептеулерден бастап логистика, монтаждау және ғимараттарды пайдалану бойынша мамандандырылған деректер базалары қарастырылған (Рис. 4.412). Әрбір осындай жүйе кәсіби мамандар командасы арқылы арналы пайдаланушы интерфейсі немесе деректер базасына сұрау арқылы басқарылады, мұнда енгізілген мәндер бойынша қабылданған барлық шешімдердің жиынтығының артында жүйе менеджері немесе бөлім басшысы тұрады, ол өз контрагенттері алдында енгізілген деректердің заңдылығы мен сапасына жауап береді, олар басқа жүйелерді қызмет көрсетеді.

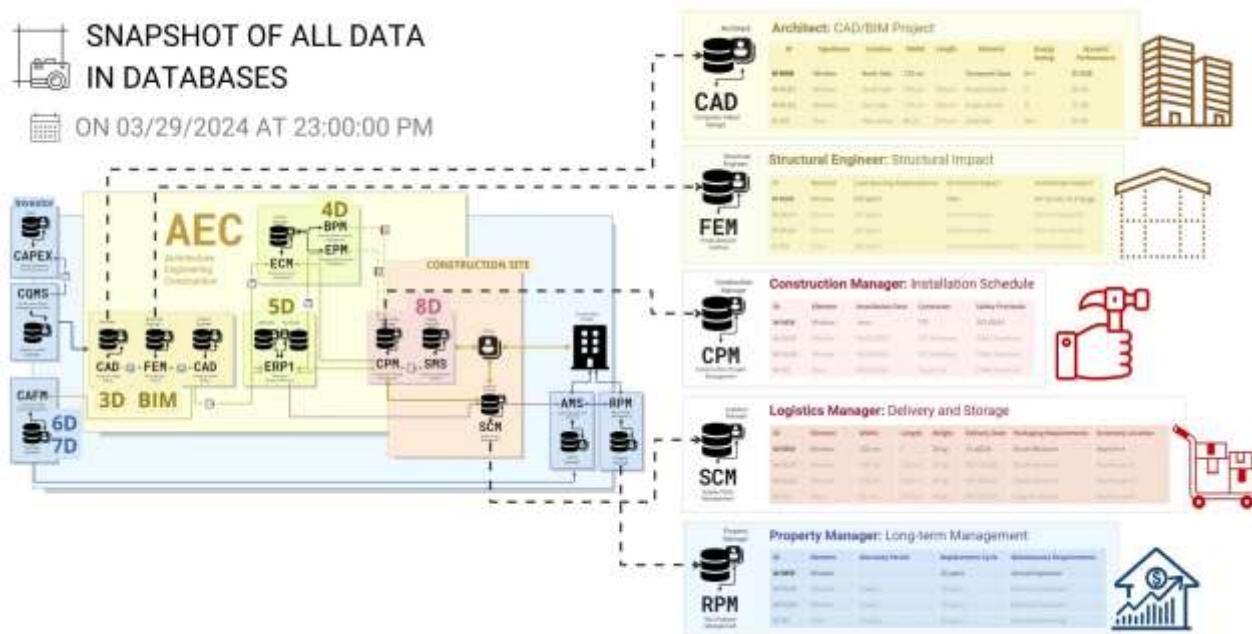


Рис. 4.412 Бір және сол субъекттің әртүрлі жүйелерде бірдей идентификаторы бар, бірақ тек осы жүйеде маңызды болатын әртүрлі атрибуттары бар.

Логикалық және физикалық деңгейде құрылымдық талаптар мен деректерді үйімдасқан түрде жинағаннан кейін, біз әртүрлі құжаттар мен жүйелерден алынған деректердің бұрын жинақталған талаптарға сәйкестігін автоматты түрде тексеру процесін баптауды қалдырамыз.

Деректерді тексеру және тексеру нәтижелері

Жүйеге келіп түсетін барлық жаңа деректер – тапсырыс берушіден, архитектордан, инженерден, проработан, логисттен немесе жылжымайтын мүлік басқарушысынан – бұрын қалыптасқан талаптарға сәйкестігін тексеруден өтуі тиіс (Рис. 4.49). Валидация процесі критикалық маңызды: деректердегі кез келген қателіктер дұрыс есептеулерге, кестедегі кешіктіруге және тіпті қаржылық шығындарға әкелуі мүмкін. Мұндай тәуекелдерді минимизациялау үшін жүйелі, қайталараптын, итерациялық деректерді тексеру процедурасын үйімдастыру қажет.

Жүйеге келіп түсетін жаңа деректерді – құрылымдық емес, мәтіндік немесе геометриялық – тексеру үшін оларды әлсіз құрылымдық немесе құрылымдық форматқа түрлендіру қажет. Содан кейін тексеру процесінде деректер талап етілетін атрибуттардың толық тізіміне және олардың рұқсат етілген мәндеріне сәйкестігіне тексерілуі тиіс.

Әртүрлі деректер типтерін: мәтін, суреттер, PDF құжаттары және CAD (BIM) аралас деректерді құрылымдық формада түрлендіруді "Деректерді құрылымдық формада түрлендіру" бөлімінде егжеттегжейлі қарастырдық.

Мысал ретінде CAD (BIM) жобасынан алғынған кестені келтіруге болады (Сур. 4.411). Ол полуструктурланған геометриялық деректер мен жобаның объектілері бойынша құрылымдық атрибуттың ақпаратты қамтиды (Сур. 3.114) – мысалы, «Терезелер» класындағы элемент.-

Тексеруді жүргізу үшін атрибут мәндерін (Сур. 4.411) сарапшылар тарағынан талап ретінде анықталған эталонды шекті мәндермен салыстырамыз (Сур. 4.49). Нәтижелік салыстырмалы кесте (Сур. 4.413) деректерді CAD (BIM) қосымшаларынан тыс пайдаланар алдында қандай мәндердің рұқсат етілгенін, ал қандайларының түзетуді қажет ететінін түсінуге мүмкіндік береді.-



Сур. 4.413 Нәтижелік тексеру кестесі жаңа «Терезелер» класындағы объектінің атрибут мәндеріне назар аударуды қажет ететінін көрсетеді.

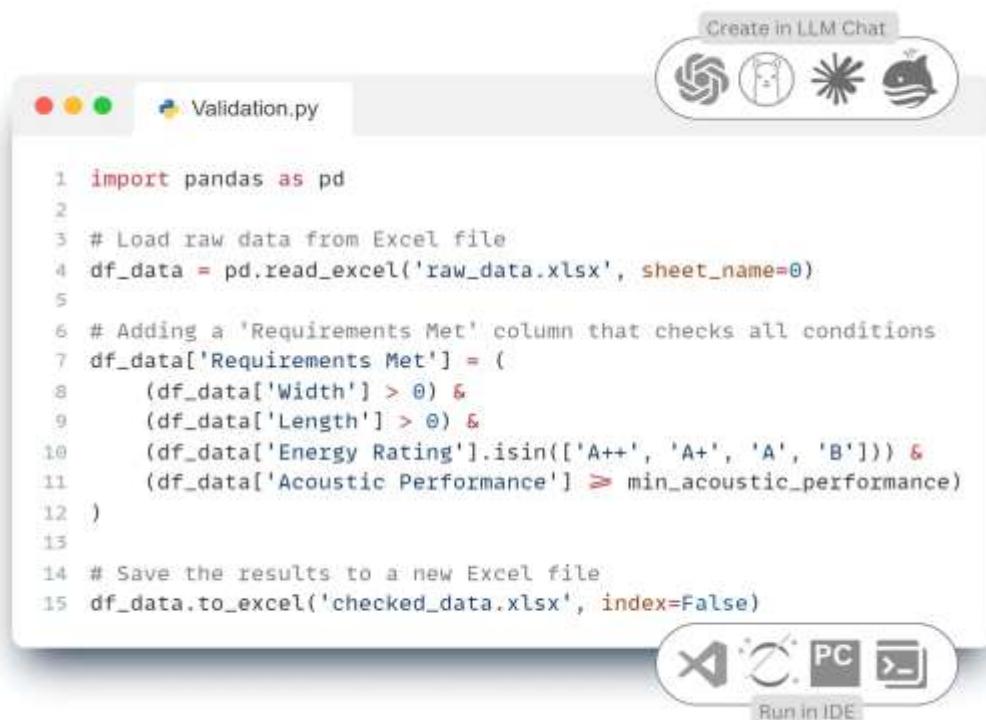
Мұндай шешімді жүзеге асыра отырып, біз бұрын "Pandas: Деректерді талдаудың алмастырылмайтын құралы" бөлімінде айтылған Pandas кітапханасын пайдаланып, CAD (BIM) файлы (RVT, IFC, DWG, NWS, DGN) арқылы алғынған кестелік файлдан деректерді тексереміз (Сур. 4.411), талаптардан алғынған басқа кестелік файлмен (Сур. 4.49) салыстыра отырып.-

Кодты алу үшін LLM-ге raw_data.xlsx файлын жүктеп, деректерді тексеріп, нәтижені жаңа checked_data.xlsx файлына сақтау қажеттігін сипаттауымыз керек (Сур. 4.413).-

💡 Pandas кітапханасын ескермей, LLM арқылы код аламыз:

raw_data.xlsx файлының кестесін тексеру үшін код жазыңыз және келесі тексеру ережелерімен тексеріңіз: 'Width' және 'Length' бағандарының мәндері нөлден үлкен, 'Energy Rating' ['A++', 'A+', 'A', 'B'] тізіміне кіреді, ал 'Acoustic Performance' - кейінірек көрсетілетін айнымалы - тексеру нәтижелік бағанын қосу арқылы, және нәтижелік кестені жаңа Excel файлына checked_data.xlsx сақтаңыз ↗

- LLM-нің жауабы Python кодын қысқаша мысалмен сипаттайты, оны кейінгі промптармен нақтылауға және толықтыруға болады:



The screenshot shows a software window titled "Validation.py". At the top, there are icons for "Create in LLM Chat", a neural network, a terminal, a star, and a whale. Below the title bar is a toolbar with icons for file operations (New, Open, Save, etc.) and a "Run in IDE" button.

```

1 import pandas as pd
2
3 # Load raw data from Excel file
4 df_data = pd.read_excel('raw_data.xlsx', sheet_name=0)
5
6 # Adding a 'Requirements Met' column that checks all conditions
7 df_data['Requirements Met'] = (
8     (df_data['Width'] > 0) &
9     (df_data['Length'] > 0) &
10    (df_data['Energy Rating'].isin(['A++', 'A+', 'A', 'B'])) &
11    (df_data['Acoustic Performance'] >= min_acoustic_performance)
12 )
13
14 # Save the results to a new Excel file
15 df_data.to_excel('checked_data.xlsx', index=False)

```

Сур. 4.414 LLM моделінің генерациялаған коды CAD (BIM) жобасын атрибуттарға қойылатын талаптарға сәйкестігін тексереді.

LLM тілдік моделінің генерациялаған коды кез келген танымал IDE немесе онлайн құралда пайдалануға болады: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагині, Thonny, Wing IDE, IntelliJ IDEA Python плагині, JupyterLab немесе танымал онлайн құралдар Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

Кодты орындау (Сур. 4.414) W-OLD1, W-OLD2, D-122 (және басқа элементтер) CAD (BIM) деректер базасынан атрибуттарға қойылатын талаптарға сәйкес келетінін көрсетеді: ені мен ұзындығы нөлден үлкен, ал энергия тиімділігі класы 'A++', 'A', 'B', 'C' тізімінің бір мәні болып табылады (Сур. 4.415).-

Бізге қажетті және жақында қосылған W-NEW элементі, солтүстік жағындағы жаңа «Терезе» классын білдіреді, талаптарға сәйкес келмейді (атрибут "Requirements Met"), себебі оның ұзындығы нөлге

тең (0.0 мәні 'Width'>0 ережесіне сәйкес қабылданбайды) және энергия тиімділігі класы көрсетілмен.

ID	TypeName	Location	Width	Length	Material	Energy Rating	Acoustic Performance	Requirements Met	
0	W-NEW	Window	North Side	120	0.0	Tempered Glass	A++	35	False
1	W-OLD1	Window	South Side	100	140.0	Double Glazed	A++	30	True
2	W-OLD2	Window	East Side	110	160.0	Single Glazed	B	25	True
3	D-122	Door	Main Entry	90	210.0	Solid Oak	B	30	True

Рис. 4.415 Тексеріс верификациация процесінен өтпеген объектілерді анықтайды және нәтижелерге 'False' немесе 'True' мәндери бар жаңа атрибут қосады.

Сол сияқты, біз жобаның барлық элементтерінің (объектілердің) және әрбір жүйе, кесте немесе деректер базасы үшін қажетті атрибуттардың келісімділігін тексереміз, біз әртүрлі мамандардан алған барлық деректерде (Рис. 4.41) терезе қосу процесінде.

Нәтиже кестесінде тексеру нәтижелерін визуализациялау үшін түспен ерекшелену ынғайлыш: жасылмен тексеруден өткен атрибуттар, сарымен - критикалық емес ауытқулар, ал қызылмен - критикалық сәйкесіздіктер (Рис. 4.416).

Тексерудің нәтижесінде (Рис. 4.416) біз атрибуттарға қойылатын талаптарға сәйкестігі тексерілген сенімді және тексерілген элементтердің тізімін аламыз. Тексерілген элементтер осы элементтердің барлық жүйелер үшін жарияланған стандарттар мен спецификацияларға сәйкес келетіндігіне сенімділік береді, олар «Терезе» класын немесе кез келген басқа классты қосу процесіне қатысады (деректерді тексеруді автоматтандыру және автоматтандырылған ETL процесін құру туралы толығырақ біз "ETL автоматтандыруы және деректерді тексеру" бөлімінде сөйлесеміз).

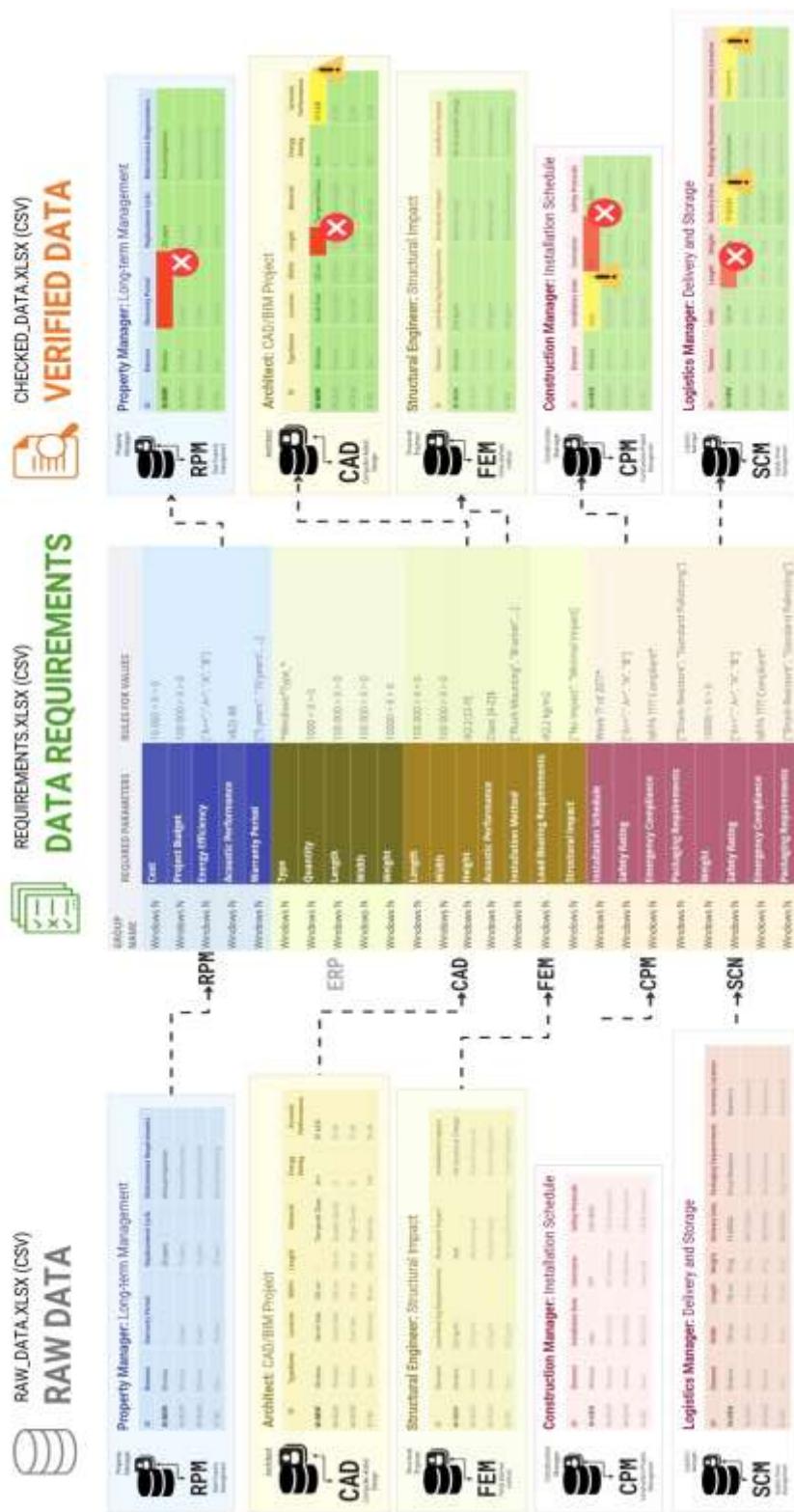


Рис. 4.416 Барлық жүйелер үшін жүргізілген тексеру нәтижесі компания талаптарына сәйкес келмейтін деректерді анықтауға мүмкіндік береді.

Тексеруден өткен объектілер, әдетте, ерекше назар аударуды қажет етпейді. Олар кедергісіз келесі өңдеу және басқа жүйелермен интеграциялау кезеңдеріне өтеді. «Сапалы» элементтерден айырмашылығы, тексеруден өтпеген элементтер ең үлкен қызығушылық тудырады. Мұндай ауытқулар туралы ақпарат критикалық маңызды: оны тек кестелік есептер түрінде ғана емес, сонымен қатар әртүрлі визуализация құралдарын пайдалана отырып жеткізу қажет. Тексерудің нәтижелерін графикалық түрде көрсету деректер сапасының жалпы жағдайын жылдам бағалауға, проблемалық аймақтарды анықтауға және параметрлерді түзету немесе нақтылау бойынша жедел шаралар қабылдауға көмектеседі.

Тексеру нәтижелерін визуализациялау

Визуализация - тексеру нәтижелерін интерпретациялаудың ең маңызды құралы. Дәстүрлі жинақ кестелерінен басқа, ол ақпараттық панельдер, диаграммалар және жобаның элементтері тексеру статусына қарай топтастырылған автоматты түрде жасалатын PDF құжаттарын қамтуы мүмкін. Түстік кодтау мұнда қосымша рөл атқара алады: жасыл сәтті тексерілген объектілерді, сары - қосымша назар аударуды қажет ететін элементтерді, ал қызыл - критикалық қателер немесе негізгі деректердің жоқтығы анықталғандарды білдіреді.

Біздің мысалымызда (Рис. 4.41) біз әр жүйеден деректерді кезең-кезеңімен талдаймыз: CAD (BIM) және жылжымайтын мүлікті басқарудан логистика мен монтаж кестелеріне дейін (Рис. 4.416). Аудит нәтижесінде әр маман үшін автоматты түрде жеке хабарламалар немесе есептік құжаттар, мысалы, PDF форматында (Рис. 4.417) жасалады. Егер деректер дұрыс болса, маманға қысқа хабарлама жіберіледі: «Ынтымақтастық үшін раҳмет». Сәйкесіздіктер анықталған жағдайда, тексеруден өтпеген элементтердің идентификаторлары, атрибуттары және мәндері көрсетілген егжетегейткелі есеп жіберіледі.--



Рис. 4.417 Валидация және автоматты есептік құжаттарды жасау деректердегі кемшіліктерді іздеу және түсіну процесін жеделдетеді, деректерді жасайтын маман үшін.

Автоматтандырылған тексеру процесінің арқасында - деректерде қате немесе бос орын анықталған сәтте, жауапты тұлғаға, яғни тиісті объектілер мен олардың атрибуттарын жасау немесе өңдеу үшін, чат хабарламасы, электрондық пошта немесе PDF құжаты түрінде дерек хабарлама жіберіледі (Рис. 4.418), тексеруден өтпеген элементтер мен атрибуттардың сипаттамаларының тізімімен.-



Рис. 4.418 Автоматты тексеру нәтижелері бойынша есептер қателерді түсінуді жеңілдетеді және жобалық деректерді толтыру жұмысын жеделдетеді.

Мысалы, егер жылжымайтын мүлік басқару жүйесіне (құрылымдалғаннан кейін) "Кепіл мерзімі" атрибуты дұрыс толтырылмаған құжат келсе, жылжымайтын мүлік менеджері тексеру және түзету қажет атрибуттардың тізімімен хабарлама алады.

Сол сияқты, монтаж кестесіндегі немесе логистика деректеріндегі кез келген кемшіліктер автоматты түрде есептің қалыптасуына және, мысалы, тиісті маманға тексеру нәтижелерімен чат немесе электрондық пошта арқылы хабарлама жіберуге әкеледі.

PDF құжаттары мен тексеру нәтижелері графиктерінен басқа, элементтердің жетіспейтін атрибуттарын бөлектеу арқылы пайдаланушыларға 3D геометрияларын визуалды түрде пайдалану мүмкіндігін беретін интерактивті 3D модельдер (Рис. 7.16, Рис. 7.212) мен бақылау панельдерін (дашбордтарды) жасау мүмкіндігі бар, бұл жобадағы элементтердің сапасы мен толықтырын бағалауға мүмкіндік береді.–

Тексеру нәтижелерін автоматты түрде жасалатын құжаттар, графиктер немесе бақылау панельдері түрінде визуализациялау деректерді интерпретациялауды айтарлықтай жеңілдетеді және жобаның қатысушылары арасындағы тиімді өзара әрекеттесуді қамтамасыз етеді.

Әртүрлі жүйелер мен ақпарат көздерінен деректерді автоматты тексеру процесін күнделікті өмірде саналы шешім қабылдаумен салыстыруға болады. Құрылым саласындағы компаниялар бастапқы деректердің сенімділігінен бастап, олардың жобаның мерзімдеріне, құнына және сапасына әсеріне дейінгі көптеген айнымалыларды ескеретіндей, адам да, мысалы, тұрғын үй таңдау кезінде, көлік қолжетімділігі, инфрақұрылым, құны, қауіпсіздік, өмір сапасы сияқты факторларды weights. Барлық осы факторлар біздің өмірімізді құрайтын соңғы шешімдерді қабылдау үшін критерийлер жүйесін

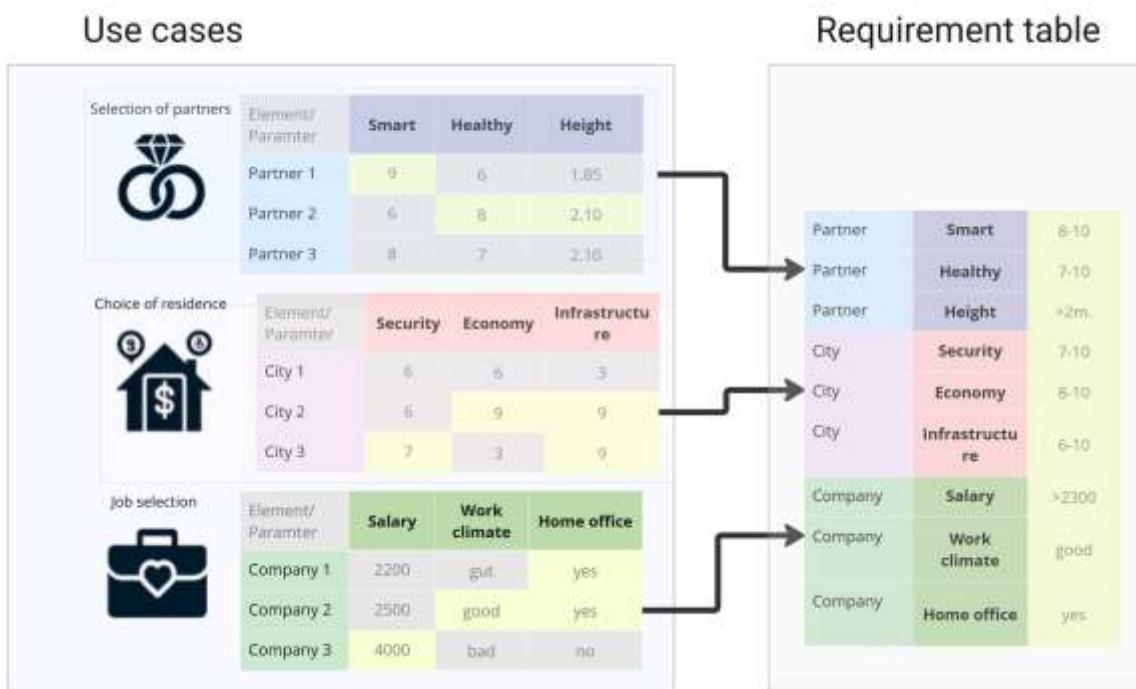
қалыптастырады.

Деректер сапасын тексеруді адамның өмірлік қажеттіліктерімен салыстыру

Деректердің сапасын бақылау әдістері мен құралдарының түрақты дамуына қарамастан, ақпараттың белгіленген талаптарға сәйкестігі - негізгі принцип өзгеріссіз қалады. Бұл принцип бизнес немесе күнделікті өмірде жетілген басқару жүйесінің негізіне енгізілген.

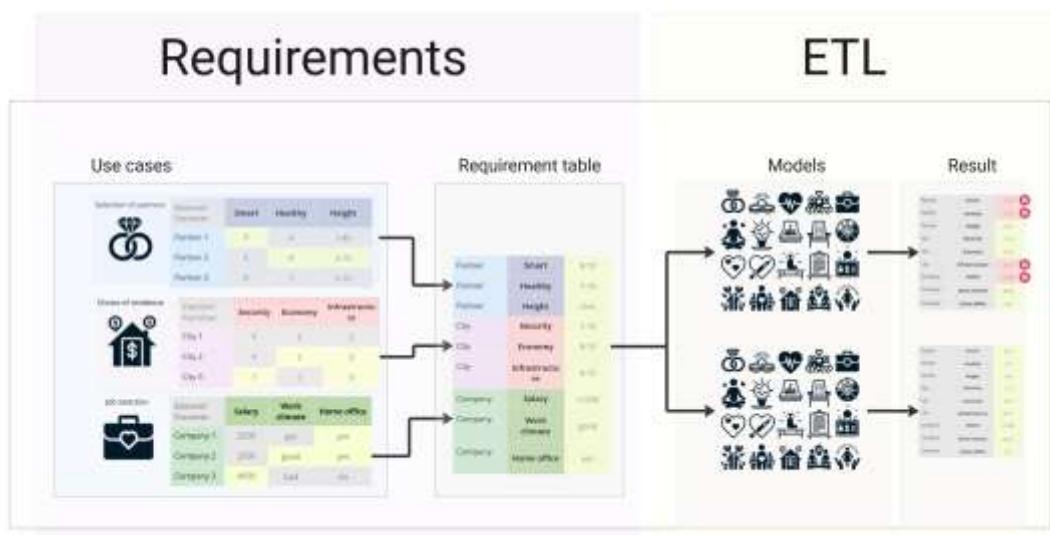
Деректерді итеративті тексеру процесі күнделікті әр адам кездесетін шешім қабылдау процесіне үқсас. Екі жағдайда да біз жинақталған тәжірибеге, қолда бар деректерге және жаңа ақпаратқа сүйенеміз. Қоғытеген өмірлік және кәсіби шешімдер - стратегиялықтан тұрмыстыққа дейін - деректер негізінде қабылдануда.

Мысалы, тұрғын үй немесе өмір серігін таңдау кезінде біз интуитивті түрде санамызда баламаларды салыстыру үшін критерийлер мен сипаттамалар кестесін қалыптастырамыз (Рис. 4.419). Бұл сипаттамалар - адам тұлғасының жеке қасиеттері немесе жылжымайтын мүлік объектісінің параметрлері болсын - соңғы шешімге әсер ететін атрибуттар болып табылады.



Сур. 4.419 Тұратын, жұмыс істейтін немесе серіктес таңдау жеке атрибуттарға қойылатын талаптарға негізделеді.

Структурланған деректерді пайдалану және талаптарды формализациялау тәсілі (Сур. 4.420) кәсіби қызметте де, жеке өмірде де негізделген және саналы таңдау жасауға ықпал етеді.

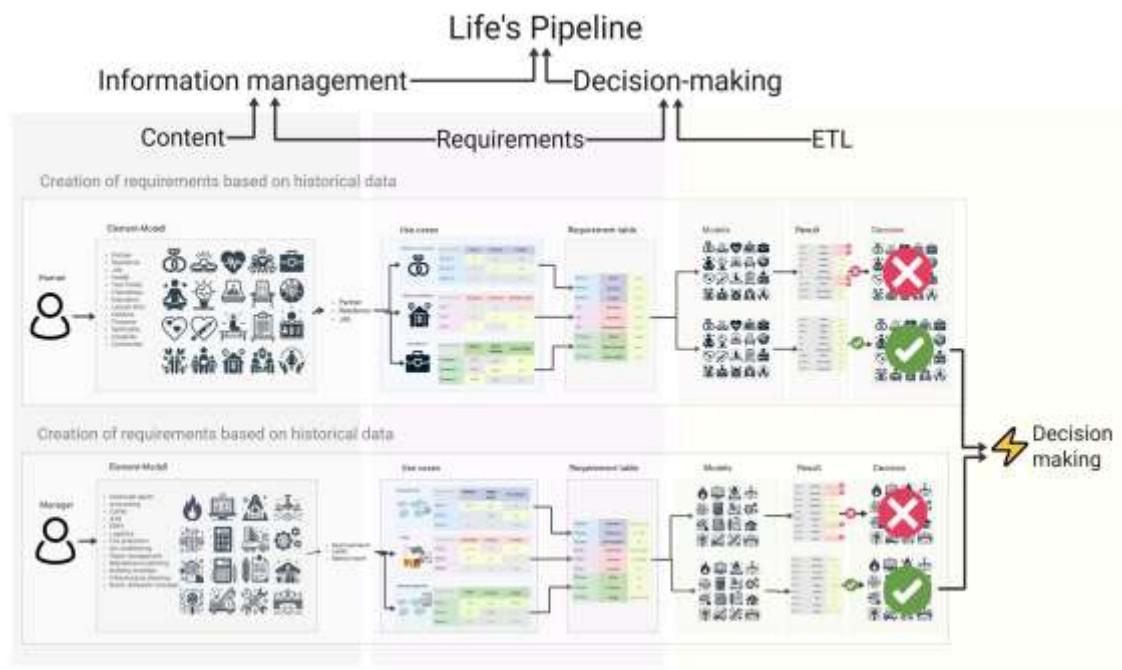


Сур. 4.420 Талаптарды формализациялау өмірлік және іскерлік шешімдерді жүйелендіруге мүмкіндік береді.

Деректерге негізделген шешім қабылдау тәсілі тек бизнес құралы емес. Ол күнделікті өмірге организақтық түрде интеграцияланған, деректерді өңдеу жалпы кезеңдеріне сәйкес келеді (Сур. 4.421), ETL (Extract, Transform, Load) процесіне үқсас, ол біз деректерді құрылымдау кезінде осы бөлімнің басында қарастырған болатынбыз және жетінші бөлімде тапсырмаларды автоматтандыру контекстінде егжей-тегжейлі қарастырамыз:

- Деректер негізі (Extract): кез келген салада – жұмыс немесе жеке өмірде – біз ақпарат жинаімыз. Бизнес саласында бұл есептер, көрсеткіштер, нарық деректері; жеке өмірде – жеке тәжірибе, жақындардың кеңестері, пікірлер, бақылаулар.
- Бағалау критерийлері (Transform): жиналған ақпарат алдын ала анықталған критерийлер негізінде интерпретацияланады. Жұмыста – бұл тиімділік көрсеткіштері (KPI), бюджет шектеулері және нормалар; жеке өмірде – бағасы, орналасу ынғайлылығы, сенімділік, харизма және т.б. параметрлер.
- Болжам жасау және тәуекелдерді талдау (Load): соғы кезеңде трансформацияланған деректерді талдау және мүмкін болатын салдарды салыстыру негізінде шешім қабылданады. Бұл бизнес-процестерге үқсас, мұнда деректер бизнес логикасы мен тәуекелдер сүзгісінен өтеді.

Біз қабылдайтын шешімдер – таңертеңгі ас не жеу сияқты қарапайым таңдаулардан бастап, мансап немесе өмір серігін таңдау сияқты маңызды өмірлік оқиғаларға дейін – негізінен деректерді өңдеу мен бағалаудың нәтижесі болып табылады.



Сур. 4.421 Бизнес және жалпы өмір – деректерге негізделген шешімдер сериясы, мұнда шешім қабылдау үшін пайдаланылатын деректердің сапасы негізгі фактор болып табылады.

Біздің өміріміздегі барлық нәрсе өзара байланысты, және тірі организмдер, соның ішінде адам, табиғат заңдарына бағына отырып, өзгермелі жағдайларға бейімделіп, эволюцияланады, сондай-ақ адамдық процестер, деректерді жинау және талдау әдістері осы табиғи принциптерді көрсетеді. Та-биғат пен адам қызметі арасындағы тығыз байланыс тек біздің табиғатқа тәуелділігімізді емес, сонымен қатар миллиондаған жылдар бойы эволюциямен жетілдірілген заңдарды деректер архитектурасын құруда, процестер мен шешім қабылдау жүйелерін әзірлеуде қолдануға деген үмтүлышсымызды растайды.

Жаңа технологиялар, әсіресе құрылыш саласында, адамзаттың табиғаттан шабыт алып, ең жақсы, тұрақты және тиімді шешімдер жасауға үмтүлұның жарқын мысалы.

Келесі қадамдар: деректерді дәл есептеулер мен жоспарларға айналдыру

Осы бөлімде біз құрылымдалмаған деректерді құрылымдық форматқа айналдыру, деректер модельдерін әзірлеу және құрылыш жобаларында ақпараттың сапасын тексеру процестерін ұйымдастыруды қарастырдық. Деректерді басқару, олардың стандартталуы және класификациясы – жүйелі тәсілді және бизнес талаптарын нақты түсінуді талап ететін негізгі процесс. Осы бөлімде қарастырылған әдістемелер мен құралдар обьектінің өмірлік циклы бойында әртүрлі жүйелер арасында сенімді интеграцияны қамтамасыз етуге мүмкіндік береді.

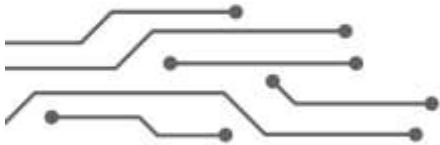
Осы бөлімнің қорытындысында, қарастырылған тәсілдерді күнделікті міндеттерінде қолдануға көмектесетін негізгі практикалық қадамдарды атап өтейік:

- Талаптарды жүйелендіруден бастаңыз

- Жобаларыңыз бер процестеріңіздің негізгі элементтері үшін атрибуттар мен параметрлердің реестрін жасаңыз
- Эр атрибут үшін шекті мәндерді құжаттаңыз
- Процестер мен сыныптар, жүйелер және атрибуттар арасындағы өзара байланыстарды блок-схемалар арқылы визуализациялаңыз (мысалы, Miro, Canva, Visio бағдарламаларында)
- Деректерді автоматтандыруды жүзеге асырыңыз
- Процестерде жиі қолданылатын құжаттарыңыздың қайсысын OCR кітапханалары арқылы цифрландыруға және кестелік форматқа ауыстыруға болатынын тексеріңіз
- CAD (BIM) жүйелерінен деректерді алу үшін кері инженерия құралдарымен таныссыңыз
- Жұмысыңызда жиі қолданылатын құжаттар немесе форматтардан деректерді автоматты түрде кестелік формаға алу үшін конфигурациялауға тырысыңыз
- Эртүрлі деректер форматтары арасында автоматты түрлендірүлерді орнатыңыз
- Классификация үшін білім базасын құрыңыз
- Сала стандарттарымен келісілген ішкі классификаторды әзірлеңіз немесе бар классификаторды пайдаланыңыз
- Эртүрлі классификация жүйелері арасындағы өзара байланыстарды құжаттаңыз
- Командаңызben элементтердің бірегей идентификациясы мен анық классификациясын қолдану тақырыбын талқылаңыз
- Деректерді автоматты тексеру процесін құра бастау – команда ішінде жұмыс істейтіндер мен сыртқы жүйелерге берілетіндер үшін

Мұндай тәсілдерді қолдану деректердің сапасын едәуір арттыруға және олардың кейінгі өнделуін және трансформациясын жеңілдетуге мүмкіндік береді. Кітаптың келесі бөлімдерінде біз құрылымдалған және дайындалған деректерді автоматтандырылған есептеулер, құнды бағалау, күнтізбелік жоспарлау және құрылыш жобаларын басқару үшін қалай қолдануға болатынын қарастырамыз.





В БӨЛІМ ШЫҒЫН ЖӘНЕ УАҚЫТ ЕСЕПТЕУЛЕРІ: ДЕРЕКТЕРДІ ҚҰРЫЛЫС ПРОЦЕСТЕРИНЕ ЕҢГІЗУ

Бесінші бөлім деректерді пайдалану арқылы құрылымдың құнын есептеу және жоспарлау бойынша практикалық аспектілерге арналған. Сметаларды құрастырудың ресурс әдісі мен есептеу процестерін автоматтандыру егжей-тегжейлі талданады. CAD (BIM) модельдерінен көлемдерді автоматтандырылған алу (Quantity Take-Off) әдістері мен оларды есептеу жүйелерімен интеграциялау қарастырылады. Уақыт параметрлерін жоспарлау және құрылымдың құнын басқару үшін 4D және 5D моделдеу технологиялары зерттеледі, олардың қолданылу мысалдарымен. Нерухом мүліктің тұрақтылығын, пайдалануын және қауіпсіздігін бағалауға кешенді тәсілді қамтамасыз ететін 6D-8D кеңейтілген ақпараттық қабаттарының талдауы ұсынылады. Заманауи экологиялық талаптар мен стандарттар контекстінде құрылымдың жобаларының көміртегі ізін және ESG көрсеткіштерін есептеу әдістемелері егжей-тегжейлі қарастырылады. Құрылымды процестерін басқаруда дәстүрлі ERP және PMIS жүйелерінің мүмкіндіктері мен шектеулері сыни түрғыдан бағаланады, олардың бағалау ашықтығына әсері талданады. Жабық шешімдерден ашық стандарттарға және құрылымды процестерінің тиімділігін арттыруға қабілетті деректерді талдаудың икемді құралданына көшу перспективалары болжанады.

ТАРАУ 5.1.

ҚҰРЫЛЫС ЖОБАЛАРЫНЫҢ ШЫҒЫНДАРЫ ЖӘНЕ СМЕТАЛАРЫН ЕСЕПТЕУ

Құрылыш негіздері: көлем, шығын және уақытты бағалау

Құрылыш саласындағы компанияның тұрақтылығын анықтайдын көптеген бизнес-процестердің арасында, мыңдаған жылдар бұрынғыдай, элементтердің дәл санын, жобаның құнын және орындау мерзімдерін бағалау процестері ерекше маңызға ие (Рис. 5.11).

Жазудың дамуы шаруашылық операцияларын есепке алу, сауданың дамуы және ерте қоғамдардағы ресурстарды басқару қажеттілігі сияқты факторлар кешенінің нәтижесі болды. Алғашқы заңды маңызы бар құжаттар – материалдардың құны мен еңбек ақысының есептеулері бар балшықтан жасалған тақталар – сауда мен құрылыш контекстінде қолданылды. Бұл тақталар құрылыш нысандарын салу кезінде тараптардың міндеттемелерін тіркеп, келісімдер мен тауар-қаржы катынастарының дәлелдері ретінде сақталды.

Мыңдаған жылдар бойы бағалау тәсілі дерлік өзгермеді: есептеулер қолмен, сметалық инженердің тәжірибесі мен интуициясына сүйене отырып жүргізілді. Алайда модульдік ERP жүйелері мен CAD құралдарының пайда болуымен, дәстүрлі бағалау тәсілі тез трансформациялануда. Заманауи цифрлық технологиялар уақыт пен құнды есептеудің негізгі процестерін толық автоматтандыруға мүмкіндік береді, бұл құрылыш жобаларының ресурс жоспарлауын дәлдігін, жылдамдығын және ашықтығын арттыруға мүмкіндік береді.

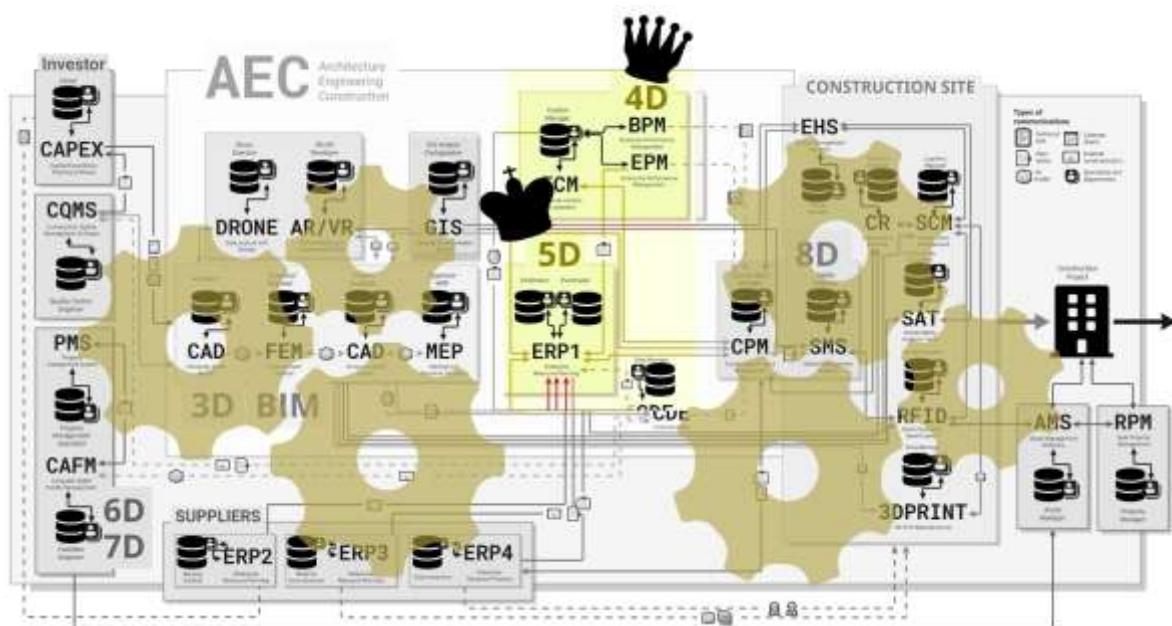


Рис. 5.11 Әртүрлі жүйелердің арасында бизнесте ең маңызды құралдар көлем, құн және уақыт көрсеткіштеріне жауап беретін құралдар болып табылады.

Құрылыш компанияларының негізгі назары жұмыстардың мерзімдері мен құны туралы дәл деректерге шоғырланған. Бұл көрсеткіштер, өз кезегінде, пайдаланылатын материалдардың көлемі мен еңбек шығындарына байланысты, ал олардың ашықтығы рентабельділікке әсер етеді. Алайда есептеу процесстерінің күрделілігі мен олардың жеткіліксіз ашықтығы жиі жобалардың қымбаттауына, мерзімдердің бұзылуына және тіпті компаниялардың банкроттығына әкеледі.

KPMG-нің "Таныс мәселелер – жаңа тәсілдер" (2023) есебіне сәйкес, тек 50% құрылыш жобалары уақытында аяқталады, ал 87% компаниялар капиталдық жобалардың экономикасына бақылаудың артқанын атап өтеді. Негізгі мәселелер білікті кадрлардың жетіспешілігі мен тәуекелдерді болжаудың күрделілігімен байланысты.

Тарихи деректер құрылыш компаниясының өмір бойы өткен жобалардың құны мен уақытын есептеу туралы мәліметтерді жинау барысында жиналады және әртүрлі жүйелердің (ERP, PMIS, BPM, EPM және т.б.) деректер базасына енгізіледі.

Сапалы тарихи деректердің болуы құрылыш ұйымының негізгі бәсекелестік артықшылығы болып табылады, бұл оның өміршеңдігіне тікелей әсер етеді.

Құрылыш және инженерлік компанияларда смета және калькуляция бөлімдері жобалық есептеулер туралы тарихи деректерді жинау, сақтау және жақарту үшін құрылады. Олардың негізгі функциясы – компанияның тәжірибесін жинақтау және жүйелеу, бұл уақыт өте келе жаңа жобалардың көлемі, мерзімі және құны бойынша бағалаудың дәлдігін арттыруға мүмкіндік береді. Мұндай тәсіл өткен жобалардың тәжірибесі мен нәтижелеріне сүйене отырып, болашақ есептеулердегі қателіктерді минимизациялауға көмектеседі.

Жобалардың сметалық құнын есептеу әдістері

Калькуляция мамандарының жұмысында әртүрлі бағалау әдістері қолданылады, олардың әрқайсысы нақты деректер түріне, ақпараттың қолжетімділігіне және жобаның егжей-тегжейлілігіне бағытталған. Ең кең таралғандары мыналарды қамтиды:

- Ресурс әдісі: жобаның сметалық құнын барлық қажетті ресурстарды, мысалы, материалдар, жабдықтар және еңбек, егжей-тегжейлі талдау негізінде бағалау. Бұл әдіс әрбір тапсырманы орындау үшін қажетті барлық тапсырмалар мен ресурстардың тізімін талап етеді, содан кейін олардың құнын есептеу. Бұл әдіс жоғары дәлдікпен ерекшеленеді және сметаларды құру кезінде кеңінен қолданылады.
- Параметрлік әдіс: жобаның параметрлері негізінде құнын бағалау үшін статистикалық модельдерді пайдаланады. Бұл құрылыш алаңының немесе жұмыс көлемінің бірлік құнын талдауды және осы құндарды жобаның нақты жағдайларына бейімдеуді қамтуы мүмкін. Әдіс, егжей-тегжейлі ақпарат әлі қолжетімді болмаған ерте кезеңдерде әсіресе тиімді.
- Бірлік көрсеткіштер әдісі (бірлік құн әдісі): жобаның сметалық құнын өлшем бірлігі (мысалы, шаршы метр немесе текше метр) бойынша есептейді. Бұл әртүрлі жобалардың немесе олардың бөліктерінің құнын жылдам және ыңғайлы салыстыру мен талдауды қамтамасыз етеді.
- Сараптамалық бағалаулар (Дельфи әдісі): сарапшылардың пікірлеріне негізделген, олар өз тәжірибесі мен білімдерін жобаның құнын бағалау үшін пайдаланады. Бұл тәсіл нақты баставпқы деректер болмаған кезде немесе жоба бірегей болғанда пайдалы.

Айта кету керек, параметрлік әдіс пен сараптамалық бағалаулар машиналық оқыту модельдеріне бейімделуі мүмкін. Бұл оқыту деректер жиынтығына негізделген жобаның құны мен мерзімдерін автоматты түрде болжауға мүмкіндік береді. Мұндай модельдердің қолдану мысалдары «Машиналық оқытуды пайдалану арқылы жобаның құны мен мерзімдерін анықтау» тарауында толығырақ қарастырылады (Рис. 9.35).-

Дегенмен, әлемдік тәжірибеде ең танымал және кеңінен қолданылатын әдіс ресурс әдісі болып қала береді. Бұл тек сметалық құнын дәл бағалауды қамтамасыз етіп қана қоймай, сонымен қатар құрылыш алаңындағы жеке процестердің және бүкіл жобаның ұзақтығын есептеуге мүмкіндік береді (толығырақ «Құрылыш графиктері және 4D-жобаның деректері» тарауында).

Құрылыш саласындағы сметалар мен есептеулерді ресурс әдісімен жасау

Ресурсқа негізделген шығындарды есептеу – бұл басқарушылық есептің әдісі, мұнда жобаның құны барлық пайдаланылған ресурстардың тікелей есебіне негізделеді. Құрылыш саласында бұл тәсіл барлық материалдық, еңбек және техникалық ресурстарды, жұмыстарды орындау үшін қажеттілерді егжей-тегжейлі талдау мен бағалауды талап етеді.

Ресурс әдісі бюджет жоспарлауда жоғары мөлшерде ашықтық пен дәлдікті қамтамасыз етеді, себебі смета құрастырылған сәттегі ресурстардың нақты бағаларына бағытталған. Бұл, әсіресе, экономикалық жағдайдың тұрақсыздығы кезінде, бағалардың ауытқулары жобаның жалпы құнына елеулі әсер етуі мүмкін болғанда маңызды.

Келесі тарауларда ресурс әдісі бойынша калькуляция процесін егжей-тегжейлі қарастырамыз.

Құрылыш саласындағы принциптерін жақсы түсіну үшін мейрамханадағы кешкі ас құнының есебімен аналогия жүргіземіз. Мейрамхана менеджері кешкі жоспарлағанда, қажетті өнімдердің тізімін құрастырады, әр тағамның дайындалу уақытына назар аударады, содан кейін шығындарды қонақтар санына көбейтеді. Құрылыш саласында процесс үқсас: жобаның әр элементі (объектілері) үшін постатейлі сметалар қалыптастырылады, ал жобаның жалпы құны барлық шығындардың қосындысымен анықталады - категориялар бойынша қорытынды сметада.

Ресурсқа негізделген тәсілдің негізгі және бастапқы кезеңі компанияның бастапқы деректер базасын қалыптастыру болып табылады. Калькуляцияның бірінші кезеңінде компанияның құрылыш жобалары аясында бар барлық заттар, материалдар, жұмыстар мен ресурстардың құрылымдық тізімі жасалады - қоймадағы шегеден бастап, адамдардың біліктілігі мен сағаттық ставкаларына дейін. Бұл ақпарат «Құрылыш ресурстары мен материалдарының деректер базасы» деп аталатын біртұтас жүйеге енгізіледі – атаулары, сипаттамалары, өлшем бірліктері мен ағымдағы бағалары туралы мәліметтерді қамтитын кестелік реестр. Дәл осы база барлық кейінгі ресурстарды есептегулер үшін – шығындар мен жұмыстардың мерзімдері бойынша негізгі ақпарат көзі болып табылады.

Құрылыш ресурстарының деректер базасы: құрылыш материалдары мен жұмыстар каталогы

Ресурстар мен материалдардың деректер базасы – құрылыш жобасында пайдаланылуы мүмкін әр элемент туралы егжей-тегжейлі ақпаратты қамтиды – тауар, өнім, материал немесе қызмет, оның атауы, сипаттамасы, өлшем бірлігі және бірлік құны, құрылымдық түрде бекітілген. Бұл кестеде жобаларда қолданылатын әртүрлі отын мен материалдардан бастап, әртүрлі категориялар бойынша сағаттық ақысы бар мамандардың егжей-тегжейлі тізімдеріне дейін барлығы табылуы мүмкін.

Database of resources

	1st grade potatoes		Sand lime bricks
	1 kg \$2,99		1 pcs \$1
	Black Angus marble beef		JCB 3CX backhoe loader
	1 kg \$26,99		1 h \$150
	Broccoli		Laborer of the 1st category
	1 pcs \$1,99		1 h \$30

Рис. 5.12 Ресурстар кестесі – материал мен қызметті сипаттайтын, бірлік құнын көрсететін ингредиенттер тізімі.

Ресурстар базасы интернет-дүкеннің тауарлар каталогына ұқсас, мұнда әр тауардың өз атрибуттарының егжей-тегжейлі сипаттамасы бар. Бұл сметашыларға нақты құрылыш процестерінің калькуляциясын есептеу үшін қажетті ресурстарды (интернет-дүкенге тауарларды қосу кезінде) таңдаудың жеңілдетеді.

Ресурстар базасын мейрамхананың аспаздық кітабындағы барлық ингредиенттердің тізімі ретінде де қарастыруға болады. Әр құрылыш материалы, жабдық және қызмет рецепттерде қолданылатын ингредиенттерге ұқсас. Ресурстар базасы – құрылыш материалдары мен қызметтердің егжей-тегжейлі тізімі, олардың бірлік құны: дана, метр, сағат, литр және т. б.

Жаңа элементтерді "Құрылыш ресурстарының базасы" кестесіне екі жолмен қосуға болады - қолмен (Рис. 5.13) немесе компанияның қорларды басқару жүйелерімен немесе жеткізушілердің деректер базаларымен интеграциялау арқылы автоматты түрде.

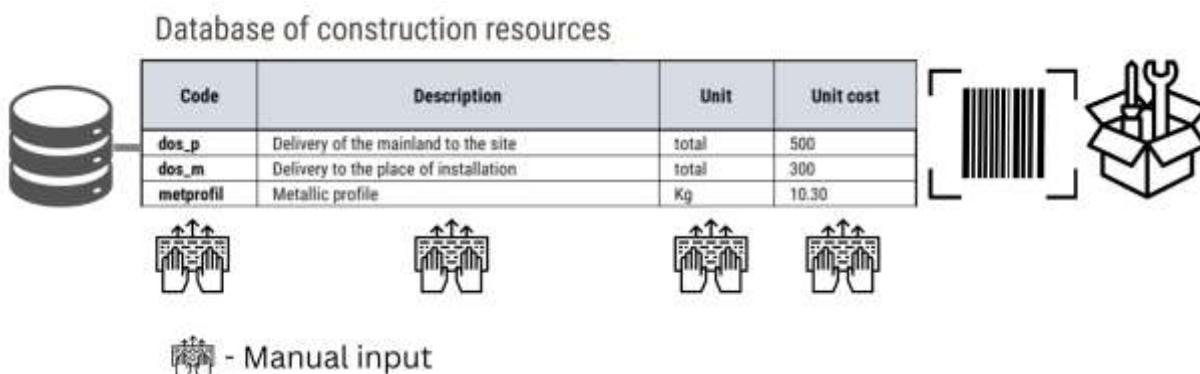


Рис. 5.13 Ресурстар базасы қолмен толтырылады немесе басқа деректер базаларынан автоматты түрде деректерді қабылдайды.

Орташа көлемдегі типтік құрылыш компаниясы құрылыш жобаларында пайдалануға болатын мындаған, кейде он мындаған элементтерді қамтитын деректер базасын пайдаланады. Бұл деректер кейіннен автоматты түрде келісімшарттар мен жобалық құжаттамада жұмыстар мен процестердің құрамын дәл сипаттау үшін қолданылады.

Нарықтық жағдайлардың, мысалы, инфляцияның өзгеруіне ілесу үшін, ресурстар базасындағы әр өнімнің (тауар немесе қызмет) "бірлік құны" атрибуты тұрақты түрде қолмен немесе басқа жүйелерден немесе онлайн платформалардан өзекті бағаларды автоматты түрде жүктеу арқылы жаңартылып отырады (Рис. 5.13).

Ресурс бірлігінің құнын жаңарту ай сайын, тоқсан сайын немесе жыл сайын жүзеге асырылуы мүмкін - бұл ресурс сипатына, инфляцияға және сыртқы экономикалық климатқа байланысты. Мұндай жаңарту есептеулер мен бағалаулардың дәлдігін сақтау үшін қажет, себебі осы базалық элементтер сметашылардың жұмысын бастау нүктесі болып табылады. Актуалды деректер негізінде сметалар, бюджеттер және нарықтың нақты жағдайларын көрсететін графиктер қалыптасады, бұл кейінгі жобалық есептеулердегі қателіктердің тәуекелдерін азайтады.

Ресурс базасына негізделген жұмыстардың құнын есептеу және сметаларды жасау

"Құрылыш ресурстарының базасын" (Рис. 5.13) минималды бірліктермен толтырғаннан кейін, әр процесс немесе құрылыш алаңындағы жұмыс үшін белгілі бір өлшем бірліктері бойынша калькуляцияларды жасауға кірісуге болады: мысалы, бір текше метр бетон, бір шаршы метр гипсокартон қабырғасы, бір метр бордюр немесе бір терезені орнату.

Мысалы, 1 м² (Рис. 5.14) кірпіш қабырғаны түрфызы үшін, алдыңғы жобалардың тәжірибесіне сүйене отырып, шамамен 65 кірпіш (сущность "Силикатный кирпич") қажет, әрқайсысының құны \$1 (атрибут "Штука құны"), бұл жалпы \$65 құрайды. Сонымен қатар, кірпіштерді жұмыс аймағына жақын орналастыру үшін 10 минут бойы құрылыш техникасы (сущность "Погрузчик JCB 3CX") қажет. Техниканың жалға алу құны сағатына 150 доллар, 6 минут пайдалану шамамен 15 доллар тұрады. Бұдан басқа, кірпішті төсеу үшін мердігердің 2 сағаттық жұмысы қажет, оның сағаттық ставкасы 30 доллар, жалпы сома - 60 доллар.

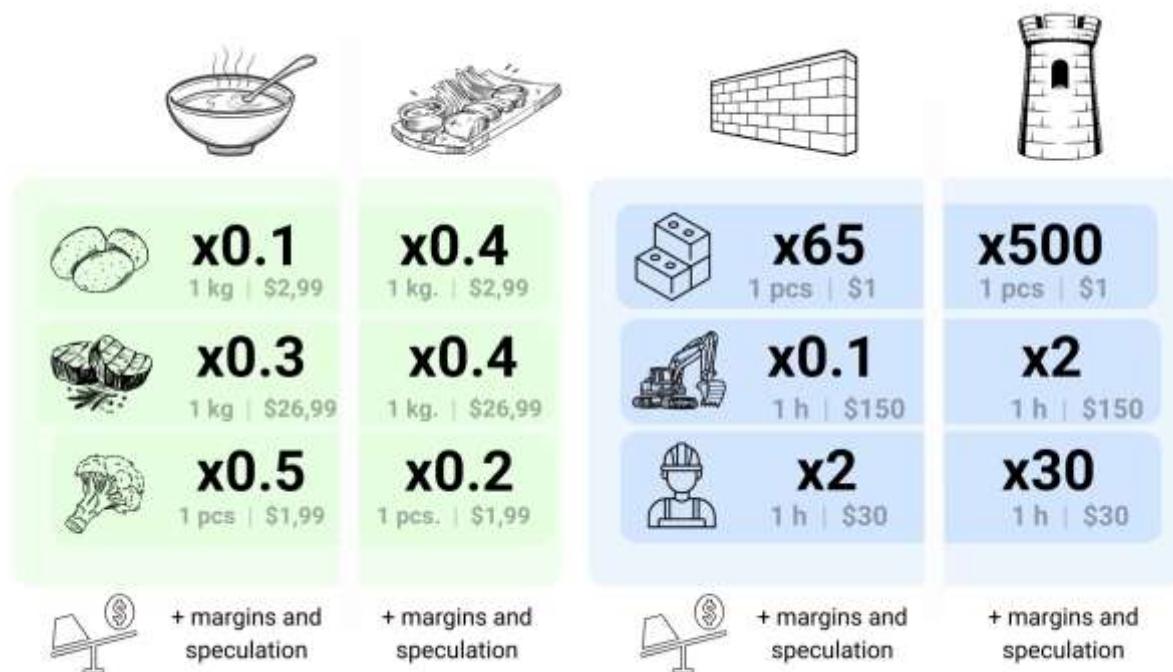


Рис. 5.14 Шығындарды есептеу құрылыш жұмыстарын және процестерді орындау үшін қажетті құрылыш материалдары мен қызметтердің ежей-тегжелі тізімін қамтиды.

Калькуляция құрамдары (осылайша "рецепттер" деп аталатын) компанияның бірдей жұмыстарды орындау барысында жинақтаған тарихи тәжірибесіне негізделеді. Бұл практикалық тәжірибе, әдетте, құрылыш алаңынан алынған кері байланыс арқылы жинақталады. Атап айтқанда, прораб құрылыш алаңында тікелей ақпарат жинайды, нақты енбек шығындарын, материалдардың жұмсауын және технологиялық операцияларды орындау ерекшеліктерін тіркейді. Кейін, сметалық бөліммен ынтымақтастықта, бұл ақпарат итерациялық түрде өндөледі: процестердің сипаттамасы

нақтыланады, ресурстардың құрамы түзетіледі, ал калькуляциялар соңғы жобалардың нақты де-ректерімен жаңартылады.

Рецептте тағамды дайындау үшін қажетті ингредиенттер мен олардың мөлшері сипатталатын болса, сметалық ведомства нақты жұмыс немесе процесс үшін қажетті барлық құрылым материалдары, ресурстар мен қызметтердің ежей-тегжейлі тізімі беріледі.

Тұрақты орындалатын жұмыстар жұмысшыларға, бригадирлерге және сметашыларға ресурстардың қажетті мөлшерін: материалдар, отын, жұмыс уақыты және басқа параметрлерді, жұмыс бірлігін орындау үшін қажет (Рис. 5.15) анықтауға мүмкіндік береді. Бұл деректер сметалық жүйелерге кесте түрінде енгізіледі, мұнда әрбір тапсырма мен операция ресурстық базаның минималды элементтері арқылы сипатталады (тұрақты жаңартылатын бағалармен), бұл есептеулердің дәлдігін қамтамасыз етеді.

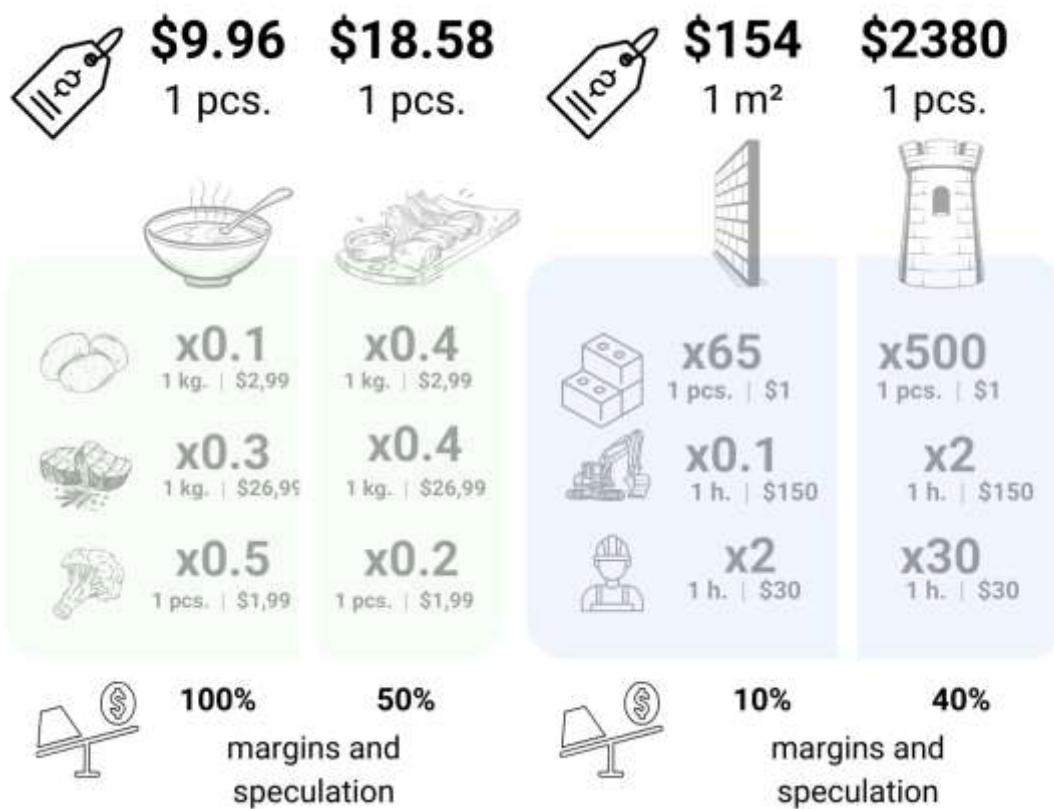


Рис. 5.15 Әрбір жұмыс үшін бірлік бағалар жиналады, мұнда көлем атрибуты оның санына көбейтіледі және пайда пайызын қосады.

Әрбір процесс немесе жұмыс (калькуляция объектісі) жалпы құнын алу үшін, құн атрибуты оның санына және коэффициенттерге көбейтіледі. Коэффициенттер әртүрлі факторларды, мысалы, жұмыстың күрделілігін, аймақтық ерекшеліктерді, инфляция деңгейін, әлеуетті тәуекелдерді (күтілетін жанама шығындар пайызы) немесе спекуляцияларды (қосымша пайда коэффициенті)

ескере алады.

Сметашы, аналитик ретінде, проработың тәжірибесі мен ұсыныстарын стандартты сметаларға айналдырады, құрылыш процестерін ресурстық элементтер арқылы кесте түрінде сипаттайды. Негізінде, сметашының міндегі - құрылыш алаңынан келіп түскен ақпаратты параметрлер мен коэффициенттер арқылы жинақтап, құрылымдау.

Осылайша, жұмыс бірлігінің (мысалы, шаршы немесе куб метр, немесе бір құрылғыны монтаждау) соңғы құны тек материалдар мен еңбекке тікелей шығындарды ғана емес, сонымен қатар компанияның үстеме шығындарын, жанама шығындарды, сақтандыруды және басқа факторларды (Рис. 5.16) қамтиды.

Бізге (рецепттер) калькуляцияларда бағалардың өзектілігі туралы алаңдаудың қажеті жоқ, себебі нақты бағалар әрқашан "ресурстар базасында" (ингредиенттер кестесінде) көрсетіледі. Калькуляция деңгейінде кестеге автоматты түрде (мысалы, элемент коды немесе оның бірегей идентификаторы бойынша) ресурстар базасынан деректер жүктеледі, олар сипаттаманы және бірлік үшін ағымдағы құнын жүктейді, олар өз кезегінде құрылыш материалдарының онлайн платформаларынан немесе интернет-дүкендерінен автоматты түрде жүктелуі мүмкін. Сметашыға жұмыс немесе процесті "ресурстар саны" атрибуты мен қосымша факторлар арқылы сипаттау ғана қалады.

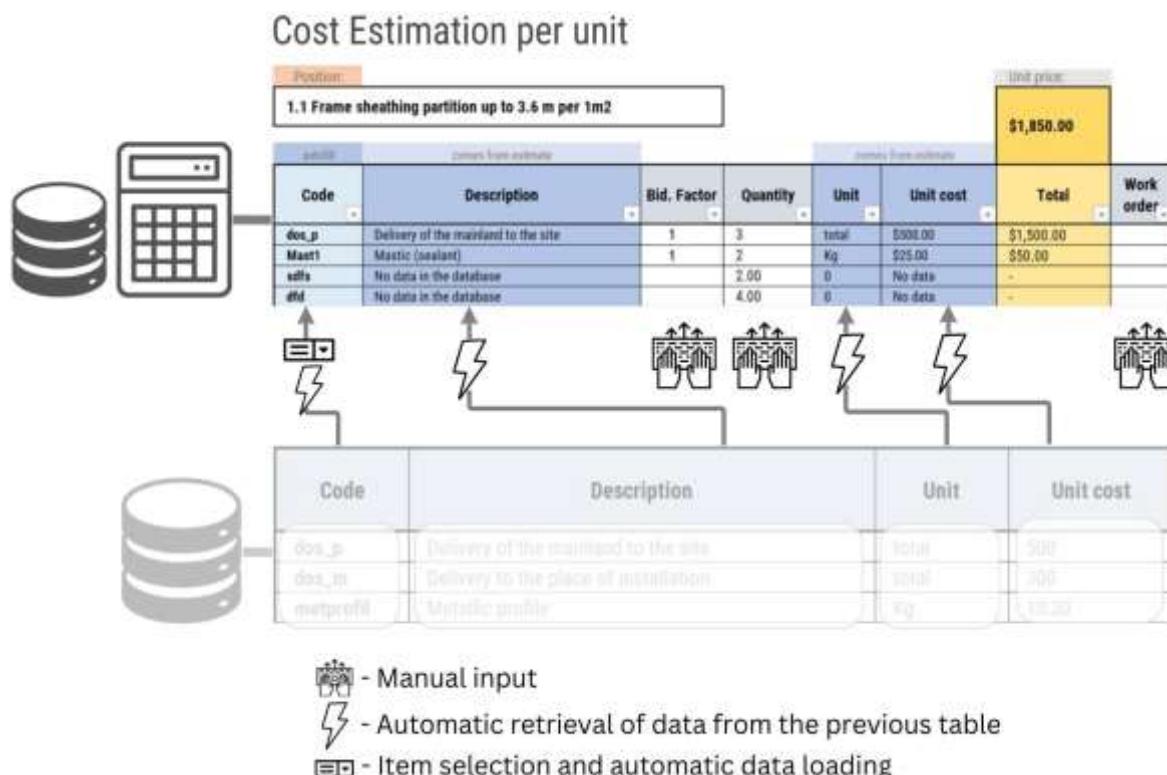


Рис. 5.16 Жұмыстың бірлік құнын есептеу кезеңінде тек қажетті ресурстардың саны атрибуттары толтырылады, қалғанының бәрі автоматты түрде ресурстар базасынан жүктеледі.

Жұмыс құнының калькуляциялары стандартты жобалардың шаблондық кестелері түрінде сақталады, олар құрылыш ресурстары мен материалдарының деректер базасымен тікелей байланысты. Бұл шаблондар қайталанатын жұмыстарды орындаудың стандартталған рецептері болып табылады, болашақ жобалар үшін есептеулерде компанияның барлық деңгейінде біркелкілікті қамтамасыз етеді.

Деректер базасындағы кез келген ресурстың құны өзгерген жағдайда (Рис. 5.13) – бұл қолмен немесе нарықтық бағаларды жүктеу арқылы автоматты түрде (мысалы, инфляция жағдайында) – жаңартулар дереу барлық байланысты калькуляцияларда көрініс табады (Рис. 5.16). Бұл ресурстың базада тек өзгерістер енгізу жеткілікті екенін білдіреді, ал калькуляция мен смета шаблондары ұзақ уақыт бойы өзгеріссіз қалады. Мұндай тәсіл кез келген баға өзгерістерінде есептеулердің тұрақтылығы мен қайталаңыштығын қамтамасыз етеді, олар тек салыстырмалы түрде қарапайым ресурстың кестеде (Рис. 5.13) ескеріледі.–

Әр жаңа жоба үшін стандартты калькуляция шаблонының көшірмесі жасалады, бұл өзгерістер енгізуге және қызмет түрлерін арнайы талаптарға сәйкес түзетуге мүмкіндік береді, компанияда қабылданған түпнұсқа шаблонды өзгертпей. Мұндай тәсіл есептеулерді бейімдеуде икемділікті қамтамасыз етеді: құрылыш алаңының ерекшеліктерін, тапсырыс берушінің тілектерін, тәуекел немесе рентабельділік коэффициенттерін еске кере отырып, және мұның бәрі компания стандарттарын бұзбай. Бұл компанияға пайдаланыс, тапсырыс берушілердің талаптарын қанағаттандыру және бәсекеге қабілеттілігін сақтау арасында тепе-тендік табуға көмектеседі.

Кейбір елдерде осындай калькуляция шаблондары онжылдықтар бойы жинақталып, ұлттық деңгейде стандартталады және құрылыш жұмыстарының құнын есептеу жүйесінің мемлекеттік стандарттарының бір бөлігіне айналады (Рис. 5.17).

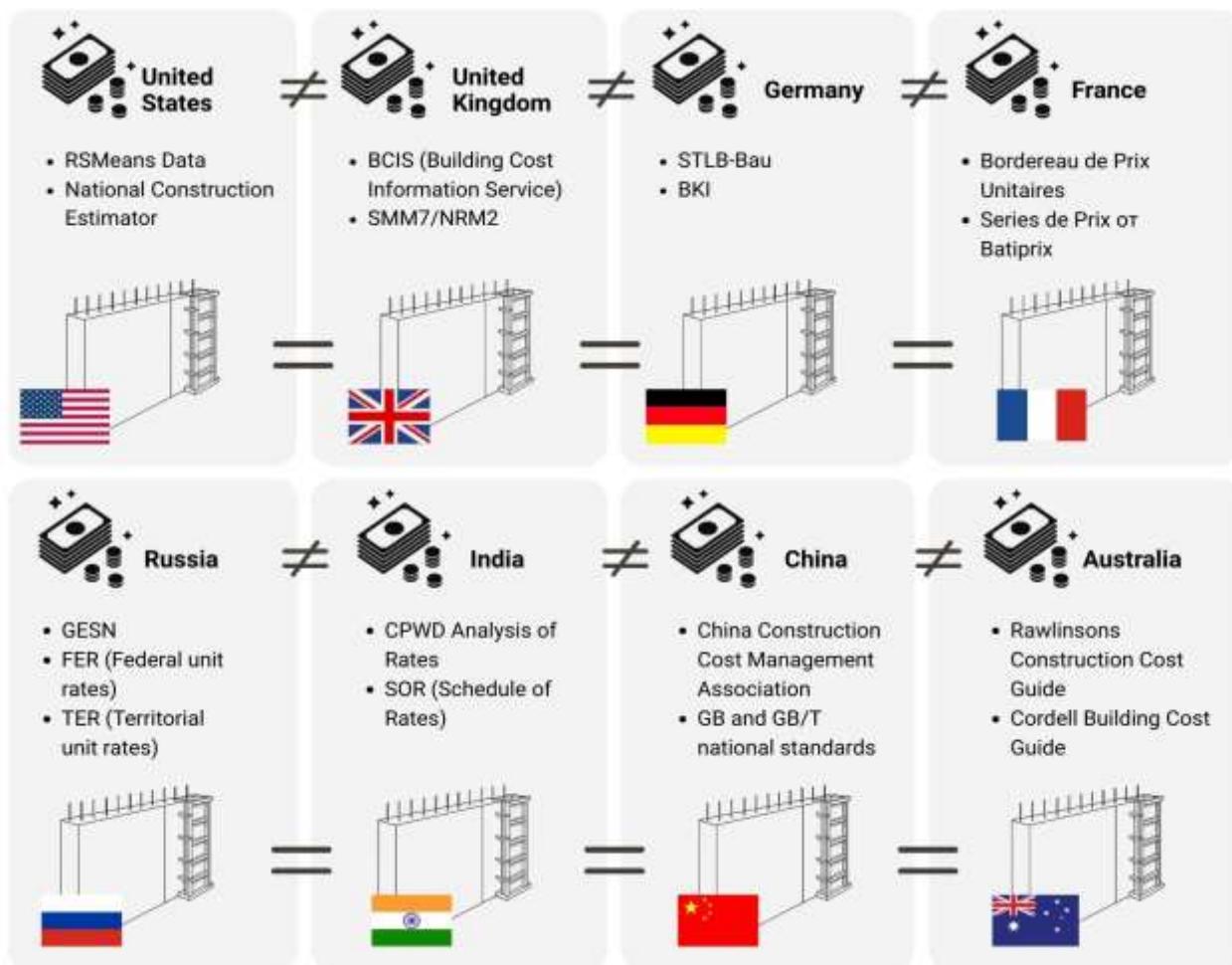


Рис. 5.17 Әлемнің әртүрлі елдерінде бірдей элементтің калькуляциялары үшін өздерінің калькуляция ережелері, құрылыш жұмыстары үшін рецептілер мен нормативтер жинақтары бар.

Мұндай стандартталған сметалық ресурс базалары (Рис. 5.17) мемлекеттік қаржыландырумен жүзеге асырылатын жобаларда нарықтың барлық қатысушылары тарапынан міндетті түрде қолданылуы тиіс. Мұндай стандарттау тапсырыс берушіге бағалардың қалыптасуында, келісімшарттық міндеттемелерде ашықтық, салыстырмалылық және әділдікті қамтамасыз етеді.

Жобаның жалпы құнын есептеу: сметалардан бюджетке дейін

Мемлекеттік және салалық сметалық нормативтер әртүрлі елдердің құрылыш практикасында әртүрлі рөл атқарады. Кейбір мемлекеттер бірдей нормативтерді қатаң сақтауды міндеттесе, да-мыған экономикалардың көпшілігі икемді тәсілді қабылдайды. Нарықтық экономикасы бар елдерде мемлекеттік құрылыш нормативтері әдette тек бастапқы нұкте ретінде қызмет етеді. Құрылыш компаниялары бұл стандарттарды өз операциялық модельдеріне бейімдейді немесе оларды толық қайта өндейді, өздерінің қызмет ерекшеліктерін ескеретін коэффициенттермен толықтырады. Бұл түзетулер корпоративтік тәжірибелі, ресурстарды басқарудың тиімділігін және көбінесе

компанияның спекулятивті пайдасын ескеретін факторларды көрсетеді.

Нәтижесінде бәсекелестік деңгейі, нарықтық сұраныс, мақсатты маржиналдылық және тіпті нақты тапсырыс берушілермен қарым-қатынастар стандартты нормалардан елеулі ауытқуларға әкелуі мүмкін. Бұл практика нарықтың икемділігін қамтамасыз етеді, бірақ сонымен бірге әртүрлі мердігерлердің ұсыныстарын ашық салыстыруды қыыннатады, құрылыш саласында калькуляция кезеңінде спекулятивті баға белгілеу элементін енгізеді.

Жекелеген жұмыс түрлері мен процестер үшін есептеу шаблондары дайындалғаннан кейін – немесе, жиі кездесетіндей, мемлекеттік сметалардан (Рис. 5.17) көшірілген, нақты компанияның «ерекшеліктерін» көрсететін коэффициенттермен толықтырылған – соңғы кезеңде әр позицияның құнын жаңа жобадағы жұмыс көлеміне немесе процестерге сәйкес көбейту ғана қалады.

Жаңа құрылыш жобасының жалпы құнын есептеу кезінде негізгі кезең барлық калькуляция баптары бойынша шығындарды жинақтау болып табылады, оларды жобадағы осы позициялардың көлеміне көбейту.

Жобаның жалпы құнын қалыптастыру үшін, біздің қарапайым мысалымызда, біз бір шаршы метр қабырғаның құрылыш құнын есептеуден бастаймыз және оның есептеу құнын (мысалы, "1 m^2 стандартты қабырғалық элементтерді орнату" жұмысы) жобадағы қабырғалардың жалпы шаршы метрімен (мысалы, CAD жобасындағы немесе проработың есептеулеріндегі "Аудан" немесе "Саны" атрибуты) көбейтеміз.

Сол сияқты, жобаның барлық элементтері үшін құнды есептейміз (Рис. 5.18): жұмыс бірлігінің құнын алып, оны осы жобадағы нақты элементтің немесе оның тобының көлеміне көбейтеміз. Сметашы тек жобадағы осы элементтердің, жұмыстардың немесе процестердің көлемін немесе санын енгізуі керек. Бұл құрылыш сметасын автоматты түрде қалыптастыруға мүмкіндік береді.

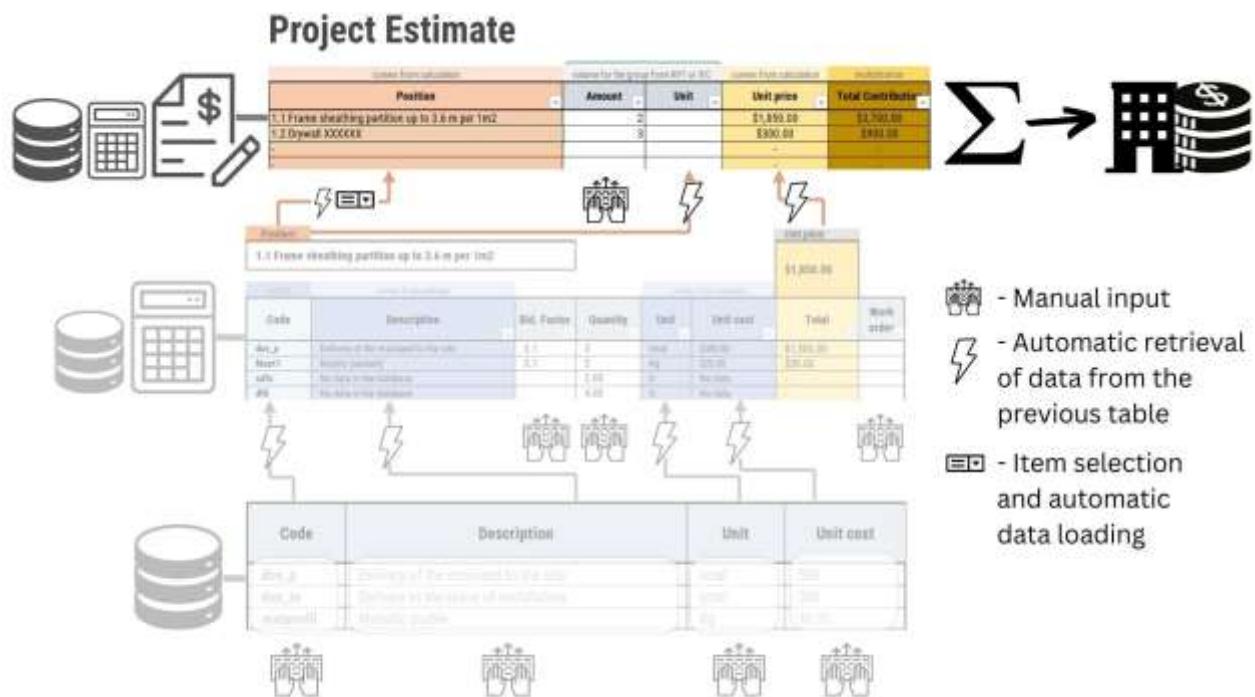


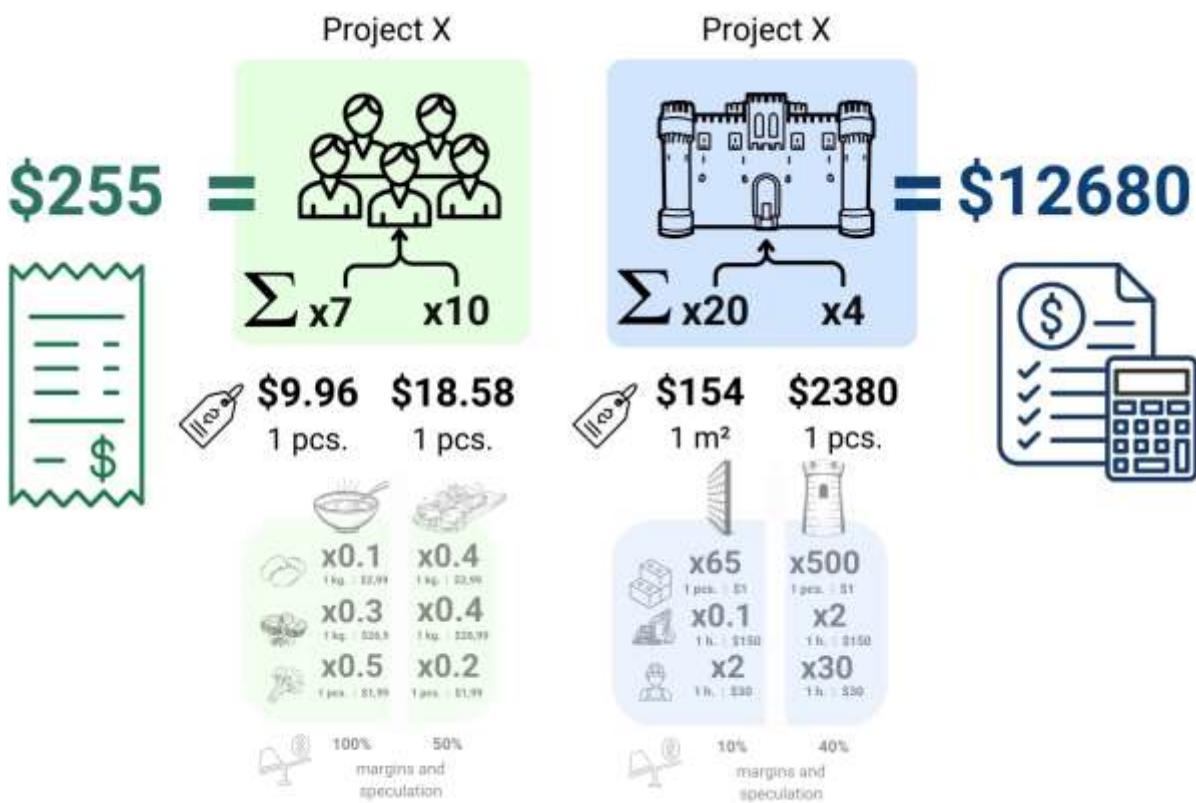
Рис. 5.18 Смета жасау кезеңінде біз тек жұмыс көлемін енгіземіз.

Калькуляциямен бірдей, осы деңгейде біз автоматты түрде дайын есептелген позицияларды (калькуляция шаблонынан немесе жаңа, шаблоннан көшірілген және өндөлген) жүктейміз, олар автоматты түрде жұмыс бірлігінің ағымдағы құнын алып келеді (ресурстар базасынан автоматты түрде жаңартылып отырады (Рис. 5.18 төменгі кесте)). Сондықтан ресурстық база немесе калькуляциялық кестелердегі деректер өзгерген жағдайда – сметадағы деректер автоматты түрде ағымдағы құнға жаңартылады, калькуляцияларды немесе сметаны өзгерту қажеттілігі жоқ.

Ресторан контекстінде іс-шараның соңғы құны дәл осылай есептеледі және барлық кешкі ас құнына тең, мұнда әр тағамның құны, қонақтар санына көбейтілп, жалпы шот құнына қосылады (Рис. 5.19). Және құрылыш саласындағыдан, ресторандардағы тағам дайындау рецептері онжылдықтар бойы өзгермеуді мүмкін. Ингредиенттердің бағасы әр сағат сайын өзгеруі мүмкін.

Ресторан иесі әр тағамның құнын порция мен адамдар санына көбейтіп, іс-шараның жалпы құнын анықтағандай, смета менеджері жобаның барлық компоненттерінің құнын жинақтап, құрылыш сметасын алуға тырысады.

Осылайша, жобадағы әрбір жұмыстың соңғы құны анықталады (Сурет 5.19), ол осы жұмыспен байланысты обьектінің атрибутивті көлеміне көбейтілгенде - жұмыс топтарының құнын береді, одан кейін бүкіл жобаның соңғы құны алынады.



Сурет 5.19 Соңғы смета әр элементтің жұмыс құны атрибутын оның көлемі атрибутына қосу арқылы есептеледі.

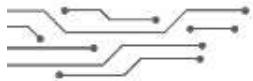
Жобаның соңғы құны (Сурет 5.18) жобаның қаржылық көрінісін білдіреді, бұл тапсырыс берушілерге, инвесторларға немесе қаржы үйымдарына жобаны жүзеге асыру үшін қажетті жалпы бюджет пен қаржылық ресурстарды кез келген күні, ағындағы бағаларды ескере отырып, түсінуге мүмкіндік береді.

Егер ресурстық базаларды, есептеулер мен сметаларды (процестер рецептілері) құру процестері он мындаған жылдар бойы жетілдіріліп, жартылай автоматтандырылған болса және мемлекеттік деңгейде жазылған болса, соңғы сметаның соңғы кезеңі үшін көлем мен элементтердің санын автоматты түрде алу бүгінгі күні жобаның құндылық және уақыттық атрибуттарын есептеу процестеріндегі, сондай-ақ жобаның жалпы бюджетінде тар орын болып қала береді.

Мындаған жылдар бойы көлемдерді есептеудің дәстүрлі әдісі плоские сзыбалар арқылы көлемдік және сандық сипаттамаларды қолмен өлшеу әдістері болды. Цифрлық дәуірдің басталуымен компаниялар көлемдер мен сан туралы ақпаратты CAD модельдерінде бар геометриялық деректерден автоматты түрде алуға болатынын анықтады, бұл мынжылдықтар бойы сандық деректерді алу әдістеріне революция жасады.

Заманауи процестерді бағалау және сметалық іс-әрекеттер CAD дерекқорларынан көлемдік және

сандық атрибуттарды автоматты түрде алу мүмкіндігін қарастырады, оларды есептеу процесіне жүктеп, жобалау кезеңінің кез келген сатысында жобалар топтарының ағымдағы көлемдерін алуға мүмкіндік береді.



ТАРАУ 5.2. КӨЛЕМДІ АЛУ ЖӘНЕ СМЕТАЛАР МЕН КҮНТІЗБЕ ЖОСПАРЛАРЫН АВТОМАТТАНДЫРУ

3D-ден 4D және 5D-ге өту: көлемдік және сандық параметрлерді пайдалану

Ресурстар арқылы сипатталған процестермен калькуляция кестелері қолда бар болғанда (Сурет 5.18), келесі қадам немесе сан параметрлерін автоматтты түрде алу болып табылады, олар есептеулер мен соңғы сметаны құру үшін қажет.

Жобаның көлемдік сипаттамалары - мысалы, қабырғалар немесе жабындар - CAD дерекқорларынан автоматтты түрде алынуы мүмкін. CAD бағдарламаларында жасалған параметрлік объектілер геометриялық ядро құралдары арқылы ұзындық, ені, ауданы, көлем және басқа параметрлердің сандық мәндеріне түрлендіріледі. 3D геометрия негізінде көлемдерді алу процесі келесі, алтыншы бөлімде (Сурет 6.33) CAD (BIM) бойынша жұмысқа арналған бөлімде қарастырылады. Көлемдерден басқа, бір типті элементтердің саны да CAD моделінің дерекқорынан объектілерді категориялар мен қасиеттер бойынша сұзу және топтастыру арқылы алынуы мүмкін. Бұл параметрлер, топтастыруды жүргізуге мүмкіндік беретін, ресурс калькуляциялары арқылы элементтерді есептеулермен, соңғы сметамен және жобаның бюджетімен байланыстырудың негізі болып табылады.

Осылайша, 3D (CAD) моделінен алынған деректер моделі 4D және 5D деп аталатын жаңа параметрлер қабаттарымен толықтырылады. 4D (уақыт) және 5D (құн) атрибуттарының жаңа қабаттарында 3D геометриялық деректері көлемдік атрибуттардың мәндерін алу үшін дереккөз ретінде пайдаланылады.

- 4D – құрылымдық операцияларының орындалу ұзақтығы туралы мәліметтерді 3D элементтерінің параметрлеріне қосатын ақпараттық қабат. Бұл деректер жұмыс кестелерін жоспарлау және жобаның іске асырылу мерзімдерін басқару үшін қажет.
- 5D – деректер моделінің келесі деңгейі, мұнда элементтер құндық сипаттамалармен толықтырылады. Осылайша, геометриялық ақпаратқа қаржылық аспект қосылады: материалдардың, жұмыстардың және жабдықтардың құны, бұл бюджет есептеулерін жүргізуге, рентабельділікті талдауға және құрылымдық процесінде шығындарды басқаруға мүмкіндік береді.

Құн туралы деректер және 3D, 4D және 5D атрибуттік қасиеттердің жобаның объектілер тобына қатысты сипаттамалары модульдік ERP, PIMS жүйелерінде (немесе Excel-ұқсас құралдарда) калькуляцияларға ұқсас сипатталады және жекелеген топтардың, сондай-ақ жобаның жалпы бюджетінің құнын автоматтты түрде есептеу үшін пайдаланылады.

5D атрибуттары және CAD-тан атрибуттардың көлемдерін алу

Құрылымдық жобасының соңғы сметасын дайындау кезінде, біз алдыңғы тарауларда қарастырған (Рис. 5.18), жобаның әр категориясының элементтері үшін көлем атрибуттары не қолмен жиналады, не CAD бағдарламалары ұсынған көлем атрибуттарының спецификацияларынан алынады.-

Дәстүрлі қолмен көлемдерді есептеу әдісі, құрылыш бригадирі мен сметашы мыңдаған жылдар бойы қағазда сзықтар түрінде ұсынылған сызбаларды, соңғы 30 жылда – PDF (PLT) немесе DWG цифрлық форматтарында талдайтынын білдіреді. Олар кәсіби тәжірибеге сүйене отырып, жұмыстар мен қажетті материалдардың көлемдерін өлшайді, көбінесе сызғыш пен транспортирулайтын пайдалана отырып. Бұл әдіс елеулі күшжігер мен уақытты талап етеді, сондай-ақ егжей-тегжейлерге ерекше назар аударуды қажет етеді.

Мұндай әдіспен жұмыстардың көлемін анықтау жобаның ауқымына байланысты бірнеше құннен бірнеше айға дейін уақыт алуы мүмкін. Сонымен қатар, барлық өлшемдер мен есептеулер қолмен орындалғандықтан, адам қателігі қаупі бар, бұл деректердің дәлсіздігіне әкелуі мүмкін, бұл кейіннен жобаның уақыт пен құнды бағалаудағы қателіктерге әсер етеді, олар үшін компания жауапты болады.

CAD дерекқорларын пайдалануға негізделген заманауи әдістер көлемдерді есептеуді айтартылады жеңілдетеді. CAD модельдерінде элементтердің геометриясы көлем атрибуттарын қамтиды, олар автоматты түрде есептеліп (геометриялық ядро арқылы (Рис. 6.33)) және кесте түрінде ұсынылуы немесе экспортталуы мүмкін.-

Мұндай сценарииде сметалық бөлім CAD жобалаушысынан жобаның элементтерінің сандық және көлемдік сипаттамалары туралы деректерді сұрайды. Бұл деректер кесте түрінде экспортталады немесе калькуляциялық базаларға тікелей интеграцияланады – бұл Excel, ERP немесе PMIS жүйелері болуы мүмкін. Мұндай процесс көбінесе ресми сұраудан емес, тапсырыс беруші (инициатор) мен архитектор мен сметашы арасындағы қысқа диалогтан басталады. Төменде құнделікті коммуникациядан автоматты есептеулер үшін құрылымдық кесте қалай қалыптасатынын көрсеттің қаралайым мысал келтірілген (QTO):

- Тапсырыс беруші – «Гимаратқа екінші қабатпен бірдей конфигурацияда тағы бір қабат қосқым келеді»
- Архитектор (CAD) – «Үшінші қабатты қосамыз, конфигурация екінші қабатпен бірдей». Осы хабарламадан кейін жаңа CAD жобасының нұсқасын сметашыға жібереді.
- Сметчик автоматты түрде топтастыру және есептеу жүргізеді (ERP, PMIS, Excel) – «Excel кестесі арқылы QTO ережелерімен (ERP, PMIS) жобаны өткізіп, жаңа қабат үшін санаттар бойынша көлемдерді аламын және сметаны құрамын».



- Нәтижесінде мәтіндік диалог топтастыру ережелерімен кесте құрылымына трансформацияланады:

Элемент	Санат	Қабат
Перекрытие	OST_Floors	3
Колонна	OST_StructuralColumns	3
Лестничный марш	OST_Stairs	3

- Сметчиктің QTO ережелері бойынша жобалауышының CAD моделін автоматты топтастыру процесінен кейін және көлемдерді ресурс калькуляциясына автоматты түрде көбейту (Рис. 5.18) нәтижесінде тапсырысшыға жіберілетін келесі нәтижелер алынады:-

Элемент	Объём	Қабат	Баға бірлігіне	Жалпы құны
Перекрытие	420 м ²	3	150 €/м ²	63 000 €
Колонна	4 шт.	3	2450 €/шт.	9 800 €
Лестничный марш	2 шт.	3	4 300 €/шт.	8 600 €
ИТОГЫ:	—	—	—	81 400 €

- Тапсырысшы – «Рахмет, көп, бірнеше бөлмені қысқарту керек». Және цикл бірнеше рет қайталанады.

Мұндай сценарий келісімдер кезеңінде, тапсырысшының жедел кері байланыс күттін кезеңінде бірнеше рет қайталануы мүмкін. Алайда, практикада мұндай процестер күндер немесе тіпті апталар бойы созылуы мүмкін. Бұғаңға таңда автоматты топтастыру және есептеу ережелерін енгізу арқасында, бұрын көп уақытты қажет ететін әрекеттер бірнеше минутқа қысқаруы тиіс. Топтастыру ережелері арқылы автоматтандырылған көлемдерді алу тек есептеулер мен сметаларды қалыптастыруды жеделдетіп қана қоймай, адам факторының минимизациясы арқасында қателіктер ықтималдығын төмендетіп, жобаның құнын бағалауда ашықтық пен дәлдікті қамтамасыз етеді.

Егер CAD жүйесінде 3D модельді жасаған кезде сметалық бөлімнің талаптары ес-керілсе (бұл практикада әлі де сирек кездеседі), ал элементтердің атаулары, идентификаторлары және олардың классификациялық белгілері сметалық топтар мен сыйнаптардың құрылымымен сәйкес параметрлер ретінде берілсе, онда көлемдік атрибуттарды сметалық жүйелерге қосымша түрлендірусіз автоматты түрде беру мүмкін болады.

CAD-тан көлемдік атрибуттарды автоматты түрде спецификация кестелері түрінде алу, жеке жұмыстар мен жобаның жалпы құны туралы өзекті деректерді жедел алуға мүмкіндік береді (Рис. 5.21). Компания сметаны соңғы өзгерістерді ескере отырып, есептеу немесе калькуляциялық жүйе бары-

сында тек CAD файлын жобалық көлемдермен жаңартып, жылдам қайта есептей алады, бұл барлық кейінгі есептеулердің жоғары дәлдігі мен келісімділігін қамтамасыз етеді.

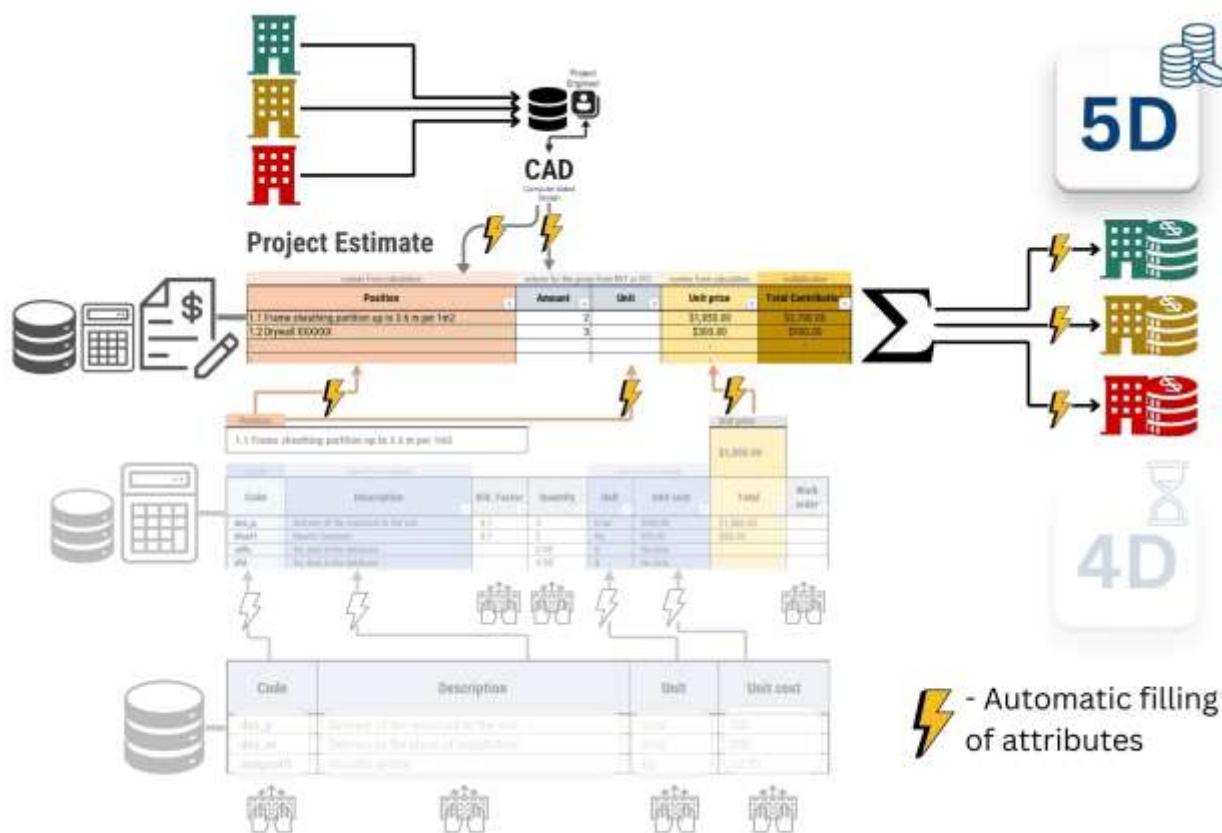


Рис. 5.21 CAD-тан кестелер немесе деректер базалары арқылы алынған көлемдік атрибуттар сметада автоматты түрде енгізіліп, жобаның жалпы құнын бірден есептеуге мүмкіндік береді.

Капиталдық жобалардың күрделілігі артып келе жатқан жағдайда, толық бюджет есептеу және жобалардың жалпы құнын талдау осындаи сценарий бойынша (Рис. 5.21) негізделген шешімдер қабылдау үшін негізгі құралға айналады.-

Accenture-дің «Капиталдық жобалар арқылы үлken құндылық жасау» (2024) зерттеуі бойынша, жетекші компаниялар тарихи ақпаратты пайдалана отырып, цифрлық бастамаларға деректерді талдауды белсенді түрде интеграциялауда. Зерттеу көрсеткендей, операторлар иелері нарықтық тенденцияларды болжау және жобалау басталмай тұрып коммерциялық өміршеңдікті бағалау үшін үлken деректер аналитикасын қолдануда. Бұл, бар жобалар портфелінен деректер қоймаларын талдау арқылы жүзеге асырылады. Сонымен қатар, операторлар иелерінің 79%-ы жобалардың тиімділігін бағалау және нақты уақыт режимінде операциялық шешімдер қабылдау үшін «сенімді» болжамдық аналитиканы енгізуде.

Заманауи тиімді құрылымын жобаларын басқару жобалау кезеңдерінде және жобалауға дейінгі процестерде үлken ақпарат көлемін өңдеу және талдаумен тығыз байланысты. Деректер қоймаларын, ресурстық калькуляцияларды, болжамдық модельдерді және машиналық оқытууды пайдалану тек

есептеулердегі тәуекелдерді минимизациялауға ғана емес, сонымен қатар жобаның қаржыландыруы бойынша стратегиялық шешімдер қабылдауға мүмкіндік береді. Деректер қоймалары мен болжамдық модельдер туралы, калькуляцияларды толықтыратын, кітаптың тоғызыншы бөлімінде толығырақ айтылатын болады.

CAD жобаларынан сметаларды құрастыру үшін қажетті көлемдік параметрлерді автоматты түрде алу QTO (Quantity Take-Off) құралдары арқылы жүзеге асырылады. QTO құралдары жобаның барлық обьектілерін арнайы элемент идентификаторлары немесе элемент атрибуттарының параметрлері бойынша топтастыру арқылы жұмыс істейді, CAD базасында жасалған спецификациялар мен кестелерді пайдалана отырып.

QTO Көлемді алу: жобалық деректерді атрибуттар бойынша топтастыру

Құрылыш саласындағы көлемдік параметрлер мен материалдар санын есептеу (QTO – Quantity Take-Off) жобаны жүзеге асыру үшін қажетті элементтердің сандық сипаттамаларын алу процесін білдіреді. Практикада QTO жиі жартылай қолмен процесс болып табылады, ол әртүрлі дереккөздерден: PDF құжаттарынан, DWG форматындағы сызбалардан және CAD цифрлық модельдерінен деректер жинауды қамтиды.

CAD дереккөрларынан алынған деректермен жұмыс істегендеге, сандық бағалау процесі (QTO) сүзгілеу, сұрыптау, топтастыру және агрегациялау операцияларының тізбегі ретінде жүзеге асырылады. Модель элементтері сыныптар, категориялар және типтер бойынша таңдалады, содан кейін олардың сандық атрибуттары – көлем, алан, ұзындық, немесе сан – есептеу логикасына сәйкес қосылады.

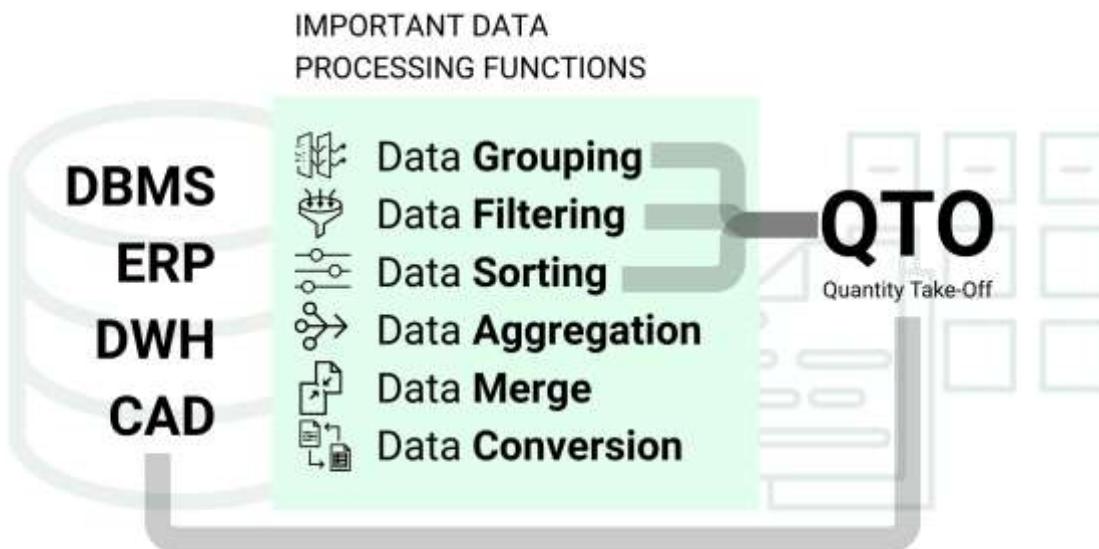


Рис. 5.22 Деректерді топтастыру және сүзгілеу – дереккөрлар мен деректер қоймаларына қолданылатын ең танымал функциялар.

QTO процесі (сүзгілеу және топтастыру) деректерді жүйелендіруге, спецификацияларды қалыптастыруға және сметаларды, сатып алуларды және жұмыстардың орындалу кестелерін есептеу үшін

бастапқы ақпаратты дайындауға мүмкіндік береді. QTO негізі – элементтерді өлшенетін атрибуттар түрі бойынша класификациялау. Эр элемент немесе элементтер тобы үшін сәйкес сандық өлшеу параметрі таңдалады. Мысалы:

- Ұзындық атрибуты (бордюрылар – метрмен)
- Алаң атрибуты (гипсокартон жұмыстары – шаршы метрмен)
- Көлем атрибуты (бетон жұмыстары – текше метрмен)
- Сан атрибуты (терезелер – дана бойынша)

Көлемдік сипаттамалардан басқа, геометрия негізінде математикалық түрде генерирленген, QTO топтастырылғаннан кейін есептеудерде жиі артық шығын коэффициенттері қолданылады (мысалы, логистика мен монтаж кезінде 10% ескеру үшін 1,1) – бұл материалдардың шығынын, монтаж, сақтау немесе тасымалдау ерекшеліктерін ескеретін түзету мәндері. Бұл материалдардың нақты шығынын дәл болжауға мүмкіндік береді және құрылыш алаңында жетіспеушілік немесе артық қордан аулақ болуға көмектеседі.

Автоматты көлемдік есептеу процесі (QTO) дәл калькуляциялар мен сметаларды жасау үшін қажет, көлемдік сипаттамаларды табу процесінде адам факторының әсерін азайтады және материалдарды тапсыру кезінде артық немесе жетіспеушілікті болдырмайды.

QTO процесінің мысалы ретінде, CAD дерекқорынан белгілі бір категория, элементтер класстары бойынша көлемдер спецификациясын көрсету қажет болған кезде кең таралған жағдайды қарастырайық. CAD жобасының қабырғалар категориясындағы барлық элементтерді типтері бойынша топтастырып, әр тип үшін көлем атрибуттарын қосып, QTO көлемдер кестесі түрінде нәтижені ұсынуымыз керек.-

CAD типтік жобасының мысалында (Рис. 5.23) CAD дерекқорындағы қабырғалар категориясындағы барлық элементтер қабырға типтері бойынша, мысалы, "Lamelle 11.5", "MW 11.5" және "STB 20.0" бойынша топтастырылған және метрлік текше түрінде анықталған көлем атрибуттарына ие.

Жобалаушылар мен есептеу мамандары арасында орналасқан менеджердің мақсаты – таңдалған категориядағы элементтер бойынша автоматтандырылған көлемдер кестесін алу. Бұл тек нақты жоба үшін ғана емес, сондай-ақ ұқсас модель құрылымына ие басқа жобаларға қолдануға болатын әмбебап түрде болуы тиіс. Бұл тәсілді масштабтауға және деректерді қайталап пайдалануға мүмкіндік береді, күш-жігерді дублирлеуден аулақ болуға көмектеседі.

Тәжірибелі прорабтар мен сметчиктер әр сызықты мұқият өлшеп, қағаз немесе PDF жоспарларында өлшеу үшін сызығышпен қаруланған дәүірлер өтті - бұл дәстүр мынжылдықтар бойы өзгермеген. 3D модельдеудің дамуы арқасында, әр элементтің геометриясы автоматты түрде есептелетін көлемдік атрибуттармен тікелей байланысты болды, көлемдер мен QTO санын анықтау процесі автоматтандырылды.

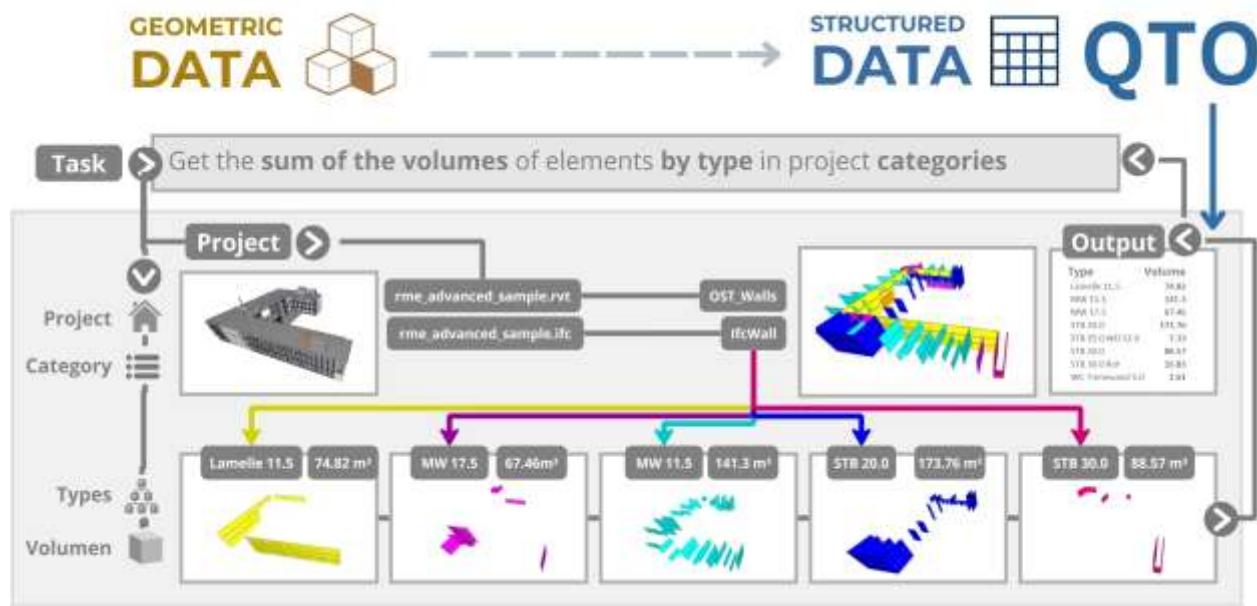


Рис. 5.23 Жобадан көлем және QTO сандық атрибуттарын алу элементтерді топтастыру мен сүзу процесін білдіреді.

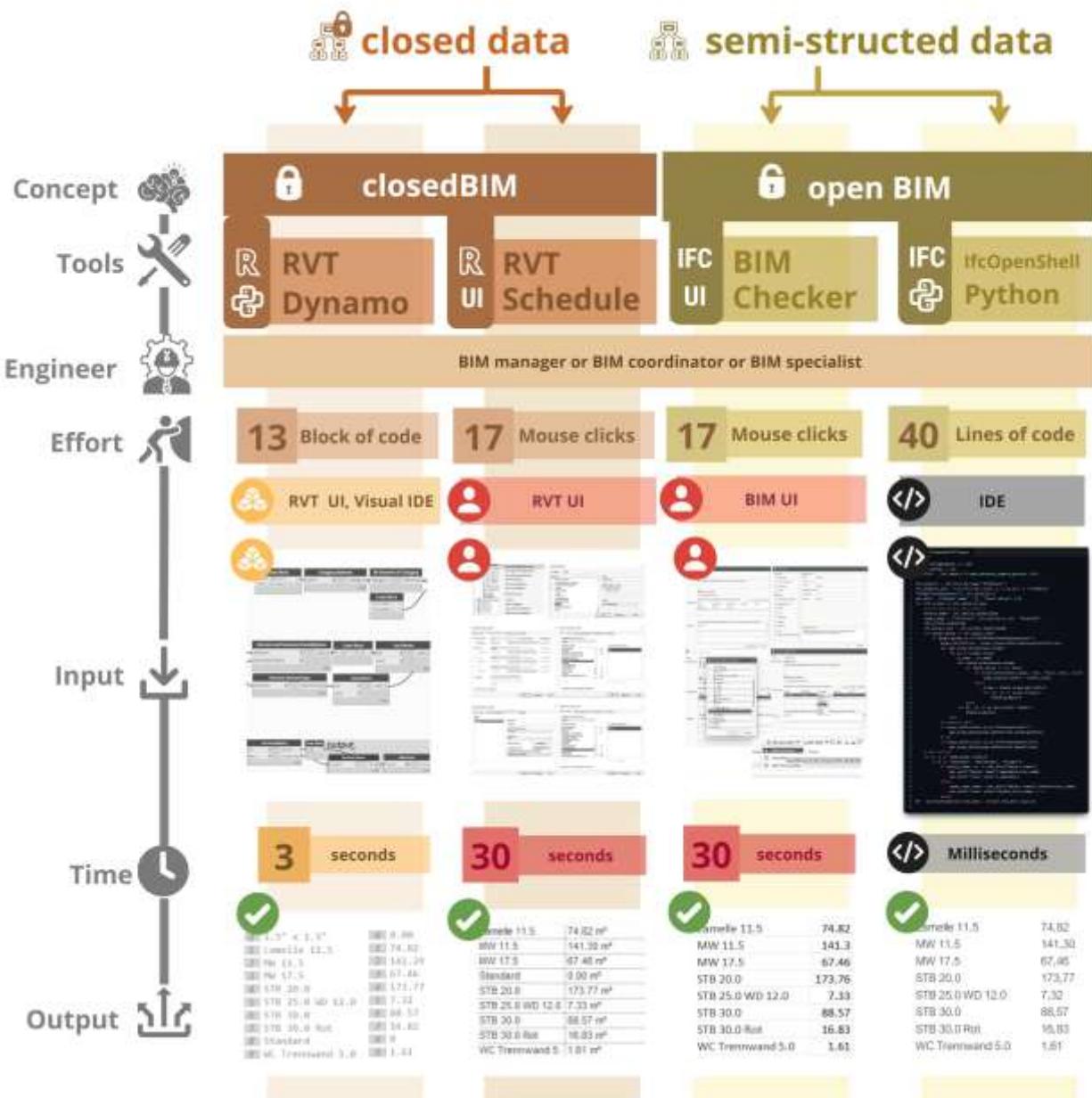
Біздің мысалымызда міндегі "жобадағы қабырғалар категориясын таңдап, барлық элементтерді типі бойынша топтастырып, көлем атрибуттари туралы ақпаратты құрылымдалған, кестелік форматта ұсыну" болып табылады, бұл кестені басқа мамандар калькуляция, логистика, жұмыс кестелері және басқа бизнесс жағдайлары үшін пайдалана алады.

CAD деректерінің жабық сипатының салдарынан, бүгінгі күні әр маман CAD дерекқорына тікелей қол жеткізе алмайды (қол жеткізу мәселелері мен шешімдері туралы кітаптың алтыншы бөлімінде толығырақ). Сондықтан көптеген мамандар open BIM және closed BIM концепцияларына негізделген арнайы BIM құралдарына жүгінуге мәжбүр. Мамандандырылған BIM құралдарымен немесе CAD бағдарламасының ортасында тікелей жұмыс істегендеге, QTO (Quantity Take-Off) нәтижелер кестесін әртүрлі тәсілдермен қалыптастыруға болады – бұл қолмен интерфейс немесе бағдарламалық автоматтандыру қолданылуына байланысты.

Мысалы, CAD (BIM) бағдарламасының пайдаланушы интерфейсін пайдалана отырып, дайын көлемдер кестесін алу үшін шамамен 17 әрекет (түймелерді басу) орындау жеткілікті (Сур. 5.24). Алайда, бұл үшін пайдаланушы модель құрылымын және CAD (BIM) бағдарламалық қамтамасыз ету функцияларын жақсы түсініү қажет.

- Егер автоматизация бағдарламалық код немесе CAD бағдарламалары ішіндегі плагиндер мен API құралдары арқылы жүзеге асырылса, көлемдер кестесін алу үшін қолмен орындалатын әрекеттер саны азаяды, бірақ қолданылатын кітапхана немесе құралға байланысты 40-тан 150 жолға дейін код жазу қажет:
- IfcOpSh (ашық BIM) немесе Dynamo IronPython (жабық BIM) – CAD форматынан немесе CAD бағдарламасынан QTO кестесін алу үшін бар болғаны ~40 жол код қажет.
- IFC_js (ашық BIM) – IFC-моделінен көлемдік атрибуттарды алу үшін шамамен 150 жол код талап етеді.

- CAD (BIM) интерфейс құралдары – қолмен, 17 тышқанды басу арқылы, сол нәтижені алуға мүмкіндік береді.

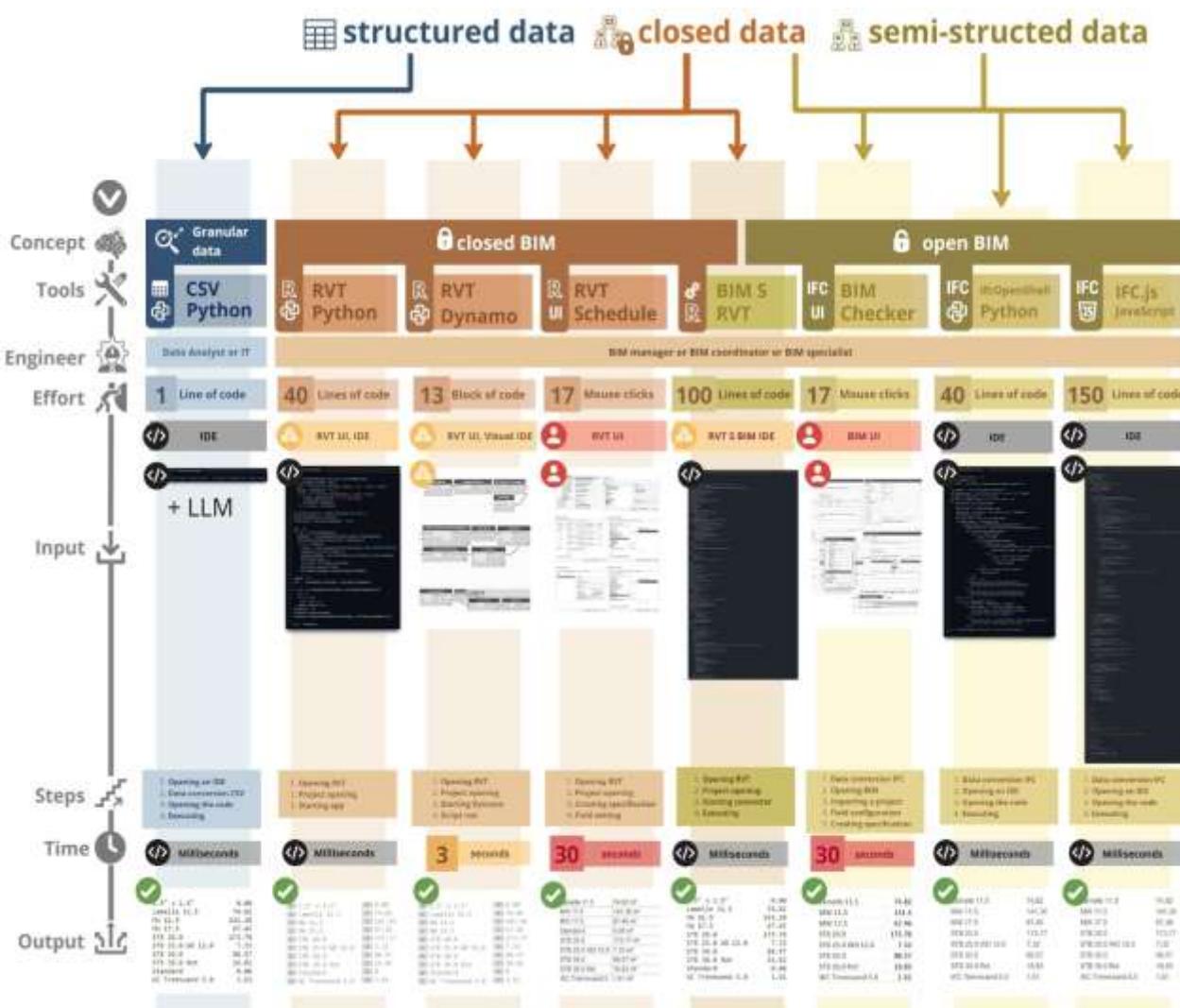


Сур. 5.24 Жобалаушылар мен CAD (BIM) менеджерлері QTO кестелерін жасау үшін 40-тан 150 жол код немесе он шақты түймені басу пайдаланады.

- Нәтижесінде, қорытынды бірдей – элементтер тобы үшін көлемдер атрибуттарымен құрылымдалған кесте. Айырмашылық тек еңбек шығындары мен пайдаланушының техникалық дайындық деңгейінде (Сур. 5.24). Заманауи құралдар, қолмен көлемдерді жинаумен салыстырғанда, QTO процесін айтартықтай жеделдетеді және қателіктер ықтималдығын төмендетеді. Олар жобаның моделінен деректерді тікелей алуға мүмкіндік береді, бұрынғыдан сыйзбалар бойынша көлемдерді қолмен қайта есептеуді қажет етпейді.

Қолданылатын әдіс – ашық BIM немесе жабық BIM болсын – жобаның элементтері үшін бірдей QTO кестесін алуға болады (Сур. 5.24). Алайда, CAD- (BIM-) пайдаланушылар жобалық деректермен жұмыс істегендегі, вендорлар ұсынған арнайы құралдар мен API-ларға тәуелді болады (Сур. 3.213). Бұл қосымша тәуелділік деңгейлерін тудырады және уникалды деректер схемаларын үрленуді талап етеді, сонымен қатар деректерге тікелей қолжетімділікті шектейді. –

CAD деректерінің жабықтығы QTO кестелерін алу және басқа параметрлерді автоматизациялауды, сондай-ақ сыртқы жүйелермен интеграцияны қыннадады. Деректер базаларына тікелей қол жеткізу құралдары мен CAD жобасының деректерін көрінештерлеуде құралдары арқылы ашық құрылымдық dataframe форматына аудару арқылы бір жол кодпен бірдей QTO кестесін алуға болады (Сур. 5.25 – гранулирленген деректер нұсқасы). –



Сур. 5.25 Әр түрлі құралдар жобаның атрибутивті кестелері түрінде бірдей нәтижелер береді, бірақ әртүрлі еңбек шығындарымен.

CAD жобаларынан ашық құрылымдық деректерді пайдалану кезінде, "CAD (BIM) деректерін құрылымдық формаға түрлендіру" бөлімінде айтылғандай, топтастыру, QTO процесі айтартылғатай жеңілдейді.

Ашық құрылымдық деректерді немесе CAD модельдеріне тікелей қол жеткізуге негізделген тәсілдер BIM аббревиатурасымен байланысты маркетингтік шектеулерден босатылған. Олар басқа салаларда ұзақ уақыт бойы қолданылып келе жатқан тексерілген құралдарға сүйенеді (Сур. 7.310 ETL процесі).-

McKinsey зерттеуіне сәйкес, «Ашық деректер: Инновацияларды және өнімділікті ағынды ақпарат арқылы ашу» [102], 2013 жылы жүргізілген, ашық деректерді пайдалану электр энергетикасы обьектілерін жобалау, инжинириング, сатып алу және құрылыш саласында жылына 30-50 миллиард доллар үнемдеу мүмкіндіктерін жасай алады. Бұл құрылыш капиталды шығындарының 15%-дық үнемдеуін білдіреді.

Ашық құрылымдалған (гранулирленген) деректермен жұмыс істеу ақпаратты іздеу мен өңдеуді жеңілдетеді, мамандандырылған BIM платформаларына тәуелділікті азайтады және патенттелген жүйелерді немесе CAD форматтарынан параметрлік және күрделі деректер модельдерін пайдаланбай автоматтандыруға жол ашады.

QTO автоматтандыруы LLM және құрылымдық деректерді пайдалану арқылы

Нысансыз деректерді құрылымдалған формаға ауыстыру әртүрлі процестердің тиімділігін айтартықтай арттырады: бұл деректерді өңдеуді жеңілдетеді (Сур. 4.11, Сур. 4.12) және талаптарды айқын әрі түсінікті етіп, валидация процесін жеделдетеді, бұған дейінгі тарауларда айтылған. Сол сияқты, CAD (BIM) деректерін құрылымдалған ашық формаға ауыстыру (Сур. 4.112, Сур. 4.113) атрибуттарды топтастыру процесін және QTO процесін жеңілдетеді.—

QTO атрибуттар кестесі құрылымдалған формада болғандықтан, құрылымдалған CAD деректерін пайдалану кезінде біз біртұтас деректер модельінде жұмыс істейміз (Сур. 5.25), бұл жобаның деректер модельдері мен топтастыру ережелерін бірдей стандартқа келтіру қажеттілігін жояды. Бұл деректерді бір немесе бірнеше атрибут бойынша тек бір код жолымен топтастыруға мүмкіндік береді. Бұған қарағанда, open BIM және closed BIM жүйелерінде, мұнда деректер полуструктурланған, параметрлік немесе жабық форматтарда сақталады, өңдеу ондаған немесе тіпті жүздеген код жолдарын және геометрия мен атрибутивті ақпаратпен өзара әрекеттесу үшін API пайдалануды талап етеді.-

- ❷ Бір атрибут бойынша құрылымдалған жобаның QTO топтастыруының мысалы. Кез келген LLM чатында (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе кез келген басқа) мәтіндік сұрау:

Менде CAD жобасы DataFrame түрінде бар – жобаның деректерін сұзгіден өткізіп, «Type» параметрі тек «Type 1» мәнін қамтитын элементтерді алшы. ↵

- LLM-нің жауабы, үлкен ықтималдықпен, Python кодында Pandas пайдалану арқылы болады:



Сур. 5.26 LLM көмегімен жазылған бір код жолы CAD жобасын «Тип» атрибуты бойынша топтастыруға және қажетті элементтер тобын алуға мүмкіндік береді.

Екі өлшемді DataFrame-нің қарапайым құрылымы арқасында LLM-ге деректер схемасы мен мөделін түсіндірудің қажеті жоқ, бұл интерпретация кезеңдерін қысқартады және соңғы шешімдерді жасауды жеделдетеді. Бұрын тіпті қарапайым код жазу үшін бағдарламалау тілдерін үйрену қажет болса, қазіргі заманғы тілдік модельдер (LLM) құрылымдалған деректермен жұмыс істегенде процестің логикасын кодқа автоматты түрде түрлендіруге мүмкіндік береді.

Автоматтандыру және LLM тілдік модельдері CAD (BIM) деректерін топтастыру және өңдеу бойынша жұмыс істейтін мамандарды бағдарламалау тілдерін немесе BIM құралдарын үйрену қажеттілігінен толық босата алады, мәтіндік сұраулар арқылы тапсырмаларды шешуге мүмкіндік береді.

Сол сұрау – «қабырғалар» санатындағы барлық элементтерді топтастыру және әр тип бойынша көлемдерді есептеу (Сур. 5.25) – CAD (BIM) ортасында 17 рет интерфейсте басу немесе 40 код жолын жазуды талап етсе, ашық деректерді өңдеу құралдарында (мысалы, SQL немесе Pandas) қарапайым және интуитивті сұрау ретінде көрінеді.

- Бір жолмен Pandas-ты пайдалану арқылы:

```
df[df['Category'].isin(['OST_Walls'])].groupby('Type')['Volume'].sum()
```

Кодтың түсіндірмесі: df (DataFrame) ішінен «Категория» атрибут-бағаны «OST_Walls» мәніне тең элементтерді алыңыз, алынған элементтерді «Type» атрибут-бағаны бойынша топтастырыңыз және топталған элементтердің «Объём» атрибутын қосыңыз.

- CAD-тан алынған құрылымдық жобаны SQL арқылы топтастыру:

```
SELECT Type, SUM(Volume) AS TotalVolume FROM elements WHERE Category = 'OST_Walls' GROUP BY Type;
```

- LLM құралдары арқылы жобаның деректер базасына топтастыру сұрауын қарапайым мәтіндік өтініш - промпт ретінде жазуға болады (Рис. 5.27):-

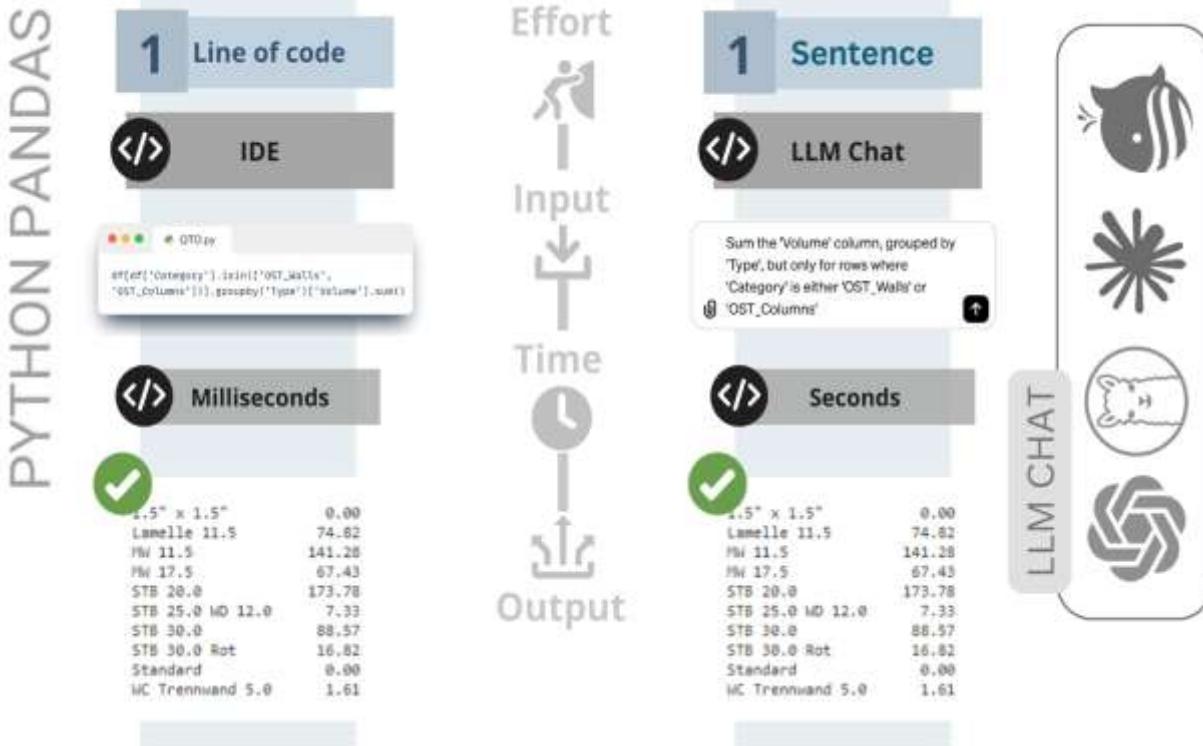


Рис. 5.27 SQL, Pandas және LLM-ді пайдалану арқасында деректерді өңдеуді автоматтандыру енді бірнеше код жолдары мен мәтіндік сұраулар арқылы мүмкін.

CAD деректерінен QTO алу LLM құралдарын (ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok) пайдалану арқылы атрибуттық ақпаратты, сандық және көлемдік деректерді алу дәстүрлі әдістерін түбейгілі өзгертерді.

Енді жобаның менеджерлері, калькуляция мамандары немесе логистика саласындағы мамандар, жобалау бойынша терең білімдері болмаса да және CAD- (BIM-) вендорларының арнайы бағдарламалық құралдарына ие болмаса да, CAD деректер базасына қол жеткізгенде, бірнеше секунд ішінде қабырға немесе басқа объектілер категориясының жалпы көлемін тек сұрау жазу немесе айту арқылы ала алады.

Мәтіндік сұрауларда (Рис. 5.28) LLM моделі пайдалануышының өтінішін өндейді, бұл белгілі бір функцияны бір немесе бірнеше параметрге - кесте бағаналарына қолдану үшін. Нәтижесінде пайдалануышы LLM-пен қарым-қатынаста жаңа параметр бағанын жаңа мәндермен немесе топтастырудан кейін бір нақты мәнді алады.

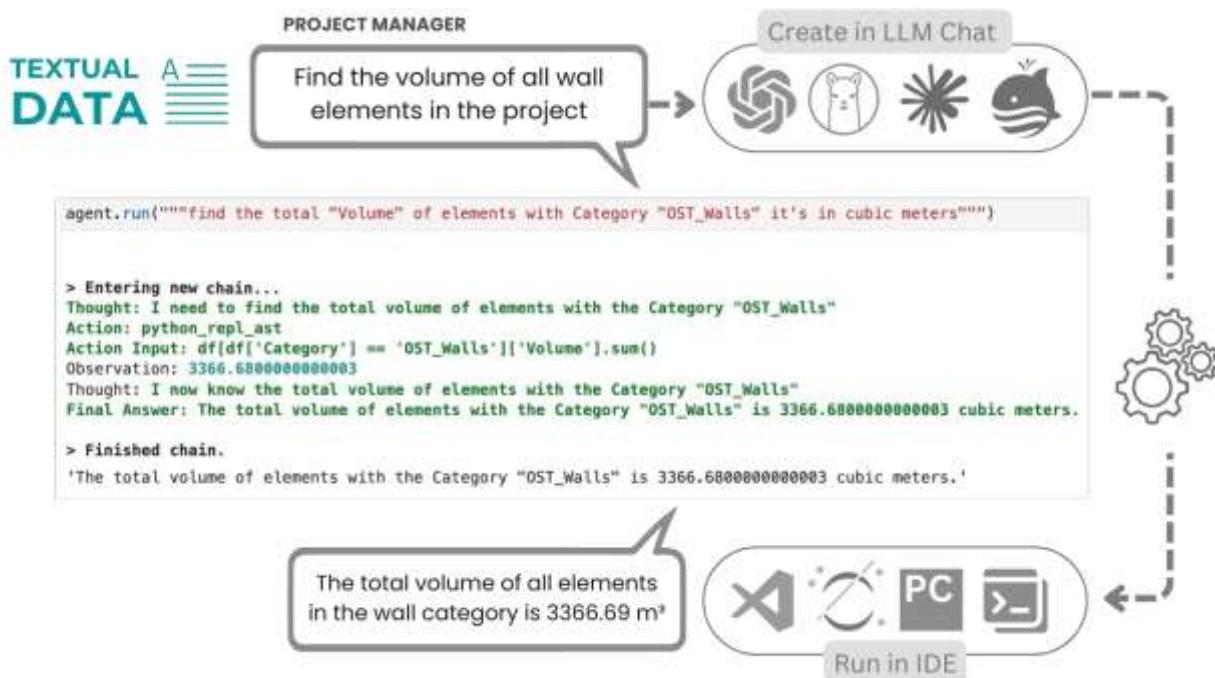


Рис. 5.28 LLM моделі құрылымдық деректермен жұмыс істей отырып, мәтіндік сұраудың контекстінен пайдалануышының қандай топтастыру мен атрибуттар туралы сұрағанын түсінеді.

Егер тек бір элементтер тобы үшін көлемдік көрсеткіштерді алу қажет болса, CAD моделінің деректеріне қарапайым QTO-сұрауын орындау жеткілікті. Алайда, бұқіл жобаның бюджетін немесе сметасын есептеу кезінде, көптеген элементтер топтарынан тұратын жобада, барлық элемент түрлері (класстары) үшін сандық сипаттамаларды алу жиі қажет, мұнда әрбір элемент категориясы жеке өндөледі - сәйкес атрибуттар бойынша топтастырумен.

Сметашылар мен бағалаушылардың тәжірибесінде әртүрлі объектілер үшін топтастыру және есептеу бойынша жеке ережелер қолданылады. Мысалы, терезелер әдетте қабаттар немесе аймақтар бойынша топтастырылады (топтастыру параметрі - Level, Rooms атрибуты), ал қабырғалар - материал немесе құрылымдық түрі бойынша (параметр Material, Type). Топтастыру процесін автоматтандыру үшін, мұндай ережелер алдын ала топтастыру ережелері кестесі түрінде сипатталады. Бұл кестелер жобадағы әрбір элемент тобы үшін есептеулерде қандай атрибуттарды пайдалану көрктігін анықтайтын конфигурациялық шаблондар ретінде қызмет етеді.

Excel кестесіндегі топтар үшін ережелерді пайдалана отырып, жобаның QTO есебі

Нақты құрылымың жобаларында бір топ элементтер ішінде бірнеше атрибут бойынша агрегацияларды орындау қажеттілігі жиі туындаиды. Мысалы, «Терезелер» категориясымен жұмыс істегендегі (мұнда Category атрибуты OST_Windows немесе IfcWindows сияқты мәндерді қамтиды), элементтерді тек типі бойынша ғана емес, сонымен қатар, тиісті атрибутта көрсетілген жылу өткізгіштік деңгейі сияқты қосымша сипаттамалар бойынша да топтастыруға болады. Мұндай көпөлшемді топтастыру нақты топ үшін дәл нәтижелер алуға мүмкіндік береді. Сол сияқты, қабырғалар немесе төбелер категориялары бойынша есептеулер жүргізгенде, материал, деңгей, қабат, отқа тәзімділік және басқа параметрлер сияқты атрибуттардың еркін комбинацияларын фильтрлер немесе топтастыру критерийлері ретінде пайдалануға болады (Рис. 5.29).

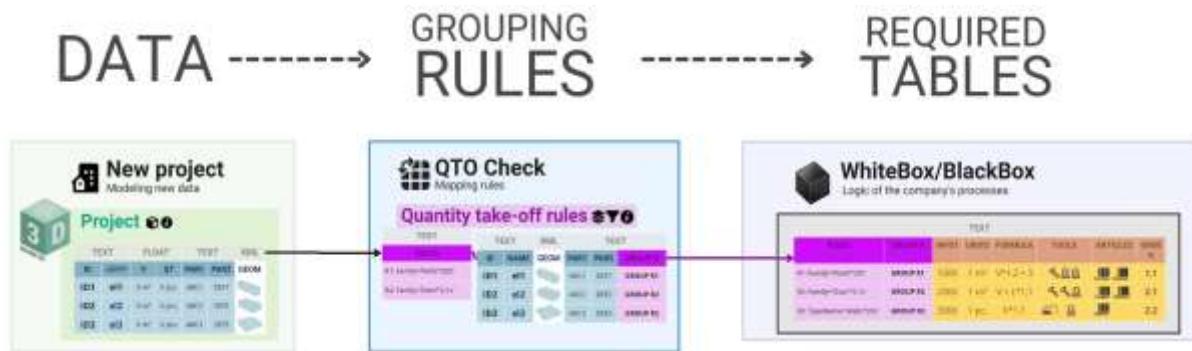


Рис. 5.29 Жобадағы әрбір топ немесе категория үшін бір немесе бірнеше критерийден тұратын топтастыру формуласы бар.

Мұндай топтастыру ережелерін анықтау процесі «Деректер талаптарын жасау және сапасын тексеру» бөлімінде сипатталған деректер моделдерімен жұмыс істеу процесіне үқсас (Рис. 4.45). Мұндай топтастыру және есептеу ережелері қажетті шарттарды ескере отырып, категориялық элементтердің сандық немесе көлемдік атрибуттарының автоматты түрде есептелуінің дәлдігі мен реlevанттылығын қамтамасыз етеді.-

- ❷ Келесі код мысалы жобалар кестесін фильтрлейді, нәтижесінде алынған деректер жиынтығы тек Category атрибут-бағаны «OST_Windows» немесе «IfcWindows» мәндерін қамтитын және Type атрибут-бағаны «Type 1» мәнін қамтитын элементтерді ғана көрсетеді:

Менде жоба DataFrame бар - деректерді фильтрлеңіз, сонда жиынтықта тек «Категория» атрибуты «OST_Windows» немесе «IfcWindows» мәндерін қамтитын және Type бағаны «Type 1» мәнін қамтитын элементтер қалуы керек ↵

■ LLM-нің жауабы:

```
1 df[(df['Category'].isin(['OST_Windows', 'IfcWindows'])) & (df['Type'].str.contains("Type 1"))]
```

Рис. 5.210 Excel формуласына үқсас бір код жолы жобаның барлық элементтерін бірнеше белгілер бойынша топтастыруға мүмкіндік береді.

CAD деректерін құрылымдық ашық форматтарға (Рис. 4.113) аудыстырғаннан кейін алынған кодты жоғарыда аталған танымал IDE (интеграцияланған даму ортасы) - PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагини, Thonny, Wing IDE, IntelliJ IDEA Python плагині, JupyterLab немесе танымал онлайн құралдар: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker - оффлайн режимінде іске қосуға болады.-

■ «Терезелер» категориясы бойынша тек белгілі бір жылу өткізгіштік мәнімен жобаның QTO DataFrame-ін алу үшін, біз LLM-ге келесі сұранысты қолдана аламыз:

Менде жоба DataFrame бар - деректерді фильтрлеңіз, сонда жиынтықта тек «Категория» атрибуты «OST_Windows» немесе «IfcWindows» мәндерін қамтитын жазбалар және ThermalConductivity бағаны 0,5 мәнін қамтитын элементтер қалуы керек

■ LLM-нің жауабы:

```
1 df[(df['Category'].isin(['OST_Windows', 'IfcWindows'])) & (df['ThermalConductivity'] == 0.5)]
```

Рис. 5.211 Pandas Python сұраныстарының өте қарапайым тілі кез келген жобалар үшін QTO жүргізуға мүмкіндік береді.

LLM-ден алынған жауапта (Рис. 5.211) екі критерийді біріктіру үшін логикалық шарт "&" пайдаланылады: жылу өткізгіштік мәні және екі категорияның біріне жату. "isin" әдісі "Категория" атрибут бағанының мәні берілген тізімде бар-жоғын тексереді.-

Элементтер тобының көп саны бар жобаларда, топтастыру логикасы әртүрлі - жобаның әрбір категориясы үшін (мысалы: терезелер, есіктер, жабындар) жеке топтастыру ережелерін орнату қажет, олар қосымша коэффициенттер немесе атрибуттарды есептеу үшін қорытынды формулаларды қамтуы мүмкін. Бұл формулалар (Рис. 5.212 атрибут „formel“, мысалы x-значение саны және у-объем тобы) және коэффициенттер әрбір топтың ерекше сипаттамаларын ескереді, мысалы:-

- материалдың көлеміне қосымша% артықшылықтарды ескеру
- материалдың фиксиленген қосымша саны
- есептеулердегі ықтимал тәуекелдер мен қателіктерге байланысты түзетулер формулалар түрінде

Топтастыру және фильтрация ережелері әрбір элемент категориясы үшін параметрлер формулалары түрінде қалыптасқаннан кейін, оларды жолдық кесте түрінде сақтау мүмкін – мысалы, Excel форматында (Рис. 5.212). Бұл ережелерді құрылымдалған түрде сақтау жобалық деректерді алу, фильтрациялау және топтастыру процесін толық автоматтандыруға мүмкіндік береді. Көптеген жеке сұрауларды қолмен жазудың орнына, жүйе параметрлер кестесін оқып, модельге (жобаның жалпы датафреймі (Рис. 4.113)) сәйкес ережелерді қолданады, жобаның әрбір элемент категориясы үшін қорытынды QTO-кестелерін қалыптастырады.-

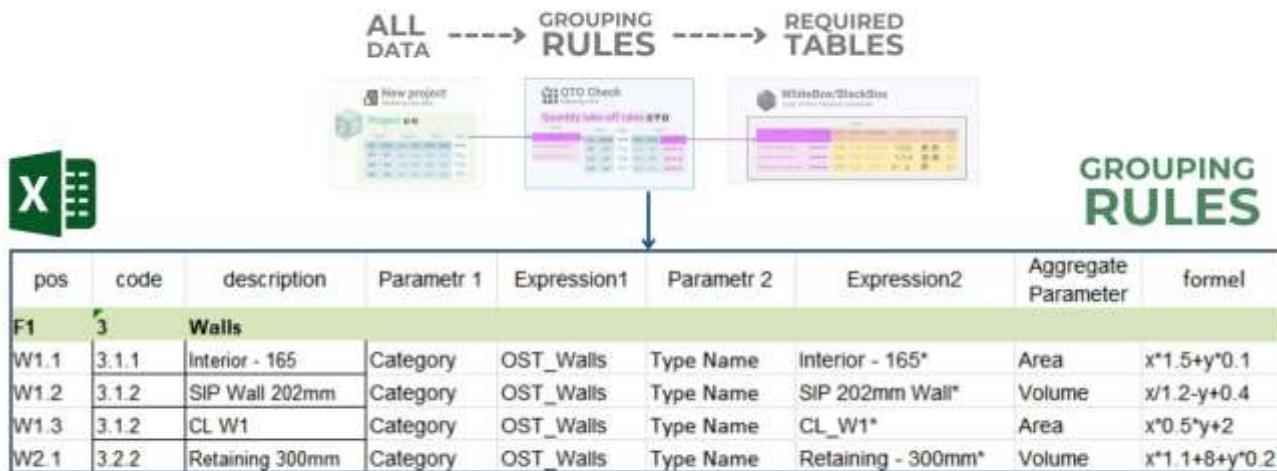


Рис. 5.212 QTO атрибуттарын топтастыру кестесі жобаның элементтерін топтастыру ережелерін орнатады, әрбір категория үшін нақты жалпы сан мен көлемді қамтамасыз етеді.

Жиналған ережелер жобаны топтастыруға және барлық қажетті есептеулерді орындауға мүмкіндік береді, соның ішінде көлем атрибуттарын түзету. Нәтижесінде көлемдер «нақты көлемге» келтіріледі, олар есептеулер мен калькуляциялар үшін пайдаланылады, CAD моделінің жобалашу кезеңінде бастапқыда болған көлемдер емес.

Жобаның барлық көлемдік QTO кестелерін автоматты түрде жасау процесінде қосымша бағдарлама топтастыру ережелері кестесінің барлық категориялары бойынша өтуі, топтастыру атрибуттарын алу, барлық жобалық элементтерді олар бойынша топтастыру және осы топ үшін көлем атрибутын агрегаттау, сонымен қатар оны нақтылаушы факторға немесе коэффициентке көбейтуі тиіс.

LLM-ден осындай шешім үшін код жазуды сұраймыз, мұнда код екі кестені жүктеуі керек - топтастыру ережелері кестесін (Рис. 5.212) және жобаның деректер кестесін (Рис. 4.113), содан кейін топтастыру ережелерін қолдану, элементтерді берілген ережелер бойынша топтастыру, агрегатталған мәндерді есептеу және нәтижелерді жаңа Excel файлына сақтау.--

❷ LLM чатында мәтіндік сұраныс жіберіңіз:

Маған «basic_sample_project.xlsx» файлына жобалық деректерді оқу үшін код қажет, содан кейін «Grouping_rules_QTO.xlsx» файлына ережелерді оқу, олар бойынша барлық деректерді 'Parameter 1' және 'Parameter 2' бойынша топтастыру, 'Aggregate Parameter' бойынша агрегаттау, 'Expression2' бойынша фильтрациялау, 'Formel1' бойынша есептеулерді орындау және QTO кестесін 'QTQ_table2.xlsx' файлына сақтау ↴

■ LLM-нің жауабы:

```

1 import pandas as pd
2
3 # Reading CSV and Excel files with project data and grouping rules respectively
4 df = pd.read_csv('../input/rac_basic_sample_project.xlsx', low_memory=False)
5 df_qto = pd.read_excel('../input//Grouping_rules_QTO.xlsx', header=0, skiprows=[1])
6
7 # Cleaning numerical columns in the dataframe by removing non-numeric characters
8 numeric_cols = ['Area', 'Volume', 'Depth', 'Width', 'Length', 'Perimeter']
9 df[numeric_cols] = df[numeric_cols]
10 df.apply(lambda col: pd.to_numeric(col.astype(str).str.extract('(\d+\.\d*)')[0]).fillna(0))
11
12 # Define a function to process each row in the grouping rules dataframe
13 def process_row(row):
14     # Group by specified parameters and calculate sum and count for the aggregate parameter
15     group = df.groupby([row["Parameter 1"], row["Parameter 2"]])[row["Aggregate Parameter"]].agg(['sum', 'count'])
16     group = group[group.index.get_level_values(1).str.contains(row["Expression2"])]
17
18     # Perform calculation defined in the rules
19     group['After Calculation'] = eval(row["formel"], {'x': group['sum'], 'y': group['count']})
20
21     return group.reset_index()
22
23 # Apply the function to each row in the grouping rules dataframe and concatenate the results
24 dfgroups = pd.concat([process_row(row) for _, row in df_qto.iterrows()]).reset_index(drop=True)
25
26 # Save the processed data to an Excel file using the OpenPyXL engine
27 dfgroups.to_excel('QTO_table2.xlsx', engine='openpyxl', index=False)

```

Рис. 5.213 Python Pandas коды CAD жобасынан жобалық деректерді өндейді, оларды Excel ережелеріне сәйкес топтастырады.

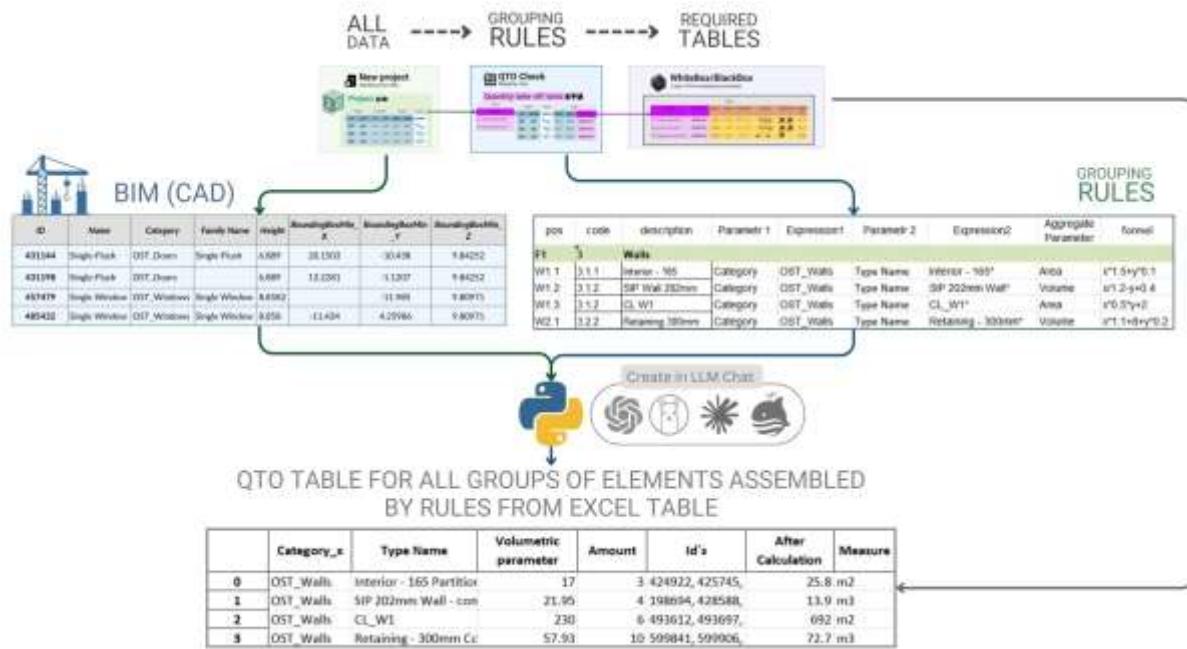
Кодты орындаудың соңғы нәтижесі (Рис. 5.213) топ-субъектілер кестесі болады, ол тек бастапқы CAD- (BIM-) моделінен алынған көлем атрибуттарын ғана емес, сонымен қатар калькуляциялар мен сметаларды дұрыс жасау үшін барлық талаптарды ескере отырып, жаңа нақты көлем атрибутын да қамтиды (Мысал Рис. 5.214).-

	Category_x	Type Name	Volumetric parameter	Amount	Id's	After Calculation	Measure
0	OST_Walls	Interior - 165 Partition	17	3 424922, 425745,		25.8 m2	
1	OST_Walls	SIP 202mm Wall - con	21.95	4 198694, 428588,		13.9 m3	
2	OST_Walls	CL_W1	230	6 493612, 493697,		692 m2	
3	OST_Walls	Retaining - 300mm Cc	57.93	10 599841, 599906,		72.7 m3	

Рис. 5.214 "Есептеуден кейін" атрибуты кодты орындағаннан кейін жинақталған кестеге қосылады, ол автоматты түрде нақты көлемді есептейді.

Алынған кодты (Сурет 5.213) жоғарыда аталған танымал IDE-лердің бірінде іске қосып, оны кез кел-

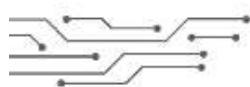
ген бар жобаларға немесе жаңа келіп түскен жобаларға (RVT, IFC, DWG, NWS, DGN және т.б.) қолдануға болады, бұл бірнеше жобалар немесе, мүмкін, әртүрлі форматтардағы жүздеген жобалар болуы мүмкін, құрылымдалған формаға келтірілген (Сурет 5.215).-



Сурет 5.215 Автоматты құрылыш деректерін топтастыру процесі BIM (CAD) деректерін QTO кестелерімен Excel электрондық кестесіндегі ережелер арқылы байланыстырады.

Параметрленген және бапталған көлемдік деректерді жинау процесі (Сурет 5.215) жобаның элементтері туралы сандық атрибуттар мен көлемдер туралы деректерді жинауды толық автоматтандыруға мүмкіндік береді, соның ішінде құнды бағалау, логистика, жұмыс кестелері және көміртек ізін есептеу мен басқа аналитикалық міндеттер.

Жобаның элементтерін белгілі бір белгілері бойынша оңай үйімдастыру және топтастыру құралдарын зерттегеннен кейін, біз енді топтастырылған және сұзілген жобаларды компанияның әртүрлі есептеулері мен бизнес-сценарийлерімен интеграциялауға дайынбыз.

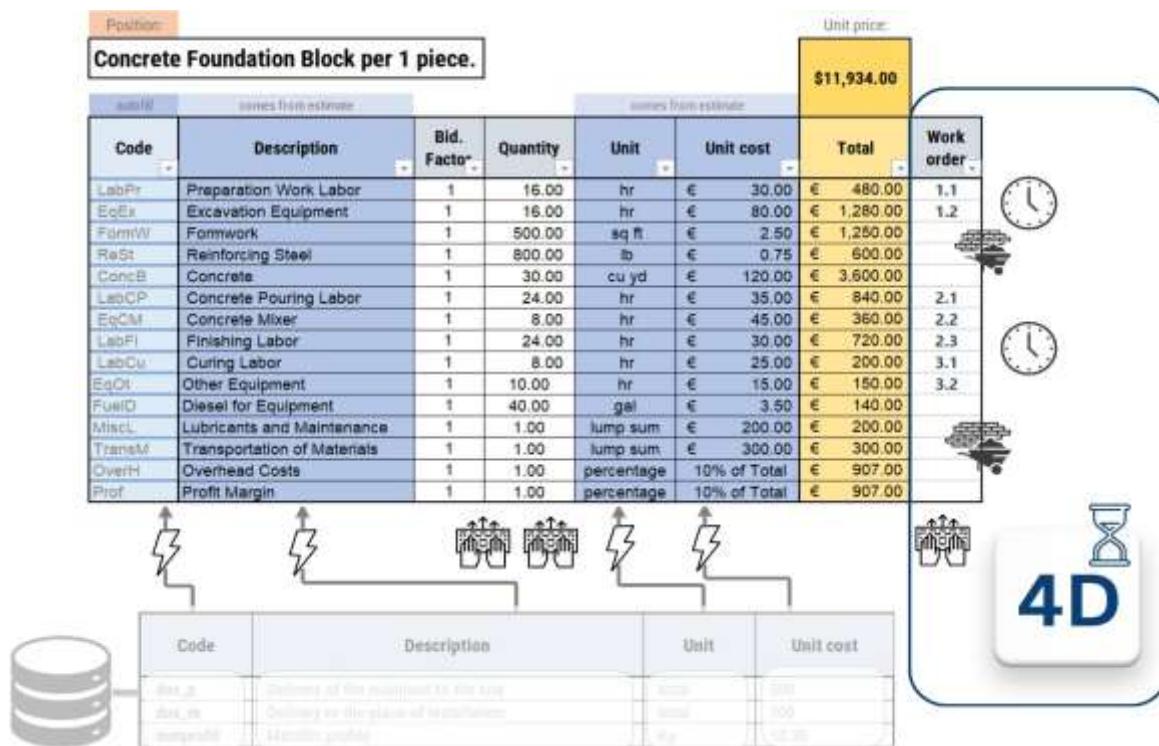


ТАРАУ 5.3. 4D, 6D-8D ЖӘНЕ СО₂ ГАЗЫНЫҢ ШЫҒАРЫЛУЫН ЕСЕПТЕУ

4D-модель: құрылыш сметаларына уақытты интеграциялау

Құнды есептеуден басқа, құрылыш саласындағы жобалық деректерді қолданудың негізгі бағыттарының бірі уақыт параметрлерін анықтау болып табылады – жеке құрылыш операциялары үшін де, бүкіл жоба үшін де. Автоматтандырылған мерзімдерді есептеу және жұмыс кестесін қалыптастыру үшін негіз ретінде жиі ресурс әдісі мен оған байланысты есептеу деректер базасы пайдаланылады, бұл «Құрылыш жобалары үшін есептеулер мен сметалар» атты алдыңғы тарауда ежей-тегжейлі қарастырылған.

Ресурс тәсілі аясында тек материалдық шығындарға емес, сонымен қатар уақыт ресурстары да ескеріледі. Калькуляцияларды құрастыру кезінде әр процесс үшін жұмыс тәртібі атрибуты (Сурет 5.31 – «Work order» параметрі) тағайындалуы мүмкін, сондай-ақ осы процесті орындаумен байланысты уақыт пен құн көрсетіледі. Бұл параметрлер, әдетте, нарықтық бағасы жоқ және тікелей сатып алуға жатпайтын операцияларды сипаттау үшін өте маңызды – мысалы, құрылыш техникасын пайдалану, жұмысшылардың жүктемесі немесе логистикалық процестер (әдетте, сағаттармен көрсетіледі). Мұндай жағдайларда құн сатып алу бөлімімен емес, тікелей орындаушы компаниямен ішкі нормативтер немесе өндірістік ставкалар негізінде анықталады (Сурет 5.31).



Сурет 5.31 Ресурс әдісіндегі жұмыстарды есептеу жұмыс уақытының шығындарын қамтиды.

Осылайша, калькуляция деңгейіндегі есептеулерге тек отын мен материалдарға (сатып алу құны) ғана емес, сонымен қатар құрылыш алаңындағы машинисттердің, техниканың және көмекші

жұмысшылардың жұмыс уақыты да енгізіледі. Келтірілген мысалда (Сурет 5.31) шығындар кестесі фундамент блоктарын орнату құнын есептеуді қамтиды, оған жұмыс кезеңдері, мысалы, дайындау, каркас орнату және бетон құю, сондай-ақ қажетті материалдар мен еңбек шығындары кіреді. Сонымен қатар, жеке операциялар, мысалы, дайындау жұмыстары, материалдық шығындары болмауы мүмкін, бірақ адам-сағаттармен көрсетілген елеулі уақыт шығындарын қамтуы мүмкін.

Құрылыш алаңындағы жұмыстардың ретін жоспарлау үшін калькуляция кестесіне қолмен мысалы «Жұмыс тапсырысы» атрибуты қосылады (Сурет 5.31). Бұл атрибут тек уақытпен (сағат, күн) өлшемнен тіптесін элементтер үшін қосымша бағанда көрсетіледі. Бұл атрибут жұмыс коды, сипаттамасы, саны, өлшем бірлігі (параметр «Бірлік») және шығындарды толықтырады. Нөмірленген реттілік (параметр «Жұмыс тапсырысы») жұмыстардың орындалу тәртібін анықтауға мүмкіндік береді және оны кесте құру кезінде пайдалануға болады.

Құрылыш графигі және оны калькуляция деректері негізінде автоматтандыру

Құрылыш кестесі – бұл жобаны жүзеге асыру аясында орындалуы тиіс жұмыстар мен процестердің жоспарының визуалды көрінісі. Ол әрбір тапсырма- жұмыстың ресурстық есептеулер негізінде жасалады (Сурет 5.31), мұнда ресурстардың құнынан басқа, уақыт пен реттілік бойынша да жазылады.

Орташа тәсілдерден айырмашылығы, мұнда уақыт есептеулері материалдар немесе жабдықтарды орнату үшін стандартты сағат санына негізделсе, ресурстық әдісте жоспарлау калькуляцияға енгізілген нақты деректерге негізделеді. Еңбек шығындарына қатысты смета позициялары жұмыс кезеңінде ресурстарды пайдалану жағдайларын ескере отырып, қолданылатын күнтізбеге сүйенеді. Калькуляция деңгейінде өнімділік коэффициенттері арқылы өнімді сағаттарды түзету (Сурет 5.31 параметр «Bid. Factor») жұмыстардың орындалу мерзімдеріне әсер ететін өнімділік айырмашылықтарын және маусымдық ерекшеліктерді ескеруге мүмкіндік береді.

Құрылыш кестесінің Гант диаграммасында процестің басталу және аяқталу даталарын анықтау үшін, біз калькуляциядағы фундаменттік блоктардың әр элементі үшін уақыт көлемі атрибутының мәндерін алғып, оларды блоктардың санына (бұл жағдайда бетон фундаменттік блоктарының санына) көбейтеміз. Бұл есептеу әр тапсырманың ұзақтығын береді. Содан кейін біз осы ұзақтықтарды жобаның басталу датасынан бастап уақыт шкаласына орналастырамыз, нәтижесінде әр тапсырманың қашан басталып, аяқталатының көрсететін визуалды көрініс аламыз. «Жұмыс тапсырысы» параметрі процестерде бізге жұмыстың параллель («Жұмыс тапсырысы» мысалы 1.1-1.1) немесе тізбектей (1.1-1.2) орындалатынын түсінуге қосымша мүмкіндік береді.

Гант диаграммасы – бұл жобаларды жоспарлау және басқару үшін графикалық құрал, ол тапсырмаларды уақыт шкаласында көлденең жолақтар түрінде көрсетеді. Әр жолақ тапсырманың орындалу ұзақтығын, оның басталуы мен аяқталуын көрсетеді.

Жұмыс кестесі немесе Гант диаграммасы жобаның жетекшілері мен жұмысшыларына құрылыш кезеңдерінің қашан және қандай реттілікпен орындалуы тиіс екенін нақты түсінуге көмектеседі, ресурстарды тиімді пайдалануды және мерзімдерді сақтауды қамтамасыз етеді.

Уш бетон фундаменттік блокты орнату бойынша жұмыс күнтізбелік жоспарын жоғарыдағы калькуляциядан есептеулерді пайдалана отырып, 2024 жылдың 1 мамырында орнатуды жоспарлауды

сұраймыз.

Калькуляцияны LLM-ге жіберу үшін, біз калькуляция кестесін XLSX форматында жүктей аламыз немесе калькуляцияның суретін JPEG форматында LLM чатына тікелей енгізе аламыз (Сурет 5.31). LLM сурет кестесін визуализациялау үшін кітапхананы өзі табады және кестедегі жұмыстардың уақыт атрибуттарын олардың көлеміне көбейтіп, барлық деректерді кестеге жинайды. -

- ❷ LLM-ге мәтіндік сұраныс жіберініз:

3 фундаменттік блокты орнату үшін уақытқа қатысты кестеден алынған мәндерді пайдалана отырып, Гант диаграммасы түрінде жұмыс графигін жасаңыз (Рис. 5.31 JPEG форматында қосымшада). Эр блок бойынша жұмыстар тізбектей орындалады. Жұмыстарды 2024 жылдың 1 мамырынан бастау керек.

- ❸ LLM-нің жауабы:

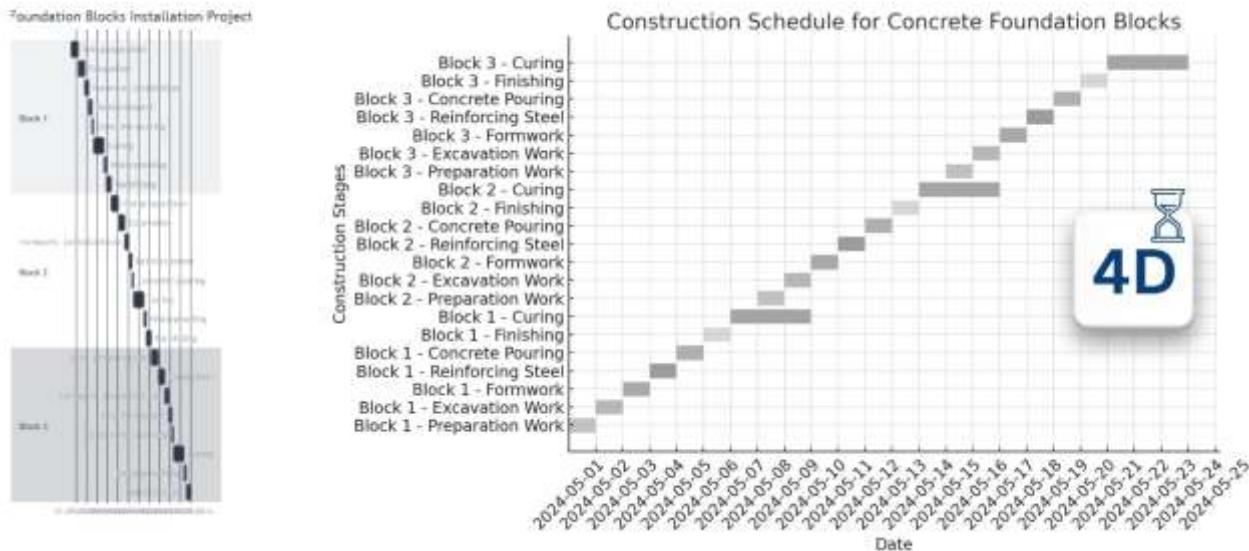


Рис. 5.32 бірнеше LLM арқылы автоматты түрде жасалған Гант диаграммасы, промпт шарттарына сәйкес үш бетон блоктың құрылышы кезеңдерін көрсетеді.

Алынған график (Рис. 5.32) уақыт диаграммасын білдіреді, мұнда әрбір көлденең жолақ фундаменттік блок бойынша жұмыстардың белгілі бір кезеңіне сәйкес келеді және операциялардың тізбегін (Work order параметрі) көрсетеді, мысалы, дайындық, жер жұмыстары, опалубка орнату, арматуралау, бетон құю және соңғы өндөу, яғни есептеулерде уақыт параметрлері мен тізбегі толтырылған процестер.

Мұндай график (Рис. 5.32) жұмыс күндері, ауысымдар немесе жұмыс уақыты нормаларына байланысты шектеулерді ескермейді, тек процесті концептуалды визуализациялау үшін арналған. Нақты график, жұмыстардың параллельдігін көрсететін, сәйкес промпттар немесе чат ішіндегі қосымша нұсқаулармен толықтырылуы мүмкін.

3D геометрияның көлемдеріне негізделген бір шығын калькуляциясын (Рис. 5.31) пайдалана отырып, автоматтандырылған сметалар арқылы жобаның құнын бағалауға және бір уақытта әртүрлі

жобалық нұсқалар үшін уақыт сипаттамаларын кестелер немесе графиктер түрінде есептеуге болады (Рис. 5.33).-

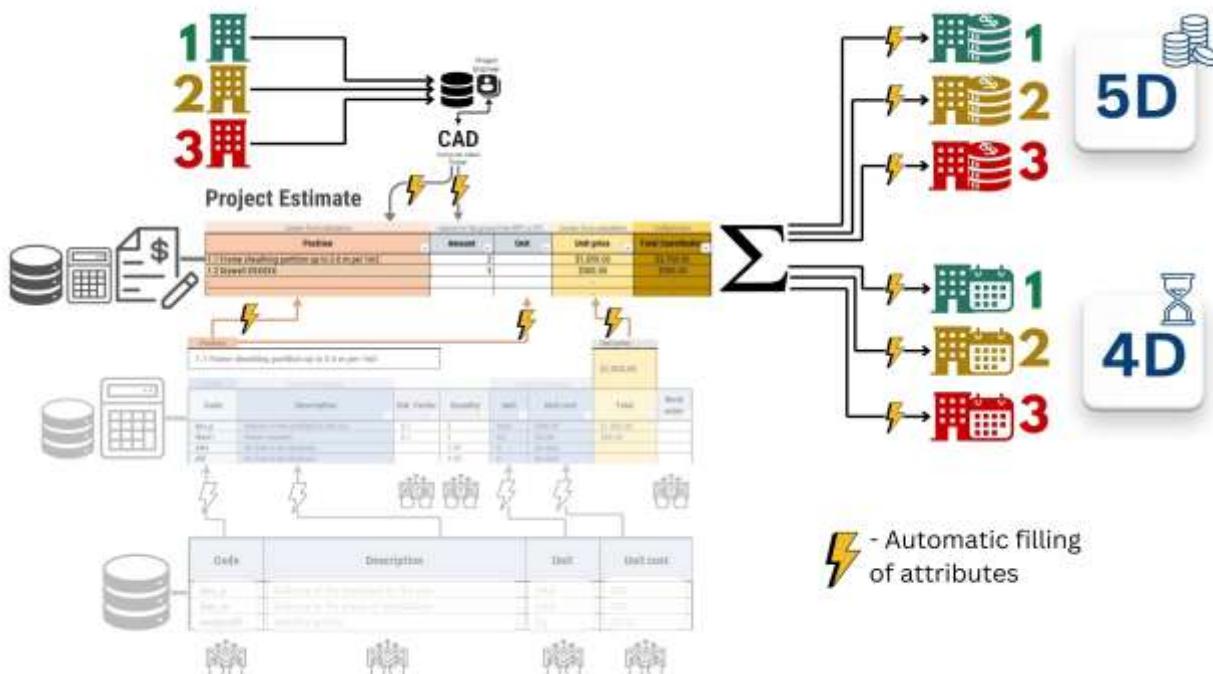


Рис. 5.33 Автоматты есептеу, әртүрлі жобалық нұсқалар үшін шығындар мен уақытты бірден және автоматты түрде болжауға мүмкіндік береді.

Заманауи модульдік ERP жүйелері (Рис. 5.44) CAD модельдерінен деректерді жүктей отырып, уақытты есептеудің осындағы автоматтандырылған әдістерін пайдаланады, бұл шешім қабылдау процесін едәуір қысқартады. Бұл нақты бағаларды ескере отырып, жұмыс графиктерін дәл жоспарлауға және жобаны жүзеге асыру үшін барлық тапсырмаларды орындауға қажетті жалпы уақытты есептеуге мүмкіндік береді.-

6D-8D кеңейтілген атрибуттық қабаттары: энергия тиімділігінен қауіпсіздікті қамтамасыз етуге дейін

6D, 7D және 8D – бұл ақпараттық моделдеудің кеңейтілген деңгейлері, олардың әрқайсысы жобаның кешенді ақпараттық моделіне қосымша атрибуттар қабаттарын енгізеді, олардың негізі 3D модельдің көлемі мен санына негізделген атрибуттар болып табылады. Әрбір қосымша деңгей, мысалы, жылжымайтын мүлкіті басқару жүйелері (PMS), объектілерді автоматтандырылған басқару (CAFM), құрылыш жобаларын басқару (CPM) және қауіпсіздік басқару жүйелері (SMS) сияқты басқа жүйелерде кейінгі топтастыру немесе идентификациялау үшін қажетті арнайы параметрлерді енгізеді.



Рис. 5.34 6D, 7D және 8D атрибуттары деректердің ақпараттық моделінде жобаның әртүрлі аспекттерін, энергия тиімділігінен бастап қауіпсіздікке дейін, қарастыруды кеңейтеді.

- 6D-де жобаның деректер базасына (немесе датафреймге (Рис. 4.113)) геометриялық және көлемдік атрибуттармен қатар экологиялық тұрақтылық туралы ақпарат (атрибуттар бағандар) қосылады. Бұл энергия тиімділігі, көміртекі іздері, материалдарды қайта өңдеу мүмкіндігі және экологиялық таза технологияларды қолдану сияқты ақпаратты қамтиды. Бұл деректер жобаның қоршаған ортаға әсерін бағалауға, жобалық шешімдерді оңтайландыруға және тұрақты даму мақсаттарына (ESG) қол жеткізуге мүмкіндік береді.-
- 7D атрибуттары ғимаратты пайдалану басқару үшін қажетті атрибуттармен толықтырылады. Бұл техникалық қызмет көрсету кестелері, компоненттердің қызмет ету мерзімі, техникалық құжаттама және жөндеу тарихы туралы деректер. Мұндай ақпарат жиынтығы модельді эксплуатация жүйелерімен (CAFM, AMS) интеграциялауға мүмкіндік береді, қызмет көрсетуді, жабдықты ауыстыруды тиімді жоспарлауға және обьектінің өмірлік циклы бойында қолдауды қамтамасыз етеді.
- 8D қосымша атрибуттің қабат - құрылымың кезеңінде де, кейінгі пайдалану кезінде де қауіпсіздікпен байланысты ақпаратты қамтиды. Модельге персоналдың қауіпсіздігін қамтамасыз ету шаралары, төтенше жағдайларда әрекет ету нұсқаулықтары, эвакуация жүйелеріне және өрт қауіпсіздігіне қойылатын талаптар енгізіледі. Бұл деректердің цифрлық модельге интеграциясы тәуекелдерді алдын ала ескеруге және еңбек қорғау мен қауіпсіздік талаптарын ескере отырып, архитектуралық, инженерлік және үйімдастырушылық шешімдерді әзірлеуге көмектеседі.

4D-дан 8D-ға дейінгі қабаттар құрылымдық кесте түрінде қосымша атрибуттар ретінде, 3D-модельдің толтырылған атрибуттарына, мысалы, атауы, категориясы, түрі және көлемдік сипаттамаларына қосылған толтырылған мәндер ретінде ұсынылады (Рис. 5.35). 6D, 7D және 8D атрибуттің қабаттардағы мәндер қайта өңдеу пайызы, көміртек ізі, кепілдік мерзімі, ауыстыру циклі, орнату күні, қауіпсіздік протоколдары және т.б. сияқты қосымша мәтіндік және сандық деректерді қамтиды.-

3D:

ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
W-NEW	Window	120 cm	-	90%	1622 kgCO ₂ e	8 years	20 years	Annual	mon	XYZ Windows	ISO 45001

6D:

ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
W-OLD1	Window	100 cm	140 cm	90%	1522 kgCO ₂ e	8 years	15 years	Biannual	08/22/2024	XYZ Windows	OSHA Standard

7D:

ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
W-OLD2	Window	110 cm	160 cm	90%	1522 kgCO ₂ e	-	15 years	Biannual	08/24/2024	???	OSHA Standard

8D:

ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
D-122	Door	90 cm	210 cm	100%	1322 kgCO ₂ e	15 years	25 years	Biennial	08/25/2024	Doors Ltd.	OSHA Standard

Рис. 5.35 6D-8D атрибуттің қабаттарды геометриялық және көлемдік атрибуттарды қамтитын ақпараттық деректер моделіне қосады, ол 3D-модельден алынған.

Біздің жаңа тереземіз үшін (Рис. 4.41) W-NEW идентификаторы бар элемент (Рис. 5.35) келесі 3D-8D атрибуттарына ие болуы мүмкін: –

3D-атрибуттар – CAD жүйелерінен алынған геометриялық ақпарат:

- "Тип атауы" - "Терезе" элементі
- "Ені" - 120 см
- Элементтің "Bounding Box" нұктелерін немесе оның "геометриясын BREP / MESH" жеке атрибут ретінде қосуға болады.

6D атрибуттары - экологиялық тұрақтылық:

- "Қайта өндеу көрсеткіші" - 90%
- "Көміртек ізі" - 1622 кг CO₂

7D атрибуттары - обьекттің басқару деректері:

- "Кепілдік мерзімі" - 8 жыл
- "Ауыстыру циклі" - 20 жыл
- "Техникалық қызмет көрсету" - жыл сайын қажет

8D атрибуттары - ғимараттарды қауіпсіз пайдалану мен эксплуатацияны қамтамасыз ету:

- "Терезе" "Орнатылды" - "XYZ Windows" компаниясымен
- "Қауіпсіздік стандарты" - ISO 45001 талаптарына сәйкес

Дерекқорға немесе деректер жиынына (Рис. 5.35) жазылған барлық параметрлер әртүрлі бөлімдердегі мамандарға топтастыру, іздеу немесе есептеулер үшін қажет. Мұндай көп өлшемді обьектілерді атрибуттар негізінде сипаттау жобаның өмірлік циклі, эксплуатациялық талаптары және жобалау, құрылыш және эксплуатация кезінде қажетті көптеген басқа аспектілер туралы толық түсінік алуға мүмкіндік береді.

СО₂ бағалау және құрылымың жобаларындағы көмірқышқыл газының шығарындыларын есептеу

6D кезеңіндегі құрылымың жобаларындағы тұрақтылығы тақырыбымен қатар, қазіргі құрылым саласында жобалардың экологиялық тұрақтылығына ерекше назар аударылуда, мұнда негізгі аспектілердің бірі жобаның өмірлік цикл кезеңдерінде (мысалы, өндіріс және монтаждау кезінде) пайда болатын көмірқышқыл газы CO₂ шығарындыларын бағалау және азайту болып табылады.-

Құрылым материалдарының көміртегі шығарындыларын бағалау және есептеу – бұл жобада қолданылатын элементтердің немесе элементтер тобының көлемдік атрибуттарын тиісті шығарындылар коэффициентіне көбейту арқылы жалпы көміртегі шығарындыларын анықтау процесі.

Құрылым жобаларын бағалау кезінде көміртегі шығарындыларын ескеру, экологиялық, әлеуметтік және басқарушылық (ESG) критерийлерінің кең ауқымының бір бөлігі ретінде кешенді талдауға жаңа деңгей қосады. Бұл, әсіресе, тапсырыс беруші-инвестор үшін LEED® (Энергия мен экологияны басқарудағы көшбасшылық), BREEAM® (Құрылым зерттеу үйімінің экологиялық бағалау әдісі) немесе DGNB® (Неміс тұрақты құрылым қоғамы) сияқты тиісті сертификаттарды алу кезінде маңызды. Бұл сертификаттардың бірін алу объектінің нарықтық тартымдылығын едәуір арттырып, пайдалануға енгізуі жеңілдетеді және тұрақтылыққа (ESG) бағытталған жалға алушылардың талаптарына сәйкестікті қамтамасыз етеді. Жобаның талаптарына байланысты HQE (Haute Qualité Environnementale, француз экологиялық құрылым стандарты), WELL (WELL Building Standard, пайдаланушылардың денсаулығы мен жайлышының бағытталған) және GRESB (Global Real Estate Sustainability Benchmark, халықаралық жылжымайтын мүлік тұрақтылығы рейтингі) сияқты стандарттар да қолданылуы мүмкін.

Экологиялық, әлеуметтік және басқарушылық ESG (экологиялық, әлеуметтік және басқарушылық) – бұл компанияның ішіндегі және сыртындағы бизнес әсерін бағалау үшін қолданылатын кең принциптер жиынтығы.

ESG, 2000-шы жылдардың басында қаржы қорлары тарапынан экологиялық, әлеуметтік және басқарушылық критерийлер туралы инвесторларға ақпарат беру үшін әзірленген, компаниялар мен жобаларды, соның ішінде құрылымы жобаларын бағалаудың негізгі көрсеткішіне айналды. Ең ірі консалтингтік компаниялардың зерттеулеріне сәйкес, экологиялық, әлеуметтік және басқарушылық факторларды (ESG) ескеру құрылымың саласының ажырамас бөлігіне айналуда.

EY (2023) «Көміртегі нейтралдығына жол» есебіне сәйкес, ESG принциптерін белсенді түрде енгізетін компаниялар ұзақ мерзімді тәуекелдерді төмендетіп қана қоймай, сонымен қатар бизнес модельдерінің тиімділігін арттырады, бұл жаһандық нарықтардың трансформациясы жағдайында ерекше маңызды [103]. PwC «ESG туралы хабардарлық» есебінде компаниялардың ESG факторларының маңыздылығы туралы хабардарлық деңгейі 67%-дан 97%-ға дейін өзгеретінін, ал үйымдардың көпшілігі бұл тенденцияларды болашақтағы тұрақты даму үшін кілтті деп санайтынын атап етеді [104] және бизнестің көпшілігі ESG принциптерін интеграциялау бойынша мүдделі тараپтар тарапынан едәуір қысым сезінетінін көрсетеді.

Осылайша, құрылымың жобаларына ESG принциптерін интеграциялау тек LEED, BREEAM, DGNB сияқты халықаралық тұрақтылық сертификаттарын алуға ғана емес, сонымен қатар компаниялардың саладағы ұзақ мерзімді тұрақтылығы мен бәсекеге қабілеттілігін қамтамасыз етуге де ықпал етеді.

Құрылымың жобасының жалпы көміртек ізіне әсер ететін ең маңызды факторлардың бірі - құрылымы материалдары мен элементтерінің өндірісі мен логистикасының кезеңдері. Объектте қолданылатын материалдар, әдетте, CO₂ шығарындыларына шешуші әсер етеді, әсіресе жобаның өмірлік циклінің бастапқы кезеңдерінде - шикізатты өндіруден бастап құрылымы алаңына жеткізуге дейін.

Құрылымы элементтерінің санаттары немесе түрлері бойынша шығарындыларды есептеу үшін әртүрлі материалдарды өндіру нәтижесінде пайда болатын CO₂ мөлшерін көрсететін сілтеме коэффициенттерін пайдалану қажет. Мұндай материалдарға бетон, кірпіш, қайта өнделген болат, алюминий және басқалары жатады. Бұл мәндер, әдетте, UK ICE 2015 (Көміртек және энергияның инвентаризациясы) және US EPA 2006 (АҚШ қоршаған ортаны қорғау агенттігі) сияқты беделді көздер мен халықаралық деректер базаларынан алынады. Келесі кестеде (Рис. 5.36) бірнеше кең таралған құрылымы материалдары үшін базалық шығарындылар коэффициенттері келтірілген. Эрбір материал үшін екі негізгі параметр көрсетілген: CO₂-нің нақты шығарындылары (килограмм материалға) және көлемді массасына айналдыру коэффициенттері (куб метрге килограмм), оларды жобалық модельге интеграциялау және QTO деректерін топтастыруға байланысты.



Carbon Emitted in Production		UK ICE Database (2015) USEPA (2006)	UK ICE Database (2015) USEPA (2006)	Coefficient m³ to kg
Material	Abbreviated	Process Emissions (kg CO2e/ kg of product) (K1)	Process Emissions (kg CO2e/ kg of product) (K2)	Kg / m³ (K3)
Concrete	Concrete	0.12	0.12	2400
Concrete block	Concrete_block	0.13**	0.14	2000
Brick	Brick	0.24	0.32	2000
Medium density fiberboard (MDF)	MDF	0.39*	0.32	700
Recycled steel (avg recy content)	Recycled_steel	0.47	0.81	7850
Glass (not including primary mfg.)	Glass	0.59	0.6	2500
Cement (Portland, masonry)	Cement	0.95	0.97	1440
Aluminum (virgin)	Aluminum	12.79	16.6	2700

Рис. 5.36 Әртүрлі құрылымы материалдарын өндіру кезінде бөлінетін көміртек мөлшері, UK ICE және US EPA деректер базасына сәйкес.

Жобаның жалпы CO₂ шығарындыларын есептеу үшін, 4D және 5D калькуляциялары сияқты, әрбір объектілер тобының атрибуттарының көлемін анықтау қажет. Бұл көлемдерді QTO құралдарын пайдалана отырып, куб метрмен алуға болады, бұл туралы Quantity take-off бөлімінде егжейтегжейлі қарастырылған. Содан кейін алынған көлемдер "CO₂ технологиялық шығарындылары" атрибуты үшін әрбір материалдар тобының сәйкес коэффициенттеріне көбейтіледі.

- ❷ CAD (BIM) жобасынан элементтердің түрлері бойынша көлемдер кестесін автоматты түрде шығарып, барлық жобалық деректерді топтастырайық, бұл алдыңғы тарауларда жасалғандай. Бұл тапсырманы орындау үшін LLM-ге жүгінеміз.

CAD (BIM) жобасындағы DataFrame кестесін "Object Name" (немесе "Type") бағанының параметрі бойынша топтастырып, әр топтағы элементтердің санын көрсетіп, сондай-ақ типтегі барлық элементтердің "Volume" параметрін жинақтауды сұрай.

2 LLM-нің жауабы:

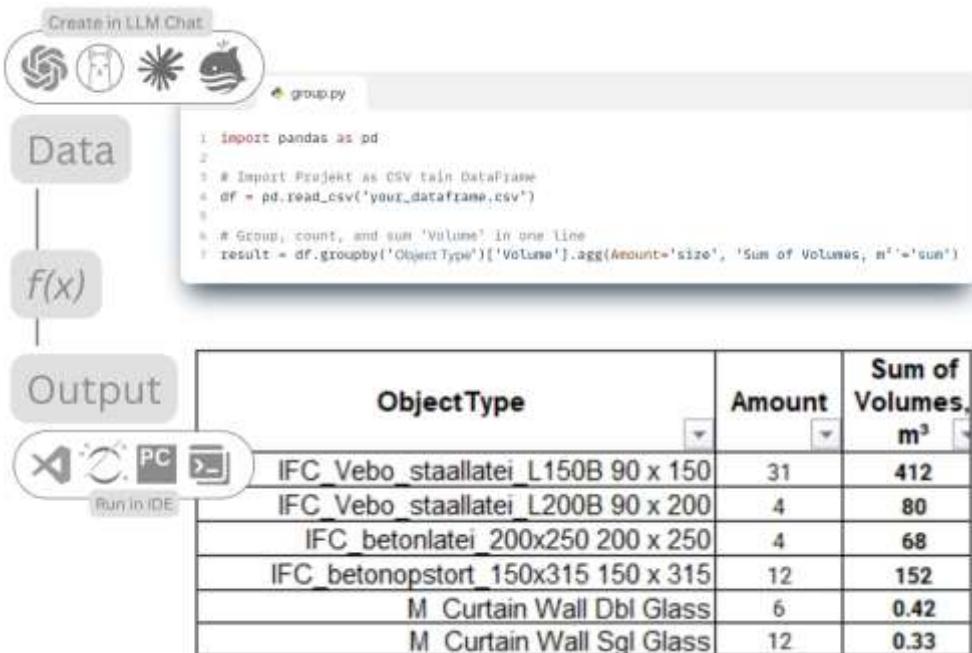


Рис. 5.37 LLM-да сгенерлен код бізге жобаның объектілерін типі (ObjectType) бойынша топтастырып, "Volume" атрибутын жинақтады.

Жобаның жалпы CO₂ шығарындыларын автоматты түрде есептеу үшін, кестедегі деректерді автоматты түрде сәйкестендіруді орнату немесе элементтердің түрлерін (Рис. 5.37) шығарындылар коэффициенттері кестесіндегі (Рис. 5.36) сәйкес материалдар түрлерімен қолмен байланыстыру жеткілікті. Шығарындылар коэффициенттері мен формулалары, сондай-ақ CAD (BIM) форматтарынан көлемдерді алу және CO₂ анықтау автоматизациясы үшін код GitHub-тан "CO₂_calculating-the-embodied-carbon_DataDrivenConstruction" сұрауымен табуға болады.--

Осылайша, CAD деректер базасынан QTO элементтерін топтастырғаннан кейін деректерді интеграциялау CO₂ шығарындыларын (Рис. 5.38) әртүрлі жобалау нұсқалары үшін автоматты түрде есептеуге мүмкіндік береді. Бұл әртүрлі материалдардың әртүрлі нұсқалардағы әсерін талдауға және тапсырыс берушінің CO₂ шығарындылары деңгейіне қойылатын талаптарға сәйкес келетін шешімдерді таңдауға мүмкіндік береді, ғимаратты пайдалануға беру кезінде тиісті сертификат алу үшін.-

CO₂ шығарындыларын бағалау құрылым компаниясының ESG рейтингін (мысалы, LEED сертификаттау) алу процесінде жобаның топталған элементтерінің көлемдерін коэффициенттерге көбейту арқылы жүзеге асырылады.

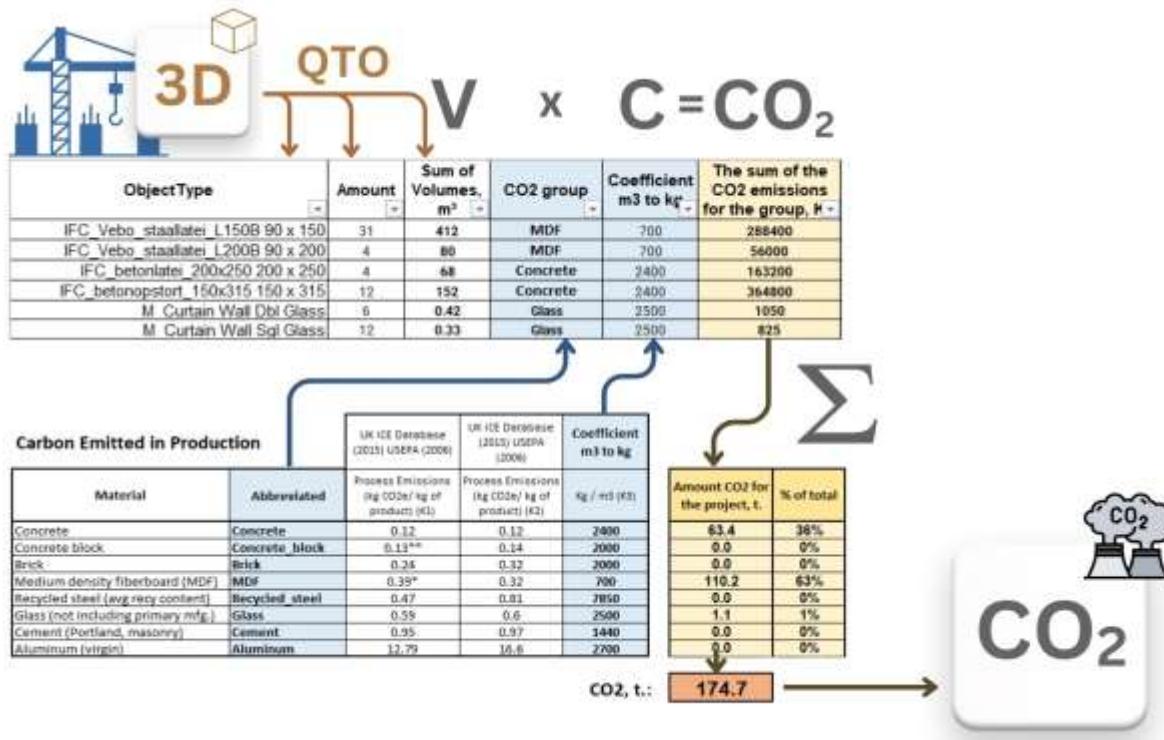


Рис. 5.38 CAD деректер базасынан QTO топтарын интеграциялау CO₂ шығарындыларының соңғы көлемдерін бағалауда дәлдік пен автоматтандыруды қамтамасыз етеді.

Элементтер топтарының көлемдерін анықтау арқылы біз материалдарды бақылау және логистика, мониторинг және сапаны басқару, энергия тұтынуды модельдеу және талдау, сондай-ақ жобаның жеке элементтері мен жалпы жобаның жаңа атрибутивтік статусын (кестедегі параметр) алу үшін көптеген басқа міндеттерді орындауға мүмкіндік аламыз.

Егер компаниядағы осындағы есептеу процестерінің саны арта бастаса, автоматтандыру қажеттілігі және есептеу нәтижелерін компанияның процестеріне және деректерді басқару жүйелеріне енгізу мәселесі туындаиды.

Кешенді шешімдердің күрделілігіне байланысты, құрылымында жұмыс істейтін орта және ірі компаниялар осындағы автоматтандыруды ERP (немесе PMIS) жүйелерін әзірлеуші компанияларға аутсорсингке береді. Әзірлеуші компаниялар ірі клиенттер үшін материалдар мен ресурстарды есептеу сияқты көптеген ақпараттық қабаттарды басқару үшін біртұтас кешенді модульдік жүйе жасайды.



ТАРАУ 5.4. ҚҰРЫЛЫС ERP ЖӘНЕ PMIS ЖҮЙЕЛЕРИ

Құрылыш ERP жүйелері есептеулер мен сметалар мысалында

Модульдік ERP жүйелері әртүрлі атрибутивтік (ақпараттық) қабаттарды және деректер ағындарын біртұтас кешенді жүйеге біріктіреді, бұл жобаларды басқарушыларға ресурстарды, қаржыны, логистиканы және жобаның басқа аспекттерін бір платформа аясында синхронды түрде басқаруға мүмкіндік береді. Құрылыш ERP жүйесі құрылыш жобаларының "мины" рөлін атқарады, автоматтандыру арқылы қайталараптың процестерді жөнілдетіп, құрылыш процесінің барлық кезеңінде ашықтық пен бақылауды қамтамасыз етеді.

Құрылыш ERP жүйелері (Enterprise Resource Planning) құрылыш процесінің әртүрлі аспекттерін басқару және оңтайландыру үшін арналған кешенді бағдарламалық шешімдер болып табылады. Құрылыш ERP жүйелерінің негізінде шығындарды есептеу және жұмыс кестелерін құру модульдері жатыр, бұл оларды ресурстарды тиімді жоспарлаудың маңызды құралы етеді.

ERP жүйесінің модульдері пайдаланушыларға деректерді енгізуге, өңдеуге және талдауға мүмкіндік береді, бұл жобаның әртүрлі аспекттерін құрылымдық түрде қамтиды, олар материалдық және еңбек шығындарын, жабдықты пайдалануды, логистиканы, кадрлық ресурстарды, байланыстарды және басқа құрылыш қызметтерін басқаруды қамтуы мүмкін.

Жүйенің функционалдық блоктарының бірі бизнес логикасын автоматтандыру модулі – BlackBox/WhiteBox болып табылады, ол процестерді басқарудың орталығы рөлін атқарады.

BlackBox/WhiteBox ERP жүйесін пайдаланатын мамандарға, алдын ала басқа пайдаланушылар немесе әкімшілер конфигурациялаған бизнес аспекттерін икемді түрде басқаруға мүмкіндік береді. ERP жүйелерінің контекстінде BlackBox және WhiteBox терминдері жүйенің ішкі логикасының мөлдірлігі мен бақылау деңгейлерін білдіреді:

- BlackBox («қара жәшік») – пайдалануши жүйемен интерфейс арқылы өзара әрекеттеседі, процестердің ішкі логикасына қол жеткізбей. Жүйе алдын ала белгіленген, соңғы пайдаланушыдан жасырын ережелер негізінде есептеулер жүргізеді. Пайдалануши деректерді енгізеді және нәтиже алады, ішкі есептеулерде қандай атрибуттар немесе коэффициенттер қолданылғанын білмейді.
- Ақ жәшігі («ақ қорап») - процестердің қисынын қараша, конфигурациялау және өзгертулер үшін қол жетімді. Жетілдірілген пайдаланушылар, әкімшілер немесе интеграторлар деректерді өңдеудің алгоритмдерін, есептеу ережелері мен өзара әрекеттесу сценарийлерін қолмен орната алады.

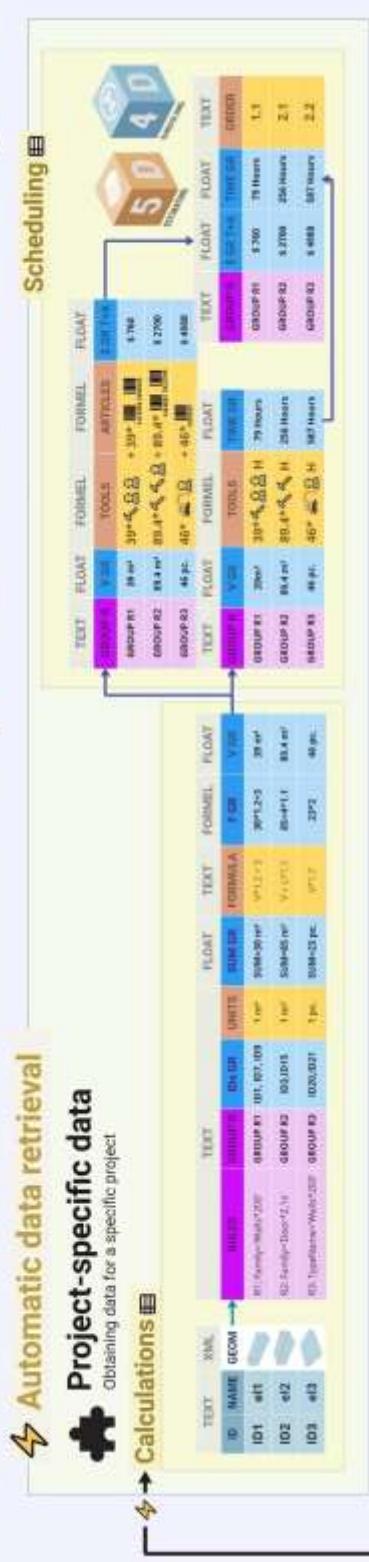
Enterprise Resource Planning ERP

↳ Automatic data retrieval

Project-specific data

Obtaining data for a specific project

↳ Calculations



WhiteBox/BlackBox



Logic of the company's processes

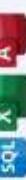


SQL

X

A

Logic of the company's processes



WhiteBox

BlackBox

SQL

X

A

WhiteBox

BlackBox

SQL

Мысалы, тәжірибелі пайдаланушы немесе әкімші ережені орнататын жағдай: сметада қандай атрибуттарды бір-біріне көбейту керек немесе белгілі бір функция бойынша топтастырылған, сонымен қатар түпкілікті нәтиже жазылуы керек. Кейіннен, инженерлер, мысалы, инженерлер, мысалы, ERP-ге жаңа деректерді пайдаланушы интерфейсі арқылы жүктейді және дайын бағалар, графикалар немесе техникалық сипаттамаларды жазасыз немесе логиканың техникалық мәліметтерін түсінбестен алады.

Алдыңғы тарауларда есептеулер мен логика модульдері LLM-мен өзара әрекеттесу түрғысынан қаралды. ERP жүйелерінің ішінде мұндай есептеулер мен түрлендірулер түймелер мен формалар интерфейсінің артында жасырылған модульдер ішінде пайдаланылады.

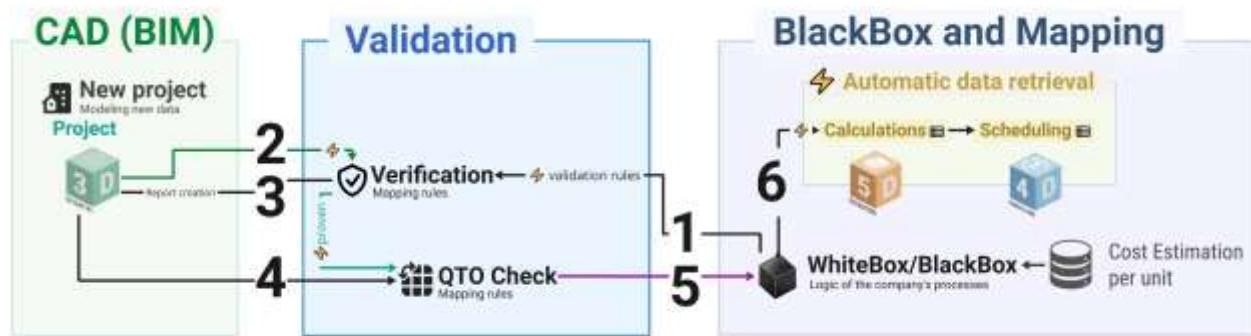
Келесі мысалда (5.41-сурет), Blackbox / WhiteBox модулі ERP жүйесінің әкімшісі Qto топтағы сметалардың төлсипаттарын салыстыру үшін ережелерді орнатыңыз. Осы теңшелген (Менеджер немесе әкімші), Blackbox / WhiteBox модулінің пайдаланушысы (бағалаушы немесе инженер) ERP пайдаланушы интерфейсі арқылы сан немесе көлемнің атрибутын қолмен қосып, дайын сметалар мен жұмыс графикін автоматты түрде қосады. Осылайша, есептеу процестері және алдыңғы тарауларда, ERP-ді қолданып, алдын-ала тарауларда ескеріліп, еденге бұрылады, автоматтандырылған конвейер.

Бұл жартылай автоматтандырылған процесті CAD (BIM) модельдерінен, мысалы, ARP жобасын көзделетін ERP модуліне қосу, мысалы, ARP модуліне тиғизетін синхрондалған механизмге және жобалау кезеңінде кез-келген жобаның құнын, біз ERP-дегі CAD моделін жүктей аламыз.-

Автоматтандырылған деректер ағынын құру үшін (5.42-сурет) CAD (BIM) және ERP жүйелері арасында «Талаптар және деректер сапасының сапасы» тарауындағы негізгі процестер мен деректердің негізгі процестері мен талаптарын құрылымдау қажет. ERP-дегі бұл процесс үқсас кезеңдерге бөлінеді:-

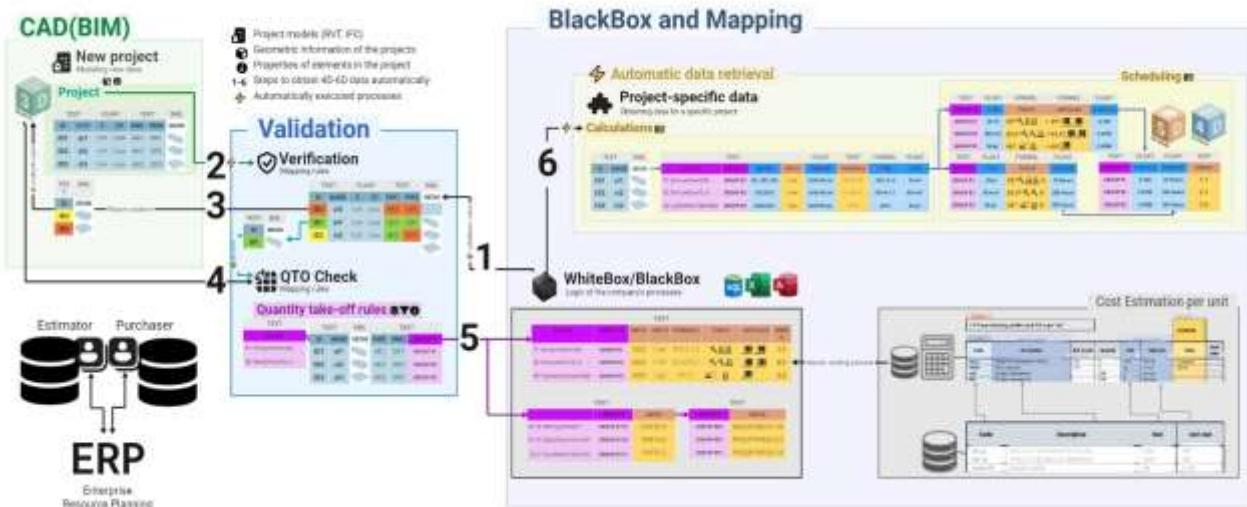
- ERP жүйесіне енудің дәлдігін қамтамасыз етуде маңызды рөл атқаратын тексеру ережелерін (1) құру (1). Тексеру ережелері жүйеге қойылатын талаптарды тек осы элементтерді, тек осы элементтерді тексеруге мүмкіндік беретін сұзгілер ретінде қызмет етеді. «Талаптар құру және мәліметтер сапасын тексеру және тексеру» тарауындағы тексеру және тексеру туралы толығырақ оқыныз.
- Содан кейін, ERP ішіне тексеру процесі (2) пайдаланылады, бұл жобаның барлық элементтері олардың атрибуттары мен құндылықтары дұрыс жасалғанын растайды және келесі өңдеу сатыларына дайын.
- Егер толық емес атрибутивті деректермен проблемалар туындаса, есеп (3), ал жоба түзету нұсқауларымен бірге келесі итерацияға дайын болғанға дейін жіберіледі.
- Жобаның деректері расталғаннан кейін, олар бұрын құрылған ережелерге сәйкес (ақ жәшіг / қара жәшіг) бойынша, субъектілер мен ресурстарға арналған сандардың атрибуттарын құруға арналған басқа ERP (4) модульде қолданылады.
- Салыстыру ережелері бойынша топтастырылған мәліметтер немесе QTO есептеулермен автоматты түрде интеграцияланған (мысалы, шығындар мен уақыт) (5).
- ERP жүйесінің соғы кезеңінде пайдаланушы, QTO кестесінен процесс кестелерінен (мысалы, есептелген мақалалар) өзгеріп отырған пайдаланушы әрбір субъектілердің әр

тобы үшін және тұтастай алғанда жобаның нәтижелерін автоматты түрде жасайды.

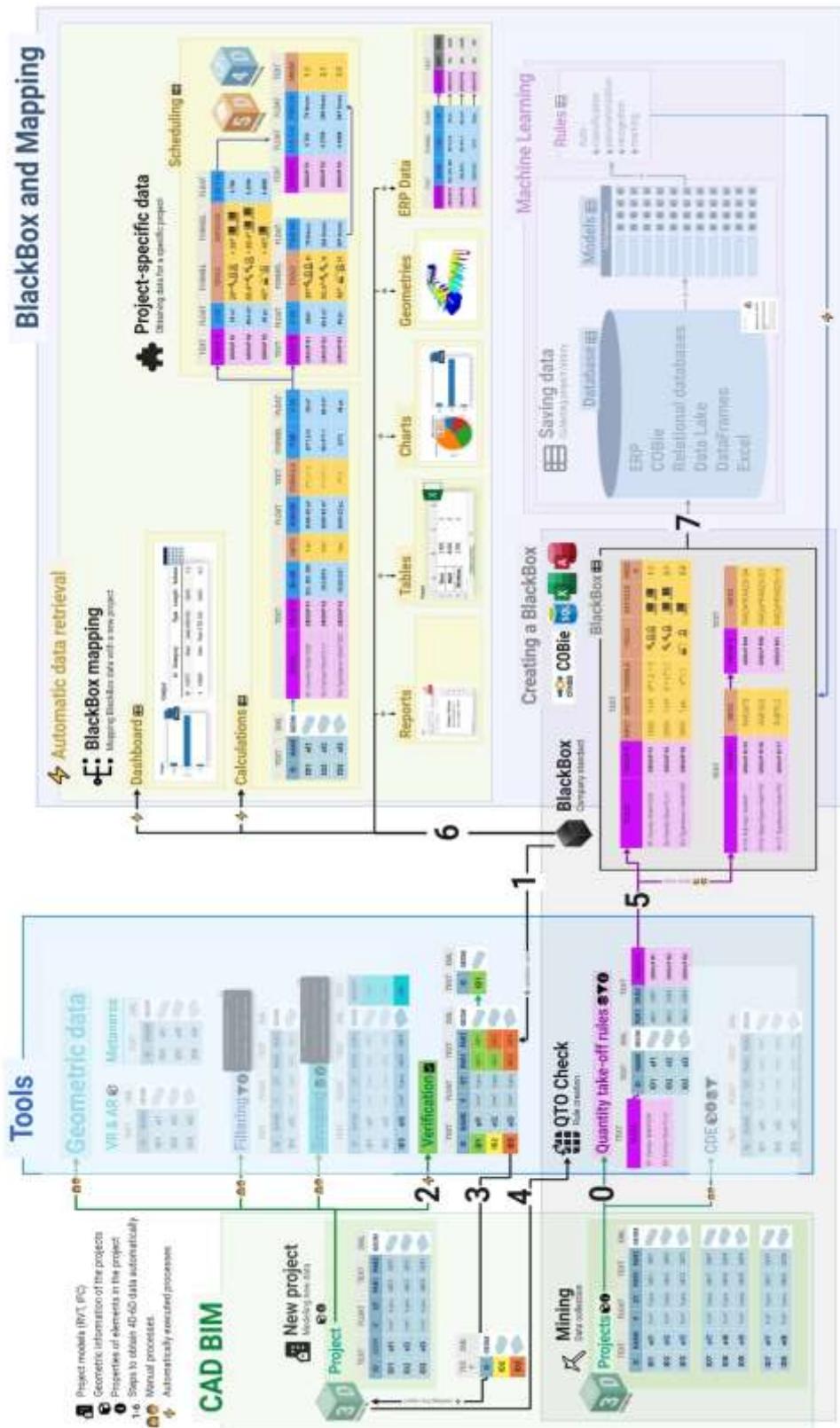


Күріш. 5.42 Валдау ережелерін (1) құра отырып, CAD (BIM) жүйесінің сәuletі (1) жұмыстың өзіндік құны мен кестелерін автоматты есептеуге дейін (5-6).

Модульдік ERP жүйесінде процестер пайдаланушы интерфейсі бар бағдарламалық жасақтама көмегімен біріктірілген. Интерфейстің артында ішкі бөлік, онда құрылымдалған кестелер деректерді өңдейді, менеджер немесе әкімші алдын-ала іске асырылатын түрлі әрекеттерді жүзеге асырады. Нәтижесінде, пайдаланушы алдын-ала орнатылған және конфигурацияланған автоматика логикасының арқасында (қара жәшігіндегі / ақ жәшігіндегі модульдерде), өз міндеттеріне сәйкес жартылай дайындалған құжаттарды алады.



Күріш. 5.43 ERP жүйесі менеджерлер мен пайдаланушыларға жаңа мәліметтер алу үшін мамандар кестелерінің арасында жылжиды.



Күріш. 5.44 ERP жүйесі аналитикалық құралдармен біріктірілген және компаниядағы шешім қабылдау процесін автоматтандырады.

Сол сияқты, ERP жүйелеріндегі процестер, басынан бастап - түпкілікті есептеуге дейін (5-суреттің 1-6 кезеңі) - бұл өзара байланысты қадамдар тізбегі, бұл, сайып келгенде, ашықтықты, тиімділігін және жоспарлаудың дәлдігін қамтамасыз етеді.-

Қазіргі заманғы құрылыш жүйелеріне тек құндылықтар мен терминдерді есептеу модульдері ғана емес, сонымен қатар, алдын-ала жасалған басқа модульдер, сонымен қатар, жұмыс істеу функцияларын, келісімшарттарды, жеткізілім тізбегін және логистиканы басқару, сонымен қатар басқа бизнес-жүйелермен және платформалармен интеграциялау. ERP интеграцияланған аналитикалық құралдар пайдаланушыларға жобаның негізгі көрсеткіштерін (KPI - KPI - негізгі көрсеткіштері) бақылау үшін ақпараттық панельдер құруды автоматтандыруға мүмкіндік береді. Бұл құрылыш жобасының барлық аспектілерін орталықтандырылған және жүйелі басқаруды, бір платформада көптеген өтініштер мен жүйелерді біріктіруге тырысады.

Болашақта ERP талдаулары дәлдікті арттырумен бірге қолданылып, дәлдікті арттыру және жобаның болашақ атрибуттарын есептеу процесін оңтайландыру үшін қолданылады. ERP жүйелерінен алынған және жиналған мәліметтер мен атрибуттар (5.44-сурет) Болашақта ықтимал модельдерді құру үшін негіз болады (5.44) болады, мысалы, ықтимал кідірістер, тәуекелдер немесе, мысалы, материалыдар құнының өзгеруі мүмкін болжамды модельдерді жасауға негіз болады.-

ERP баламасы ретінде, құрылыш индустриясы көбінесе PMIS (жобаларды басқарудың ақпараттық жүйесі) - жеке құрылыш алаңындағы міндеттерді егжей-тегжейлі бақылау үшін жобаларды басқару жүйесі.

PMIS: ERP мен құрылыш алаңы арасындағы аралық буын

Компанияның барлық бизнес-процестер тізбегін қамтитын ERP-ге қарағанда, PMIS нақты жобаны басқаруға бағытталған, белгіленген мерзімдер, бюджеттер, ресурстар және құжаттама жүргізуі қамтамасыз ету.

PMIS (жобаларды басқарудың ақпараттық жүйесі) - бұл жобаның барлық аспектілері бойынша жоспарлау, бақылау, талдау және есеп беруге арналған құрылыш жобаларын басқаруға арналған бағдарлама.

PMIS сізге құжаттарды, графиктерді, бюджеттерді, бюджеттерді басқаруға және PMIS ERP-ге қатысты қайталараптын шешім болып көрінуге мүмкіндік береді, бірақ негізгі айырмашылық бақылау деңгейі болып табылады:

- ERP компанияның бизнестің бизнес-процестеріне бағытталған: корпоративтік деңгейде шығындар, келісімшарттар, сатып алулар, қызметкерлер және ресурстар.
- PMIS жеке жобаларды басқаруға, өзгерістерді, өзгерістерді өзгертуді, өзгерістерді өзгертуді, есеп беруді және оларды үйлестіруді қамтамасыз етуге бағытталған.

Көптеген жағдайларда, бұл жеткілікті функционалдылығы бар ERP жүйелері және PMIS енгізу компанияның ыңғайлылығы мен қалаулары туралы мәселе туындаиды. Көптеген мердігерлер мен клиенттер PMIS-ті қажет емес, бірақ оны белгілі бір платформада жинақтағысы келетін сатушы немесе ірі клиент жүктелгендіктен.

Құрылыш жобаларында халықаралық терминологияда PLM (өнімнің өмірлік циклін басқару) сияқты көптеген жеке үғымдар бар, мысалы, өнімнің өмірлік циклін, сонымен қатар EPC және EPC-

М (инжиниринг, сатып алу және құрылыш) - құрылыш индустриясындағы келісім-шарттар.

Егер компания Жобаны басқару модульдерімен ERP қолданса, PMIS енгізу функционалдылықтың қайталайтын қосымша сілтеме болуы мүмкін. Алайда, егер процестер автоматтандырылмаса және деректер шашыраңқы болса, PMIS құралға ыңғайлыш және оңай бола алады.

Алыпсатарлық, пайда, жақындық және болмауы ERP және PMIS-тегі мөлдірлік

Интерфейстер мен процедуралардың сыртқы қарапайымдылығына қарамастан, көптеген жағдайларда Құрылыш ERP және PMIS жүйелері жабық және жағымсыз шешімдерді жабады. Мұндай жүйелер, әдетте, ішкі мәліметтер базасына және технологиялық логикаларға қол жетімділігі шектеулі бір сатушыдан алдын-ала жасалған бағдарламалық кешен түрінде жеткізіледі.

Мұндай жүйелерді дамыту және бақылау ценді (BIM) өнімдерді алуша, өйткені оларда ERP жүйелеріне қажетті мәліметтері бар мәліметтер базасы бар, өйткені олардың мәліметтер базасы, жобалық элементтердің сандық және көлемді атрибуттары. Алайда, бұл мәліметтерге ашық немесе машинаны оқылатын форматта қол жеткізудің орнына, сатушылар тек шектеулі пайдалануышы сценарийлері мен жабық өңдеу логикасын ұсынады, олар қара жәшікке арналған модульдер ішінде алдын-ала анықталған. Бұл жүйенің икемділігін азайтады және оны жобалардың нақты жағдайларына бейімдеуге жол бермейді.

Деректердің шектеулі мөлдірлігі құрылыш саласындағы сандық процестердің негізгі мәселелерінің бірі болып қала береді. Жабық мәліметтер базасының архитектурасы, құрылыш элементтерінің толық жиынтықтарына, автоматтандырудың қара белгілерінен қол жетімділіктің болмауы, автоматтандыру модулине бағдарлау және ашық интерфейстердің болмауы құжаттамалық бюрократияның тәуекелдерін едәуір арттырады.

Мұндай шектеулер шешім қабылдау процесінде «Тар орындарды» құрады, ақпаратты тексеруді қыыннатады және ERP / PMIS жүйелерінде деректерді немесе алыпсатарлықты жасыру мүмкіндіктерін ашады. Пайдаланушылар, әдетте, тек шектеулі қол жетімділікке ие болады - бұл кесілген интерфейс немесе жартылай API ба, жоқ па - бұл бастапқы деректердің бастапқы көздерімен өзара әрекеттесу мүмкіндігінсіз. Бұл, әсіресе, QCO есептеулерінде қолданылатын көлемдер, аудандар мен мөлшерлер сияқты жиынтық жобалардан жасалған параметрлер туралы маңызды.

Нәтижесінде, процестерді автоматтандыру, деректердің ашықтығы, мәміле бойынша шығындарды азайту және жаңа бизнес-модельдер құру, көптеген құрылыш компаниялары сыртқы параметрлерді басқаруға бағытталған, олар жобаның құны, түзету факторлары және есептеу әдістерімен, жобаның құны бойынша, Жинақтаушы ERP / PMIS платформаларында қолданылады. Бұл алыпсатарлық үшін топырақты жасайды, нақты өндіріс шығындарын бүрмалайды және құрылыш процесіне қатысушылардың барлық қатысушылары арасындағы сенімді азайтады.

Құрылыштағы, пайда аяқталған жобадан түсेतін түсім мен айнымалы шығындар арасындағы айырмашылық, оған жобалық, материалдар, еңбек ресурстары және объектінің жүзеге асырумен тікелей байланысты басқа да шығындар кіреді. Алайда, бұл шығындардың құнына әсер ететін негізгі фактор тек технологиялар немесе логистика ғана емес, сонымен қатар есептеулердің жылдамдығы мен дұрыстығы, сонымен қатар Қоғамдағы басқарушылық шешімдердің сапасы.

Мәселе көптеген құрылыш компанияларында құндылықтарды есептеу процесі тек клиенттер үшін ғана емес, сонымен қатар есептелген немесе қаржы бөлімдерінің құрамына кірмейтін қызметкерлер үшін де бос қалады. Мұндай жақындық ERP / PMIS жүйелеріндегі атрибуттар мен түзету коэффициенттерін өндеге айрықша құқығы бар «Қаржы сараптамасы» компаниясы бойынша артықшылықты мамандар тобының құрылуына ықпал етеді. Бұл қызметкерлер компаниялар басшыларымен бірге жобаның қаржылық логикасын нақты бақылай алады.

Сметалар, мұндай жағдайда, компанияның пайдасын ұлғайтуға және клиенттің бәсекеге қабілетті бағасын сақтаудың қажеттілігін теңдестіру, «қаржылық джюгілерге» айналады. Сонымен бірге, олар компанияның беделіне нұқсан келтірмеу үшін айқын және өрескел айла-шарттардан аулақ болуға мәжбүр. Дәл осы кезеңде коэффициенттер, бұл инфляцияланған көлемдерді немесе материалдардың құнын жасырады.

Нәтижесінде құрылыш саласында жұмыс істейтін компаниялардың тиімділігі мен рентабельдігін арттырудың негізгі схемасы автоматтандыру мен шешім қабылдау процестерін жеделдету емес, материалдар мен жұмыстардың бағаларына спекуляция жасау болып табылады (Рис. 5.45). Жұмыстар мен материалдардың құнын көтеру "сүр" бухгалтерия арқылы жабық ERP/PMIS жүйелерінде нарықтық орташа бағалардан жоғары пайыздармен немесе коэффициенттер арқылы жұмыстардың көлемдерін арттыру жолымен жүзеге асырылады (Рис. 5.16), олар "Калькуляция жасау және ресурстық база негізінде жұмыстардың құнын есептеу" бөлімінде талқыланды.-

Нәтижесінде тапсырыс беруші нақты құнды немесе жұмыстардың көлемдерін көрсетпейтін, бірақ көптеген жасырын ішкі коэффициенттердің туындысы болып табылатын есепті алады. Осылайша, субподрядшылар, жалпы подрядчикпен белгіленген төмен бағаларға сәйкес келуге тырысып, жиі арзан және сапасыз материалдарды сатып алуға мәжбүр болады, бұл құрылыш сапасын нашарлатады.

Спекулятивті пайда табу процесі, ақырында, клиенттерге, оларды жалған ақпаратпен қамтамасыз ететін, сондай-ақ орындаушыларға, оларды жаңа спекуляция модельдерін іздеуге мәжбүр ететін, зиян келтіреді.

Нәтижесінде, жоба неғұрлым ауқымды болса, деректер мен процестерді басқаруда бюрократия деңгейі соғұрлым жоғары болады. Әрбір кезең мен модульдің артында көбінесе есептеу алгоритмдеріне және ішкі процедураларға енгізілген мәлдір емес коэффициенттер мен үстемелер жасырылады. Бұл тек аудитті қыннатпайды, сонымен қатар жобаның қаржылық көрінісін едауір бұрмалайды. Ирі құрылыш жобаларында мұндай тәжірибелер жиі соғы құнның бірнеше есе (кейде он есе) артуына әкеледі, ал нақты көлемдер мен шығындар тапсырыс берушінің тиімді бақылауынан тыс қалады (Рис. 2.13 Германиядағы ірі инфрақұрылымдық жобалар бойынша жоспарланған және

нақты шығындарды салыстыру).

McKinsey & Company-ның "Құрылыш саласының цифрлық болашағын елестету" (2016) есебіне сәйкес, ірі құрылыш жобалары орташа есеппен жоспарланған мерзімнен 20% кешігіп аяқталады және бюджеттен 80% артық шығындалады [107].

Смета және бюджеттеу бөлімдері компания ішіндегі ең қорғалатын буынға айналады. Оларға қолжетімділік тіпті ішкі мамандар үшін де қатаң шектелген, ал деректер базасының логикасы мен құрылымының жабықтығы жобалық шешімдердің тиімділігін объективті бағалауды бұрмалаусыз жүзеге асыруды мүмкін емес етеді. Түсініктіліктің болмауы компанияларды процестерді оңтайландырудың орнына, "шығармашылық" цифrlар мен коэффициенттерді басқару арқылы өмір сүрге мәжбүр етеді (Рис. 5.31, Рис. 5.16 - мысалы, "Bid Factor" параметрі).-



Рис. 5.45 Есептеулер деңгейіндегі спекуляция коэффициенттері - компаниялардың негізгі пайдасы және жұмыс сапасы мен бедел арасындағы жонглирлеу өнері.

Мұның бәрі құрылыш саласында жабық ERP/PMIS жүйелерін одан әрі пайдаланудың мақсатқа сай екендігін күмән тудырады. Цифрлық трансформация және тапсырыс берушілер тарарапынан түсініктілікке қойылатын талаптардың артуы жағдайында, жобаларды ұзақ мерзімді перспективада патенттелген шешімдерге, икемділікті шектейтін, интеграцияға кедергі келтіретін және бизнестің дамуын тежейтін шешімдерге тәуелді болып қалуы екіталай.-

Құрылыш компанияларына деректер силосымен және жабық деректер базаларындағы мәлдір емес деректермен жұмыс істеу қаншалықты тиімді болса да, құрылыш саласының болашағы ашық платформаларға, машинамен оқылатын және мәлдір деректер құрылымдарына, сондай-ақ сенім принциптеріне негізделген автоматтандыруға көшуімен байланысты болады. Бұл трансформация «жоғарыдан» — тапсырыс берушілер, реттеуші органдар және қоғам тарарапынан, есеп беруді, тұрақтылықты, мәлдірлікті және экономикалық негіздемені талап етудің артуымен жүзеге асатын болады.

Жабық ERP/PMIS дәуірінің соны: құрылыш саласы жаңа тәсілдерге мүқтаж

Ондаған миллион жол кодтан тұратын ауыр модульдік ERP/PMIS жүйелерін пайдалану кез келген өзгерістерді жүзеге асыруды өте қынданатады. Сонымен қатар, компанияға алдын ала бапталған модульдер, ресурстар базасындағы он мындаған артикулдар (Сурет 5.13) және дайын калькуляциялардың мындаған болуы жағдайында жаңа платформаға көшу қымбат және ұзақ процесс болып табылады. Кодтың және ескірғен архитектуралық шешімдердің саны артқан сайын, ішкі тиімділік деңгейі де жоғарылайды, ал әр жаңа жоба жағдайды тек нашарлатады. Қөптеген компанияларда деректерді көшіру және жаңа шешімдерді интеграциялау жылдар бойы созылатын эпопеяға айналады, ол үнемі қайта өңдеулермен және шексіз компромистерді іздеумен бірге жүреді. Нәтижесінде, шектеулері бола тұра, ескі, таныс платформаларға қайта оралу жиі кездеседі.--

Германиядағы "Қара кітап" атты неміс баяндамасында (Германия) құрылыш деректерін басқарудағы жүйелік сәтсіздіктерге арналған, ақпараттың фрагментациясы және оны басқарудағы орталықтандырылған тәсілдің болмауы – тиімділіктің төмендеуінің негізгі себебі болып табылады. Стандарттау мен интеграциясыз деректер өз құндылығын жоғалтып, архивке айналады, ал басқару құралына айналмайды.

Деректер сапасының жоғалуының негізгі себебі – құрылыш жобаларын жоспарлау мен бақылаудың жеткіліксіздігі, бұл жиі шығындардың елеулі өсуіне әкеледі. "Қара кітаптың" "Шығындардың жарылысы" бөлімінде мұндай жағымсыз салдарға ықпал ететін негізгі факторлар талданады. Олардың арасында – қажеттіліктерді жеткіліксіз талдау, техникалық-экономикалық негізdemelerdің болмауы және қосымша шығындарға әкелетін келісілмеген жоспарлау бар, оларды болдырмауға болатын еді.

Зерттелген IT-екожүйесінде ескі жүйені ауыстыру, бұрыннан салынған ғимараттағы тірек колоннаны ауыстырумен салыстыруға болады. Ескі жүйені жай ғана демонтаждау және жаңасын орнату жеткіліксіз – бұл ғимараттың тұрақтылығын сақтап, төбелердің құламауын және барлық коммуникациялардың жұмысын жалғастыруын қамтамасыз ету маңызды. Міне, осы жерде қыындық жатыр: кез келген қате компанияның бұқіл жүйесіне ауыр салдарға әкелуі мүмкін.

Дегенмен, ірі ERP өнімдерінің әзірлеушілері құрылыш саласында жазылған кодтың санын өз платформаларының артықшылығы ретінде пайдалануды жалғастыруда. Арнайы конференцияларда «Мұндай жүйені қайта құру үшін 150 адам-жыл қажет» деген сөздерді естуге болады, алайда мұндай жүйелердің көп функционалдылығы негізінен деректер базалары мен кестелермен жұмыс істеуге арналған қарапайым функциялардың артында жасырынып жатыр, олар арнайы бекітілген пайдаланушы интерфейсіне оралған. Практикада «150 адам-жыл» кодының көлемі бәсекелестік артықшылықтан гөрі, жүккө айналады. Кодтың көлемі неғұрлым көп болса, қолдаудың құны соғұрлым жоғары, жаңа жағдайларға бейімделу қынырақ және жаңа әзірлеушілер мен клиенттер үшін кіру шегі жоғары болады.

Бұгінгі таңда көптеген модульдік құрылым жүйелері көлемді және ескірген "Франкенштейн-конструкт" секілді, мұнда кез келген абайсыз өзгеріс ақауларға әкелуі мүмкін. Эр жаңа модуль барынша жүктелген жүйені одан әрі құрделендіріп, оны тек тар шеңбердегі мамандарға түсінікті лабиринтке айналдырады, бұл оны қолдау мен жаңартуды одан әрі қызындатады.

Кұрделілік әзірлеушілердің өздері тарапынан да сезіледі, олар периодты түрде жаңа технологияларды ескере отырып, архитектуралық қайта қарау үшін үзілістер жасайды. Алайда, рефакторинг тұрақты түрде жүргізілсе де, құрделілік 不可避免 но артады. Мұндай жүйелердің архитектуралық мамандары артып келе жатқан құрделілікке үйреніп кетеді, бірақ жаңа пайдаланушылар мен мамандар үшін бұл жеңілмейтін кедергіге айналады. Нәтижесінде, барлық сараптама бірнеше әзірлеушінің қолында шоғырланады, және жүйе масштабталмайды. Қысқа мерзімді перспективада мұндай сарапшылар пайдалы, бірақ ұзақ мерзімді перспективада олар мәселенің бір бөлігіне айналады.

Ұйымдар «кіші» деректерді олардың үлкен аналогтарымен интеграциялауды жалғастырады, және бір ғана қосымша – қаншалықты қымбат немесе сенімді болса да – барлығын шешеді деп сенетіндер ақымақ. Фил Саймон, Conversations About Collaboration подкастының жүргізуісі.

Заңды сұрақ туындаиды: шынымен де бізге жұмыс құны мен мерзімдерін есептеу үшін мұндай көлемді және жабық жүйелер қажет пе, егер басқа салаларда осындай міндеттермен аналитикалық құралдар ашық деректер мен айқын логикамен әлдеқашан айналысса?

Қазіргі уақытта жабық модульдік платформалар құрылым саласында әлі де сұранысқа ие, ең алдымен калькуляциялық есептің ерекшеліктеріне байланысты. Мұндай жүйелер жи «сұр» немесе ашық емес схемаларды жүргізу үшін пайдаланылады, бұл тапсырыс берушіден нақты шығындарды жасыруға мүмкіндік береді. Алайда, тапсырыс берушілердің цифрлық жетілуі мен саланың «уберленген дәүірге» өтуімен, делдалдар, яғни құрылым компаниялары өздерінің ERP жүйелерімен уақыт пен құнды есептеуде маңыздылығын жоғалтады. Бұл құрылым индустриясының келбетін мәңгілікке өзгерtedі. Толығырақ кітаптың соңғы бөлімінде және «Құрылым 5.0: жасыру мүмкін болмайтын кезде қалай табыс табуға болады» тарауында қарастырылады.

Соңғы 30 жылда жинақталған мындаған ескірген легаси-шешімдер, мындаған адамжылдар инвестицияланған әзірлеу жұмыстары жедел түрде жойыла бастайды. Ашық, мөлдір және икемді деректерді басқаруға көшу – міндетті. Сұрақ тек, қандай компаниялар осы өзгерістерге бейімделе алатыны, ал қайсысы ескі модельдің тұтқыны болып қалатыны.

CAD- (BIM-) құралдары саласында да ұқсас жағдай байқалады, олардың деректері бұгінгі күні ERP/PMIS жүйелеріндегі жобалық объектілердің көлемді параметрлерін толтырады. BIM идеясы (2002 жылды әзірленген) бастапқыда біртұтас интеграцияланған деректер базасы концепциясына негізделген, алайда бұгінгі күні BIM-мен жұмыс істеу арнайы бағдарламалар мен форматтардың

тұтас жиынтығын талап етеді. Жобалауды және құрылым басқаруды жеңілдетуі тиіс нәрсе, патент-телген шешімдердің жаңа қабатына айналып, интеграцияны қыындастып, бизнестің икемділігін төмендettі.

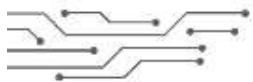
Келесі қадамдар: жобалық деректерді тиімді пайдалану

Осы бөлімде құрылымдық деректердің құрылымы жобаларының құны мен мерзімдерін дәл есептейдің негізі болып табылатынын көрсеттік. QTO, құнтізбелік жоспарлау және сметалық есептеу процестерін автоматтандыру енбек шығындарын азайтады және нәтижелердің дәлдігін айтартықтай арттырады.

Осы бөлімнің қорытындысында, қарастырылған тәсілдердің күнделікті тапсырмаларда қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн. Бұл тәсілдер әмбебап – олар компанияның цифрлық трансформациясы үшін де, есептеулермен айналысатын мамандардың күнделікті жұмысы үшін де пайдалы:

- Рутиналық есептеулерді автоматтандырыңыз
- Сіз өз жұмысыңызды қатысы бар стандартты калькуляцияларды табуға тырысыңыз
- Сіздің елдегі құрылым алаңында жұмыстарды немесе процестерді қалай есептейтінін талдаңыз (Рис. 5.17)-
- CAD жүйелерімен жұмыс істесеңіз, CAD- (BIM-) бағдарламалық қамтамасыз етудегі спецификацияларды және QTO деректерін автоматты түрде алу функцияларын зерттеңіз
- Есептеулерді автоматтандыру үшін кодтың алғашқы нұсқасын жазу үшін LLM пайдаланыңыз
- QTO үшін өз құралдарыңызды әзірлеңіз
- Көлемдерді автоматты түрде есептеу үшін скрипттер немесе кестелер жасаңыз
- Бағалауға жүйелі тәсіл үшін элементтердің санаттары мен топтарын стандарттаңыз
- Жаңа жобаларда нәтижелердің қайта өндірілуін қамтамасыз ету үшін есептеу әдістемесін құжаттаңыз
- Жобаның әртүрлі аспектілерін өз жұмысыңызды интеграциялаңыз
- Модульдік жүйелермен жұмыс істесеңіз, процестеріңізді тек схемалар немесе диаграммалар ретінде ғана емес, деректер деңгейінде – әсіресе кесте түрінде визуализациялауға тырысыңыз
- CAD базаларынан алынған деректерді есептеулермен автоматты түрде біріктіруді үйреніңіз – Python кодын пайдаланып, топтастыру, сұзу және агрегация арқылы
- Әріптестер мен клиенттерге кешенді ақпаратты ұсыну үшін QTO топтарының көрнекі визуализацияларын жасаңыз
- Бұл қадамдар автоматтандыру мен деректерді стандарттауға негізделген тұрақты есептеу жүйесін құруға көмектеседі. Мұндай тәсіл дәлдікті арттырады, есептеулермен байланысты күнделікті мәселелердегі рутинаны азайтады.
- Келесі тараулар CAD- (BIM-) өнімдерінің техникалық аспектілеріне және CAD деректер

базаларын компаниялардың бизнес-процестеріне интеграциялаудың қындығына арналған. Егер сізді қазіргі уақытта BIM-нің құрылышқа енгізілу тарихы, CAD құралдарының эволюциясы және осы технологиялармен жұмыс істеудің техникалық ерекшеліктері қызықтырмаса, кітаптың жетінші бөліміне «Деректер негізінде шешім қабылдау» дереу өтуінізге болады.



БАСПА НҰСҚАСЫМЕН МАКСИМАЛДЫ ҚОЛАЙЛЫЛЫҚ

Сіздер Data-Driven Construction-ның тегін цифрлық нұсқасын ұстап отырсыздар. Материалдарға жылдам қол жеткізу және ыңғайлы жұмыс үшін баспа басылымына назар аударуды ұсынамыз:



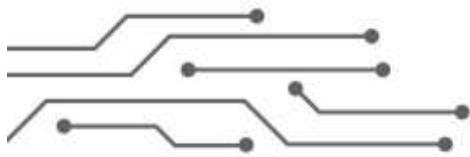
- Эрдайым қол астында: баспа форматындағы кітап сенімді жұмыс құралы болып табылады, қажетті визуализациялар мен схемаларды кез келген жұмыс жағдайында жылдам табуға және пайдалануға мүмкіндік береді.
- Суреттердің жоғары сапасы: баспа басылымындағы барлық бейнелер мен графиктер максималды сапада ұсынылған.
- Ақпаратқа жылдам қол жеткізу: ыңғайлы навигация, ескертпелер жасау, бетбелгілер қою және кітаппен кез келген жерде жұмыс істеу мүмкіндігі.



Кітаптың толық баспа нұсқасын сатып алу арқылы сіз ақпаратпен ыңғайлы және

тиімді жұмыс істеуге арналған құрал аласыз: күнделікті тапсырмаларда визуалды материалдарды жедел пайдалану, қажетті схемаларды жылдам табу және ескертпелер жасау мүмкіндігі. Сонымен қатар, сатып алуыңыз ашық білімнің таралуын қолдайды.

Кітаптың баспа нұсқасын мына жерден тапсырыс беруге болады:
datadrivenconstruction.io/books



VI БӨЛІМ

CAD ЖӘНЕ BIM: МАРКЕТИНГ, ШЫНДЫҚ ЖӘНЕ ҚҰРЫЛЫС САЛАСЫНДАҒЫ ЖОБАЛАУ ДЕРЕКТЕРІНІҢ БОЛАШАҒЫ

Кітаптың алтыншы бөлімі CAD және BIM технологияларының эволюциясын және олардың құрылыс саласындағы деректерді басқару процестеріне әсерін сини талдауға арналған. BIM концепциясының бастапқы интеграцияланған деректер базасы идеясынан қазіргі бағдарламалық қамтамасыз ету жеткізушілері тарарапынан ұсынылатын маркетингтік конструкцияларға дейінгі тарихи трансформациясы қарастырылады. Проприетарлық форматтар мен жабық жүйелердің жобалық деректермен жұмыс істеу тиімділігіне және құрылыс саласының жалпы өнімділігіне әсері бағаланады. Әртүрлі CAD жүйелерінің үйлесімділік мәселелері мен құрылыс компанияларының бизнес-процестерімен интеграциялау қындықтары егжей-тегжейлі талданады. USD сияқты қарапайым ашық деректер форматтарына көшу тенденциялары және олардың салаға ықтимал әсері қарастырылады. Жабық жүйелерден ақпарат алу үшін баламалы тәсілдер, соның ішінде кері инженерия әдістері ұсынылады. Құрылыс саласындағы жобалау және деректерді талдау процестерін автоматтандыру үшін жасанды интеллект пен машиналық оқытудың қолдану перспективалары талданады. Пайдаланушылардың нақты қажеттіліктеріне, бағдарламалық қамтамасыз ету жеткізушілерінің мүдделеріне емес, бағытталған жобалау технологияларының дамуы бойынша болжамдар формулирленеді.

ТАРАУ 6.1. ҚҰРЫЛЫС САЛАСЫНДАҒЫ ВІМ-КОНЦЕПТЕРДІҢ ПАЙДА БОЛУЫ

Алғашында осы алтыншы бөлім, CAD (BIM) тақырыбына арналған, кітаптың бірінші нұсқасында болмаған. Проприетарлық форматтар, геометриялық ядролар және жабық жүйелер тақырыптары тым техникалық, артық мәліметтермен толтырылған және, бір қарағанда, деректермен жұмыс істеуді түсінгісі келетіндер үшін пайдасты болып көрінеді. Алайда, кітаптың бірінші нұсқасына түсінктеме қосу туралы пікірлер мен өтініштер көрсеткендегі, CAD жүйелерінің ішкі механизмдері, геометриялық ядролар, форматтардың әртүрлілігі және бірдей деректерді сақтау схемаларының үйлесімсіздігімен байланысты барлық қыындықтарды түсінбей, жеткізушилер тарапынан ұсынылатын концепциялардың ақпаратпен жұмыс істеуді қыыннататынын және ашық параметрленген жобалауға көшуге кедергі келтіретінін шын мәнінде түсіну мүмкін емес. Соңдықтан бұл бөлім кітаптың құрылымында өз орнын алды. Егер CAD (BIM) тақырыбы сіз үшін басымдық болмаса, келесі бөлімге – «VII БӨЛІМ: Деректер негізінде шешім қабылдау, аналитика, автоматтандыру және машиналық оқыту» дереу өтуіңзеге болады.

ВІМ және open BIM-нің CAD жеткізушилері тарапынан маркетингтік концепт ретінде пайда болуының тарихы

90-шы жылдары цифрлық деректердің пайда болуымен компьютерлік технологиялар тек бизнес-процестерге ғана емес, жобалау процестеріне де енгізіле бастады, бұл CAD (автоматтандырылған жобалау жүйелері) және кейінірек BIM (құрылым ақпараттық моделін) сияқты концепциялардың пайда болуына әкелді.

Алайда, кез келген инновация сияқты, олар даму нүктесі емес. BIM сияқты концепциялар құрылым саласының тарихында маңызды кезең болды, бірақ ерте ме, кеш пе, олар болашақтың талаптына жақсы жауап беретін жетілдірілген құралдар мен тәсілдерге орын беруге мәжбүр болады.

CAD жеткізушилерінің әсеріне ұшырап, өздерінің жүзеге асыру қыындықтарында шатасып қалған BIM концепциясы, 2002 жылы пайда болған, рок-жұлдыз сияқты жарқырап, бірақ тез сөнү мүмкін. Себебі қарапайым: деректермен жұмыс істейтін мамандардың талаптары CAD жеткізушилері бейімделуге үлгермей, жылдам өзгеруде.

Сапалы деректердің жетіспеушілігімен бетпе-бет келген қазіргі құрылым саласының мамандары кросс-платформалық үйлесімділік пен CAD жобаларынан ашық деректерге қол жеткізуі талап етеді, бұл олардың талдауын және өндөуін жөнілдетеді. CAD деректерінің күрделілігі мен өндөуінің шатасуы құрылым процесінің барлық қатысушыларына: жобалаушыларға, жобаларды басқарушыларға, құрылым алаңындағы жұмысшыларға және, ақырында, тапсырыс берушіге теріс етеді.

Бұғынгі күні тапсырыс беруші мен инвестор толық деректер жиынтығын алудың орнына, CAD форматындағы контейнерлерді алады, олар күрделі геометриялық ядроларды, деректер схемаларын, жыл сайын жаңартылатын API құжаттамасын және CAD (BIM) деректерімен жұмыс істеу үшін арнайы бағдарламалық қамтамасыз етуді талап етеді. Сонымен қатар, жобалық деректердің көп бөлігі пайдаланылмай қалады.

Бұғынгі күні CAD деректеріне қол жеткізудің күрделілігі жобаларды басқаруда артық инженерияға әкеледі. CAD деректерімен жұмыс істейтін немесе BIM шешімдерін әзірлейтін орта және ірі компаниялар, деректерге API арқылы қол жеткізу үшін CAD шешімдерінің жеткізушілерімен тығыз қарым-қатынаста болуға мәжбүр, немесе CAD жеткізушілерінің шектеулерін айналып өту үшін қымбат SDK конвертерлерін пайдаланып, ашық деректерді алу үшін реверс-инженеринг жүргізеді.

Проприетарлық деректерді пайдалану тәсілі ескірген және қазіргі цифрлық орта талаптарына жауап бермейді. Болашақта компаниялар екі типке бөлінеді: ашық деректерді тиімді пайдаланатындар және нарықтан кететіндер.

BIM (Building Information Modeling) концепциясы құрылыш саласында 2002 жылы ірі CAD жеткізушілерінің бірі - Whitepaper BIM жариялануымен пайда болды және BOM (Bills of Materials) машина жасау концепциясымен толықтырылды, параметрлік тәсілден бастау алды. Параметрлік тәсіл жобалық деректерді жасау мен өндөде алғашқылардың бірі болып Pro-E жүйесінде жүзеге асырылды. Бұл жүйе қазіргі заманғы CAD шешімдерінің, соның ішінде құрылыш саласында қолданылатындардың прототипі болды.-

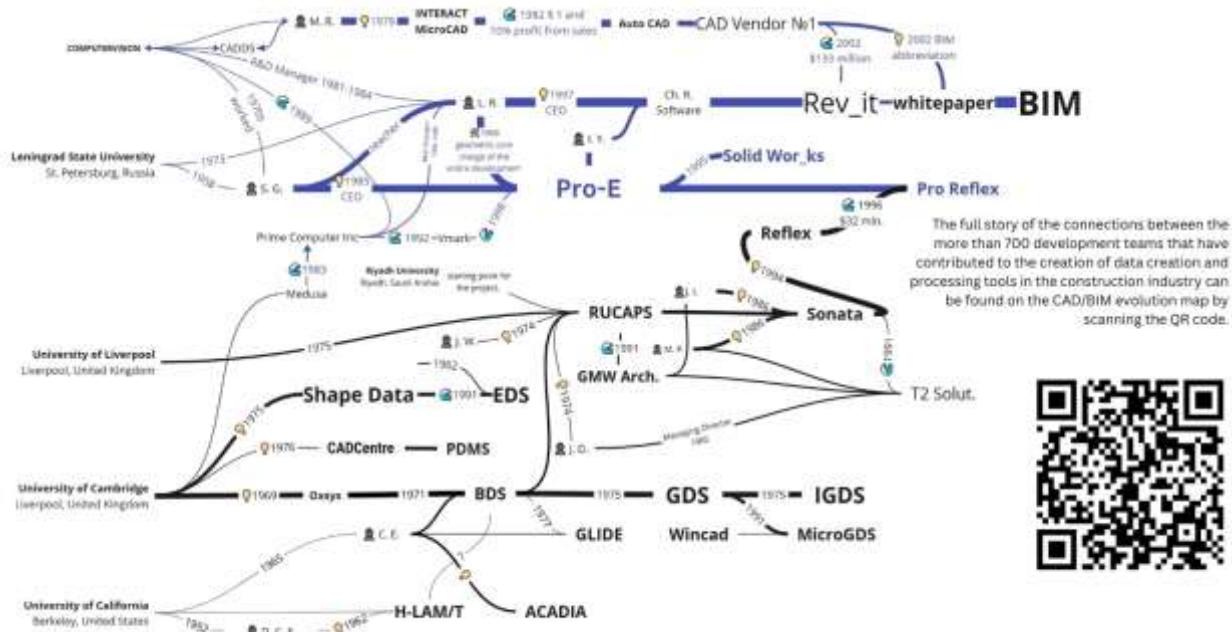


Рис. 6.11 BIM концепциясының және оған үқсас концепциялардың пайда болу тарихының картасы.

2000 жылдардың басында CAD жеткізушілерінің құралдарын насиҳаттаған АЕС журналистері мен консультанттары 2002 жылдан бастап Whitepaper BIM-ге назар аударды. Нақтырақ айтқанда, 2002-2004 жылдардағы Whitepaper BIM және 2002, 2003, 2005 және 2007 жылдары жарияланған мақалалар BIM концепциясының құрылыш саласында танымал болуында маңызды рөл атқарды.

Гимараттардың ақпараттық моделін жасау – бұл [CAD жеткізуши компанияның атауы] құрылым индустриясында ақпараттық технологияларды қолдану стратегиясы. – Whitepaper BIM, 2002

2000-жылдардың ортасына қарай «зерттеушілер» 2002 жылы CAD-вендорымен жарияланған BIM концепциясын Чарльз Истменнің BDS сияқты ертерек ғылыми жұмыстарымен байланыстыра бастады, ол GLIDE, GBM, BPM, RUCAPS сияқты жүйелердің негізі болды. Чарльз Истменнің "Building Description System" (1974) атты жаңашыл еңбегі қазіргі ақпараттық моделдеудің теориялық негіздерін қалады. "Деректер базасы" термині оның еңбегінде 43 рет кездеседі (Рис. 6.12) – басқа сөздерден, "ғимарат" сөзінен басқа, жиі кездеседі.-

Истменнің негізгі идеясы ғимарат туралы барлық ақпарат – геометриядан элементтердің қасиеттері мен олардың өзара байланыстарына дейін – біртұтас құрылымдық деректер базасында сақталуы тиіс деген болатын. Дәл осы базадан автоматты түрде сызбалар, спецификациялар, есептеулер жасауға және нормативтерге сәйкестікті талдауға болады. Истмен сызбаларды ескірген және артық коммуникация тәсілі ретінде тікелей сынға алды, ақпараттың дублиренуі, жаңартудағы қызындықтар және өзгерістер енгізілген кезде қолмен жаңартудың қажеттілігіне назар аударды. Оның орнына, ол деректер базасында біртұтас цифрлық модельді ұсынды, мұнда кез келген өзгеріс бір рет енгізіліп, барлық көріністерде автоматты түрде көрсетіледі.

Айта кету керек, Истменнің концепциясында визуализация басты орынға қойылмаған. Оның жүйесінде орталық орын ақпаратқа берілген: параметрлер, байланыстар, атрибуттар, талдау және автоматтандыру мүмкіндіктері. Сызбалар оның түсінігінде деректер базасынан алынған ақпараттың тек бір түрі болды, жобалық ақпараттың бастапқы көзі емес.

BIM бойынша алғашқы Whitepaper-да жетекші CAD-вендорында "деректер базасы" сөз тіркесі Чарльз Истменнің BDS-інде сияқты 23 рет қолданылған [60] жеті бетте және құжаттағы "Building", "Information", "Modeling" және "Design" сөздерінен кейінгі ең танымал сөздердің бірі болды. Алайда, 2003 жылға қарай аналогиялық құжаттарда "деректер базасы" термині тек екі рет кездеседі [61], ал 2000-жылдардың соңына қарай деректер базалары тақырыбы жобалық деректер талқылауынан дерлік жоғалып кетті. Нәтижесінде "визуалды және сандық талдау үшін біртұтас интеграцияланған деректер базасы" концепциясы толық жүзеге асырылмады.

Осылайша, құрылым саласы Чарльз Истменнің BDS концепциясынан, оның деректер базаларына назар аударған идеяларынан және Сэмюэль Гейсбергтің Pro-E машина жасау өніміндегі деректер базасынан автоматты түрде жобалық деректерді жаңарту туралы идеяларынан қазіргі заманғы маркетингтік BIM-ге өтті, мұнда деректерді басқару деректер базалары арқылы іс жүзінде айтылмайды, алайда дәл осы концепция бастапқы теориялық әзірлемелердің негізінде жатыр.

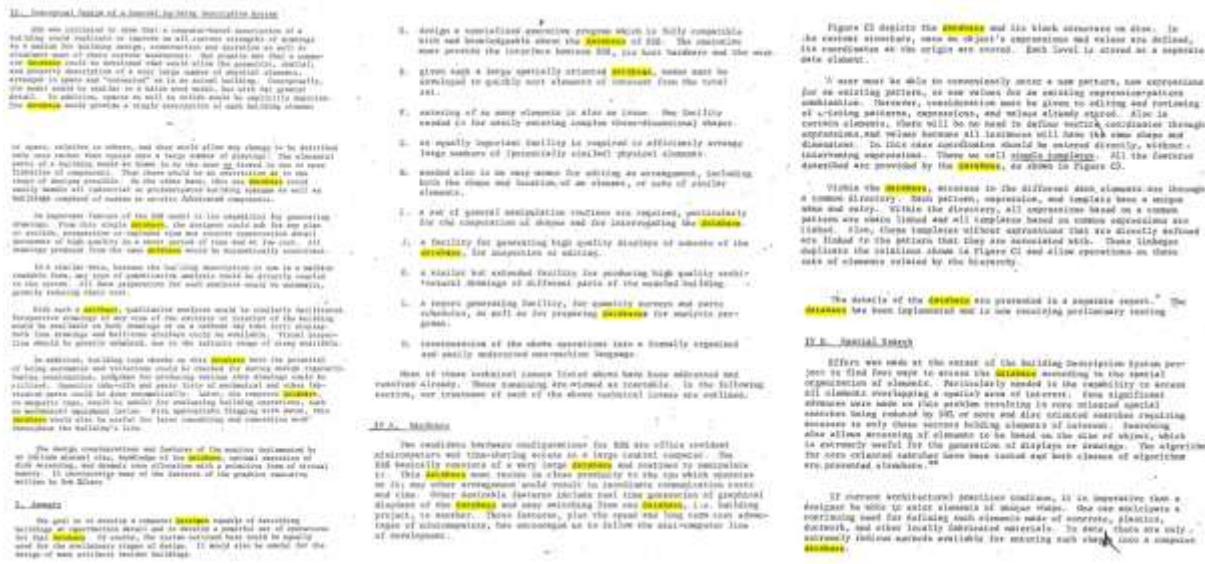


Рис. 6.12 Чарльз Истменнің 1974 жылы сипаттаған BDS концепциясында "Деректер базасы" сез тіркесі (сарымен бөлінген) 43 рет қолданылған.

BDS және оған үқсас концепциялар 2000-жылдарға дейін ғимараттардың цифрлық деректер базасы ретінде әзірленді, визуализация құралы ретінде емес. 2002 жылы BIM жобалау құралына айналды, мұнда деректер базасы екінші орынға шықты. 1990-жылдардағы BDS және оған үқсас концепциялардан BIM-ге 2010-жылдардың ортасына дейін өткенде не жоғалтық:

- Ашық деректер базалары: BDS және басқа үқсас концепциялар аналитикаға назар аударды, BIM - дизайнға.
- Деректермен жұмыс істеудің икемділігі: BDS деректер аналитикасына назар аударды, BIM - қандай деректерге негізделуі тиіс процестерге.
- Айқындық: BDS ашық интеграцияланған деректер базасы ретінде ойластырылған, ал CAD жеткізушилері BIM-да өз деректер базаларын толық жабық етіп жасаған және 20 жыл бойы меншікті форматтарды аштын кері инженерия құралдарымен сәтсіз күресуде.

Соңғы 30 жылда жобалаушылар «интеграцияланған деректер базасына» қол жеткізе алмады, ал BIM құралдарына қатысты 20 жылдық маркетингтік эйфориядан кейін құрылыш индустриясы бұл қызығушылықтың салдарын түсіне бастайды.

BIM шындығы: интеграцияланған деректер базаларының орнына жабық модульдік жүйелер

Деректерге, олардың құрылымына және біртұтас процестерге интеграциялануына назар аударудың орнына, CAD- (BIM-) жүйелерінің пайдаланушылары әрқайсысы өз ережелерін белгілейтін меншікті шешімдердің фрагменттеген жиынтығымен жұмыс істеуге мәжбүр.

- BIM туралы алғашқы ақ қағаздарда айтылған біртұтас деректер базасы миф болып қалды. Дауыс көтерген мәлімдемелерге қарамастан, деректерге қол жеткізу әлі де шектеулі және жабық жүйелер арасында бөлінген.
- BIM модельдері құрал емес, жабық экожүйе болып қалды. Ақпараттың ашық алмасуының орнына, пайдаланушылар жазылымдар үшін ақы төлеуге және меншікті API-лерді

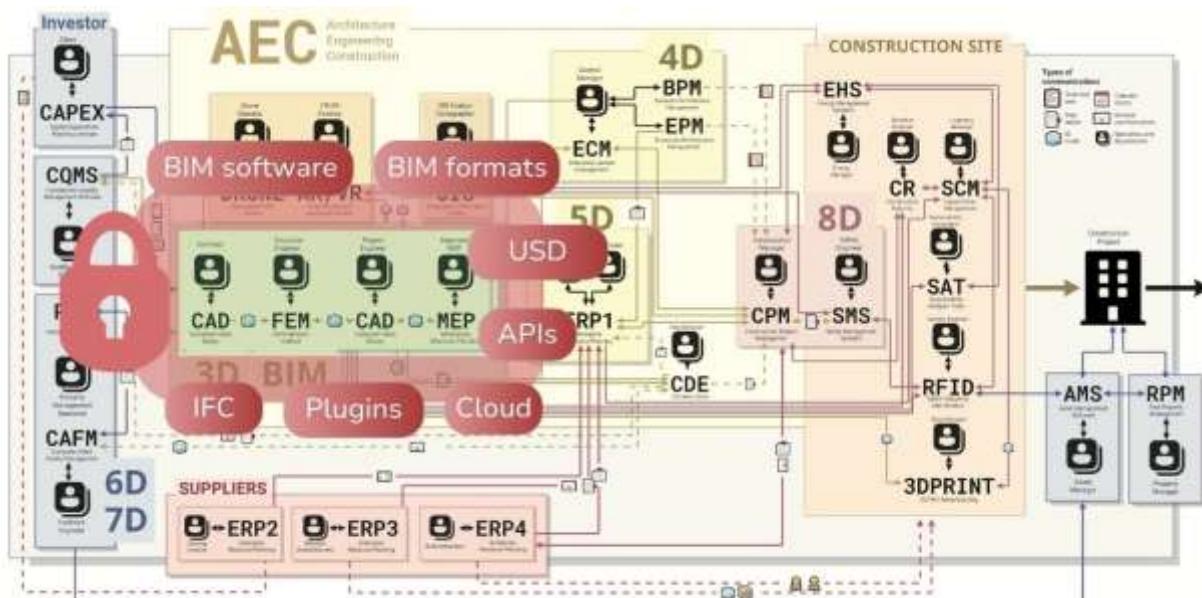
пайдалануға мәжбүр.

- Деректер жеткізушілерге тиесілі, пайдаланушыларға емес. Жобалар туралы ақпарат меншікті форматтарда немесе бұлттық қызметтерде жабық, ашық және тәуелсіз форматтарда қолжетімді емес.

Инженер-жобалаушылар мен жобаларды басқарушылар көбінесе CAD жүйелерінің деректер базасына немесе өз жобаларының деректері сақталатын форматқа қол жеткізе алмайды. Бұл ақпаратты жылдам тексеруді немесе деректердің құрылымы мен сапасына талаптарды қалыптастыруды мүмкін емес етеді. Мұндай деректерге қол жеткізу API және плагиндер арқылы біріктірілген арнайы бағдарламалық қамтамасыз етуді талап етеді, бұл құрылыш саласындағы процестердің артық бюрократиялануына әкеледі. Осы уақытта бұл деректер ондаған ақпараттық жүйелер мен жүздеген мамандармен бір уақытта пайдаланылады.-

Біз CAD (BIM) деректерін басқаруды, оларды цифрлық түрде сақтауды және өмірлік цикл мен процестерді басқару бағдарламалық қамтамасыз етуді сауды үйренуіміз керек, өйткені әрбір инженер [жобалаушы] CAD бағдарламасында бір нәрсе жасаған сайын, осы деректермен жұмыс істейтін он адам бар.

CAD жеткізушісінің бас директоры, BIM концепциясын жасаған, 2005 ж.



CAD- (BIM-) деректер базалары құрылыш бизнесінің экожүйесіндегі IT бөлімдері мен деректер менеджерлері үшін соңғы жабық жүйелердің бірі болып қала береді.

BIM-нің деректерді басқарудың толыққанды құралы емес, деректер базаларын коммерцияландыру құралы екендігі айқын болғанда, деректерді қалай қайтаруға болады деген зандың сұрақ туындаиды. Жауап – ақпараттың иесі бағдарламалық қамтамасыз етуші емес, пайдаланушы болатын ашық құрылымдық деректерді пайдалану.

Құрылыш саласындағы пайдаланушылар мен шешім әзірлеушілер, басқа экономикалық салалардағы әріптерестері сияқты, соңғы 30 жылда басым болған бағдарламалық қамтамасыз етушілердің бұлдықтың терминологиясынан біртіндеп бас тартатын болады, цифрландырудың негізгі аспектілеріне – «деректер» мен «процестерге» назар аударады.

1980-жылдардың соңында құрылыш саласындағы цифрлық технологиялардың дамуының негізгі бағыты деректерге қол жеткізу және жобаның ақпаратын басқару мәселесі ретінде қарастырылды. Алайда, уақыт өте келе акценттер өзгерді. Деректермен жұмыс істеудің ашық және қолжетімді тәсілдерін дамыту орнына, IFC форматы мен open BIM концепциясын белсенді түрде ілгерілету басталды – бұл мамандардың жобалық деректер базаларын басқару тақырыбынан назарын аудару әрекеті ретінде қарастырылды.

Құрылыш саласында IFC ашық форматының пайда болуы

IFC (Industry Foundation Classes) деп аталатын ашық формат әртүрлі CAD- (BIM-) жүйелері арасында үйлесімділікті қамтамасыз ету үшін стандарт ретінде позицияланады. Оның дамуы ең ірі CAD жеткізушилері тарарапынан құрылған және бақыланған ұйымдар аясында жүргізілді. IFC форматы негізінде 2012 жылы екі CAD компаниясы OPEN BIM маркетингтік концепциясын әзірледі.

IFC (Industry Foundation Classes) – құрылыш саласындағы деректерді алмасудың ашық стандарты, ол әртүрлі CAD- (BIM-) жүйелері арасында үйлесімділікті қамтамасыз ету үшін әзірленген.

Open BIM концепциясы CAD деректер базасынан ақпаратпен жұмыс істеуді және жүйелер арасында ақпарат алмасуды IFC ашық форматы арқылы жүзеге асыруды білдіреді.

Open BIM бағдарламасы – ... [1 CAD жеткізуши], ... [2 CAD жеткізуши] және басқа компаниялар тарарапынан бастамашы болған маркетингтік науқсан, ол AEC индустриясында OPEN BIM концепциясын жаһандық үйлестірілген ілгерілетуге, келісілген коммуникация мен бағдарламаның қатысушыларына қолжетімді жалпы брендтеуге бағытталған. – CAD жеткізушинің сайтынан, OPEN BIM бағдарламасы, 2012

IFC 1980-жылдардың соңында Мюнхен техникалық университеті тарарапынан STEP машина жасау форматын бейімдеу арқылы әзірленді, кейінірек ірі жобалау компаниясы мен ірі CAD жеткізуши тарарапынан 1994 жылы IAI (Industry Alliance for Interoperability) альянсын құру үшін тіркелді. IFC форматы әртүрлі CAD жүйелері арасында интероперабельділікті қамтамасыз ету үшін әзірленді және STEP машина жасау форматының негізінде құрылды, ол 1979 жылы CAD пайдаланушылары мен жеткізушилері тобының NIST (Ұлттық стандарттар және технологиялар институты) және АҚШ Қорғаныс министрлігінің қолдауымен жасалған IGES форматынан шыққан.-

Алайда, IFC-тің күрделі құрылымы, геометриялық ядроға тығыз тәуелділігі, сондай-ақ форматтың әртүрлі бағдарламалық шешімдермен жүзеге асырылуындағы айырмашылықтар практикалық қолдануда көптеген мәселелерге әкелді. IGES, STEP форматтарымен жұмыс істеген машина жасау

мамандары да дәл осындағы қызындықтарға — детализацияның жоғалуы, дәлдіктің шектелуі және аралық форматтарды пайдаланудың қажеттілігіне тап болды.

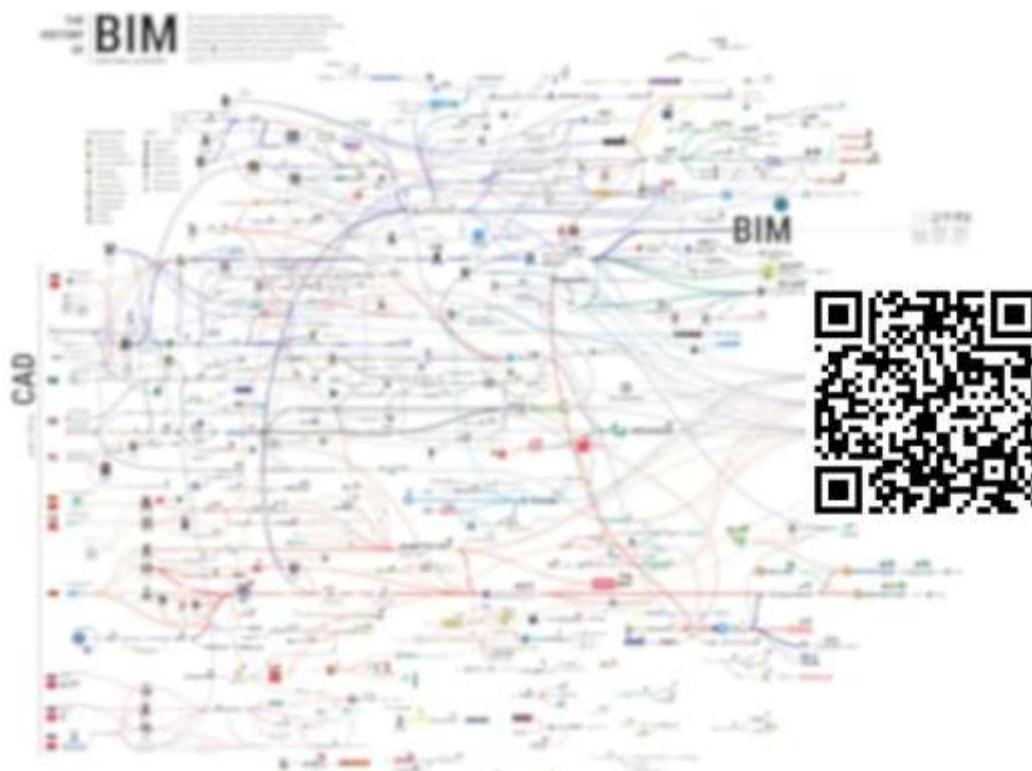


Рис. 6.14 CAD (BIM) әзірлеушілер мен өнімдер командаларының байланыс картасы.

2000 жылы IFC форматын тіркеген және IAI (кейін bS) ұйымын құрған CAD жеткізуши «Интеграцияланған жобалау және өндіріс: Артықшылықтар мен негіздеме» атты ақ қағазды жариялады. Бұл құжатта бағдарламалар арасында деректерді алмасу кезінде толық детализацияны сақтау маңыздылығы атап көрсетілді, нейтралды форматтарды, мысалы, IGES, STEP (IFC-ке үксас) пайдаланбай. Оның орнына CAD негізгі деректер базасына қосымшалардың тікелей қол жеткізуін қамтамасыз ету ұсынылды, бұл ақпараттың дәлдігін жоғалтуға жол бермеуі тиіс еді.

2002 жылы осы CAD вендоры параметрлік ВОМ өнімін сатып алады (Сурет 3.118, үшінші бөлімде толығырақ) және оның негізінде BIM концепциясын қалыптастырады. Нәтижесінде құрылышы жобаларының деректер алмасуында тек жабық CAD форматтары немесе IFC (STEP) форматы қолданылады, оның шектеулері туралы CAD-вендор 2000 жылы жазған, бұл формат құрылыш саласына енгізілген.-

Құрылыш деректерін жасау және өңдеу құралдарын жасауда қатысқан 700-ден астам әзірлеуші командастың өзара әрекеттесуінің егжей-тегжейлі тарихы «CAD (BIM) эволюциясы» картасында ұсынылған [116].

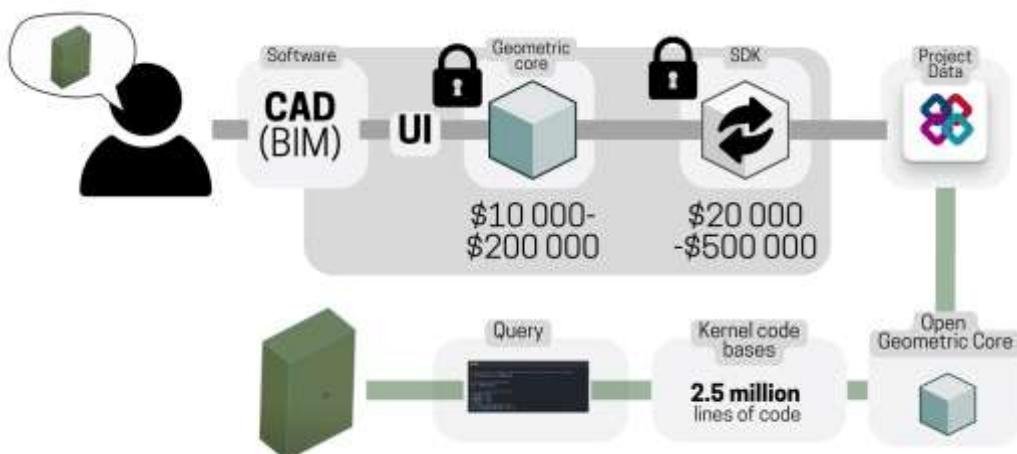
IFC ашық формасы жобаның әлементтерінің геометриялық сипаттамасынан және метаақпарат сипаттамасынан тұрады. IFC форматында геометрияны көрсету үшін CSG және Swept Solids сияқты

әртүрлі тәсілдер қолданылады: алайда BREP параметрлік көрсетілімі IFC форматында элементтердің геометриясын беру үшін жетекші стандартқа айналды, себебі мұндай формат CAD- (BIM)- бағдарламалардан экспорттау кезінде қолдау көрсетіледі және IFC-ны CAD бағдарламаларына көрі импорттау кезінде элементтерді потенциалды түрде өңдеуге мүмкіндік береді.

IFC форматының геометриялық ядроға тәуелділігі мәселесі

Көп жағдайда, егер IFC-да геометрия параметрлік түрде (BREP) берілсе, жобаның объектілерінің геометриялық қасиеттерін, мысалы, көлемін немесе ауданын визуализациялау немесе алу мүмкін болмайды, себебі тек IFC файлы болған жағдайда геометриямен жұмыс істей және оны визуализациялау үшін геометриялық ядро қажет болады (Сурет 6.15), ол бастапқыда жоқ.-

Геометриялық ядро – бұл CAD (BIM), BIM және басқа инженерлік қосымшаларда геометриялық объектілерді жасау, өңдеу және талдау үшін негізгі алгоритмдерді қамтама-сыз ететін бағдарламалық компонент. Ол 2D және 3D геометрияны құрумен қатар, оған қатысты операцияларды, мысалы, булевы операциялар, тегістеу, қырылышу, трансфор-мациялар және визуализациямен айналысады.



Сурет 6.15 Бүгінгі күні CAD бағдарламалары арқылы геометрияны жасау проприетарлық геометриялық ядролар мен SDK арқылы жүзеге асырылады, олар көбінесе CAD-вендорларға тиесілі емес.

Әр CAD бағдарламасы мен параметрлік форматтар немесе IFC форматы бойынша жұмыс істейтін кез келген бағдарламаның өзіне тән немесе сатып алынған геометриялық ядросы бар. Егер IFC-BREP форматындағы примитивті элементтермен проблемалар туындаласа, және әртүрлі геометриялық ядролармен бағдарламаларда бұл элементтер үқсас түрде көрсетілсе, геометриялық ядролардың әртүрлі қозғалтқыштарымен проблемалардан басқа, дұрыс көрсету үшін өз ерекшеліктері бар элементтер жеткілікті. Бұл мәселе 2019 жылы жарияланған «IFC бағдарламалық қамтама-сыз етуін қолдау бойынша референс зерттеу» халықаралық зерттеуінде егжей-тегжейлі қарастырылған [117].

Бірдей стандартталған деректер жиынтығы қарама-қайшы нәтижелер береді, при этом жалпы заңдылықтардың аздығы байқалады, және стандартты [IFC] қолдауында елеулі проблемалар табылды, мүмкін, стандартты модель деректерінің жоғары күрделілігіне байланысты. Мұнда стандарттардың өздері де кінәлі, себебі олар жиі кейбір мәліметтерді анықтамай, жоғары дәрежеде еркіндік пен әртүрлі мүмкін интерпретацияларды қалдырады. Олар обьектілерді ұйымдастыру мен сақтау кезінде жоғары күрделілікті қабылдайды, бұл тиімді әмбебап түсінуге, бірегей жүзеге асыруларға және деректерді келісілген модельдеуге ықпал етпейді [117]. – IFC бағдарламалық қамтамасыз етуін қолдау бойынша референс зерттеу, 2021.

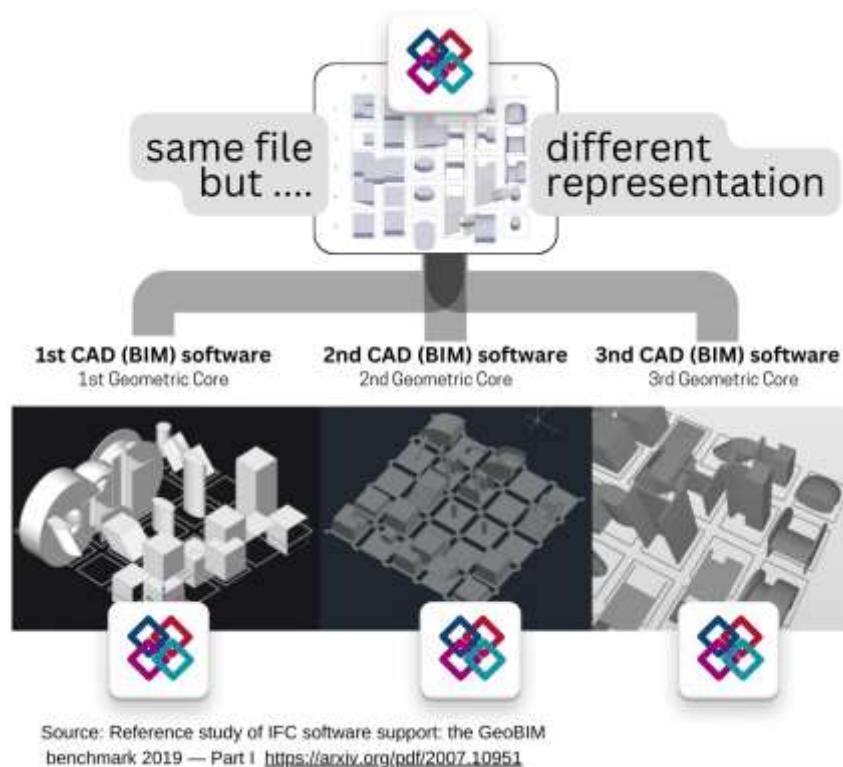


Рис. 6.16 Әртүрлі геометриялық ядролар бір және сол геометрияны параметрлік түрде сипаттауда әртүрлі көрініс береді, материалдар [117] негізінде.

“Белгілі жағдайларды” дұрыс түсіну тек IFC әзірлеумен айналысадын арнағы ұйымдардың ақылы мүшелеріне қолжетімді. Соның нәтижесінде, IFC-нің белгілі ерекшеліктері туралы маңызды білім алғысы келетіндер ірі CAD жеткізушилерімен ынтымақтастық орнатуға тырысады немесе өз зерттеулері арқылы ерекшеліктерді сапалы есепке алуға жетуге тырысады.

Сіз IFC форматы арқылы деректерді импорттау және экспорттау мәселесіне тап боласыз және әріптес жеткізушилерден: “Неліктен IFC файлына параметрлік бөлмелерді беру туралы ақпарат осылай беріледі? Ашық спецификацияда бұл туралы ештеңе айтылмаған” деп сұрайсыз. “Көбірек білетін” европалық жеткізушилердің жауабы: “Иә, айтылмаған, бірақ рұқсат етілген”. – CAD әзірлеушісімен 2021 жылғы сұхбаттан [118].

IFC геометрияны параметрлік примитивтер арқылы сипаттайды, бірақ оның құрамында кіріктірілген ядро жоқ – бұл рөлді CAD бағдарламасы атқарады, ол геометриялық ядро арқылы геометрияны компиляциялады. Геометриялық ядро математикалық есептеулерді жүргізеді және қылыштарды анықтайды, ал IFC тек оның интерпретациясы үшін деректерді ұсынады. Егер IFC-де дұрыс емес қырлар болса, әртүрлі геометриялық ядролары бар бағдарламалар оларды не елемеуі, не қателіктер көрсетуі мүмкін, ядроға байланысты.

Нәтижесінде, IFC форматымен жұмыс істеу үшін негізгі сұраққа жауап беру қажет, оған нақты жауап табу қыын - қандай құрал, қандай геометриялық ядромен пайдалану көрек, жобаның бастапқы CAD бағдарламасында болған деректердің сапасын алу үшін?

Деректер сапасының мәселелері мен IFC форматының күрделілігі жобалық деректерді автоматтандыру, талдау және өңдеу процестерінде тікелей пайдалануға мүмкіндік бермейді, бұл әзірлеушілерді деректерге “сапалы” қолжетімділікпен жабық CAD шешімдерін пайдалануға мәжбүр етеді [63], бұл туралы 1994 жылы IFC-ны тіркеген жеткізуши де жазған [65].

IFC параметрлерін геометриялық ядроға көрсету және генерациялау ерекшеліктері тек геометриялық ядролармен жұмыс тәжірибесі бар үлкен әзірлеуші командалармен жүзеге асырылуы мүмкін. Соңдықтан IFC форматының ерекшеліктері мен күрделілігі қазіргі практикада негізінен CAD жеткізушилеріне тиімді және “adopt, extend, destroy” бағдарламалық қамтамасыз ету жеткізушилерінің стратегиясымен көп ұқсастықтары бар, өйткені стандарттың өскен күрделілігі фактісінде нарықтағы шағын ойыншылар үшін кедергілерді тудырады [94].

Мұндай стратегиядағы ірі жеткізушилердің стратегиясы ашық стандарттарды бейімдеу, өздерінің кеңейтулері мен функцияларын қосу арқылы пайдаланушыларды өз өнімдеріне тәуелді ету және бәсекелестерді ығыстыру болуы мүмкін.

IFC форматы, әртүрлі CAD- (BIM-) жүйелері арасында әмбебап көпір болуға арналған, шын мәнінде әртүрлі CAD платформаларының геометриялық ядролары арасындағы үйлесімділік мәселелерінің индикаторы ретінде қызмет етеді, STEP форматы сияқты, ол бастапқыда пайда болған.

Нәтижесінде, бүгінгі таңда IFC онтологиясын толыққанды және сапалы жүзеге асыру ірі CAD жеткізушилерінің қолынан келеді, олар барлық сущностарды және олардың ішкі геометриялық ядро-сымен маппингін қолдануға елеулі ресурстар сала алады, ал IFC стандарты үшін мұндай ядро жоқ. Ирі вендорлар техникалық ерекшеліктердің егжей-тегжейлерін келісу мүмкіндігіне ие, бұл тіпті IFC форматымен айналысатын ұйымдардың ең белсенді қатысушысына да қолжетімді болмауы мүмкін.

Өздерінің геометриялық ядросы жоқ шағын тәуелсіз командалар мен ашық кодты жобалар үшін интероперабельді форматтарды дамытуға ұмтылу үлкен мәселе болып табылады. Онсыз деректерді платформалар арасында алмасу кезінде барлық нәзік және ерекшеліктерді ескеру практикалық түрфыдан мүмкін емес.

IFC параметрлік форматы мен open BIM концепциясының дамуы барысында құрылыш саласында деректер мен процестерді басқаруда онтология мен семантиканың рөлі туралы пікірталастар белсенді түрде жүріп жатыр.

Құрылыш саласында семантика және онтология тақырыбының пайда болуы

1990-шы жылдардың соңында семантикалық интернетті құру идеялары мен IFC форматты дамытуға арналған ұйымдардың күш-жігерінің арқасында семантика мен онтология құрылыш саласында стандарттаудың негізгі элементтерінің бірі ретінде 2020-шы жылдардың ортасына қарай талқылануда.

Семантикалық технологиялар – бұл әртүрлі деректер массивтерін унификациялау, стандарттау және модификациялау, сондай-ақ күрделі іздеуді жүзеге асыру.

Семантикалық деректерді сақтау үшін OWL (Web Ontology Language) онтологиялар тілі қолданылады, ол RDF-триплеттер (Resource Description Framework) графтары түрінде ұсынылады. OWL деректердің графтық модельдеріне жатады, олардың түрлері туралы біз «Деректер модельдері: деректердегі қатынастар мен элементтер арасындағы байланыстар» бөлімінде толығырақ айттық.-

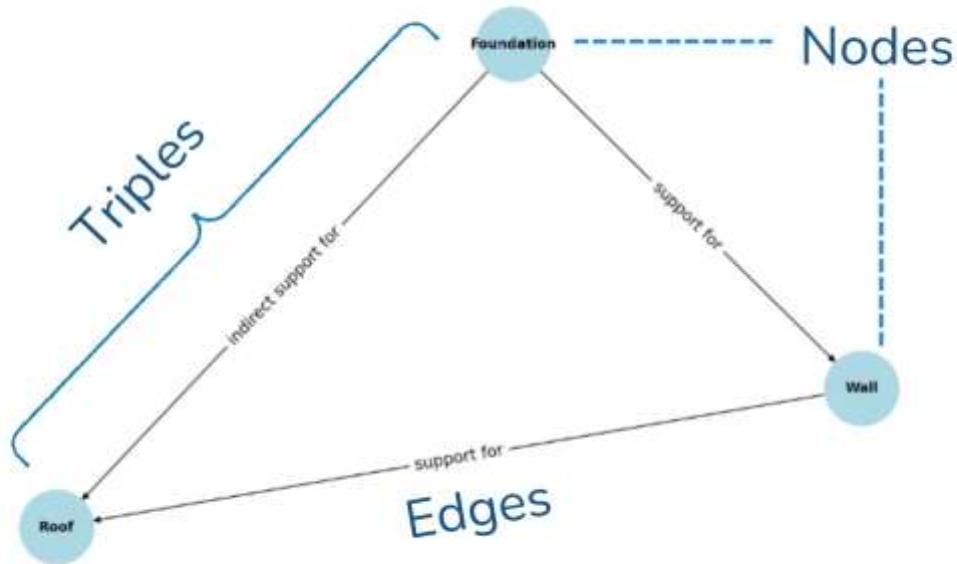


Рис. 6.17 RDF деректер моделі: құрылыш элементтері арасындағы қатынастарды иллюстрациялайтын түйіндер (Nodes), байланыстар (Edges) және тройкалар (Triples).

Теориялық түрғыдан, логикалық шығару ризонерлері (автоматты логикалық шығару бағдарламалары) онтологияларға негізделген жаңа мәлімдемелер алуға мүмкіндік береді. Мысалы, егер құрылыш онтологиясында "негіз – бүл қабырға опора" және "қабырға – бүл шатырға опора" деп жазылса, ризонер автоматты түрде "негіз – бүл шатырға опора" деген қорытынды жасай алады.-

Мұндай механизм деректерді талдауды оңтайландыру үшін пайдалы, себебі ол барлық тәуелділіктерді ашық түрде жазып шығуды болдырмайды. Алайда, ол жаңа білімді жасамайды, тек бұрыннан белгілі фактілерді анықтап, құрылымдайды.

Семантика өздігінен жаңа мағына немесе білімді жасамайды және бүл аспектте деректерді сақтау мен өндеудің басқа технологияларынан асып кетпейді. Реляциялық базалардан деректерді триплеттер түрінде ұсыну оларды мағыналы етпейді. Кестелерді графтық құрылымдарға ауыстыру деректер модельдерін унификациялау, ыңғайлыш іздеу және қауіпсіз редактирлеу үшін пайдалы болуы мүмкін, бірақ бүл деректерді "ақылды" етпейді – компьютер олардың мазмұнын жақсырақ түсінбейді.

Деректердегі логикалық байланыстарды күрделі семантикалық технологияларды қолданбай-ақ ұйымдастыруға болады (Рис. 6.18). Дәстүрлі реляциялық деректер базалары (SQL), сондай-ақ CSV немесе XLSX форматтары ұқсас тәуелділіктерді құруға мүмкіндік береді. Мысалы, бағандық деректер базасында "шатыр тірегі" деген өрісті қосып, қабырға жасалған кезде шатырды фундаментпен автоматты түрде байланыстыруға болады. Мұндай тәсіл RDF, OWL, графтар немесе ризонерлерді пайдаланбай-ақ жүзеге асырылады, деректерді сақтау мен талдауда қарапайым әрі тиімді шешім болып табылады.

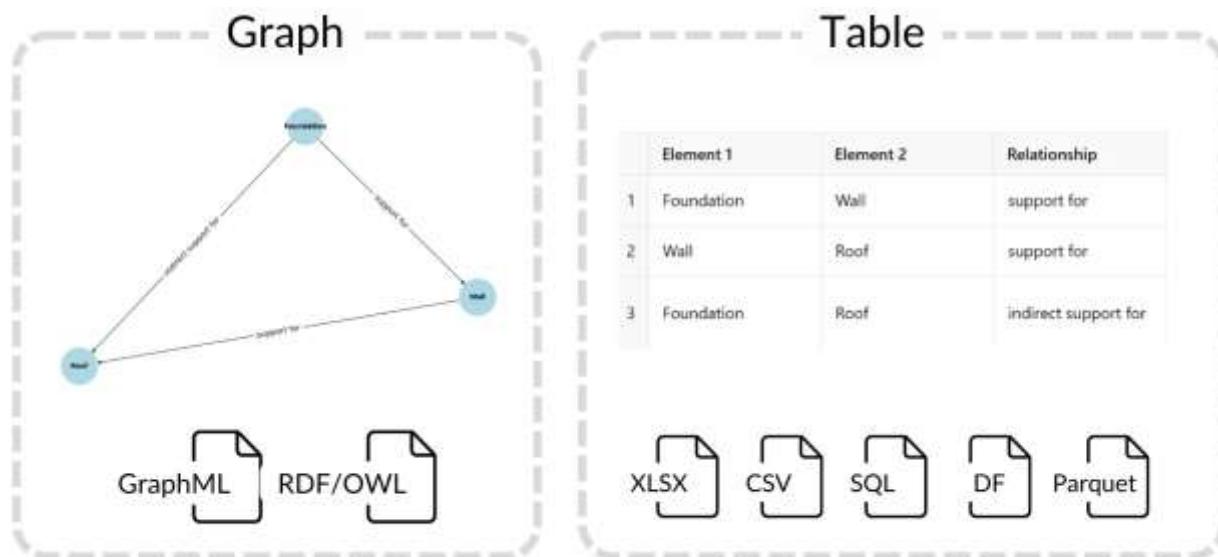


Рис. 6.18 Графтық және кестелік деректер модельдерінің бірдей логикалық байланыстарды қөрсету салыстырмасы.

Бірқатар ірі құрылыш компаниялары мен IFC форматтарын әзірлеумен айналысатын үйым семантикалық веб концепциясын, 1990-шы жылдардың соңында перспективалы болып көрінген, қабылдауды құрылыш саласындағы стандарттарды әзірлеуге елеулі әсер етті.

Алайда парадокс мынада, семантикалық веб концепциясы, бастапқыда интернет үшін арналған, тіпті өзінің туған ортасында да кеңінен таралмады. RDF және OWL әзірленгеніне қарамастан, толыққанды семантикалық веб бастапқы жоспар бойынша пайдаланған жоқ, және оның құрылуы енді ықтимал емес.

Неліктен семантикалық технологиялар құрылыш саласында күтілген нәтижелерді ақтамайды

Басқа салалар семантиканы пайдалану технологияларының шектеулеріне тап болды. Ойын индустриясында ойын объектілерін және олардың өзара әрекеттерін онтологиялар арқылы сипаттау әрекеттері өзгерістердің жоғары динамикасы себебінен тиімді болмады. Нәтижесінде, XML және JSON сияқты қарапайым деректер форматтары алгоритмдік шешімдермен бірге артықшылықа ие болды. Мұндай жағдай жылжымайтын мүлік саласында да орын алды: терминологиядағы аймақтық айырмашылықтар мен нарықтағы жиі өзгерістердің салдарынан онтологияларды пайдалану тым күрделі болып шықты, ал қарапайым деректер базалары мен RETS сияқты стандарттар деректер алmasу міндеттерін жақсы орындады.

Техникалық қындықтар, мысалы, разметканың күрделілігі, қолдаудың жоғары еңбек сыйымдылығы және әзірлеушілердің төмен мотивациясы семантикалық вебтің басқа экономикалық секторларда енгізілуін тежеді. RDF (Resource Description Framework) жаппай стандартқа айналмады, ал онтологиялар тым күрделі және экономикалық түрғыдан тиімсіз болып шықты.

Нәтижесінде, жаһандық семантикалық вебті құру идеясы жүзеге аспады. Дегенмен, онтологиялар мен SPARQL сияқты технологияның жекелеген элементтері корпоративтік шешімдерде қолданылды, ал деректердің біртұтас құрылымын құрудың бастапқы мақсаты орындалмады.

Компьютерлер контенттің мағынасын түсінетін интернет концепциясы техникалық түрғыдан күрделі және коммерциялық түрғыдан тиімсіз болып шықты. Сондықтан, бұл идеяны қолдаған компаниялар уақыт өте келе оны жекелеген пайдалы құралдарға дейін қысқартты, RDF және OWL-ды жалпы интернет үшін емес, тар мамандандырылған корпоративтік қажеттіліктер үшін қалдырыды. Соңғы 20 жылдағы Google Trends талдауы (Рис. 6.19) семантикалық вебтің даму перспектиналары, мүмкін, енді қалмағанын бағалауға мүмкіндік береді.-

Қажетсіз сущностарды көбейтудің қажеті жоқ. Егер қандай да бір құбылысты түсіндіретін бірнеше логикалық түрғыдан қарама-қайши емес түсініктемелер болса, оларды бірдей жақсы түсіндіретін болса, онда тең жағдайларда ең қарапайымын таңдау керек.— Оккамның бритвасы.

Мұнда логикалық сұрақ туындаиды: неге үштік, ризаонерлер және SPARQL-ды құрылыс саласында қолдану керек, егер деректерді танымал құрылымдық сұраулар (SQL, Pandas, Apache®) арқылы өңдеуге болады? Корпоративтік қосымшаларда SQL деректер базасымен жұмыс істеудің стандартты әдісі болып табылады. SPARQL, керінше, күрделі графтық құрылымдар мен мамандандырылған бағдарламалық қамтамасыз етуді талап етеді және Google-дағы трендтер бойынша әзірлеушілердің қызығушылығын тартпайды.

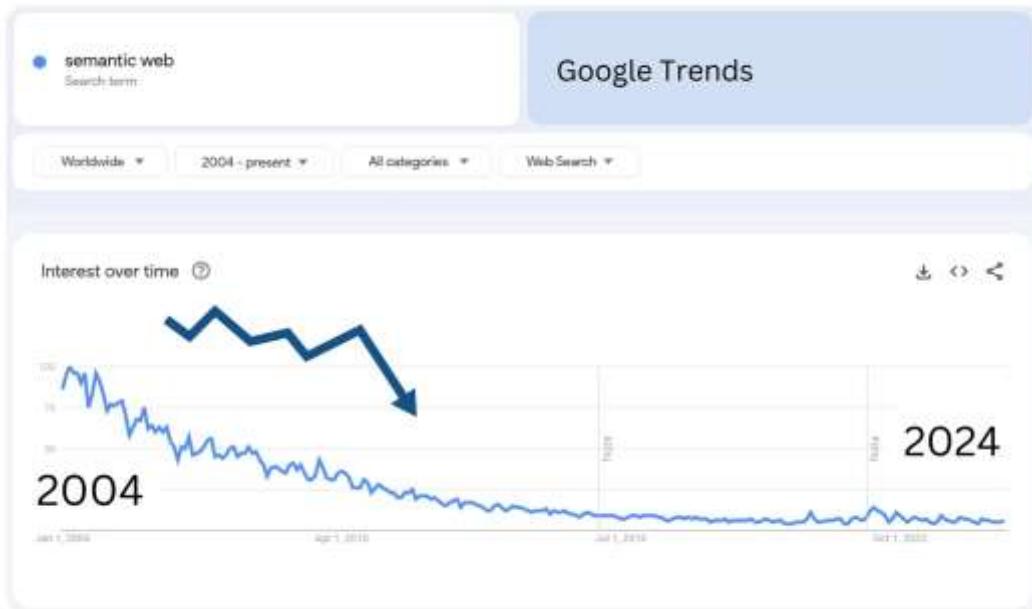


Рис. 6.19 «Семантикалық интернет» сұрауларына қызығушылық Google статистикасы бойынша.

Графтық деректер базалары мен класификация ағаштары кейбір жағдайларда пайдалы болуы мүмкін, бірақ олардың қолданылуы күнделікті тапсырмалардың көпшілігі үшін әрдайым негізделмеліді. Нәтижесінде, білім графтарын құру және семантикалық веб технологияларын пайдалану тек

әртүрлі көздерден деректерді унификациялау немесе күрделі логикалық қорытындыларды жүзеге асыру қажет болғанда ғана мағынаға ие.

Кестелерден графтық деректер модельдеріне көшу ақпаратты іздеуді жақсартуға және ақпарат ағынын унификациялауға мүмкіндік береді, бірақ деректерді машиналар үшін мағыналы етпейді. Сұрақ семантикалық технологияларды қолдану қажеттігінде емес, олар қай жерде шын мәнінде пайда әкелетініне байланысты. Онтология, семантика және графтық деректер базаларын компанияның енгізбестен бұрын, бұл технологияларды сәтті қолданған компанияларды және оларды қолданудағы сәтсіздіктерді анықтаңыз.

Амбициозды күтулерге қарамастан, семантикалық технологиялар құрылыш саласында деректерді құрылымдаудың әмбебап шешіміне айналмады. Практикада бұл технологиялар әмбебап шешімге әкелмей, тек жаңа қындықтарды қосып, семантикалық интернет концепцияның жүзеге аспаған амбицияларын қайталауда, мұнда күтулер шындықты едәүір асып түсті.



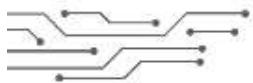
Рис. 6.110 Құрылыш процестеріндегі геометрия және ақпарат: күрделі CAD және BIM жүйелерінен аналитика үшін қарапайым деректерге дейін.

IT саласындағы семантикалық вебтің сәтсіздіктері жаңа технологиялардың (улкен деректер, IoT, машиналық оқыту, AR/VR) пайда болуымен өтелсе, құрылыш саласында мұндай себептер жоқ.

Жобаның элементтері арасындағы деректердің өзара байланысын беру үшін концепцияларды қолданудағы проблемалардан басқа, деректердің қолжетімділігі мәселесі де сақталады. Құрылыш саласында әлі күнге дейін жабық жүйелер басым, бұл деректермен жұмыс істеуді, ақпарат алмасуды және процестердің тиімділігін арттыруды қынданатады.

Деректердің жабық табиғаты құрылыш саласындағы цифровық шешімдердің дамуына кедергі келтіретін негізгі кедергілердің бірі болып табылады. IT индустриясынан айырмашылығы, мұнда

ашық және унификацияланған деректер форматтары стандартқа айналса, CAD (BIM) секторында әрбір бағдарламалық қамтамасыз ету өз форматын пайдаланады, бұл жабық экожүйелерді құрып, пайдаланушыларды жасанды түрде шектейді.



ТАРАУ 6.2. ЖАБЫҚ ЖОБА ФОРМАТТАРЫ ЖӘНЕ ИНТЕРОПЕРАБЕЛІК МӘСЕЛЕЛЕР

Жабық деректер және өнімділіктің төмендеуі: CAD (BIM) саласының дағдарысы

CAD жүйелерінің меншікті табиғаты әрбір бағдарламаның өз деректер форматына ие болуына әкелді, ол не жабық және сырттан қолжетімсіз - RVT, PLN, DWG, NDW, NWD, SKP, не полуструктурированная форма, жеткілікті күрделі конверсиялау процесі арқылы қолжетімді - JSON, XML (CPIXML), IFC, STEP және ifcXML, IfcJSON, BIMJSON, IfcSQL, CSV және т.б.

Әртүрлі деректер форматтары, бірдей жобалар туралы бірдей деректерді сақтау үшін қолданылтын, тек құрылымымен ғана емес, сонымен қатар әзірлеушілердің қолданбалардың үйлесімділігін қамтамасыз етуі үшін ескеруі қажет ішкі белгілеудің әртүрлі нұсқаларын да қамтиды. Мысалы, 2025 жылғы CAD форматы 2026 жылғы CAD бағдарламасында ашылады, бірақ осы жобаны 2025 жылға дейінгі CAD бағдарламаларының барлық нұсқаларында ашу мүмкін емес.

Деректер базаларына тікелей қолжетімділік бермей, құрылыш саласындағы бағдарламалық қамтамасыз ету жеткізуіші жиі өзіне тән ерекше формат пен оған арналған құралдарды жасайды, оларды маман (инженер-жобалаушы немесе деректер менеджери) деректерге қол жеткізу, импорттау және экспорттау үшін пайдалану қажет.

Нәтижесінде, базалық CAD (BIM) және онымен байланысты шешімдердің (мысалы, ERP/PMIS) жеткізуішілері өнімдерді пайдалану бағасын үнемі көтеріп отырады, ал қарапайым пайдаланушылар деректерді форматтар бойынша беру кезеңінде "комиссия" төлеуге мәжбүр: қосылу, импорттау, экспорттау және пайдаланушылар өздері жасаған деректермен жұмыс істеу үшін.

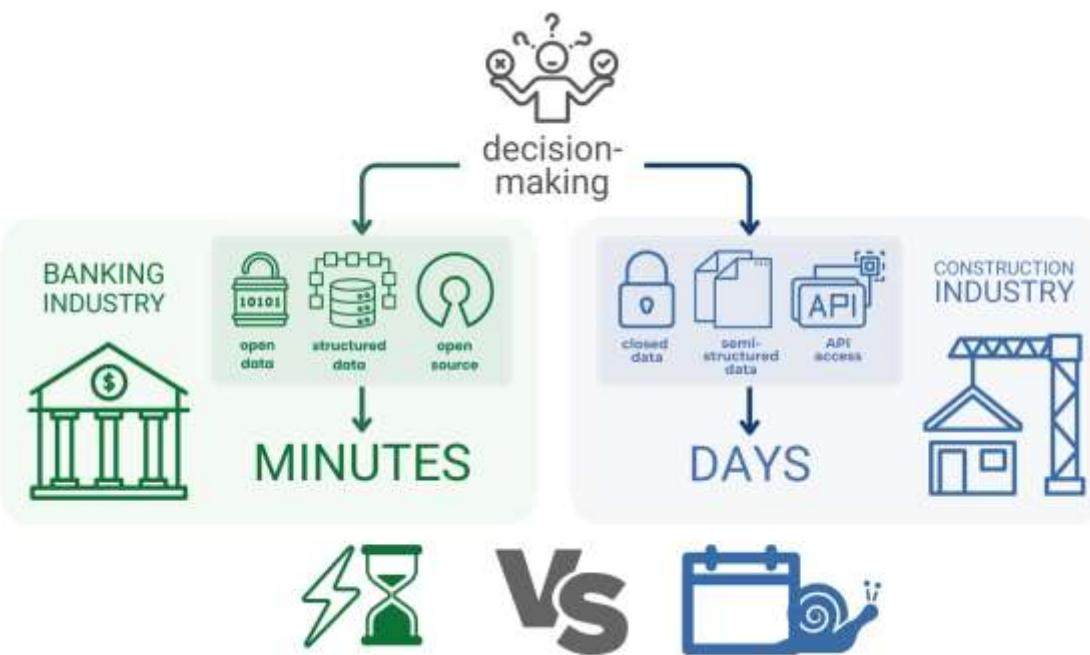
2025 жылды таныма CAD- (BIM-) өнімдеріндегі бұлтты сақтау жүйесіне қолжетімділік құны транзакция үшін 1 долларды құрайды, ал орташа компаниялар үшін құрылыш ERP өнімдеріне жазылым жылына бес- және алты таңбалы сандарға жетеді.

Қазіргі құрылыш бағдарламалық қамтамасыз етуінің мәні автоматизация немесе тиімділікті арттыру емес, инженерлердің нақты тар мамандандырылған бағдарламалық қамтамасыз етуді түсіну қабілеті құрылыш жобасының деректерін өңдеу сапасы мен құнына, сондай-ақ құрылыш саласын жүзеге асыратын компаниялардың пайdasы мен ұзак мерзімді өміршенідігіне әсер етеді.

CAD жүйелерінің деректер базаларына қолжетімділіктің болмауы, олар ондаған басқа жүйелер мен жүздеген процестерде қолданылады, және, соның салдарынан, жеке мамандар арасында сапалы коммуникацияның болмауы құрылыш саласын экономиканың ең тиімді емес секторларының біріне айналдырды.

Соңғы 20 жыл ішінде CAD- (BIM-) жобалаудың, жаңа жүйелердің (ERP), жаңа құрылыш технологиялары мен материалдарының пайда болуымен, құрылыш саласының жалпы өнімділігі 20%-ға төмендеді, ал деректер базаларына қолжетімділік мәселелері жоқ барлық экономикалық секторлардың

жалпы өнімділігі 70%-ға (өндөу өнеркәсібінде 96%-ға) өсті.-



Жобалық деректердің оқшаулануы мен күрделілігінің салдарынан, құрылыш саласында ондаған бөлімдер мен жүздеген процестерге тәуелді, шешім қабылдау жылдамдығы басқа салалармен салыстырғанда бірнеше есе төмен.

Дегенмен, CAD шешімдері арасында интероперабельділікті қамтамасыз етудің баламалы тәсілдерінің жекелеген мысалдары да бар. 2018 жылы басталған SCOPE жобасымен Еуропаның ең ірі құрылыш компаниясы классикалық CAD- (BIM-) жүйелерінің логикасынан тыс шығудың жолын көрсетеді. IFC-ді бағындыруға немесе меншікті геометриялық ядроларға сүйенуге тырыспай, SCOPE әзірлеушілері әртүрлі CAD бағдарламаларынан деректерді алу үшін API және SDK кері инженериясын пайдаланады, оларды OBJ немесе CPIXML сияқты нейтралды форматтарға түрлендіреді, біргей Open Source геометриялық ядро OCCT негізінде, және оларды құрылыш және жобалау компанияларының жүздеген бизнес-процестерінде қолданады. Дегенмен, идеяның прогрессивтілігіне қарамастан, мұндай жобалар тегін геометриялық ядролардың шектеулері мен күрделілігіне тап болады және олар әлі де бір компанияның жабық экожүйесінің бөлігі болып қалады, бұл моновендорлық шешімдердің логикасын қайта жаңғыртады.

Жабық жүйелердің шектеулері мен деректер форматтарының айырмашылықтары, сондай-ақ олардың унификациясы үшін тиімді құралдардың болмауы салдарынан CAD форматтарымен жұмыс істейтін компаниялар әртүрлі деңгейдегі құрылымдылық пен жабықтықтағы елеулі деректер көлемінің жинақталуымен бетпе-бет келеді. Бұл деректер тиісті түрде пайдаланылмай, архивтерде жоғалып, уақыт өте келе мәңгілікке ұмытылып, пайдаланусыз қалады.

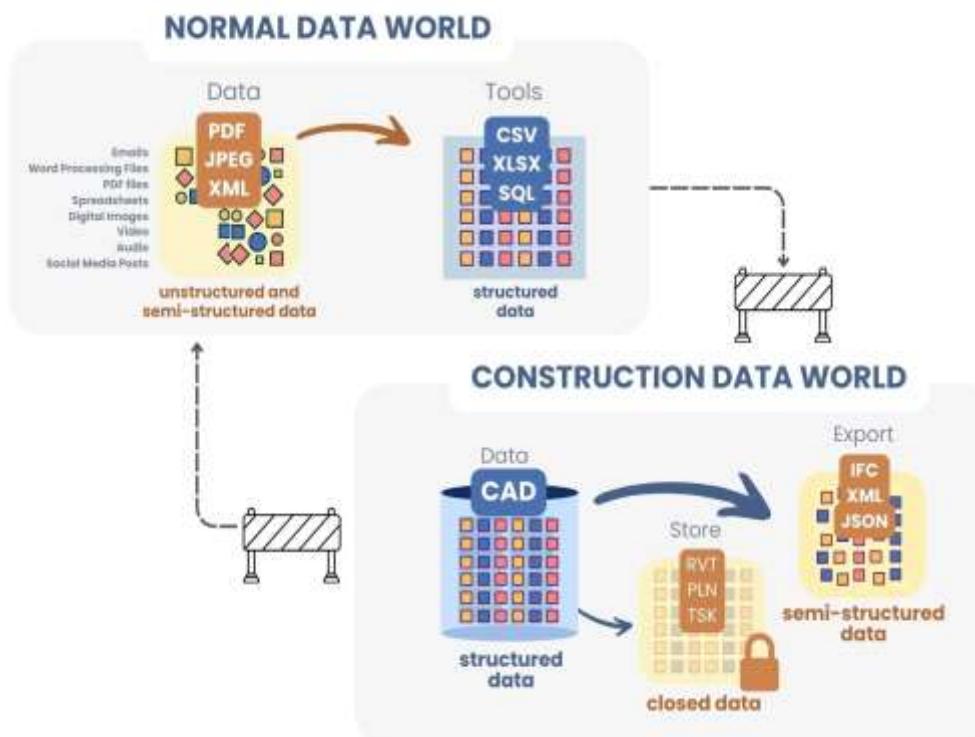
Жобалау кезеңінде елеулі күш-жігермен алынған деректер, олардың күрделілігі мен жабықтығы салдарынан, әрі қарай пайдалануға қолжетімсіз болып қалады.

Нәтижесінде, соңғы 30 жыл бойы құрылыш саласындағы әзірлеушілер бір және сол мәселені қайта-қайта шешуге мәжбүр: әр жаңа жабық формат немесе меншікті шешім бар ашық жабық CAD

жүйелерімен интеграция қажеттілігін туындалады. Әртүрлі CAD- және BIM-шешімдері арасында интероперабельділікті қамтамасыз етуге бағытталған тұрақты әрекеттер деректер экожүйесін күрделендіруден басқа, оның қарапайымдылығы мен стандартталуына ықпал етпейді.

CAD жүйелері арасындағы интероперабельділік туралы миф

1990-шы жылдардың ортасында CAD ортасындағы интероперабельділікті дамытудағы негізгі бағыт DWG меншікті форматын бұзы болды – бұл Open DWG альянсының жеңісімен аяқталып, құрылыш саласы үшін ең танымал сызба форматын фактілік түрде ашты, – ал 2020-шы жылдардың ортасына қарай акцент өзгерді. Құрылыш индустриясында жабық CAD жүйелері (closed BIM), IFC форматы және ашық шешімдер (open BIM) арасында «көпірлер» құруға бағытталған көптеген әзірлеуші командалар пайда болды. Мұндай бастамалардың көпшілігінің негізінде IFC форматы мен ОССТ геометриялық ядросын пайдалану жатыр, бұл әртүрлі платформалар арасында техникалық байланыс орнатуға мүмкіндік береді. Бұл тәсіл деректер алмасуды едәуір жақсартуға және бағдарламалық құралдардың үйлесімділігін арттыруға қабілетті перспективалық бағыт ретінде қарастырылады.



Басқа салалар ашық деректермен жұмыс істегенде, құрылыш саласы жабық немесе нашар құрылымдалған CAD (BIM) форматтарымен жұмыс істеуге мәжбүр.

Мұндай тәсілдің тарихи параллельдері бар. 2000-жылдары әзірлеушілер, ең ірі графикалық редакторлар вендорының (2D әлемі) үстемдігін жеңу үшін, оның меншікті шешімі мен тегін Open Source баламасы GIMP арасында үздіксіз интеграция жасауға тырысты. Сол кезде де, бұғынгі құрылыш саласында да, жабық және ашық жүйелерді біріктіру, сонымен қатар бағдарламалық қамтамасыз етудің күрделі параметрлері, қабаттары мен ішкі логикасын сақтау туралы сөз болды.-

Алайда, пайдаланушылар шын мәнінде қарапайым шешімдерді іздеді – тегіс, ашық деректер, бағдарламалардың қабаттары мен параметрлерінің артық күрделілігінен босатылған (CAD-дағы геометриялық ядроға аналог). Пайдаланушылар артық логикадан босатылған қарапайым және ашық деректер форматтарына үмтүлді. Графикада мұндай форматтар JPEG, PNG және GIF болды. Бұғынгі күні олар әлеуметтік желілерде, веб-сайттарда, қосымшаларда қолданылады – олар платформалар мен бағдарламалық қамтамасыз етушілерге қарамастан оңай өндөледі және интерпретацияланады.

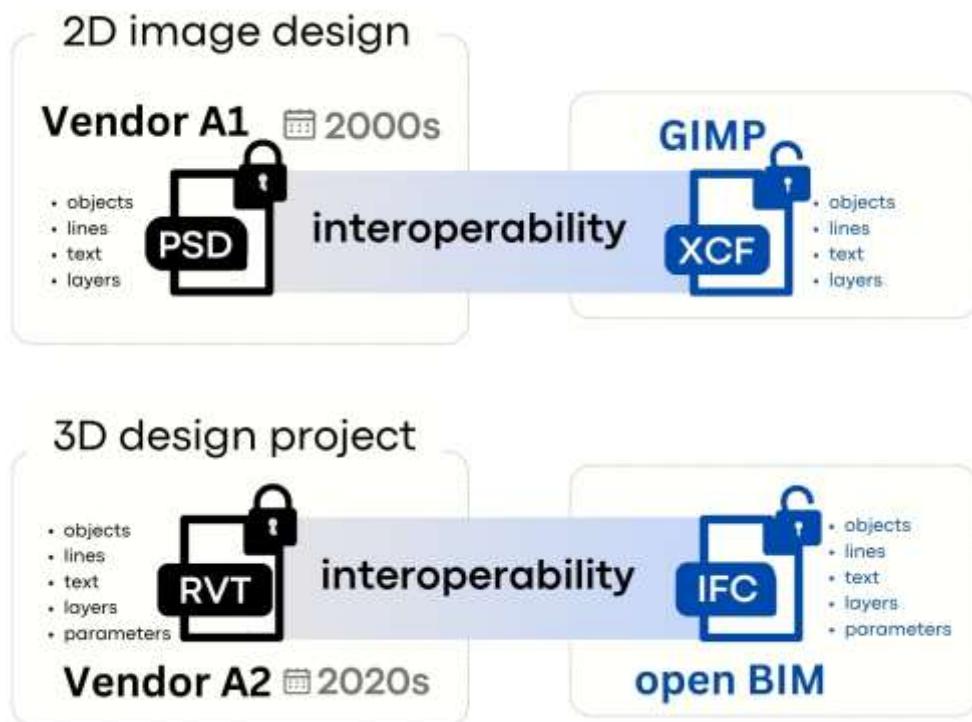


Рис. 6.23 Деректер форматтарының құрылыш саласындағы өзара алмастырылуы 2000-жылдары танымал меншікті өнімді вендор мен Open Source GIMP біріктіруге тырысудан бастау алады.

Нәтижесінде, бұғынгі күні кескіндер индустриясында ешкім жабық форматтарды, мысалы, PSD немесе ашық XCF-ті Facebook және Instagram сияқты әлеуметтік желілерде немесе веб-сайттардағы контент ретінде пайдаланбайды. Оның орнына, көптеген тапсырмаларда JPEG, PNG және GIF сияқты тегіс және ашық форматтар қолданылады, оларды пайдалану оңай және кең үйлесімділікпен қамтамасыз етеді. JPEG және PNG сияқты ашық форматтар кескіндерді алмасу стандарттарына айналды, олардың әмбебаптығы мен кең қолдауы арқасында, оларды әртүрлі платформаларда пайдалануды жөнілдетеді. Осындағы өтпелі процесс басқа алмасу форматтарында да байқалады, мысалы, видео және аудио, мұнда MPEG және MP3 сияқты әмбебап форматтар қысу тиімділігі мен кең үйлесімділігімен ерекшеленеді. Стандартизацияға осындағы көшу контент пен ақпаратты алмасуды және воспроизведение-ді жөнілдетіп, оларды әртүрлі платформалардағы барлық пайдаланушылар үшін қолжетімді етті.

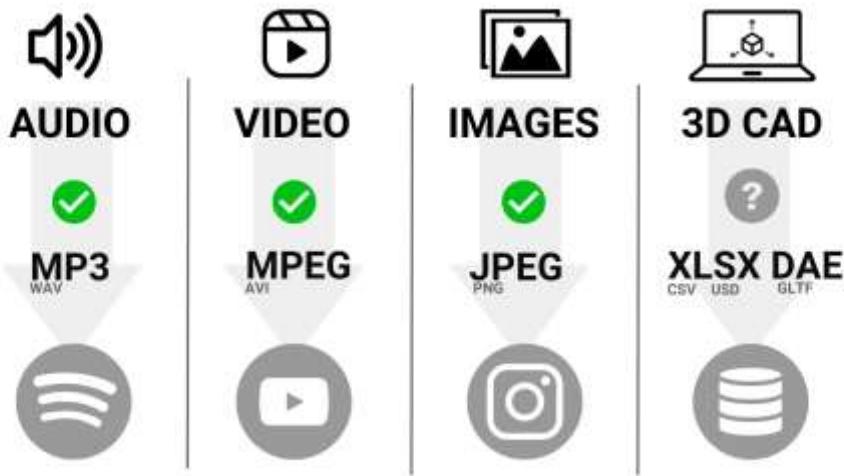


Рис. 6.24 Күрделі өндіу функциялары жоқ қарапайым форматтар деректерді алмасу және пайдалану үшін танымал болды.

3D-моделдеу саласында да үқсас процестер жүріп жатыр. USD, OBJ, glTF, DAE, DXF, SQL және XLSX сияқты қарапайым және ашық форматтар CAD (BIM) ортасынан тыс деректер алмасу жобалағында жиі қолданылады. Бұл форматтар геометрия мен метадеректерді қамтитын барлық қажетті ақпаратты сақтайды, BREP, геометриялық ядролар немесе нақты вендорлардың ішкі классификаторлары сияқты күрделі құрылымдармен жұмыс істеуді қажет етпейді. HWC, SVF, SVF2, CPIXML және CP2 сияқты жетекші бағдарламалық қамтамасыз етушілер ұсынатын меншікті форматтар да үқсас функцияларды орындаиды, бірақ ашық стандарттардан айырмашылығы, жабық болып қалады.

Ескерту керек (және бұл бұрынғы тарауда атап өткеніміздей, тағы да еске салу қажет), 2000 жылы CAD-өндіруші, 1994 жылы IFC форматында тіркелген BIM бойынша ақ қағазды жасаған, осындай идеяны – IGES, STEP және IFC сияқты аралық нейтралды және параметрлік форматтардан бас тартуды – қолдады. 2000 жылғы «Интеграцияланған жобалau және өндіріс» [65] ақ қағазында CAD-деректер базасына нативті қол жеткізуің маңыздылығын атап өтті, аралық трансляторлар мен параметрлік форматтарды пайдаланбай, ақпараттың толықтығы мен дәлдігін сақтау үшін.

Құрылыш саласына CAD деректер базаларына қол жеткізу құралдары туралы келісу, немесе олардың мәжбүрлі кері инженериясы, немесе CAD (BIM) платформаларынан тыс пайдалануға арналған жалпы қарапайым деректер форматын қабылдау қажет. Мысалы, Орталық Еуропа мен неміс тілдес аймақтардағы көптеген ірі компаниялар құрылыш секторында ERP жүйелерінде CPIXML форматында жұмыс істейді [121]. Бұл XML-дің бір түрі болып табылатын меншікті формат, CAD (BIM) жобасының деректерін, геометриялық және метадеректерді біртұтас үйымдастыру қарапайым құрылымға біріктіреді. Сондай-ақ, ірі құрылыш компаниялары SCOPE жобасы сияқты жаңа меншікті форматтар мен жүйелерді құруда.

Параметрлік CAD форматтарының жабық логикасы немесе IFC (STEP) параметрлік файлдарының күрделілігі көптеген бизнес процестерде артық болып табылады. Пайдаланушылар BREP геометриясын құрудың артық логикасынан, геометриялық ядроларға тәуелділіктен және нақты CAD және

BIM өнімдерінің ішкі классификацияларынан босатылған, USD, CPIXML, XML&OBJ, DXF, glTF, SQLite, DAE&XLSX сияқты қарапайым және тегіс форматтарды іздейді (Рис. 6.25).-

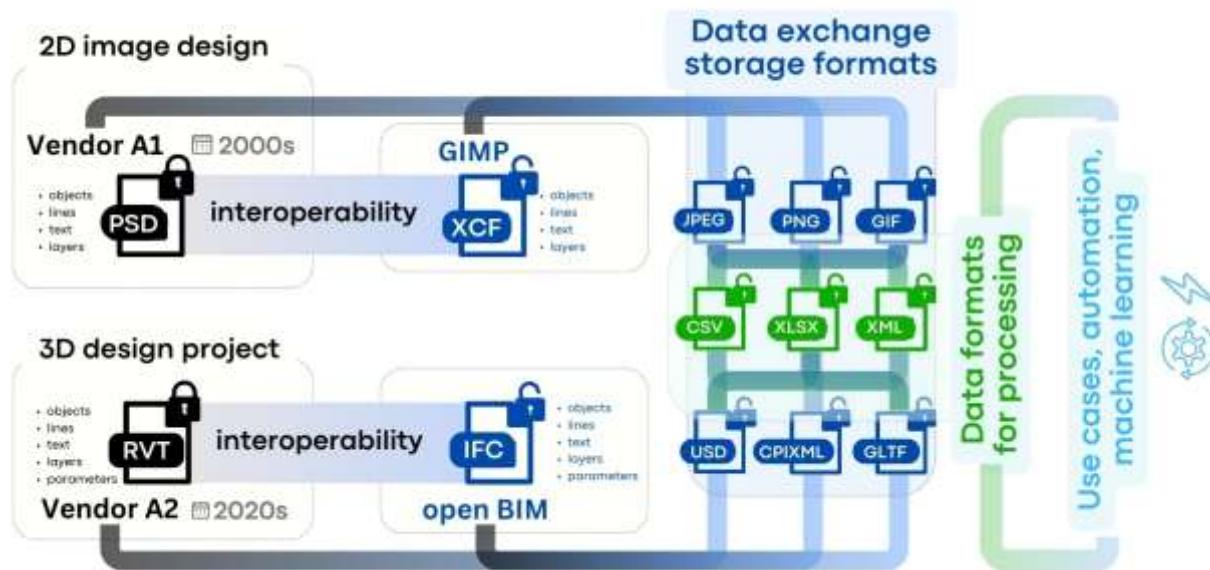


Рис. 6.25 Көптеген қолдану жағдайлары үшін пайдаланушылар вендорлық бағдарламаларға тәуелді емес, мүмкіндігінше қарапайым форматтарды таңдайды.

JPEG, PNG және GIF сияқты қарапайым кескін форматтарының пайда болуы, вендорлардың ішкі қозғалтқыштарының артық логикасынан босатылған, графиканы өңдеу және пайдалану үшін мындаған үйлесімді шешімдердің дамуына ықпал етті. Бұл әртүрлі қосымшалардың пайда болуына әкелді: ретушь және фильтрация құралдарынан Instagram, Snapchat және Canva сияқты әлеуметтік желілерге дейін, мұнда ғұл қарапайым деректерді нақты бағдарламалық жасақтама әзірлеушісіне байланбай пайдалану мүмкін болды.

CAD жобалау форматтарының стандартталуы мен қарапайымдалуы құрылыш жобаларымен жұмыс істеуге арналған көптеген жаңа ыңғайлы және тәуелсіз құралдардың пайда болуын ынталандырады.

Жабық геометриялық ядроларға негізделген вендорлық қосымшалардың күрделі логикасынан бас тарту және қарапайым элементтер кітапханаларына негізделген әмбебап ашық форматтарға көшу деректермен жұмыс істеудің икемді, ашық және тиімді тәсілдерін қалыптастыруға негіз болады. Бұл сондай-ақ құрылыш процесінің барлық қатысушыларына – жобалаушылардан тапсырыс берушілер мен эксплуатациялық қызыметтерге дейін ақпаратқа қол жеткізууді ашады.

Дегенмен, жоғары ықтималдықпен, алдағы жылдары CAD жеткізушилері CAD деректер базаларына интероперабельдік және қолжетімділік туралы талқылауда акценттерді қайтадан өзгертуге тырысады. Сөз "жаңа" концепциялар туралы болады – мысалы, гранулирленген деректер, интеллектуалды графтар, "федеративті модельдер", ғұлттық репозиторийлердегі цифрлық егіздер – сондай-ақ BIM және open BIM жолын жалғастыратын салалық альянстар мен стандарттарды құру туралы. Тартымды терминологияға қарамастан, мұндай бастамалар қайтадан пайдаланушы-

ларды проприетарлық әкожүйелерде ұстап тұру құралдарына айналуы мүмкін. 2023 жылдан бастап USD (Universal Scene Description) форматының "жаңа стандарт" ретінде CAD (BIM) арасындағы платформалық өзара әрекеттесуді белсенді түрде насиҳаттауы – соның бір мысалы.

USD және гранулирленген деректерге көшу

2023 жылы AOUSD альянсының пайда болуы құрылыш саласында маңызды бетбұрыс болып табылады. Біз CAD жеткізушилері тарапынан құрылыш деректерімен жұмыс істеуде бірнеше маңызды өзгерістер арқылы жаңа шындықтың басталғанын байқаймыз. Бірінші маңызды өзгеріс CAD деректерін қабылдаудағы өзгерістермен байланысты. Концептуалды жобалаудың алғашқы кезеңдеріне қатысатын мамандар, CAD ортада жобаны жасау тек бастапқы нұктесін барған сайын түсінеді. Жобалау процесінде қалыптасқан деректер уақыт өте келе объектілерді талдау, пайдалану және басқару үшін негіз болады. Бұл деректердің дәстүрлі CAD құралдарынан тыс жүйелерде қолжетімді және пайдалануға жарамды болуы керек екенін білдіреді.

Осыған параллель, жетекші әзірлеушілердің тәсілдерінде революция болып жатыр. BIM концепциясын және IFC форматтарын жасаған жетекші CAD жеткізуши стратегиясында күтпеген бетбұрыс жасады. 2023 жылдан бастап компания деректерді жеке файлдарда сақтау тәсілінен бас тартып, гранулирленген (нормаланған және құрылымдалған) деректермен жұмыс істеуге, дата-центрлік тәсілге көшуді қолға алуда.

Жеткізушилер басқа салалардың тарихи тенденцияларына ілесуде: көпшілік пайдаланушыларға жабық CAD форматтары (PSD-ға ұқсас) немесе IFC параметрлік файлдары (GIMP-те қабат логикасымен ұқсас) қажет емес. Оларға CAFM (құрылыш Instagram), ERP (Facebook) және Excel кестелері мен PDF құжаттарымен толтырылған мындаған басқа процестерде пайдалануға болатын объектілердің қарапайым бейнелері қажет.

Құрылыш саласындағы қазіргі тенденциялар параметрлік және құрделі форматтардан USD, GLTF, DAE, OBJ (метаинформациямен бірге гибридті, сондай-ақ жеке құрылымдалған немесе әлсіз құрылымдалған форматтарда) сияқты әмбебап және тәуелсіз форматтарға біртіндеп көшу үшін жағдайлар туғызады. 1990-шы жылдардың ортасында IFC-ны белсенді түрде насиҳаттаған тарихи көшбасшылар, бүгінгі таңда USD жаңа форматын ашық түрде насиҳаттауда [93], оның қарапайымдылығы мен әмбебаптығын атап өтуде (Рис. 6.26). USD-ның өнімдерге жаппай енгізілуі, GLTF-пен үйлесімділігі және Blender, Unreal Engine және Omniverse сияқты құралдарға белсенді интеграциясы деректермен жұмыс істеудің жаңа парадигмасының басталуының әлеуетін көрсетеді. Еуропалық ERP-да қолданылатын CPIXML сияқты жергілікті шешімдердің танымалдылығымен қатар, USD-ның Орталық Еуропадағы позицияларын күштейтуі мүмкін. IFC форматтарын дамытуға арналған ұйымдар USD-ға өз стратегияларын бейімдеуде [126], бұл тек өзгерістің 不可避免性 iн растайды.-

Technical Specifications				Comparison / Notes
File Structure	Monolithic file	Uses ECS and linked data	IFC stores all data in one file; USD uses Entity-Component-System and linked data for modularity and flexibility	
Data Structure	Complex semantics, parametric geometry	Flat format, geometry in MESH, data in JSON	IFC is complex and parametric; USD is simpler and uses flat data	
Geometry	Parametric, dependent on BREP	Flat, MESH (triangular meshes)	IFC uses parametrics; USD uses meshes for simplified processing.	
Properties	Complex structure of semantic descriptions	Properties in JSON, easy access	Properties in USD are easier to use thanks to JSON	
Export/import	Complex implementation, dependent on third-party SDKs	Easy integration, wide support	USD integrates more easily and is supported in many products	
Format Complexity	High, requires deep understanding	Low, optimized for convenience	The time required to understand the structure of the file and the information stored in it.	
Performance	Can be slow when processing large models	High performance in visualization and processing	USD is optimized for speed and efficiency. Simulations, machine learning, AI, smart cities will be held in the Nvidia Omniverse	
Integration with 3D Engines	Limited	High, designed for graphics engines	USD excels with native support for real-time visualization platforms.	
Support outside CAD Software	BlenderBIM, IfcOpenShell	Unreal Engine, Unity, Blender, Omniverse	USD is widely supported in graphics tools	
Cloud Technology Support	Limited	Well-suited for cloud services and online collaboration	USD is optimized for cloud solutions	
Ease of Integration into Web Applications	Difficult to integrate due to size and complexity	Easy to integrate, supports modern web technologies	USD is preferable for web applications	
Change Management	Versions through separate files	Versioning built into the format core	IFC handles changes via separate files, while USD embeds versioning directly into its structure	
Collaboration Support	Supports data exchange between project participants	Designed for collaborative work on complex scenes	USD provides efficient collaboration through layers and variations	
Learnability	Steep learning curve due to complexity	Easier to master thanks to a clear structure	USD is easier to learn and implement	

Рис. 6.26 IFC және USD форматтарының техникалық сипаттамаларын салыстыру.

Осы түрғыдан алғанда, USD әлеуетті түрде де-факто стандартқа айналуы мүмкін, қазіргі CAD- (BIM-) форматтарының күрделілігімен және олардың интерпретациясының геометриялық ядроларға тәуелділігімен байланысты көптеген шектеулерді женуге уәде береді.

Параметрлік және күрделі CAD форматтары мен IFC орнына, USD, glTF, DAE, OBJ сияқты деректердің қарапайым форматтары, элементтердің метаақпаратымен CSV, XLSX, JSON, XML форматтарында құрылыш саласында өз орнын табады, себебі олардың қарапайымдылығы мен икемділігі бар.

Құрылыш саласындағы қазіргі өзгерістер, бір қарағанда, ескірген IFC-дан заманауи USD-ға көшүімен байланысты технологиялық серпіліс ретінде көрінеді. Дегенмен, 2000 жылы IFC-ны әзірлеген CAD-

вендордың өз проблемалары мен деректер базасына қол жеткізу қажеттілігі туралы жазғанын ес-керу қажет, ал қазір олар жаңа стандарт - USD-ға көшүін белсенді түрде насиҳаттауда.

«Ашық деректер» USD және CAD вендорларының насиҳаттап жатқан гранулирленген деректерді басқару бойынша «жаңа» концепцияларының артында вендорлардың жобалық деректерді басқаруды монополиялауға деген ниеті болуы мүмкін, мұнда пайдаланушылар форматты таңдау корпоративтік мұдделермен, нақты қажеттіліктермен емес, байланысты болады.

Негізгі фактілерді талдау [93] көрсеткендегі, бұл өзгерістердің басты мақсаты - пайдаланушылардың ыңғайлылығы емес, вендорлардың мүддесінде экожүйелер мен деректер ағындарын бақылауды сақтау. 40 жыл бойы CAD деректер базаларына қол жеткізуді қамтамасыз ете алмаған вендорлар үшін бұл маңызды.

Мүмкін, дәл қазір компанияларға вендорлардың жаңа концепцияларын күтуден бас тартып, деректерге бағытталған даму жолына назар аударудың уақыты келді. Деректерге қол жеткізу мәселелерінен босап, кері инженерия құралдары арқылы сала жаңа концепцияларды таңдамай, заманауи, тегін және ыңғайлышты деректермен жұмыс істеу және талдау құралдарына өз бетінше көшеді.

CAD (BIM) Maturity Levels: From Stage 0 to Structured Data

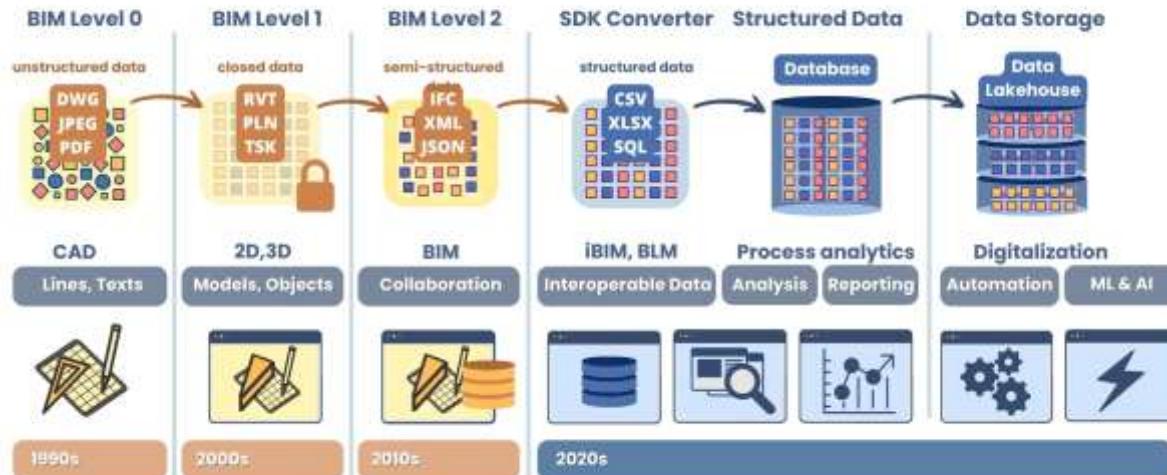


Рис. 6.27 CAD (BIM) жетілу деңгейі: құрылымдалмаған деректерден құрылымдалған деректер мен қоймалардың деңгейіне дейін.

Деректер базаларына қол жеткізу, ашық деректер және форматтар құрылымы саласында стандартқа айналуы сөзсіз, вендорлардың бұл процесті баяулатуға деген әрекеттеріне қарамастан – бұл тек уақыт мәселесі (Рис. 6.27). Бұл көшү қарқыны, егер көбірек мамандар ашық форматтармен, деректер базаларымен жұмыс істеу құралдарымен және CAD жүйелеріне тікелей қол жеткізуді үйімдастыруға мүмкіндік беретін кері инженерияның қолжетімді SDK-ларымен танысса, едәуір артуы мүмкін.-

Болашақ – ашық, унификацияланған және аналитикаға қолжетімді деректерде. Вендорлардың шешімдеріне тәуелділіктен құтылып, жабық экожүйелердің тұтқынына айналмау үшін құрылыш және жобалау компанияларына ерте ме, кеш пе, деректерге толық бақылауды қамтамасыз ететін форматтар мен шешімдерді таңдау арқылы ашықтық пен тәуелсіздікке ставка жасау қажет.

Бұгін құрылыш саласында жасалатын деректер болашақта бизнес шешімдерін қабылдаудағы негізгі ресурсқа айналады. Олар құрылыш компанияларының дамуы мен тиімділігін қамтамасыз ететін стратегиялық «отын» рөлін атқарады. Құрылыш саласының болашағы – деректермен жұмыс істей білу, ал форматтарды немесе деректер моделін таңдауда емес.

Ашық форматтар USD, glTF, DAE, OBJ және меншікті параметрліk CAD форматтарының арасындағы айырмашылықты түсіну үшін, визуализация мен жобаларды есептеудің ең курделі және маңызды элементтерінің бірі – геометрия мен оның қалыптасу процестерін қарастыру маңызды. Геометриялық деректердің құрылыш саласындағы аналитика мен есептеулердің негізі қалай қалыптасатынын түсіну үшін, геометрияны генерациялау, оның түрлендірілуі және сақталу механизмдерін тереңірек зерттеу қажет.



ТАРАУ 6.3. ҚҰРЫЛЫС САЛАСЫНДАҒЫ ГЕОМЕТРИЯ: СЫЗБАЛАРДАН КУБОМЕТРЛЕРГЕ

Қашан сызықтар ақшаға айналады немесе құрылыштыларға геометрия не үшін қажет

Құрылыш саласындағы геометрия – бұл тек визуализация ғана емес, сонымен қатар дәл сандық есептеулердің негізі. Жобаның моделінде геометрия элементтердің параметрлер тізімдерін (Сурет 3.116) маңызды көлемдік сипаттамалармен, мысалы, ұзындық, алаң және көлем сияқты көрсеткіштермен толықтырады. Бұл көлемдік параметрлердің мәндері геометриялық ядролар арқылы автоматты түрде есептеледі және сметалар, графиктер мен ресурстық модельдер үшін бастапқы нүктес болып табылады. Біз кітаптың бесінші бөлімінде және «Құрылыш жобаларының құнын есептеу және сметалар» бөлімінде талқылағанымыздай, CAD-модельдеріндегі объектілердің көлемдік параметрлері қазіргі заманғы ERP, PMIS жүйелерінің негізін құрайды. Геометрия жобалау кезеңінде ғана емес, жобаны жүзеге асыруды басқаруда, мерзімдерді бақылауда, бюджеттеуде және пайдалану кезінде де негізгі рөл атқарады. Египет пирамидаларын салу кезінде жобаның дәлдігі локоть пен кубит сияқты ұзындық өлшемдеріне байланысты болса, бұғынгі күні CAD бағдарламаларындағы геометрияның интерпретациясының дәлдігі нәтижеге, яғни бюджет пен мерзімдерге, мердігерлерді таңдау мен жеткізу логистикасына тікелей әсер етеді.

Жоғары бәсекелестік және шектеулі бюджет жағдайында геометрияға тікелей байланысты көлемдік есептеулердің дәлдігі өмір сүру факторы болып табылады. Заманауи ERP жүйелері CAD және BIM модельдерінен алғынған дұрыс көлемдік сипаттамаларға тікелей тәуелді. Сондықтан элементтердің дәл геометриялық сипаттамасы – бұл тек визуализация емес, құрылыш құнын және мерзімдерін басқарудың негізгі құралы.

Тарихи түрғыдан алғанда, геометрия инженерлік өзара әрекеттесудің негізгі тілі болды. Папиустың беттеріндегі сызықтардан бастап, цифрлық модельдерге дейін – сызбалар мен геометриялық көріністер жобалаушылар, құрылыш бригадирлері мен сметашылар арасында ақпарат алмасы құралы қызметін атқарды. Компьютерлер пайда болғанға дейін есептеулер қолмен, сызғыштар мен транспортерлер арқылы жүргізілді. Бұғынгі күні бұл міндет көлемдік модельдеу арқасында автоматандырылған: CAD бағдарламаларының геометриялық ядролары сызықтар мен нүктelerді үш өлшемді денелерге түрлендіреді, одан кейін барлық қажетті сипаттамалар автоматты түрде алынады.

CAD бағдарламаларында жұмыс істегендеге, есептеулер үшін геометриялық элементтерді жасау CAD (BIM) бағдарламаларының пайдалануши интерфейсі арқылы жүзеге асырылады. Нүктeler мен сызықтарды көлемдік денелерге түрлендірудің үшін геометриялық ядро пайдаланылады, ол геометрияны көлемдік модельдерге түрлендірудің негізгі міндетін атқарады, одан кейін аппроксимациядан кейін элементтің көлемдік сипаттамалары автоматты түрде есептеледі.

Сызықтардан көлемдерге: қалайша алаң мен көлем деректерге айналады

Инженерлік практикада көлемдер мен алаңдар аналитикалық немесе параметрлік модельдер, мысалы, NURBS (гетерогенді рационалды B-сплайндар) арқылы сипатталған геометриялық беттер негізінде есептеледі, BREP (шекаралық элементтерді көрсету) шенберінде.

NURBS (Нон-Униформды Рационалды B-Splines) – бұл қисықтар мен беттердің математикалық сипаттамасы, ал BREP – объектінің толық үш өлшемді геометриясын сипаттау құрылымы, оның шекараларын NURBS пайдалану арқылы анықтауға болады.

BREP және NURBS-тың дәлдігіне қарамастан, олар қуатты есептеу ресурстарын және күрделі алгоритмдерді талап етеді. Алайда, мұндай математикалық дәл сипаттамалар бойынша тікелей есептеулер көбінесе есептеу жағынан күрделі, сондықтан практикада дерлік әрқашан тесселяция қолданылады – беттерді үшбұрыштар торына түрлендіру, бұл кейінгі есептеулерді жеңілдетеді. Тесселяция – бұл күрделі бетті үшбұрыштарға немесе полигондарға бөлу. CAD/CAE ортасында бұл әдіс визуализация, көлемдерді есептеу, соқтығысу іздеу, MESH сияқты форматтарға экспорттау және соқтығысу анализі үшін қолданылады. Табиғаттан мысал – аралардың ұялары, мұнда күрделі форма реттелген торға бөлінеді (Рис. 6.31). -

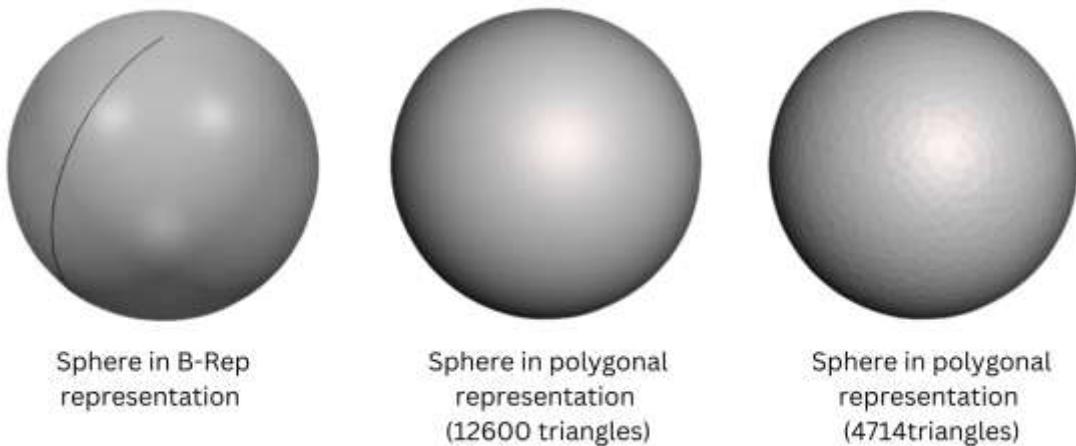


Рис. 6.31 Параметрлік сипаттамадағы BREP және полигондық көрсетілімдегі бір және сол сфера, үшбұрыштардың әртүрлі саны бар.

CAD-да қолданылатын BREP (NURBS) геометрияның негізі моделі емес. Бұл әдіс шенберлер мен рационалды сплайндарды көрсету үшін ыңғайлы құрал ретінде жасалған және геометрия туралы деректерді сақтау көлемін минимизациялау үшін арналған. Алайда, оның шектеулері бар – мысалы, винттік сызықтар мен беттердің негізінде жатқан синусоиданы дәл сипаттау мүмкін еместігі, сондай-ақ күрделі геометриялық ядроларды қолдану қажеттілігі.

Үшбұрышты торлар мен параметрлік фигуralардың тесселяциясы, керісінше, қарапайымдылық, тиімді жадты пайдалану және үлкен деректер көлемдерін өңдеу қабілетімен ерекшеленеді (Рис. 6.32). Бұл артықшылықтар күрделі және қымбат геометриялық ядроларды, сондай-ақ олардың ішінде орналасқан он миллиондаған код жолдарын есептеулерде пайдаланбай-ақ жұмыс істеуге мүмкіндік береді.

Көптеген құрылыш жағдайларында көлемдік сипаттамалардың қалай анықталуы маңызды емес – параметрлік модельдер (BREP, IFC) немесе полигондар (USD, glTF, DAE, OBJ) арқылы. Геометрия әрқашан аппроксимация түрінде қалады: NURBS немесе MESH арқылы болсын, бұл әрқашан формаға жуықтау сипаттамасы.

Полигондар немесе BREP (NURBS) түрінде берілген геометрия белгілі бір дәрежеде тек үздіксіз форманың жуықтау сипаттамасы болып табылады. Френель интегралдары дәл аналитикалық өрнекке ие болмағандай, геометрияның полигондар немесе NURBS арқылы дискретизациясы әрқашан жуықтау болып табылады, дәл сол сияқты үшбұрышты MESH.

BREP форматындағы параметрлік геометрия негізінен деректердің минималды көлемі маңызды болғанда және оны өңдеу мен көрсету үшін ресурстарды көп қажет ететін және қымбат геометриялық ядроларды пайдалануға мүмкіндік болғанда қажет. Бұл көбінесе CAD бағдарламаларын әзірлеушілерге тән, олар өз өнімдерінде MCAD вендорларының геометриялық ядроларын пайдаланаңды. Сонымен қатар, тіпті осы бағдарламалардың ішінде, BREP модельдері визуализация және есептеулер үшін тесселяция процесінде жиі үшбұрыштарға түрлендіріледі (PSD файлдарының JPEG форматына жеңілдетілуі сияқты).

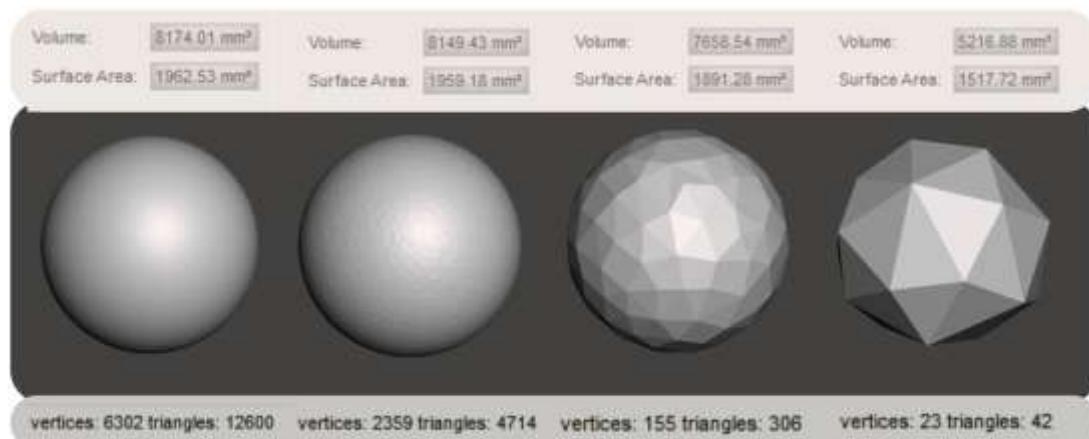


Рис. 6.32 Әртүрлі полигон саны бар фигуralардың көлемдік сипаттамаларының айырмашылығы.

Полигональды MESH және параметрлік BREP өз артықшылықтары мен шектеулеріне ие, бірақ олардың мақсаты бір – пайдалануышының тапсырмаларын ескере отырып, геометрияны сипаттау. Сонында геометриялық модельдің дәлдігі тек оның ұсынылу әдісіне ғана емес, сонымен қатар нақты тапсырманың талаптарына да байланысты.

Көптеген құрылыш тапсырмаларында параметрлік геометрия мен күрделі геометриялық ядроларға қажеттілік артық болуы мүмкін.

Автоматизациялау тапсырмасының әрбір нақты жағдайында CAD әзірлеушілері өз бағдарламалық өнімдерін жылжыту мен сату мүддесінде параметрлік геометрияның маңыздылығын асыра бағаламау керек.

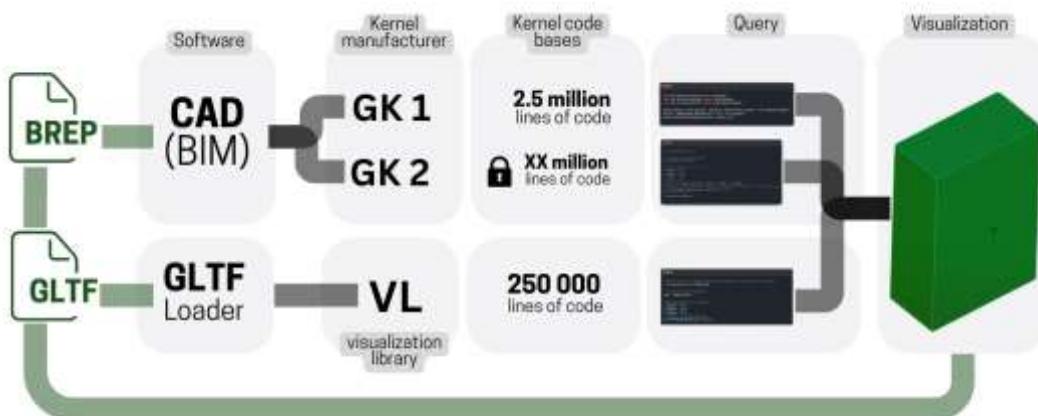
MESH, USD және полигондарға көшу: геометрия үшін тесселяцияны пайдалану

Құрылыш саласында ағынды жұмыс, жүйелерді, деректер базаларын немесе жобалық ақпарат пен элементтердің геометриясымен жұмыс істейтін процестерді автоматтандыру кезінде нақты CAD редакторларына және геометриялық ядроларға тәуелсіздікке үмтүлу маңызды.

Есептеу бөлімдерінде және құрылыш алаңында пайдаланылатын алмасу форматы нақты CAD- (BIM-) бағдарламасын білдірмейі туіс. Геометриялық ақпарат тікелей тесселяция арқылы, геометриялық ядроға немесе CAD архитектурасына байланбай ұсынылуы керек.

CAD-тан алынған параметрлік геометрия аралық дереккөз ретінде қарастырылуы мүмкін, бірақ әмбебап форматтың негізі ретінде емес. Көптеген параметрлік сипаттамалар (BREP және NURBS-ты қоса алғанда) кез келген жағдайда кейінгі өңдеу үшін полигональды MESH-ке түрлендіріледі. Егер нәтиже бірдей болса (тесселяция және полигондар), ал процесс қарапайым болса, онда таңдау айқын. Бұл графикалық онтология мен құрылымдық кестелер арасындағы таңдау сияқты: артық күрделілік сирек ақталған.-

OBJ, STL, glTF, SVF, CPIXML, USD және DAE сияқты ашық форматтар үшбұрышты торлардың әмбебап құрылымын пайдаланады, бұл оларға елеулі артықшылықтар береді. Бұл форматтар тамаша үйлесімділікке ие – оларды қолжетімді ашық кітапханалар арқылы оқу және визуализациялау оңай, күрделі мамандандырылған геометриялық ядролардың миллиондаған код жолдарын қажет етпейді. Бұл әмбебап геометриялық форматтар әртүрлі салаларда қолданылады – IKEA™-дағы ас үй жобалар құралдарынан бастап, кино және VR қосыншаларындағы объектілерді визуализациялау жүйелеріне дейін. Маңызды артықшылық – бұл форматтармен жұмыс істеу үшін көптеген тегін және ашық кітапханалардың болуы, олар көшілік платформалар мен бағдарламалау тілдеріне қолжетімді.-



Бірдей геометрияны көрсету параметрлік форматтар мен геометриялық ядроларды пайдалану арқылы немесе триангуляцияланған форматтар мен ашық визуализация кітапханалары арқылы жүзеге асырылады.

Пайдаланушылар сияқты, CAD вендорлары да геометриялық ядролардағы айырмашылықтардан туындаған параметрлік CAD форматтарын немесе ашық IFC-ны интерпретациялау мәселелеріне тап болады. Практикада барлық CAD вендорлары, ешбір ерекшеліксіз, жүйелер арасында деректерді беру үшін көрі инженерияның SDK-ларын пайдаланады, және олардың ешқайсысы интероперабельділік мақсатында IFC немесе USD форматтарына сенбейді.

CAD вендорларының альянстары ұсынған концепцияларды қолданудың орнына, CAD шешімдерінің әзірлеушілері мен пайдаланушылары нақты контексте әрбір тәсілдің артықшылықтарын түсінуге назар аударғаны жөн және қолдану жағдайына байланысты геометрияның бір немесе басқа түрін таңдауы керек. Әртүрлі геометриялық ұсынystар арасындағы таңдау – бұл дәлдік, есептеу тиімділігі және нақты тапсырманың практикалық қажеттіліктері арасындағы компромисс.

Улken вендорлар жобалау деректерін өңдеу кезінде құрылыш саласына мәжбүрлелітін геометриялық ядроларды қолданудың күрделілігі көбінесе артық болып табылады. MESH-геометриясына негізделген USD форматы сала үшін «Пандораның жәшігі» бола алады, әзірлеушілерге IFC және CAD вендорларына тән параметрліk BREP құрылымдарының шеңберінен тыс деректер алмасуды ұйымдастырудың жаңа мүмкіндіктерін ашады.

USD, DAE, gLTF, OBJ және басқа құрылымдармен танысқан кезде, күрделі параметрика мен жабық геометриялық ядроларға сүйенудің қажеті жоқ геометриялық ақпаратты тиімді беру мен пайдалануды ұйымдастыруға мүмкіндік беретін қарапайым, ашық форматтардың бар екендігі айқын болады. Мұндай тәсіл әзірлеушілер үшін техникалық кіру шегін төмендетіп қана қоймай, сонымен қатар цифровық құрылыш үшін икемді, масштабталатын және шын мәнінде ашық шешімдердің дамына ықпал етеді.

LOD, LOI, LOMD – CAD (BIM) жүйесіндегі детализацияның ерекше классификациясы

Геометриялық ұсыныс форматтарына қосымша, әртүрлі салалар әртүрлі деректер деңгейлері мен тереңдіктерін пайдаланатын әлемде CAD- (BIM-) әдістемелері өздерінің бірегей классификация жүйелерін ұсынады, олар ғимарат модельдерінің ақпараттық мазмұнына құрылымдық тәсілді қалыптастырады.

Стандартизацияға жаңа тәсілдердің бірі – графикалық және ақпараттық құрамдас бөліктердің дайындық және сенімділік дәрежесін көрсететін модельді әзірлеу деңгейлерін енгізу. CAD- (BIM-) деректермен жұмыс істеуде ақпараттық мазмұнды шектеу үшін LOD (Level Of Detail) – модельдің графикалық бөлігінің детализация деңгейі, және LOI (Level Of Information) – деректердің өндегу деңгейі пайда болды. Қосымша кешенді тәсіл үшін LOA (Level of Accuracy) – ұсынылған элементтердің дәлдігі және LOG (Level of Geometry) – графикалық ұсыныстың дәлдігін анықтау ұғымы енгізілді.

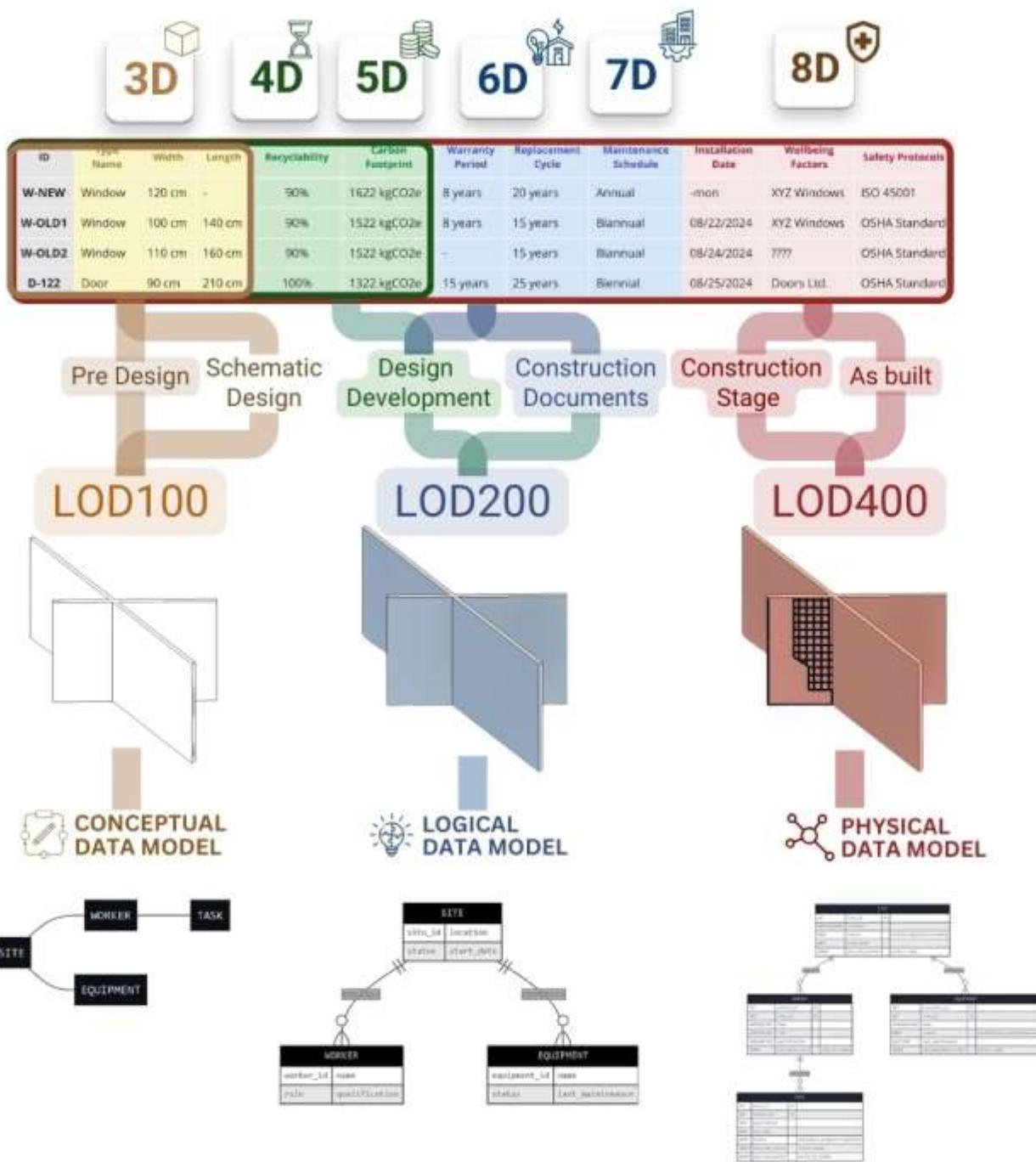
Детайлдеу деңгейлері (LOD) 100-ден 500-ге дейінгі сандармен белгіленеді, модельдің өндегу дәрежесін көрсетеді. LOD 100 – жалпы формалар мен өлшемдері бар концептуалды модель. LOD 200 – шартты детализациямен, бірақ нақты өлшемдер мен формаларды қамтиды. LOD 300 – элементтердің дәл өлшемдері, формалары мен орналасуын қамтитын егжей-тегжейлі модель. LOD 400 – элементтерді жасау және монтаждау үшін қажетті егжей-тегжейлі ақпаратты қамтиды. LOD 500 –

Құрылыш аяқталғаннан кейін объектінің нақты күйін көрсетеді және пайдалану мен қызмет көрсету үшін қолданылады. Бұл деңгейлер CAD (BIM) моделінің ақпараттеп қанығу құрылымын өмірлік циклдің әртүрлі кезеңдерінде, 3D, 4D, 5D және одан әрі сипаттайды.

Нақты жобаларда жоғары детализация (LOD400) көбінесе артық болып табылады, және LOD100 геометриясын немесе тіпті жазық сызбаларды қолдану жеткілікті, ал қалған деректерді есептеу арқылы немесе айқын геометриясы жоқ байланысты элементтерден алуға болады. Мысалы, бөлмелердің кеңістіктері мен элементтері (элементтер категориясы «Бөлмелер») визуалды геометрияға ие болмауы мүмкін, бірақ сонымен бірге көптеген бизнес-процестерді құруға негіз болатын елеулі ақпараттар мен деректер базаларын қамтуы мүмкін.

Сондықтан жобалауды бастамас бұрын қажетті детализация деңгейін нақты анықтау маңызды. 4D-7D қолдану жағдайлары үшін көбінесе DWG сызбалары мен LOD100 минималды геометриясы жеткілікті. Талаптармен жұмыс істеу барысында негізгі міндет - модельдің байлығы мен практикалық қолданылуы арасында тепе-тендікті табу.

Негізінде, CAD (BIM) деректерін деректер базасы ретінде қарастырсақ (олай болып табылады), модельдің байлығын жаңа аббревиатуралар арқылы сипаттау - бұл ақпараттық жүйелер үшін деректерді кезең-кезеңімен модельдеу, концептуалды деңгейден физикалық деңгейге дейін (Сурет 6.34), бұл кітаптың үшінші және төртінші бөлімдерінде егжей-тегжейлі қарастырылған. LOD және LOI деңгейінің әрбір ұлғаюы жаңа міндеттер үшін қажетті ақпараттың қосылуын білдіреді: есептеулер, құрылыш басқару, пайдалану және модельді қосымша ақпараттық қабаттармен (3D-8D) байыту арқылы сипатталады, бұл туралы кітаптың бесінші бөлімінде айтылған.



Сурет 6.34 Жобаның ақпараттық толтыру процесі концептуалды модельден физикалық деректер моделіне дейінгі деректерді модельдеуге үқсас.

Геометрия - бұл жобалық деректердің тек бір бөлігі, оның қажеттілігі құрылыш жобаларында әрдайым негізделмеген, ал CAD деректерімен жұмыс істеудің негізгі сұрағы модельдердің қалай визуализациялануында емес, осы модельдерден алынған деректердің CAD (BIM) бағдарламаларынан тыс қалай пайдаланылуында.

2000-шы жылдардың ортасына қарай құрылыш саласы деректерді басқару және өндөу жүйелеріндегі деректер көлемінің жедел өсуімен бетпе-бет келді, әсіресе CAD (BIM) бөлімдерінен

келген деректер. Деректер көлемінің мұндай күрт өсуі компания менеджерлерін тосыннан ұстап, олар деректердің сапасы мен басқару талаптарына дайын болмады.

CAD (BIM) жаңа стандарттары - AIA, BEP, IDS, LOD, COBie

CAD деректер базаларына ашық қол жетімділіктің болмауын және деректерді өңдеу нарығындағы шектеулі бәсекелестікті пайдалана отырып, сондай-ақ BIM-мен байланысты жаңа аббревиатураларапта қатысты маркетингтік науқандарды қолдана отырып, CAD деректерімен жұмыс істеу тәсілдерін дамытуға бағытталған ұйымдар жаңа стандарттар мен концепцияларды құра бастады, олар заңды түрде деректерді басқару практикасын жақсартуға бағытталуы тиіс.

Дегенмен, CAD (BIM) жеткізушілері мен әзірлеушілері тікелей немесе жанама түрде қолдаған барлық бастамалар жұмыс процестерін оңтайландыруға бағытталғанымен, олар әртүрлі мүдделі тараптар тараپынан лоббилинген стандарттардың пайда болына әкелді, бұл құрылыш саласын деректерді өңдеу процестерінде белгілі бір екіүштылық пен шатасуға алып келді.

Соңғы жылдары құрылыш саласында пайда болған LOD, LOI, LOA, LOG-дан басқа жаңа деректер стандарттарының кейірін атап өтейік:

- BEP (BIM орындау жоспары) - жобада CAD (BIM) интеграциялау және пайдалану әдістері мен деректерді өңдеу процестерін анықтайды.
- EIR/AIA құжаты (Тапсырыс берушінің ақпараттық талаптары) - тапсырыс беруші тендер жарияламас бұрын дайындейді және мердігерден ақпаратты дайындау және ұсыну жөніндегі талаптарды қамтиды. Ол тиісті жобада BEP негізі болып табылады.
- AIM (Актив ақпараттық моделі) - BIM процесінің бір бөлігі. Жоба аяқталғаннан кейін деректер моделі "Актив ақпараттық моделі" немесе AIM деп аталады. AIM мақсаты - жүзеге асырылған активті басқару, қызмет көрсету және пайдалану.
- IDS (Ақпараттық жеткізу спецификациясы) – құрылыш жобасының әртүрлі кезеңдерінде қандай деректер мен қандай форматта қажет екенін анықтайтын талаптарды белгілейді.
- iLOD – BIM-модельдегі ақпараттың ұсынылу деңгейі. Ол модельдегі ақпараттың қаншалықты егжей-тегжейлі және толық екенін анықтайды, негізгі геометриялық көрсетілімдерден бастап, егжей-тегжейлі спецификациялар мен деректерге дейін.
- eLOD – CAD (BIM) моделіндегі жеке элементтердің LOD деңгейі. Ол әр элементтің моделдеу дәрежесін және оған қатысты ақпаратты, мысалы, өлшемдер, материалдар, пайдалану сипаттамалары және басқа да тиісті атрибуттарды анықтайды.
- APS (Платформа қызметтері) және ірі CAD (BIM) жеткізушілерінің басқа өнімдері – байланысты және ашық деректер модельдерін жасау үшін қажетті құралдар мен инфрақұрылымды сипаттайтын.

CAD (BIM) стандарттарын енгізуіндегі жарияланған мақсаты – деректерді басқару сапасын арттыру және автоматтандыру мүмкіндіктерін кеңейту, бірақ практикада олардың қолданылуы жиі артық күрделілік пен процестердің фрагментациясына әкеледі. CAD (BIM) моделін деректер базасының

бір түрі ретінде қарастырғанда, көптеген стандарттардың бұрыннан қалыптасқан және басқа экономикалық секторларда ақпараттық жүйелермен жұмыс істеген тиімді тәсілдерді қайталайтыны айқын болады. Мұндай бастамалар қарапайымдылық пен унификацияны қамтамасыз етудің орнына, қосымша терминологиялық жүктемені туындарып, шын мәнінде ашық және икемді шешімдерді енгізуге кедері келтіреді.

Көптеген жаңа концепциялардың фактісі, олар деректерді моделдеу және тексеру процесін алмастырады, бұл процесс кітаптың алғашқы бөлімдерінде ежей-тегжейлі қарастырылған және басқа экономикалық секторларда ұзақ уақыт бойы қолданылып келеді. Құрылыш саласында стандарттау процесі көбінесе кері бағытта жүреді – жаңа деректерді сипаттау форматтары, жаңа стандарттар мен тексеру концепциялары жасалады, олар әрдайым шынайы унификация мен практикалық қолдануға әкелмейді. Нәтижесінде, деректерді өңдеуді жеңілдетудің және автоматтандырудың орнына, сала қосымша регламенттеу және бюрократия деңгейлерімен бетпе-бет келеді, бұл әрдайым тиімділікті арттыруға ықпал етпейді.

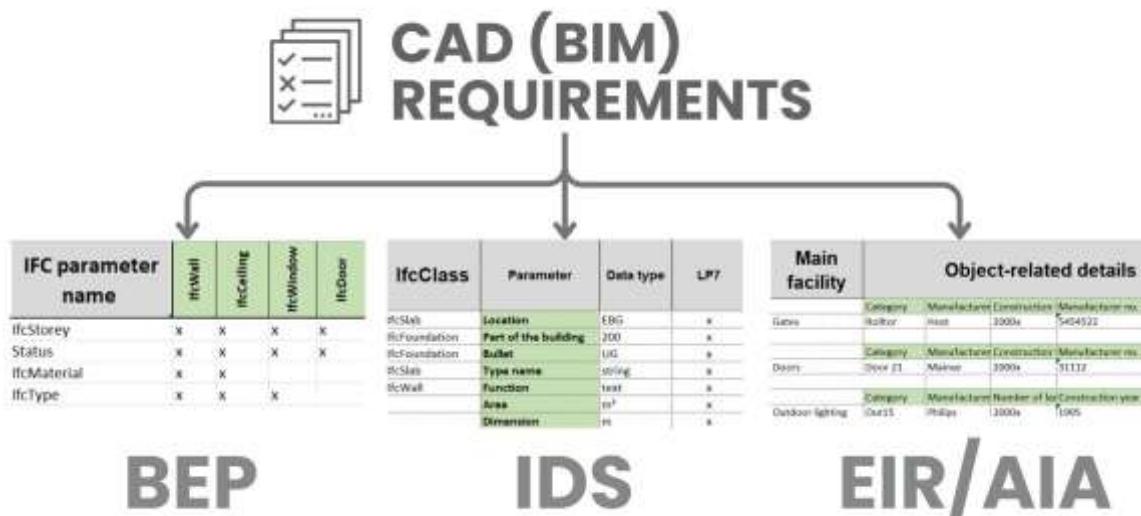


Рис. 6.31 Деректер мен ақпараттық толтыру талаптары атрибуттар мен олардың шекті мәндерін сипаттауға, кестелер арқылы көрсетіледі.

Жаңа CAD (BIM) деректерімен байланысты концепциялар деректерді өңдеуді жеңілдеді тудің орнына, интерпретация мен негізгі анықтамалар кезеңінде қосымша қысындықтар мен дауларды туындарады.

Соңғы жаңа концепциялардың бірі – 2020 жылды пайдаланып болған IDS форматы, ол open BIM концепциясында ақпараттық модельдің атрибуттық құрамына қойылатын талаптарды сипаттауға мүмкіндік береді. IDS талаптары атрибуттар мен олардың шекті мәндері туралы ақпаратты құрылымдық кесте (Excel немесе MySQL) түрінде сипаттайтыны, ол кейін полуструктуралы XML форматындағы белгілеуге айналдырылады, XML-ден IDS арнайы абревиатурасына қайта аталады.

Вендорлар мен оларды қолдайтын BIM және open BIM концепцияларының пікіріне қарамастан, құрылыш саласындағы деректермен жұмыс істеу арнайы құралдарды, мысалы, CAD және BIM қолдану арқылы ерекше екендігі, деректер форматтары мен басқару әдістері басқа салалардағы деректерді өңдеу форматтары мен концепцияларынан ерекшеленбейді.

CAD (BIM) жобаларына қойылатын талаптар мен форматтарды біртұтас талаптар кестесін пайдалану арқылы жеңілдетуге болады, атрибут-бағандармен, "Талаптарды құрылымдық формага аудару" бөлімінде ежей-тегжейлі сипатталған, бастапқыда құрылымдалған талаптарды кесте емес форматтарға аударудың қажеті жоқ (IDS бастапқыда кесте арқылы сипатталады).

Жеңілдетілген тәсіл (Рис. 6.32), бұрынғы тарауларда (Рис. 4.49, Рис. 4.416, Рис. 7.310) ежей-тегжейлі қарастырылған, объектілердің идентификаторлары, қасиеттері және шекті мәндер үшін бағандарды қамтиды, IDS-XML форматына талаптарды түрлендірудің қажеті жоқ. Мұндай тәсіл де-ректердің сапасын бақылаудың тікелей, аз ауыртпалықты және айқын механизмін қамтамасыз етеді. Ол кеңінен қолданылатын құралдарға сүйенеді: регуляярлы өрнектер (RegEx), Pandas кітапханасы және стандартты ETL құбырлары – дәл осында, басқа экономикалық секторларда деректермен жұмыс істейтін мамандар арасында кеңінен қолданылады.—

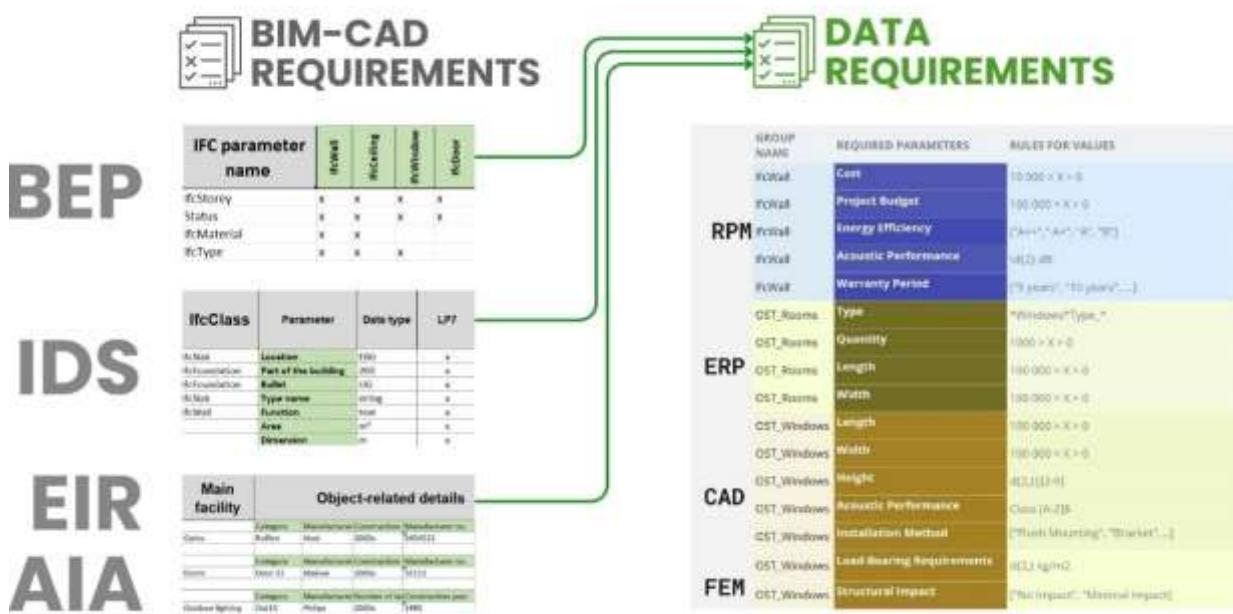


Рис. 6.32 Деректерге қойылатын талаптар басқа салаларда атрибуттар мен олардың шекті мәндерін құрылымдық сипаттауға дейін жеңілдетіледі.

Уақыт өте келе құрылыш саласында, деректердің жабықтығынан, осы әртүрлі форматтағы деректерді бақылау және басқару үшін жаңа тәсілдер мен әдістемелер пайда болуда, алайда құрылыш жобаларындағы деректер, негізінен, басқа салалардағы деректерден ерекшеленбейді. Басқа салалар деректерді өңдеудің стандартталған тәсілдерімен сәтті айналысса, құрылыш жаңа бірегей деректер форматтарын, талаптарын және олардың тексеру концепцияларын әзірлеуді жалғастыруда.

Құрылыш саласында деректерді жинау, дайындау және талдау үшін қолданылатын әдістер мен құралдар, басқа экономикалық салаларда қолданылатындардан принципті түрде ерекшеленбейі тиіс.

Салада ерекше терминологиялық экожүйе қалыптасты, ол сынни түрғыдан ойлауды және қайта бағалауды талап етеді:

- STEP форматы IFC жаңа атауымен ұсынылып, құрылыш категориясымен толықтырылады, бірақ STEP форматындағы шектеулерді ескермейді.
- Параметрлік IFC форматы деректерді беру процестерінде қолданылады, визуализация мен есептеулер үшін қажетті унификацияланған геометриялық ядроның болмауына қарамастан.
- CAD жүйелерінің деректер базаларына қолжетімділік "BIM" термині арқылы алға жылжытылуда, бұл деректер базаларының ерекшеліктері мен оларға қолжетімділік мәселелері талқыланбайды.
- Вендорлар IFC және USD форматтары арқылы интероперабельділікті алға жылжытады, көбінесе оларды практикада қолданбай, қымбат кері инженерияны пайдаланады, онымен өздері күрескен.
- LOD, LOI, LOA, LOG, BEP, EIR, AIA, AIM, IDS, iLOD, eLOD терминдері бірдей параметрлерді сипаттау үшін көнінен қолданылады, модельдеу және верификация құралдарына байланбай, басқа салаларда бұрыннан қолданылып келеді.

Құрылыш саласы, аталғандардың бәрі біртүрлі естілсе де, құрылыш саласында мүмкін екенін көрсетеді – әсіресе, егер негізгі мақсат деректерді өңдеудің әр кезеңін монетизациялау болса, арнайы қызыметтер мен бағдарламалық қамтамасыз етуді сату арқылы. Бизнес түрғысынан мұнда ешқандай айып жоқ. Дегенмен, CAD (BIM) байланысты осындағы аббревиатуралар мен тәсілдердің шын мәнінде құндылық қосып, кәсіби процестерді жеңілдететіні туралы сұрақ ашық, күйінде қалады.

Құрылыш саласында мұндағы жүйе жұмыс істейді, себебі индустрияның өзі негізгі спекулятивті пайдасын дәл осы жүйелер мен аббревиатуралардың лабиринттерінен алады. Ашық процестер мен ашық деректерге қызығушылық танытатын компаниялар өте сирек кездеседі. Бұл күрделі жағдай, тапсырыс берушілер, клиенттер, инвесторлар, банктер және жеке капитал өкілдері ақпаратты басқарудағы айқын және негізделген тәсілдерді талап етпейінше, белгісіз уақытқа дейін жалғасуы мүмкін.

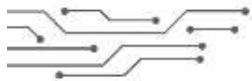
Сала артық аббревиатураларды жинақтады, бірақ олардың барлығы әртүрлі дәрежеде бір және сол процестер мен деректерге қойылатын талаптарды сипаттайтын. Олардың жұмыс процестерін жеңілдетудегі нақты пайдасы күмән тудырады.

Концепциялар мен маркетингтік аббревиатуралар келіп-кеткенімен, деректерге қойылатын талаптарды тексеру процестері бизнес-процессердің ажырамас бөлігі болып қала береді. Жаңа және жаңа мамандандырылған форматтар мен регламенттерді жасауға орнына, құрылыш саласы басқа салаларда, мысалы, қаржы, өнеркәсіп және IT-де тиімділігін дәлелдеген құралдарға назар аударуы тиіс.

Терминдер, аббревиатуралар мен форматтардың молдығы цифрлық құрылыш процестерінің терең өндөлгендігі туралы иллюзияны қалыптастырады. Алайда, маркетингтік концепциялар мен

Күрделі терминологияның артында жиі қарапайым, бірақ ыңғайсыз шындық жасырылады: деректер қолжетімсіз, нашар құжатталған және нақты бағдарламалық шешімдерге қатты байланған.

Аббревиатуралар мен форматтар шеңберінен шығу үшін CAD (BIM) жүйелеріне ақпаратты басқарудың сиқырлы құралдары ретінде емес, шын мәнінде олар не екенін — мамандандырылған деректер базалары ретінде қарau қажет. Тек осы көзқарас арқылы маркетингтің қай жерде аяқталып, ақпаратпен нақты жұмыс басталатынын түсінуге болады.



ТАРАУ 6.4. ЖОБАЛАУДЫҢ ПАРАМЕТРЛЕНДІРІЛУІ ЖӘНЕ CAD МЕН ЖҰМЫС ИСТЕУ ҮШІН LLM ПАЙДАЛАНУ

CAD (BIM) деректерінің бірегейлігі иллюзиясы: аналитика мен ашық форматтарға жол

Заманауи CAD (BIM) платформалары жобалау мен құрылым ақпаратын басқару тәсілдерін едәуір трансформациялады. Бұрын бұл құралдар негізінен сыйбалар мен үш өлшемді модельдер жасау үшін қолданылса, бүгінгі күні олар жобалық деректердің толыққанды қоймаларының функцияларын атқарады. Бірінші дереккөз концепциясы аясында параметрлік модель жиі жобаның негізгі және көбінесе жалғыз ақпарат көзіне айналуда, оның тұтастыры мен өзектілігін объектінің өмірлік циклі бойында қамтамасыз етеді.

CAD- (BIM-) платформалардың құрылым деректерін басқарудың басқа жүйелерінен негізгі айырмашылығы – ақпаратқа (бірінші дереккөз) қол жеткізу үшін мамандандырылған құралдар мен API-ларды пайдалану қажеттілігі. Бұл деректер базалары дәстүрлі мағынада әмбебап емес: ашық құрылым мен икемді интеграцияның орнына, олар нақты платформа мен форматқа қатты байланған жабық орта болып табылады.

CAD-деректермен жұмыс істеудің барлық күрделілігіне қарамастан, техникалық іске асырудан тыс тұрған маңызды сұрақ туындаиды: CAD (BIM) деректер базалары шын мәнінде не? Бұл сұраққа жауап беру үшін бағдарламалық қамтамасыз етушілер ұсынған дәстүрлі абревиатуралар мен концепциялардан шығу қажет. Оның орнына жобалық ақпаратпен жұмыс істеудің мәніне, яғни деректер мен оларды өндөу процестеріне назар аудару керек.

Құрылым саласындағы бизнес-процесс CAD- немесе BIM-құралдарында жұмыс істеуден басталмайды, керісінше, жобаның талаптарын қалыптастыру мен деректерді модельдеуден басталады. Алдымен тапсырманың параметрлері анықталады: объектілердің тізімі, олардың бастапқы сипаттамалары және нақты тапсырманы шешу кезінде ескеру қажет шекті мәндер. Тек осы параметрлер негізінде CAD- (BIM-) жүйелерінде модельдер мен элементтер жасалады.

CAD- (BIM-) деректер базасында ақпаратты жасаудан бұрынғы процесс деректерді модельдеу процесін толықтай қайталайды, ол кітаптың төртінші бөлімінде және «Деректерді модельдеу: концептуалдық, логикалық және физикалық модель» бөлімінде ежей-тегжейлі қарастырылған.-

Деректерді модельдеу процесінде сияқты, біз кейінірек деректер базасында өндегіміз келетін деректерге талаптар жасаймыз, CAD-деректер базалары үшін менеджерлер жобалау талаптарын бірнеше бағандардан тұратын кесте немесе «кілт-мән» жүптарының тізімдері түрінде жасайды. Тек осы бастапқы параметрлер негізінде API арқылы автоматты немесе қолмен, жобалаушы CAD- (BIM) деректер базасында объектілерді жасайды (немесе, дұрысы, нақтылайды), содан кейін олар бастапқы талаптарға сәйкестігін тексереді. Бұл процесс – анықтау → жасау → тексеру → түзету – итеративті түрде қайталанады, деректердің сапасы, деректерді модельдеу кезінде сияқты, мақсатты жүйе – құжаттар, кестелер немесе дэшбордтар үшін қажетті деңгейге жеткенше. -

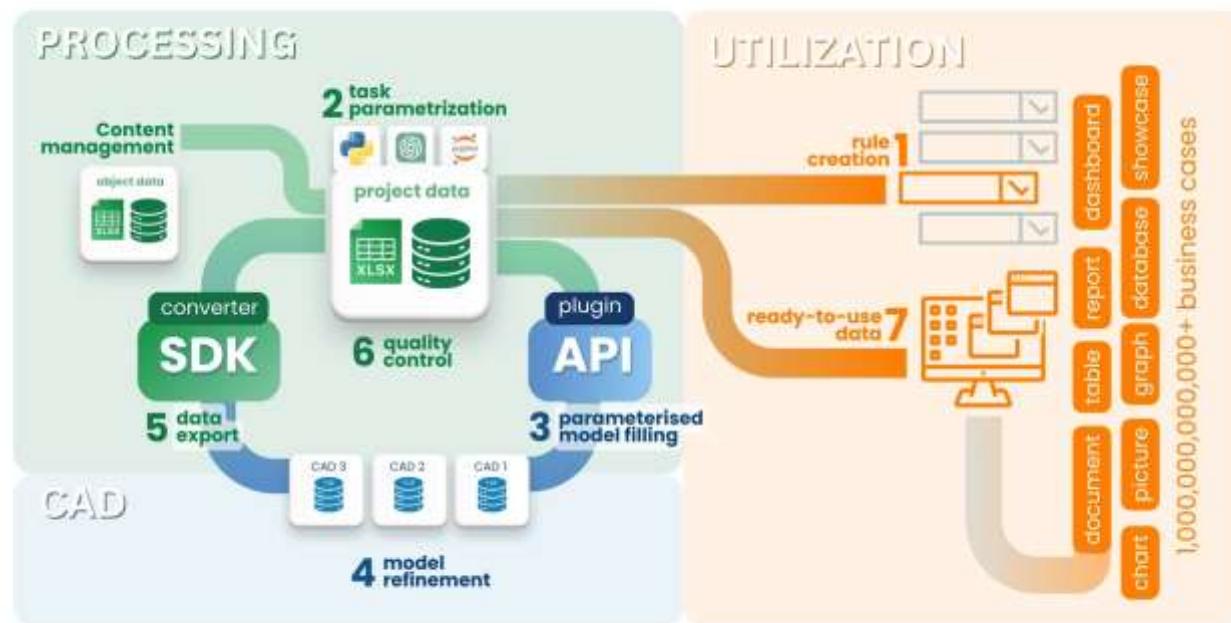


Рис. 6.41 Құрылыш жобаларын жүзеге асырудағы бизнес-процесстер үшін деректер базаларын ақпаратпен байыту циклі.

CAD (BIM) механизмін талаптар негізінде қалыптасқан «кілт–мән» жүптарының жиынтығы ретінде қарастыrsaқ, онда талқылаудың фокусы нақты бағдарламалық шешімдер мен олардың шектеулерінен деректер құрылымына, деректер модельдеріне және оларға қойылатын талаптарға ауысады. Негізінде, бұл деректер базасын параметрлермен байыту және классикалық деректерді модельдеу процесі. Айырмашылығы тек CAD-деректер базаларының жабықтығы мен қолданылатын форматтардың ерекшеліктері салдарынан бұл процесс мамандандырылған BIM-құралдарын қолданумен бірге жүреді. Сұрақ туындаиды: BIM-нің бірегейлігі неде, егер басқа экономикалық салаларда осындағы тәсілдер жоқ болса?-

Соңғы 20 жылда BIM тек деректердің бірегей көзі ретінде ғана емес, одан да көп нәрсе ретінде позицияланды. Маркетингтік түрғыдан CAD-BIM байланысы көбінесе бастапқыда интеграцияланған деректер базасымен параметрлік құрал ретінде сатылады, құрылыш объектілерінің жобалау, модельдеу және өмірлік циклын басқару процесстерін автоматтандыруға қабілетті. Алайда, шын мәнінде, BIM вендор платформасында пайдаланушыларды ұстап түру құралына айналды, деректер мен процесстерді басқарудың ынғайлы әдісі емес.

Нәтижесінде CAD- (BIM-) деректері өз платформаларында оқшауланған, жобаның ақпаратын меншікті API және геометриялық ядролар арқылы жасырады. Бұл пайдаланушыларды деректер базаларына өз бетінше қол жеткізу, деректерді шығару, талдау, автоматтандыру және вендор зекожүйелерін айналып өтіп, басқа жүйелерге беру мүмкіндігінен айырды.

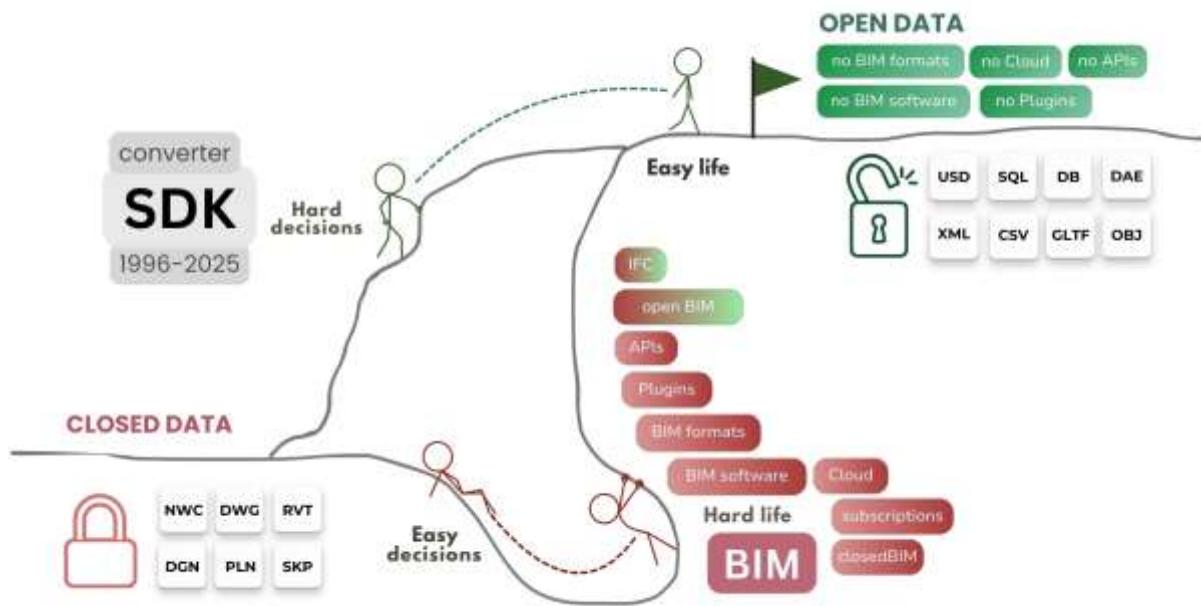


Рис. 6.42 Құрылыш саласында қазіргі заманғы форматтар күрделі геометриялық ядроларды, жыл сайын жаңартылатын API мен CAD- (BIM-) бағдарламаларына арнайы лицензияларды талап етеді.

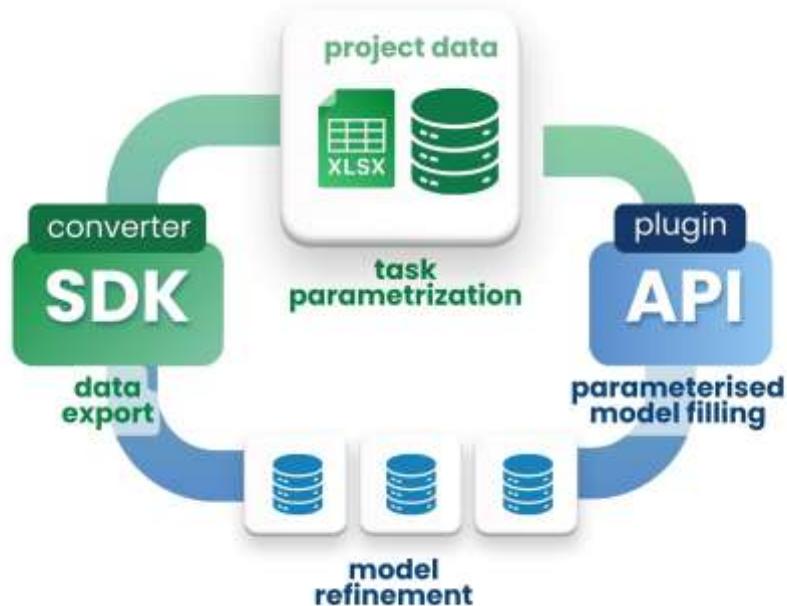
Заманауи CAD құралдарымен жұмыс істейтін компаниялар, CAD вендорларының 1995 жылдан бері күресіп келе жатқан SDK құралдарын пайдалану арқылы деректерді қайта инженериялау үшін қолданатын тәсілді қолдануы тиіс. CAD деректер базасына толық қол жеткізе отырып және қайта инженериялау құралдарын пайдалана отырып, біз атрибуттары бар жазық объектілер жиынтығын алғып, оларды геометрия мен жобалық элементтердің параметрлерін қамтитын кез келген ыңғайлы ашық форматқа экспорттай аламыз. Мұндай тәсіл ақпаратпен жұмыс істеу парадигмасын файлға негізделген архитектурадан деректерге негізделген архитектураға өзгертеуді.

- RVT, IFC, PLN, DB1, CP2, CPIXML, USD, SQLite, XLSX, PARQUET және басқа деректер форматтары бір жобаның элементтері туралы бірдей ақпаратты қамтиды. Бұл нақты формат пен оның схемасын білу деректермен жұмыс істеуге кедергі болмауы тиіс дегенді білдіреді.
- Кез келген форматтағы деректерді үшбұрышты геометрия MESH және объектілердің барлық элементтерінің қасиеттерін қамтитын бір ашық құрылымдық және гранулирленген құрылымға біріктіруге болады, геометриялық ядролардың шектеулерінсіз.
- Дата-аналитика әмбебаптықа үмтүлады: ашық деректерді пайдалана отырып, жобалық деректермен қолданылатын форматқа қарамастан жұмыс істеуге болады.
- Вендорлардың API және плагиндеріне тәуелділікті азайту: деректермен жұмыс істеу енді API-ді пайдалану дағдыларына тәуелді емес.

CAD деректері мен талаптары талдауға ыңғайлы құрылымдық форматтарға түрлендірілген кезде, әзірлеушілер деректердің нақты схемалары мен жабық экожүйелерге тәуелділіктен құтылады.

Параметрлер арқылы жобалау: CAD және BIM-нің болашағы

Дүниежүзіндегі ешбір құрылыш жобасы CAD бағдарламасында басталмаған. Сызба немесе модель CAD-да пішін алмас бұрын, олар концептуализация кезеңінен өтеді, онда болашақ объектінің негізгі идеясы мен логикасын анықтайтын параметрлерге басты назар аударылады. Бұл кезең деректерді модельдеудің концептуалдық деңгейіне сәйкес келеді. Параметрлер жобалаушының санасында ғана болуы мүмкін, алайда идеалды жағдайда олар құрылымдық тізімдер, кестелер түрінде немесе деректер базаларында сақталады, бұл жобалау процесінің ашықтығын, қайталанбалығын және әрі қарай автоматтандыруды қамтамасыз етеді. --



Жобалау процесі – бұл CAD деректер базасын сыртқы талаптар арқылы ақпаратпен толтырудың итеративті процесі.

CAD моделдеуіне кіріспес бұрын, жобаның негізін құрайтын шекаралық параметрлерді анықтау маңызды. Бұл атрибуттар, басқа талаптар сияқты, деректерді пайдалану тізбегінің соңынан жиналады және олардың көмегімен жобадақ болашақ объектілердің шектеулери, мақсаттары мен негізгі сипаттамалары анықталады.

Модельдеудің өзі, сауатты құрастырылған талаптар болған кезде, параметрлік модельдеу құралдарының көмегімен 60-100% -ға толық автоматтандырылуы мүмкін (6.43-сурет). Жоба параметрлер түрінде сипатталғаннан кейін, оның қалыптасуы техникалық мүмкін, мысалы, заманауи A CAD ортасына немесе блендерде, UE, UE, OMNISE өнімдеріне салынған VisuagePer Dynamo сияқты визуалды бағдарламалуа тілдерін қолдана отырып, бағдарламалуа тілдерін қолданады.

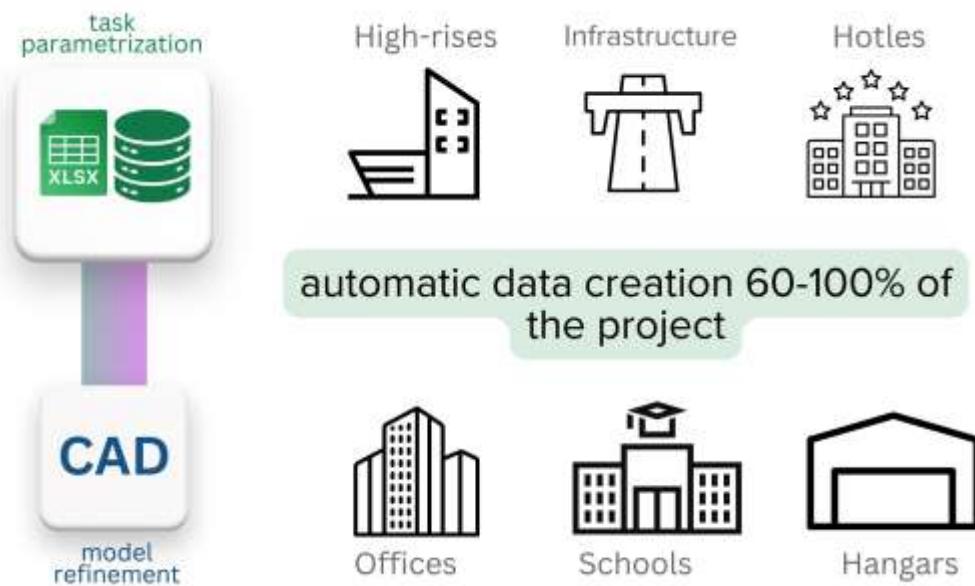


Рис. 6.44 Көптеген типтеген жобалар бүгінгі таңда параметрлік бағдарламалау құралдарының арқасында толық автоматты түрде жасалуда.

Бүгінгі таңда ірі өнеркәсіптік және типтік жобалар жобалаушылар бөлімінің қолымен емес, параметрлік құралдар мен визуалды бағдарламалау арқылы жасалуда. Бұл модельді деректерге негізделе отырып, нақты жобалаушы немесе менеджердің субъективті шешімдерінен емес, құруға мүмкіндік береді.

Мазмұн дизайннан бұрын келеді. Дизайн мазмұнсыз – бұл дизайн емес, тек декорация.

Джеффри Зельдман, веб-дизайнер және кәсіпкер

Процесс чертеж жасау немесе 3D-модельдеу кезеңінен басталмайды, ол талаптарды қалыптастырудан басталады. Дәл осы талаптар жобада қандай элементтердің қолданылатынын, басқа бөлімдер мен жүйелерге қандай деректерді беру қажеттілігін анықтайды. Тек құрылымдалған талаптардың болуы модельдерді автоматты түрде жүйелі түрде тексеруге мүмкіндік береді (мысалы, әр 10 минут сайын, жобалаушының жұмысын бұзбай).

Болашақта CAD- (BIM-) жүйесі деректер базасын толтыру үшін тек интерфейс болып қалуы мүмкін, ал модельдеу (физикалық деңгейде) қандай CAD құралында жүргізілетіні маңызды болмайды.

Дәл сол сияқты машина жасау саласында үш өлшемді модельдеу жиі қолданылады, бірақ жобалық қажетті немесе міндетті элементі болып табылмайды. Көп жағдайда классикалық 2D құжаттамасы жеткілікті – оның негізінде қажетті ақпараттық модель жасалады. Бұл модель салалық стандарттарға сәйкес құрылымдалған компоненттерден жинақталады және конструкцияны түсіну мен өндірісті ұйымдастыру үшін қажетті барлық ақпаратты қамтиды. Содан кейін оның негізінде

зауыттық ақпараттық модель қалыптасады, оған нақты өнімдер мен технологиялық карталар қосылады, олар технологиялық мамандардың қажеттіліктеріне бағытталған. Барлық процесті артық күрделіліксіз ұйымдастыруға болады, 3D графикасымен жүйені артық жүктемей, ол нақты артықшылықтар бермейтін жерлерде.

Маңызды түсіну, 3D модель мен CAD жүйесі басты рөл атқармауы тиіс – олар тек сандық және геометриялық талдау үшін құрал болып табылады. Геометриядан басқа, объектіні сипаттайтын барлық параметрлер мүмкіндігінше CAD ортасынан (BIM) тыссақталып, өндеделуі тиіс.

Параметрлер арқылы жобалау – бұл тек тренд емес, құрылыш саласының болашағы. Күрделі 3D модельдерді қолмен жасаудың орнына, жобалаушылар деректермен жұмыс істеп, оларды тексеріп, процестерді автоматтандырады, құрылыш саласын бағдарламалуа әлеміне жақындаады. Уақыт өте келе, жобалау процестері бағдарламалық қамтамасыз етуді өзірлеу принциптеріне негізделеді.

- Талаптарды жасау → Модельді жасау → Серверге жүктеу → Өзгерістерді тексеру → Pull request
- Pull request (қосылу сұранысы) шенберінде модельдің талаптарға сәйкестігін автоматты түрде тексеру жүргізіледі, бұл тексерулер жобалау басталғанға дейін немесе жобалау барысында жасалған.
- Деректердің сапасын тексергеннен кейін және өзгерістер мақұлданғаннан кейін олар жобада, жалпы деректер базасында енгізіледі немесе автоматты түрде басқа жүйелерге беріледі.

Қазіргі уақытта машина жасау саласында жобаның осындағы өзгерістері өзгеріс туралы хабарламаны қалыптастырудан басталады. Үқсас схема құрылыш саласын да күтуде: жобалау итеративті процесс ретінде жүзеге асырылатын болады, мұнда әрбір қадам параметрлік талаптармен бекітіледі. Мұндай жүйе жобалаушыларға автоматтандырылған тексерулер мен нақты талаптарға сәйкес автоматтандырылған pull request (қосылу сұрауы) жасауға мүмкіндік береді.

Болашақ жобалаушы – бұл, ең алдымен, деректер операторы, ал қолмен модельдеуші емес. Оның міндеті – жобаны параметрлік объектілермен толтыру, мұнда геометрия тек бір атрибут болып табылады.

Трансформацияда маңызды рөл атқаратын фактор – деректерді модельдеу, классификациялау және стандарттау маңыздылығын түсіну, олар кітаптың алдыңғы тарауларында егжей-тегжейлі қарастырылған. Болашақта жобалауды реттейтін нормативтік актілер кілт-мән параметрлерінің жүптары түрінде XLSX немесе XML-схемаларында рәсімделеді.

Құрылыш саласының болашағы – деректерді жинау, оларды талдау, тексеру және аналитикалық құралдар арқылы процестерді автоматтандыру. BIM (немесе CAD) – бұл соңғы мақсат емес, тек эволюция кезеңі. Мамандар деректермен тікелей жұмыс істей алатындарын түсінгенде, дәстүрлі CAD құралдарын айналып өтіп, «BIM» термині құрылыш жобасының құрылымдалған және гранулирленген деректерін пайдалану концепцияларына біртіндеп орын береді.

Трансформацияны жеделдететін негізгі факторлардың бірі – үлкен тілдік модельдердің (LLM) және оларға негізделген құралдардың пайда болуы. Бұл технологиялар жобалық деректермен жұмыс істеу тәсілін өзгертип, ақпаратқа терең API немесе вендор шешімдерін білмей-ақ қол жеткізуге мүмкіндік береді. LLM арқасында талаптарды жасау және CAD деректерімен өзара әрекеттесу процесі интуитивті және қолжетімді болады.

CAD жобалық деректерін өндеу процестерінде LLM-нің пайда болуы

CAD деректер базаларына және ашық, жеңілдетілген CAD форматтарына қол жеткізу құралдарының дамуына қосымша, жобалық деректерді өндеуде LLM құралдарының (Үлкен Тілдік Модельдер) пайда болуы революциялық өзгерістер енгізеді. Бұрын ақпаратқа қол жеткізу негізінен күрделі интерфейстер арқылы жүзеге асырылып, бағдарламалау мен API білімін талап етсе, енді деректермен табиғи тіл арқылы өзара әрекеттесу мүмкіндігі пайда болды.

Инженерлер, менеджерлер және техникалық білімі жоқ жобалаушылар жобалық деректерден қажетті ақпаратты қарапайым тілде сұрау формулировкасы арқылы ала алды. Деректер құрылымдалған және қолжетімді болған жағдайда (Сурет 4.113), LLM чатында «10 текше метрден асатын барлық қабырғаларды типтер бойынша топтастырылған кесте түрінде көрсет» деген сұрақ қою жеткілікті – модель автоматты түрде бұл сұранысты SQL немесе Pandas кодына айналдырып, нәтижелік кесте, график немесе тіпті дайын құжатты қалыптастырады.

Төменде LLM модельдерінің әртүрлі CAD- (BIM-) форматтарында ұсынылған жобалық деректермен қалай өзара әрекеттесетіні туралы бірнеше нақты мысалдар келтірлген.

- ❶ RVT форматының CAD жобасына LLM чатында сұраныс мысалы (Сурет 4.113) кестелік деректер фрейміне конвертацияланғаннан кейін:-

RVT файлына негізделген деректерді «Тип атаяу» бойынша топтастырып, «Көлем» параметрін қосу арқылы жинақта және топтадағы элементтер санын көрсет. Және, нөлдік мәндерсіз, барлығын көлденен ғистограмма түрінде көрсетіңіз.

- LLM-нің көлдененг гистограмма түріндегі жауабы (PNG форматы):

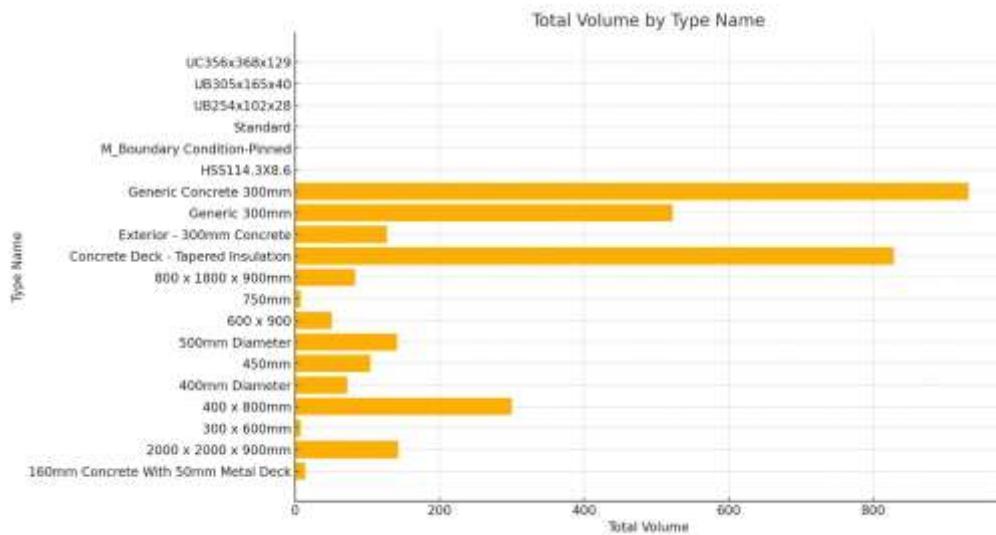


Рис. 6.45 17 рет тышқанды басу немесе 40 жол код жазу орнына плагиндерді пайдалану арқылы LLM бізге мәтіндік сұрау арқылы QTO кестесін бірден береді.

- «Қабырғалар» категориясынан жалпы алаңы мен саны бар қабырға түрлерінің QTO кестесін қалыптастыру үшін LLM-чатқа мәтіндік сұрау жасаймыз:

Датафреймдегі «Категория» параметрінде «OST_Walls» бар элементтерді ғана ал, оларды «Тип атауы» бойынша топтастырып, «Алан» бағанының мәндерін қос, санын қосып, нөлдік мәндерді жойып, кесте түрінде көрсет.

- LLM-нің жауап ретінде дайын QTO кестесі:

Type Name	Total Area	Count
CL_W1	393.12 sq m	10
Cavity wall_sliders	9.37 sq m	1
Foundation - 300mm Concrete	30.90 sq m	1
Interior - 165 Partition (1-hr)	17.25 sq m	3
Interior - Partition	186.54 sq m	14
Retaining - 300mm Concrete	195.79 sq m	10
SH_Curtain wall	159.42 sq m	9
SIP 202mm Wall - conc clad	114.76 sq m	4
Wall - Timber Clad	162.91 sq m	8

Рис. 6.46 Табиғи тілде QTO кестесін жасау CAD- (BIM-) құралдарын пайдалану арқылы алынған сапалы нәтижемен бірдей.

- IFC форматындағы жобаны кестелік датафреймге конвертациялағаннан кейін, кез келген LLM чатында үқсас мәтіндік сұрау енгіземіз:

«Parent» параметрінде Level 1 және Level 2 мәндері бар элементтерді ғана ал, «Category» параметрінде IfcSlab мәндері бар элементтерді ал, содан кейін осы элементтерді «ObjectType» параметрі бойынша топтастырып, «PSet_RVT_Dimensions Area» параметріндегі мәндерді қосып, оларды дөңгелек диаграмма түрінде көрсет.

■ IFC деректерінен элементтер тобының дайын Pie chart түріндегі LLM-нің жауабы:

For ■ Floor:127mm Slab on Grade, ■ Floor:150mm Exterior Slab on Grade,
 ■ Floor:Finish Floor - Ceramic Tile, ■ Floor:Finish Floor - Wood, and
 ■ Floor:Residential - Wood Joist with Subflooring



Рис. 6.47 IFC деректеріне сұрау нәтижесі құрылымдалған форматта кез келген график түрінде болуы мүмкін, ол деректерді түсінуге ыңғайлы.

Алынған дайын шешімдердің (Рис. 6.45 - Рис. 6.47) артында Python тілінде Pandas кітапханасын пайдалана отырып жазылған он шақты код жолы жатыр. Алынған кодты LLM чатынан көшіріп, кез келген локалды немесе онлайн IDE-де пайдаланып, LLM чатынан тыс бірдей нәтижелер алуға болады.-

Бір және сол LLM чатында біз тек 3D CAD (BIM) форматтарынан алынған жобалармен ғана емес, сонымен қатар DWG форматындағы жазық сызбалармен де жұмыс істей аламыз, оларды құрылымдалған формаға конвертациялағаннан кейін LLM чатында элементтер топтары туралы деректерді сызықтар немесе 3D геометриялар түрінде көрсету үшін сұрау жасай аламыз.

LLM және Pandas көмегімен DWG файлдарын автоматтандырылған талдау

DWG файлдарынан деректерді өңдеу процесі ақпараттың құрылымсыздығына байланысты әрқашан күрделі міндет болып табылады, ол арнайы бағдарламалық қамтамасыз етуді және жиі қолмен талдауды талап етеді. Алайда, жасанды интеллект пен LLM құралдарының дамуы арқасында бүгінгі күні бұл негізінен қолмен орындалатын процестің көптеген кезеңдерін автоматтандыру мүмкін болды. DWG сызбаларымен жұмыс істеу үшін LLM-ге (бұл мысалда ChatGPT) сұраулар тізбегін қарастырайық, олар жобамен жұмыс істеуге мүмкіндік береді:

- DWG деректерін қабаттар, ID және координаттар бойынша сұзу
- Элементтердің геометриясын визуализациялау
- Параметрлер негізінде сызбаларды автоматты түрде аннотациялау
- Қабырғалардың полилинияларын горизонтальды жазықтыққа айналдыру

- Жазық деректерден интерактивті 3D визуализацияларды жасау
- Құрылыш деректерін күрделі CAD құралдарынсыз құрылымдау және талдау

Біздің жағдайда Pipeline құру процесі LLM арқылы кодты кезең-кезеңімен генерациялаудан басталады. Алдымен тапсырманы сипаттайтын сұрау қалыптастырылады. ChatGPT Python кодын генерациялайды, ол орындалып, талданады, нәтижесі чат ішінде көрсетіледі. Егер нәтиже күтілгенге сәйкес келмесе, сұрау түзетіледі, және процесс қайталанады.

Пайплайн – бұл деректерді өңдеу және талдау үшін автоматтандырылған қадамдардың тізбегі. Мұндай процесте әр кезең деректерді кіріс ретінде қабылдайды, түрлендірлерді орындауды және нәтижені келесі қадамға жібереді.

Қажетті нәтижені алғаннан кейін код LLM-ден көшіріліп, кез келген ыңғайлы IDE-де, біздің жағдайда Kaggle.com платформасында блоктар түрінде кодқа енгізіледі. Алынған код фрагменттері деректерді жүктеуден бастап, олардың соңғы талдауына дейінгі процесті автоматтандыратын біртұтас Пайплайнға біріктіріледі. Мұндай тәсіл бағдарламалауда терен сараптама қажет етпей, аналитикалық процестерді жылдам әзірлеуге және масштабтауға мүмкіндік береді. Төменде көрсетілген барлық фрагменттердің толық кодын және сұрау мысалдарын Kaggle.com платформасында «DWG Analyse with ChatGPT | DataDrivenConstruction» сұрауымен таба аласыз.

DWG деректерімен жұмыс процесін бастаймыз, құрылымдалған түрге конвертацияланғаннан кейін (Рис. 4.113), классикалық кезеңнен – барлық сызба деректерінен, біздің міндетімізге қажетті қабырға элементтерін, нақты полилинияларды ('ParentID' параметрі сыйықтарды топтастыруға мүмкіндік береді) топтастырудан бастаймыз, олардың «Слой» параметрінде (деректер фреймінің бағанаасында) «wall» деген жолдық мән бар.

- ❷ Мұндай тапсырма үшін код алу және нәтижені сурет түрінде алу үшін LLM-ге келесі сұранысты жазу қажет:

Алдымен DWG-ден алған деректер фреймі 'Layer', 'ID', 'ParentID' және 'Point' сияқты белгілі бір бағаналарды қамтитынын тексеріңіз. Содан кейін 'Layer' бағанасынан 'wall' жолын қамтитын идентификаторларды сұзгіден өткізіңіз. 'ParentID' бағанаасынан осы идентификаторларға сәйкес элементтерді табыңыз. 'Point' бағанаасындағы деректерді тазалау және бөлу функциясын анықтаңыз. Бұл дөңгелектерді жоюды және 'x', 'y' және 'z' координаттарына мәндерді бөлуді қамтиды. Деректер графигін matplotlib көмегімен құрыңыз. Эрбір бірегей «ParentID» үшін «Point» координаттарын байланыстыратын жеке полилинияны салыңыз. Егер мүмкін болса, бірінші және соңғы нұктелерді байланыстыруды қамтамасыз етіңіз. Тақтадағы белгілер мен тақырыптарды орнатыңыз, х және у осьтері бойынша бірдей масштабтауды қамтамасыз етіңіз.

- ❸ LLM-нің жауабы Python тілінде жазылған кодты генерациялау арқылы дайын суретті көрсетеді:

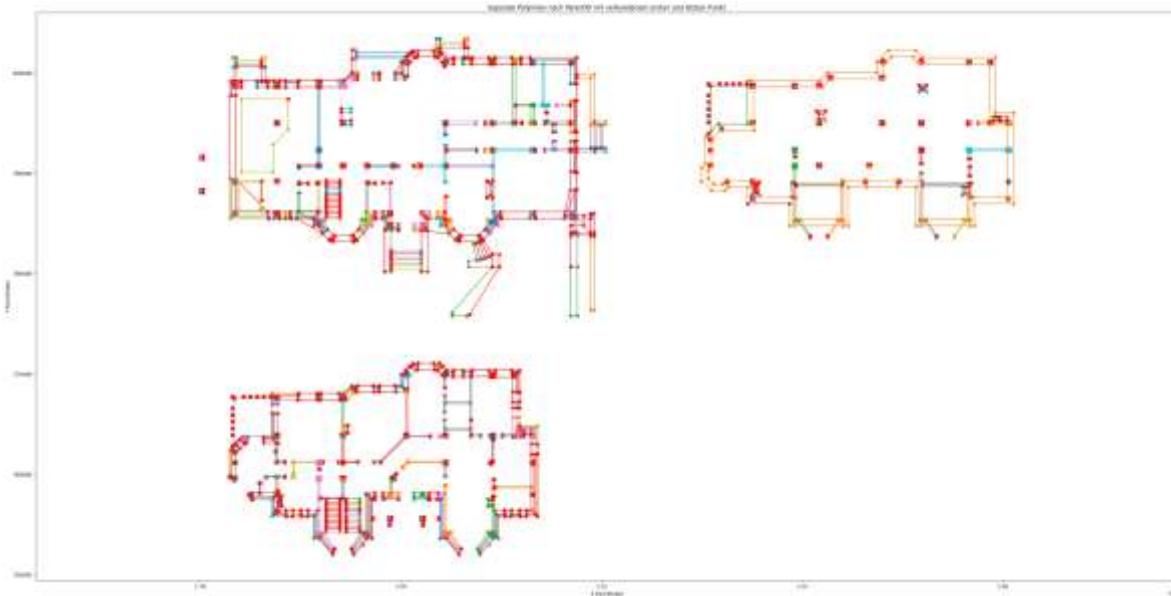


Рис. 6.48 LLM код DWG-файлының барлық «wall» қабатының сыйықтарын шығарып, олардың координаттарын тазалап, полилинияларды Python кітапханаларының бірін пайдаланып салды.

- Енді полилинияларға әрқайсысының қасиеттерінде бар аудан параметрін қосамыз (декреттер фреймінің бір бағанасында):

Енді әр полилиниядан тек бір «ParentID» алыңыз - бұл идентификаторды «ID» бағанасынан табыңыз, «Area» мәнін алыңыз, 1 000 000-та бөліп, бұл мәнді графикке қосыңыз.

- LLM-нің жауабы жаңа графикті көрсетеді, онда әр полилинияның ауданымен белгісі пайда болады:



Рис. 6.49 LLM кодын толықтырып, әр полилинияның аудан мәндерін алып, оны сзықтардың визуализациясымен бірге суретке қосады.

- Келесі қадамда әр полилинияны горизонтальды сзызыққа айналдырамыз, 3000 мм білдіктікте параллель сзызық қосамыз және оларды бір жазықтыққа біріктіріп, осылайша қабырға элементтерінің беттерін көрсету үшін.

«Layer» бағанындағы «wall» мәні бар барлық элементтерді алу қажет. Бұл ID-лерді «ID» бағанынан тізім ретінде алып, оларды датафреймнің «ParentID» бағанынан табыңыз. Барлық элементтер – бұл полилинияға біріктірілетін сзықтар. Әр сзызықтың «Point» бағанында x, y координаталарымен бірінші нүктесінің геометриясы бар. Әр полилинияны кезекпен алып, 0,0 нүктесінен горизонталь бойынша полилинияның әр сегментінің ұзындығын құрастырыңыз. Полилинияның әр сегментінің ұзындығын бір сзызыққа жинаңыз. Содан кейін дәл сондай сзықтарды 3000 жоғарыда жүргізіп, барлық нүктелерді бір жазықтықта қосыңыз.

- LLM жауп кодын шығарады, ол қабырғалардың жазықтықтағы проекцияларын құруға мүмкіндік береді.

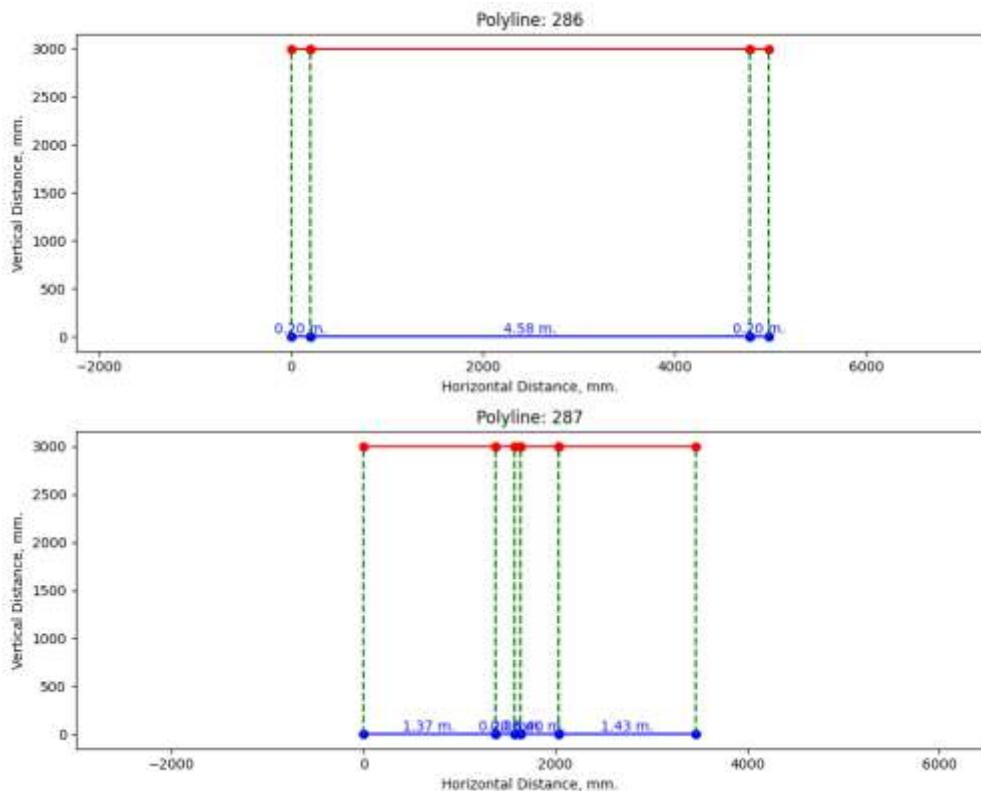


Рис. 6.410 Эр полилинияны промпттар арқылы раскладкаға айналдырамыз, ол қабырғалардың жазықтықтарын LLM чаты арқылы тікелей көрсетеді.

- Енді екі өлшемді проекциядан қабырғалардың 3D-моделіне өтеміз, полилиниялардың жоғарғы және төменгі қабаттарын қосу арқылы:

3D-де қабырға элементтерін визуализациялаңыз, полилинияларды $z = 0$ және $z = 3000$ мм биіктігінде қосу арқылы. Гимараттың қабырғаларын білдіретін замкнутую геометрияны жасау үшін Matplotlib 3D графигін пайдаланыңыз.

- LLM интерактивті 3D графикті генерациялайды, онда әр полилиния жазықтықтар жиынтығы ретінде ұсынылады. Пайдаланушы компьютерлік тышқан арқылы элементтер арасында еркін қозғалуға мүмкіндік алады, 3D режимінде модельді зерттей алды, чатты IDE-ге кодты көшіре отырып:

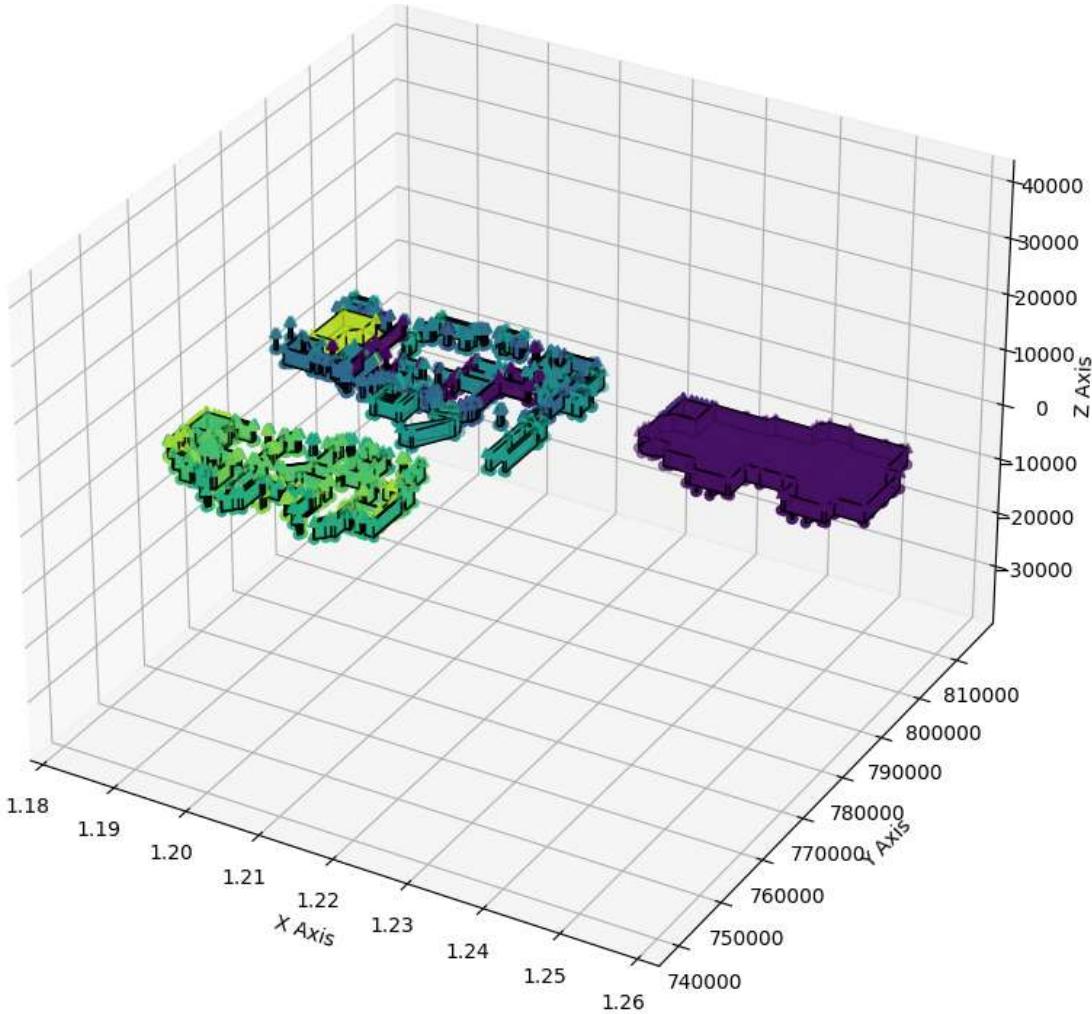


Рис. 6.411 LLM 3D көрінісінде сыйбадағы жазық сыйықтарды визуализациялау үшін код [129] құруға көмектесті, оны IDE ішіндегі 3D көрушіде зерттеуге болады.

DWG файлын бастапқы конвертациялау мен жүктеден бастап соңғы нәтижеге дейін логикалық, және қайталанатын Pipeline құру үшін, әр қадамнан кейін LLM-нің сгенерленген код блогын IDE-ге көшіру ұсынылады. Осылайша, сіз тек чаттағы нәтижені тексеріп қана қоймай, оны өзіңіздің даму ортақызыда бірден іске қосасыз. Бұл процесті кезең-кезеңімен құруға, оны түзетіп, қажетіне қарай бейімдеуге мүмкіндік береді.

Барлық фрагменттердің Pipeline коды (Рис. 6.48 - Рис. 6.411) және сұрау мысалдары «DWG Analyse with ChatGPT | DataDrivenConstruction» сұрауымен Kaggle.com платформасында табуға болады [129]. Kaggle-де кодты және қолданылған промпттарды тексеріп қана қоймай, бастапқы DWG да-тафреймдерімен бүкіл Pipeline-ды бұлттық ортада қосымша бағдарламалық қамтамасыз етусіз және IDE-ді орнатусыз тегін сынап көруге болады.--

Осы тарауда ұсынылған тәсіл DWG жобалары негізінде құжаттарды тексеру, өңдеу және генерациялауды толық автоматтандыруға мүмкіндік береді. Дамытылған пайплайн жеке сызбаларды өңдеуге де, ондаған, жүздеген және мыңдаған DWG файлдарын пакетпен өңдеуге де жарамды, әр жоба бойынша қажетті есептер мен визуализацияларды автоматтыйтуарда.

Процесті кезең-кезеңімен және ашық түрде құруға болады: алдымен CAD-файлдардан деректер автоматтыйтуарда XLSX форматына конверттеледі, содан кейін деректер фрейміне жүктеледі, одан кейін топтастыры, тексеру және нәтиже қалыптастыры жүзеге асырылады – бұның бәрі бір Jupyter ноутбуқта немесе Python-скрипtte, кез келген танымал IDE-де іске асырылады. Қажет болған жағдайда процесс жобалық құжаттаманы басқару жүйелерімен интеграциялау арқылы оңай кеңейтіледі: CAD-файлдарды автоматтыйтуарда берілген критерийлер бойынша шығарып, нәтижелерді сақтау жүйесіне қайтара алады және пайдаланушыларды нәтижелердің дайын екендігі туралы хабардар ете алады – электрондық пошта немесе мессенджерлер арқылы.

LLM чаттары мен агенттерін жобалық деректермен жұмыс істеу үшін пайдалану ма-
мандаудырылған CAD бағдарламаларына тәуелділікті азайтады және архитектуралық
жобаларды талдау мен визуализациялауды интерфейспен қолмен өзара әрекеттесуіз
– тышқанмен басу мен мәзір бойынша құрделі навигацияны есте сақтаусыз орындауға
мүмкіндік береді.

Құрылыш саласында LLM, гранулирленген құрылымдық деректер, DataFrame және колонкалыш деректер базалары туралы күн сайын көбірек естуге болады. Әртүрлі дерекқорлар мен CAD форматтарынан қалыптасқан унифицированный екі өлшемді DataFrame қазіргі заманғы аналитикалық құралдар үшін тамаша отын болады, олармен басқа экономиканың салаларында мамандар белсенді жұмыс істейді.

Автоматизация процесі айтарлықтай жеңілдейді – жабық нишалық өнімдердің API-ін зерттеу мен параметрлерді талдау немесе трансформациялау үшін құрделі скрипттер жазудың орнына, енді тапсырманы жеке мәтіндік командалар жиынтығы түрінде формулировка жасау жеткілікті, олар қажетті Pipeline немесе Workflow-процесіне жинақталады, қажетті бағдарламалау тілі үшін, олар кез келген құрылғыда тегін іске қосылады. CAD- (BIM-) құралдарынан жаңа өнімдер, форматтар, плагиндер немесе жаңартулардың шығуын күтудің қажеті болмайды. Инженерлер мен құрылышшылар деректермен тәуелсіз жұмыс істеу мүмкіндігіне ие болады, қарапайым, тегін және түсінікті құралдарды пайдалана отырып, LLM чаттары мен агенттері көмектеседі.

Келесі қадамдар: жабық форматтардан ашық деректерге көшу

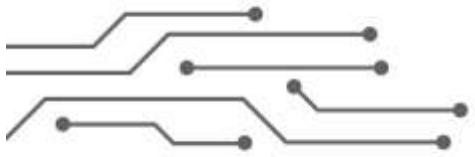
Болашақтағы жобалық деректермен жұмыс істегендеге, ешкімге де меншікті құралдардың геометриялық ядроларында терең түсінік алу немесе бірдей ақпаратты қамтитын жүздеген үйлесімсіз форматтарды зерттеу шынымен қажет болмайды. Дегенмен, ашық құрылымдық деректерге көшу неге маңызды екенін түсінбей, жаңа тегін құралдарды, ашық деректер мен тәсілдерді пайдаланудың қажеттілігін дәлелдеу қын болады, оларды бағдарламалық қамтамасыз ету вендорлары алға жылжытпайды.

Бұл бөлімде біз CAD (BIM) деректерінің негізгі ерекшеліктерін, олардың шектеулерін және мүмкіндіктерін қарастырудық, және маркетингтік уәделерге қарамастан, инженерлер мен жоба-

лаушылар күн сайын жобалық ақпаратты алу, беру және талдау кезінде қыындықтарға тап болатынын атап өттік. Бұл жүйелердің архитектурасын түсіну және LLM негізіндегі ашық форматтар мен автоматизацияның баламалы тәсілдерімен танысу, тіпті бір маманның өмірін айтартылғанда жеңілдетуге қабілетті, компанияларды айтпағанда. Осы бөлімнің қорытындысында, күнделікті тапсырмаларыңызда қарастырылған тәсілдерді қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

- Жобалық деректермен жұмыс істеу үшін құралдарыңызды кеңейтіңіз.
- Пайдаланатын CAD- (BIM-) жүйелеріңізден деректерді алу үшін қолжетімді плагиндер мен утилиталарды зерттеңіз.
- Жабық форматтардан деректерді автоматты түрде алу үшін арнайы бағдарламалық қамтамасыз етуді қолмен ашудың қажеті жоқ SDK және API-лерді зерттеңіз.
- Ашық параметрлік емес геометрия форматтарымен (OBJ, glTF, USD, DAE) және сәйкес ашық кітапханалармен жұмыс істеудің негізгі дағдыларын меңгеріңіз.
- Геометриядан бөлек, жобаның метадеректерін сақтау жүйесін CAD (BIM) шешімдерінен тыс ойластырып, талдауды және басқа жүйелермен интеграцияны жеңілдетуге тырысыңыз.
- Деректерді форматтар арасында түрлендіру мәселелерін автоматтандыру үшін LLM пайдаланыңыз.
- Жобалық ақпаратты өндөу үшін өз процестеріңізді жасаңыз.
- Модельдеу тапсырмалары мен талаптарын қарапайым және құрылымдалған форматтарда параметрлер мен олардың мәндері арқылы сипаттауды бастаңыз.
- Жиі орындалатын операциялар үшін жеке скрипттер немесе код блоктары кітапханасын жасаңыз.
- Жұмысыңызда ашық стандарттарды қолдануды насиҳаттаңыз.
- Әріптеріңіз бен серіктестеріңізге деректерді бағдарламалық қамтамасыз ету вендорларының экожүйесімен шектелмеген ашық форматтарда алмасуды ұсыныңыз.
- Құрылымдалған деректерді қолданудың артықшылықтарын нақты мысалдармен көрсетіңіз.
- Жабық форматтармен байланысты мәселелер мен мүмкін шешімдер туралы талқылауларды бастаңыз.

CAD- (BIM-) платформаларға қатысты компания саясатын өзгерте алмасаңыз да, ашық форматтар мен жұмыс істеу принциптерін жеке түсіну жұмысыңыздың тиімділігін айтартылған арттыруға мүмкіндік береді. Әртүрлі форматтардан деректерді алу және түрлендіру үшін өз құралдарыңыз бен әдістеріңізді жасап, жұмыс процестерін оңтайландырып қана қоймай, стандартты бағдарламалық шешімдердің шектеулерін айналып өтуге мүмкіндік беретін икемділікке ие боласыз.



VII БӨЛІМ

ДЕРЕКТЕРГЕ НЕГІЗДЕЛГЕН ШЕШІМ ҚАБЫЛДАУ, АНАЛИТИКА, АВТОМАТТАНДЫРУ ЖӘНЕ МАШИНАЛЫҚ ОҚУ

Жетінші бөлім құрылыш саласындағы деректер аналитикасы мен процестерді автоматтандыруға арналған. Мұнда деректер шешім қабылдаудың негізі ретінде қалай қызмет ететіні және тиімді талдау үшін ақпаратты визуализациялау принциптері түсіндіріледі. Негізгі тиімділік көрсеткіштері (KPI), инвестициялардың қайтарымын (ROI) бағалау әдістері және жобаларды мониторингтеу үшін ақпараттық панельдер құру егжей-тегжейлі сипатталады. ETL (Extract, Transform, Load) процестеріне және оларды автоматтандыруға, яғни деректерді талдау үшін құрылымдалған ақпаратқа айналдыруға мүмкіндік беретін конвейерлерді (Pipeline) қолдануға ерекше назар аударылады. Apache Airflow, Apache NiFi және n8n сияқты жұмыс процестерін оркестрациялау құралдары бағдарламаладың терең білімінсіз автоматтандырылған деректер конвейерлерін құруға мүмкіндік береді. Үлкен тілдік модельдер (LLM) деректерді талдауды жеңілдету және рутиналық тапсырмаларды автоматтандыру үшін маңызды рөл атқарады.

ТАРАУ 7.1.

ДЕРЕКТЕР АНАЛИТИКАСЫ ЖӘНЕ ДЕРЕКТЕРГЕ НЕГІЗДЕЛГЕН ШЕШІМ ҚАБЫЛДАУ

Деректерді жинау, құрылымдау, тазалау және тексеру кезеңдерінен кейін талдауға жарамды тұтас деректер массиві қалыптасты. Кітаптың алдыңғы бөлімдерінде PDF құжаттарынан, кездесулердің мәтіндік жазбаларынан, CAD модельдерінен және геометриялық деректерден алынған әртүрлі көздерді жүйелендіру және құрылымдау қарастырылды. Ақпаратты әртүрлі жүйелер мен классификаторлардың талаптарына сәйкес тексеру және сәйкестендіру, дубликаттар мен қарама-қайшылықтарды жою процесі егжей-тегжейлі сипатталған.

Осы деректер негізінде (кітаптың үшінші, төртінші бөліктері) жүргізілген барлық есептеулер – қарапайым түрлендірulerден бастап, мерзімдер, шығындар және ESG көрсеткіштерін есептеуге дейін (бесінші бөлім) – аналитикадағы ірілендірілген міндеттерді білдіреді. Олар жобаның қазіргі жағдайын түсіну, оның параметрлерін бағалау және кейінгі шешім қабылдау үшін негіз қалыптастырады. Нәтижесінде, есептеулер нәтижесінде деректер бөлшектелген жазбалар жиынтығынан басқарылатын ресурсқа айналады, ол бизнес үшін негізгі сұрақтарға жауап береді.

Алдыңғы тарауларда деректерді жинау және олардың сапасын бақылау процестері құрылыс саласына тән типтік бизнес-кейіс және процестер үшін егжей-тегжейлі қарастырылды. Бұл контексте аналитика басқа салалардағы қолданулармен көп жағынан үқсас, бірақ бірқатар ерекшеліктерге ие.

Келесі тарауларда деректерді талдаудың ірілендірілген процесі, бастапқы ақпаратты алу және оны түрлендіруден бастап, мақсатты жүйелер мен құжаттарға кейінгі беруге дейінгі автоматизация кезеңдері егжей-тегжейлі қарастырылады. Алдымен, деректер негізінде шешім қабылдаудың жеке аспектілеріне арналған теориялық бөлім ұсынылады. Содан кейін, келесі тарауларда автоматизация және ETL-Pipeline құрумен байланысты практикалық бөлім басталады.

Шешім қабылдауда деректер ресурсы

Деректер негізінде шешім қабылдау процесі жиі итеративті процесс болып табылады және әртүрлі ақпарат көздерінен жүйелі ақпарат жинаудан басталады. Табиғи айналымға үқсас, деректердің жеке элементтері мен тұтас ақпараттық жүйелер біртіндеп топырақта түседі - компаниялардың ақпараттық қоймаларында жиналады (Рис. 1.32). Уақыт өте келе, бұл деректер, түскен жапырақтар мен бұтақтар сияқты, құнды материалға айналады. Деректер инженерлері мен аналитиктердің мицелийі ақпаратты болашақта пайдалану үшін ұйымдастырып, дайындауды және түскен деректер мен жүйелерді құнды компостқа айналдырады, жаңа өскіндер мен жаңа жүйелерді өсіру үшін (Рис. 1.25).-

Әр түрлі салаларда аналитиканың кеңінен қолданылу тенденциясы жаңа дәүірдің басталуын білдіреді, мұнда деректермен жұмыс көсіби қызметтің негізіне айналады (Рис. 7.11). Құрылыс саласының мамандарына осы өзгерістерге бейімделу және деректер мен аналитика дәүіріне көшуге дайын болу маңызды.-

Деректерді кестелер арасында қолмен жылжыту және есептеулерді қолмен орындау біртіндеп өткенге айналуда, автоматизацияға, деректер ағындарын талдауға, аналитикаға және машиналық қоқытуға орын беруде. Бұл құралдар қазіргі заманғы шешім қабылдау жүйесінің негізгі элементтеріне айналуда.

McKinsey компаниясының "Перезагрузка. McKinsey-дің цифрлық технологиялар мен жаңанды интеллект дәуірінде бәсекелестікті жеңу жөніндегі нұсқаулығы" [130] атты кітабында 2022 жылы әртүрлі өнірлер, салалар және функционалдық бағыттардан 1 330 жоғары басшылардың қатысуымен жүргізілген зерттеу келтірілген. Зерттеу нәтижелері бойынша, 70% жетекшілер өз идеяларын қалыптастыру үшін озық аналитиканы пайдаланады, ал 50% – шешім қабылдау процестерін жақсарту және автоматизациялау үшін жаңанды интеллектін енгізеді.



Рис. 7.11 Деректерді талдау және аналитика - компанияда шешім қабылдау жылдамдығын арттырудың негізгі құралы.

Деректерді талдау, мицелийдің таралуына үқсас, өткен шешімдердің гумусы арқылы өтеді, жеке жүйелерді байланыстырып, менеджерлерді құнды білімдерге бағыттайты. Бұл білімдер, шіріген ағаштардан алынған қоректік заттар сияқты, компаниядағы жаңа шешімдерді қоректендіреді, тиімді өзгерістер мен сапалы ақпараттық өсімге әкеледі, бай және сау топырақтан пайда болатын жаңа бұтақтар мен өскіндерге үқсас (Рис. 1.25).-

Сандардың айтуға тиіс маңызды тарихы бар. Олар сізден анық және сенімді дауыс беруін күтеді [131].

– Стивен Фью, Деректерді визуализациялау бойынша сарапшы

Орта және шағын компанияларда ақпаратты алу және талдау үшін дайындау жұмысы бүгінгі күні өте еңбек-intensive процесс болып табылады (Рис. 7.12), XVIII ғасырдағы көмір өндірумен салыстыруға болады. Соңғы уақытқа дейін деректерді алу және дайындау жұмысы негізінен шектеулі құралдармен жұмыс істейтін, әртүрлі құрылымдалмаған, әлсіз құрылымдалған, аралас және жабық дереккөздерден деректермен айналысатын тар мамандарға арналған.

Басшылар мен шешім қабылдаушылар, әртүрлі деректер мен жүйелермен жұмыс істеу тәжірибесіне ие болмағанымен, олардың негізінде шешім қабылдауға мәжбүр. Нәтижесінде, қазіргі құрылым саласындағы деректерге негізделген шешім қабылдау соңғы онжылдықтарда автоматтандырылған процесс емес, алғашқы көмір шахталарындағы кеншілердің көп күндік қол еңбегіне ұқсас.

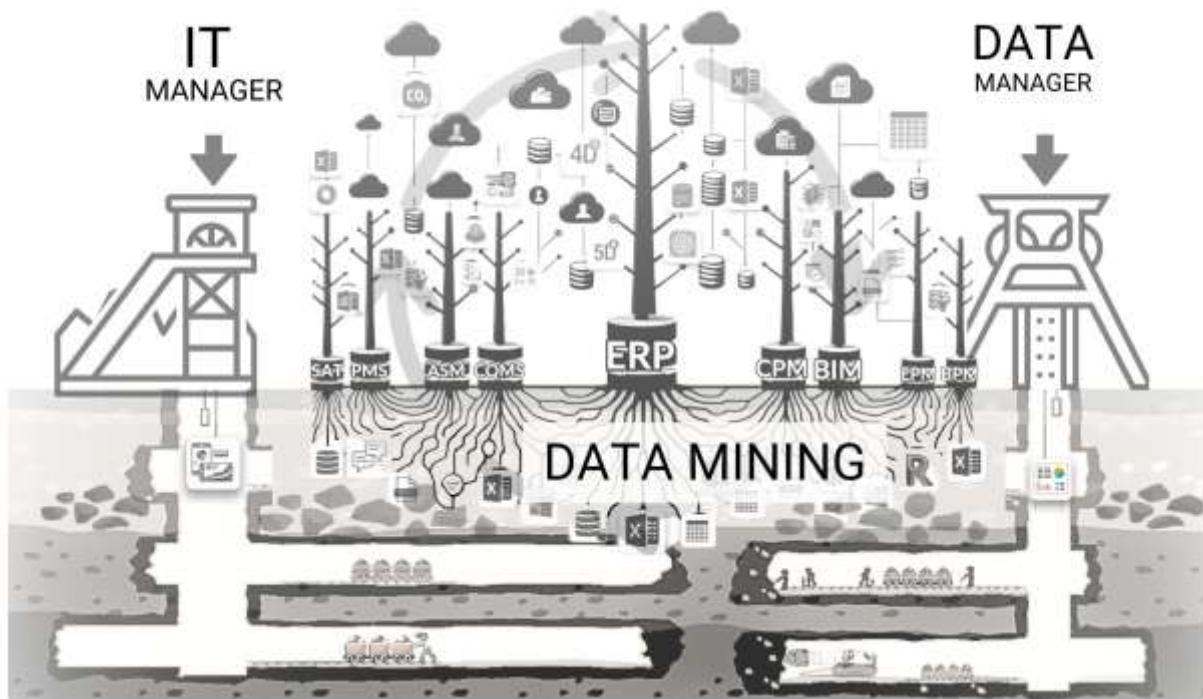


Рис. 7.12 Деректерді алу процесінде мамандар деректерді дайындаудың күрделі жолынан өтеді – тазалаудан бастап, кейінгі аналитика үшін құрылымдауға дейін.

Заманауи деректерді алу әдістері құрылым индустріясында, XII ғасырдағы кеншілердің примитивті техникасынан гөрі, сөзсіз прогрессивті, бірақ бұл әлі де күрделі және жоғары тәуекелді міндет болып табылады, ол тек ірі компаниялардың қолынан келетін елеулі ресурстар мен сараптаманы тарап етеді. Жинақталған өткен жобалардың мұрасынан деректерді алу және аналитикамен айналысадын процесс, соңғы уақытқа дейін, негізінен, онжылдықтар бойы деректерді жинап, сақтаған ірі, технологиялық дамыған компаниялармен жүзеге асырылды.

Бұрын аналитикада жетекші рөлді технологиялық жағынан жетілген компаниялар атқарды, олар онжылдықтар бойы деректерді жинады. Бүгінгі күні жағдай өзгеруде: деректер мен оларды өңдеу құралдарына қолжетімділік демократиялануда – бұрын күрделі шешімдер енді барлығына қолжетімді және тегін.

Аналитиканы қолдану компанияларға нақты уақыт режимінде дәл және негізделген шешімдер қабылдауға мүмкіндік береді. Тарихи деректердің қаржылық негізделген шешімдер қабылдауға қалай көмектесетін көрсететін практикалық мысал келтірілген:

- Жоба менеджері – «Қазір қаладағы бетонның орташа бағасы 82 €/м³, біздің сметамызда 95 €/м³».
- Сметашы – «Алдыңғы жобалар бойынша артық шығын шамамен 15% болды, сондықтан мен сақтандым».
- Деректер менеджері немесе тапсырыс беруші тарарапынан бақылауышы инженер – «Келініздер, соңғы үш тендер бойынша аналитиканы қарастырайық».

Өткен жобалардан алғынған DataFrame-ді талдау нәтижесінде:

- **Орташа нақты сатып алу бағасы: 84,80 €/м³**
- **Орташа артық шығын коэффициенті: +4,7%**
- Сметада ұсынылған ставка: ~85 €/м³

Мұндай шешім субъективті сезімдерге емес, нақты тарихи статистикаға негізделеді, бұл тәуекелдерді азайтуға және тендерлік ставканың негізділігін арттыруға мүмкіндік береді. Өткен жобалардан алғынған деректерді талдау жаңа, дәл шешімдердің «органикалық тыңайтқышы» ретінде қызмет етеді.

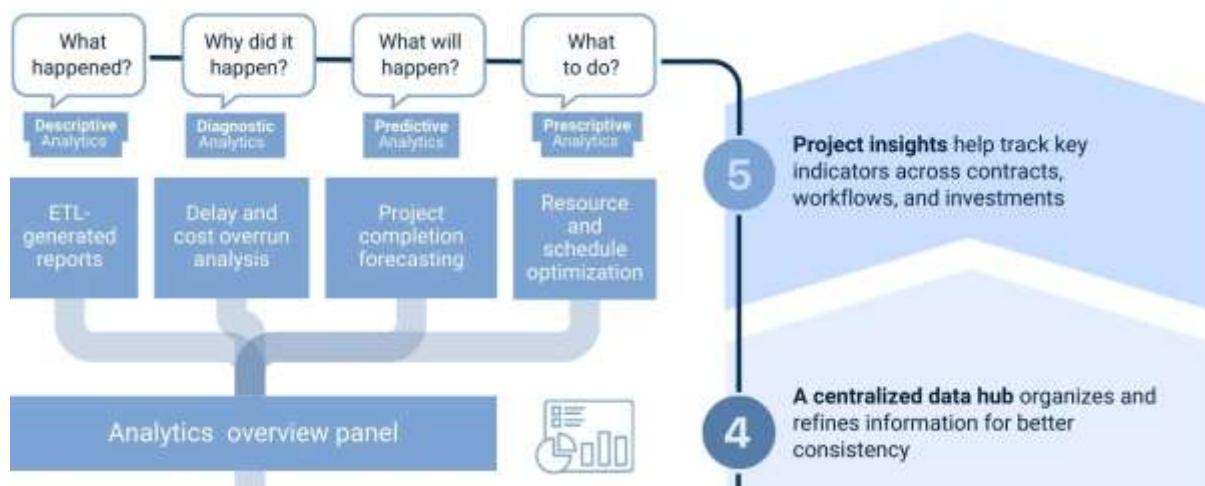


Рис. 7.13 Деректер аналитикасы үш негізгі сұраққа жауап береді: не болды, неге бұл болды және келесі не істей керек.

Басшылар мен шешім қабылдаушылар жиі әртүрлі деректер мен жүйелермен жұмыс істей қажеттілігіне тап болады, ал техникалық дайындықтары жеткіліксіз. Мұндай жағдайларда деректерді түсінуде визуализация - аналитикалық процестің алғашқы және маңызды кезеңдерінің бірі болып табылады. Ол ақпаратты көрнекі және түсінікті түрде ұсынуға мүмкіндік береді.

Деректерді визуализациялау: түсіну мен шешім қабылдаудың кілті

Қазіргі құрылымы саласында, жобалық деректердің күрделілігі мен көп деңгейлі құрылымымен сипатталатын, визуализация маңызды рөл атқарады. Деректерді визуализациялау жобалар мен инженерлерге үлкен, әртүрлі деректер көлемінде жасырын күрделі заңдылықтар мен тенденцияларды визуализациялауға мүмкіндік береді.

Деректерді визуализациялау жобаның жағдайын түсінуді жеңілдетеді: ресурстардың бөлінуі, шығындардың динамикасы немесе материалдардың пайдаланылуы. Графиктер мен диаграммалар арқылы күрделі және құрғақ ақпарат қолжетімді және түсінікті болады, бұл назар аударуды қажет ететін негізгі салаларды тез анықтауға және әлеуетті проблемаларды анықтауға мүмкіндік береді.

Деректерді визуализациялау тек ақпаратты интерпретациялауды жеңілдетпейді, сонымен қатар аналитикалық процесс пен негізделген басқарушылық шешімдер қабылдаудағы ең маңызды кезең болып табылады, «не болды?» және «қалай болды?» деген сұрақтарға жауап беруге көмектеседі (Рис. 2.25).-

Графика - логикалық тапсырмаларды шешудің визуалды құралдары. Жак Бертен, «Графика және графикалық ақпаратты өңдеу»

Кілтті шешімдер қабылдамас бұрын, жобаның басшылары визуалды деректерді, электрондық кестелердегі немесе мәтіндік хабарламалардағы құрғақ және қыын интерпретацияланатын сандарды пайдалануға ықтималдылығы жоғары.

Визуализацияланбаған деректер құрылым алаңында шашылып жатқан құрылым материалдарына ұқсас: олардың әлеуеті анық емес. Тек олардан нақты визуализация пайда болғанда, кірпіш пен бетоннан - үй, олардың қандай құндылыққа ие екендігі айқын болады. Үй салынғанша, материалдардың үйшіктен, сәнді вилладан немесе зәулім ғимараттан тұратынын айту мүмкін емес.

Компаниялар әртүрлі жүйелерден (Рис. 1.24 - Рис. 2.110) алынған деректерге, қаржылық операцияларға және кең көлемдегі мәтіндік деректерге ие. Алайда, бұл деректерді бизнес мұддесінде пайдалану көбінесе күрделі міндет болып табылады. Мұндай жағдайларда визуализация деректердің мағынасын жеткізуіндегі маңызды құралына айналады, ол ақпаратты кез келген маманға түсінікті форматта, мысалы, аспаптық панельдер, графикитер және диаграммалар түрінде ұсынуға көмектеседі.-

RwC-дің “Студенттерге тез өзгеретін бизнес әлемінде табысқа жету үшін не қажет” (2015) зерттеуі [9] табысты компаниялардың деректерді талдаумен шектелмейтінін, керісінше, шешім қабылдауды қолдау үшін интерактивті визуализация құралдарын, мысалы, графикитер, инфографика және аналитикалық панельдерді белсенді түрде пайдаланатынын атап өтеді. Есепке сәйкес, деректерді визуализациялау клиенттерге деректердің айтып тұрған тарихын графикитер, диаграммалар, дэшбордтар және интерактивті деректер модельдері арқылы түсінуге көмектеседі.

Ақпаратты визуалды графикалық формаларға, мысалы, диаграммалар мен графикитерге айналдыру процесі, деректерді (Рис. 7.14) адамның миымен түсіну мен интерпретациялауды жақсартады. Бұл жобаларды басқарушылар мен аналитиктерге күрделі сценарийлерді жылдам бағалауға және интуицияға емес, визуалды түрде танылатын тенденциялар мен паттерндерге негізделген негізделген шешімдер қабылдауға мүмкіндік береді.

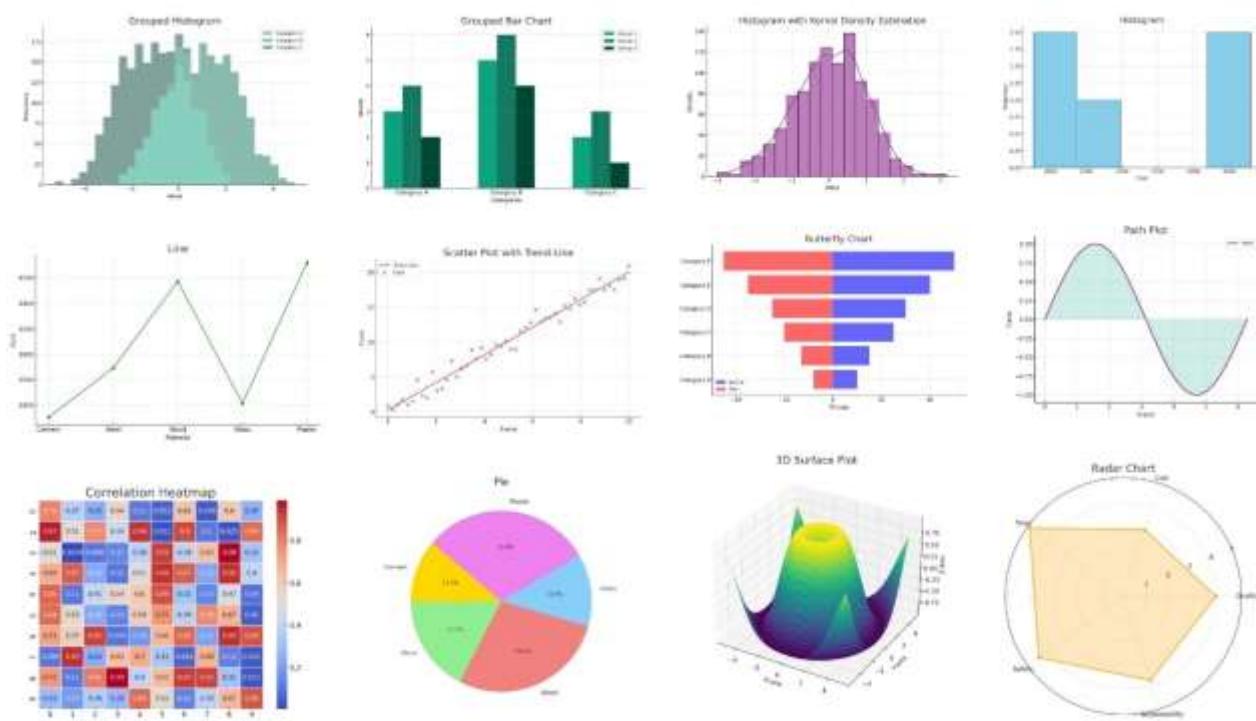


Рис. 7.14 Әр түрлі визуализация түрлері адамның миына құрғақ сандық ақпаратты жақсы түсінуге және қабылдауға көмектесу үшін жасалған.

Деректерден визуализацияларды жасау мәселелері және әртүрлі тегін визуализация кітапханаларын пайдалану келесі ETL процестеріне арналған тарауда қарастырылады.

Визуализация құрылыс саласында деректермен жұмыс істеудің ажырамас элементіне айналады – ол деректерді тек "көрумен" шектелмей, оларды басқару міндеттері контекстінде түсінуге көмектеседі. Алайда, визуализацияның шын мәнінде пайдалы болуы үшін, не нәрсені визуализациялау керектігін және жобаның тиімділігін бағалау үшін қандай метрикалардың маңызды екенін алдын ала анықтау қажет. Мұнда тиімділік көрсеткіштері, мысалы, KPI және ROI, маңызды рөл атқарады. Олардың жоқтығында, тіпті ең әдемі дашбордтар "ақпараттық шу" болып қалу қаупі бар.

Нәтижелердің тиімділігінің көрсеткіштері KPI және ROI

Қазіргі құрылыс саласында тиімділік көрсеткіштерін (KPI және ROI) басқару, сондай-ақ оларды есептер мен ақпараттық панельдер (дашбордтар) арқылы визуализациялау жобаларды басқарудың өнімділігі мен тиімділігін арттыруды маңызды рөл атқарады.

Кез келген бизнес сияқты, құрылыс саласында табысты, рентабельділік пен өнімділікті бағалау үшін метрикаларды нақты анықтау қажет. Әр түрлі процестер бойынша деректерді алған кезде, деректер негізінде жұмыс істейтін үйым алдымен негізгі KPI (Негізгі Өнімділік Көрсеткіштері) – стратегиялық және операциялық мақсаттарға жету дәрежесін көрсететін сандық көрсеткіштерді анықтауды үйренуі тиіс.

KPI есептеу үшін әдетте (Рис. 7.15) нақты және жоспарлы көрсеткіштерді қамтитын формула қолданылады. Мысалы, жобаның, қызметкердің немесе процестің жеке KPI-ін есептеу үшін нақты көрсеткіштерді жоспарлы көрсеткіштерге бөлу және алған нәтижені 100%-ға көбейту қажет.-

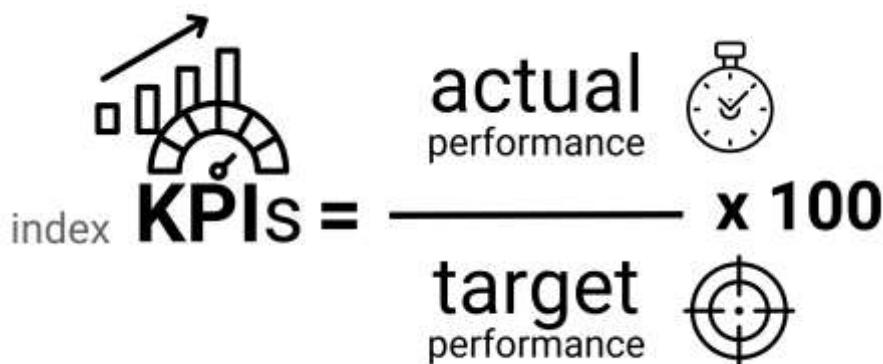


Рис. 7.15 KPI жобаның немесе процестің негізгі мақсаттарға жетудегі табысын өлшеу үшін қолданылады.

Құрылым алаңында нақты KPI метрикаларын қолдануға болады:

- Негізгі кезеңдердің (негіз, монтаж, әрлеу) орындалу мерзімдері – жұмыс жоспарларының сақталуын бақылауға мүмкіндік береді.
- Материалдардың артық шығындарының пайызы – сатып алушы басқаруға және шығындарды минимизациялауға көмектеседі.
- Жабдықтың жоспардан тыс тоқтап қалу саны – өнімділік пен шығындарға әсер етеді.

Дұрыс емес метрикаларды таңдау «не істеу керек?» деген сұрақ бойынша қате шешімдерге әкелуі мүмкін. Мысалы, егер компания тек шаршы метрдің құнына назар аударса, бірақ қайта өндеу шығындарын ескерсе, материалдарды үнемдеу сапаның нашарлауына және болашақ жобалардағы шығындардың өсуіне әкелуі мүмкін.

Мақсаттарды қою кезінде не өлшенетінін нақты анықтау маңызды. Белгісіз формулировкалар дұрыс қорытындыларға әкелмейді және бақылауды қыннадатады. Құрылым саласындағы сәтті және сәтсіз KPI мысалдарын қарастырайық.

Жақсы KPI:

- "Жылдың сонына дейін әрлеу жұмыстарының қайта өндеу пайызын 10%-ға төмендету"
- "Келесі тоқсанда сапаны төмендетпей, фасад монтажының жылдамдығын 15%-ға арттыру"
- "Жылдың сонына дейін жұмыс кестелерін оңтайландыру арқылы жабдықтың тоқтап қалу уақытын 20%-ға қысқарту"

Бұл метрикалар нақты өлшенеді, нақты мәндер мен уақыт шеңберіне ие.

Жаман KPI:

- "Біз жылдам саламыз" (Қаншалықты жылдам? "Жылдам" деген не?)
- "Біз бетон жұмыстарын сапасын арттырамыз" (Сапа қалай өлшенеді?)
- "Біз алаңдағы мердігерлермен өзара әрекеттесуді жақсартамыз" (Жақсартуды көрсететін

қандай критерийлер бар?)

Жақсы KPI – бұл өлшеуге және объективті бағалауға болатын көрсеткіш. Құрылым саласында бұл әсіресе маңызды, себебі нақты көрсеткіштерсіз жұмыс тиімділігін бақылау және тұрақты нәтижелерге қол жеткізу мүмкін емес.

KPI-дан бөлек, инвестициялардың тиімділігін бағалау үшін қосымша метрика бар: ROI (Return on Investment) – инвестициялардың қайтарымы, пайда мен салынған қаржат арасындағы қатынасты көрсетеді. ROI жаңа әдістерді, технологияларды немесе құралдарды енгізуудің қаншалықты негізделгенін бағалауға мүмкіндік береді: цифрлық шешімдер мен автоматизациядан (мысалы, Рис. 7.32) бастап, жаңа құрылым материалдарын қолдануға дейін. Бұл көрсеткіш бизнес пайдасына нақты әсерін ескере отырып, болашақ инвестициялар туралы негізделген шешімдер қабылдауға көмектеседі.-

Құрылым жобаларын басқару контекстінде ROI (инвестициялардың қайтарымы) компанияның жобада, технологияда немесе процестерді жақсартуда инвестициялардың қайтарымдылығын бағалау мақсаты болса, негізгі тиімділік көрсеткіштерінің (KPI) бірі ретінде пайдаланылуы мүмкін. Мысалы, егер жаңа құрылым басқару әдістемесі енгізілсе, ROI оның пайдадылығын қаншалықты арттырғанын көрсетуі мүмкін.

KPI және ROI көрсеткіштерін материал шығындары, еңбек сағаттары және шығындар сияқты әртүрлі көздерден жиналған деректер негізінде тұрақты түрде өлшеу жобаларды басқару басшылығына ресурстарды тиімді басқаруға және жылдам шешімдер қабылдауға мүмкіндік береді. Бұл деректерді ұзақ мерзімді сақтау болашақ тенденцияларды талдауға және процестерді оңтайландыруға мүмкіндік береді.

KPI, ROI және басқа көрсеткіштерді визуализациялау үшін әртүрлі графиктер мен диаграммалар пайдаланылады, олар әдетте бақылау панельдерінде біріктіріледі.

Ақпараттық панельдер мен дашбордтар: тиімді басқару үшін көрсеткіштерді визуализациялау

Көрсеткіштер мен метрикаларды визуализациялау үшін әртүрлі графиктер мен диаграммалар қолданылады, олар әдетте деректер витриналары мен ақпараттық панельдерде (dashboard) біріктіріледі. Мұндан панельдер жобаның немесе оның жеке бөліктерінің жағдайын орталықтандырылған түрде көрсетуге мүмкіндік береді, негізгі көрсеткіштерді (идеалында нақты уақыт режимінде) көрсетеді. Актуалды және тұрақты жаңартылып отыратын дашбордтар командаға өзгерістерге тез жауап беруге мүмкіндік береді.

Дашбордтар – бұл сандық бағалауларды визуализациялайтын құралдар, оларды жобаның барлық қатысушыларына қолжетімді және түсінікті етеді.



Рис. 7.16 КПІ басқару және ақпараттық панельдер арқылы визуализациялау - жобаның өнімділігі мен тиімділігін арттырудың кілті.

Ақпараттық панельдерді жасауға болатын танымал құралдардың бірнеше мысалдары:

- Power BI - Microsoft-тың интерактивті есептер мен ақпараттық панельдер жасауға арналған құралы.
- Tableau және Google Data Studio - код жазбай-ақ деректерді визуализациялау және ақпараттық панельдер жасау үшін қуатты құралдар.
- Plotly (Рис. 7.16, Рис. 7.212) – интерактивті графиктер жасауға арналған кітапхана, ал Dash - деректерді талдау үшін веб-қосымшалар жасауға арналған фреймворк. Оларды интерактивті бақылау панельдерін жасау үшін біріктіруге болады.--
- Python-ның көптеген кітапханалары (Рис. 7.29 - Рис. 7.211) – Python-да деректерді визуализациялау үшін көптеген ашық және тегін кітапханалар бар, мысалы, Matplotlib, Seaborn, Plotly, Bokeh және басқалар. Оларды графиктер жасау және Flask немесе Django сияқты фреймворктар арқылы веб-қосымшаға интеграциялау үшін пайдалануға болады.--
- JavaScript кітапханалары: Open Source JavaScript кітапханалары, мысалы, D3.js немесе Chart.js көмегімен интерактивті ақпараттық панельдер жасауға және оларды веб-қосымшаға интеграциялауға мүмкіндік береді.

KPI-ді бағалау және ақпараттық панельдер жасау үшін актуалды деректер мен ақпаратты жинау және талдау кестесі қажет.

Жалпы алғанда, KPI, ROI және ақпараттық панельдер құрылышында жобаларды басқаруға аналитикалық тәсілдің негізін құрайды. Олар тек ағымдағы жағдайды бақылауға және бағалауға көмектесіп қана қоймай, болашақ жоспарлау мен процестерді оңтайландыру үшін құнды ақпарат

береді – бұл процестер деректерді интерпретациялауға және дұрыс әрі уақтылы сұрақтар қоюға тікелей байланысты.

Деректерді талдау және сұрақ қою өнері

Деректерді интерпретациялау – бұл талдаудың соңғы кезеңі, онда ақпарат мағынаға ие болып, «сейлей» бастайды. Дәл осы жерде «не істеу керек?» және «қалай істеу керек?» деген негізгі сұрақтарға жауаптар формулировкасы жасалады (Рис. 2.25). Бұл кезең нәтижелерді жинақтауға, заңдылықтарды анықтауға, себеп-салдарлық байланыстарды орнатуға және визуализация мен статистикалық талдау негізінде қорытындылар жасауға мүмкіндік береді.

Мүмкін, жақын арада жаңа үлкен күрделі әлемдік мемлекеттердің тиімді азаматы ретінде толыққанды қалыптасу үшін орташа мәндерді, максимумдарды және минимумдарды есептей білу, қазіргі уақытта оқу мен жазуды менгеру сияқты қажеттілік туындастын уақыт келеді [133].

Сэмюэл С. Уилкс, 1951 жылды Американдық статистикалық ассоциацияға президенттің үндеуінен алғынған дәйексөз

Ұлыбритания үкіметі жариялаған «Деректер аналитикасы және жасанды интеллект мемлекеттік жобаларды жүзеге асыруда» (2024) атты есепке сәйкес, деректер аналитикасы мен жасанды интеллекттің енгізу жобаларды басқару процестерін айтарлықтай жақсартуға мүмкіндік береді, мерзімдер мен шығындарды болжаудың дәлдігін арттырып, тәуекелдер мен белгісіздікті азайтады. Құжатта заманауи аналитикалық құралдарды пайдаланатын мемлекеттік ұйымдар инфрақұрылымдық бастамаларды жүзеге асыруда жоғары тиімділік көрсеткіштеріне қол жеткіzetіні атап өтілген.

Бәсекелестік деңгейі жоғары және маржиналдылығы төмен қазіргі құрылымы бизнесін төртінші өнеркәсіптік революция жағдайында әскери әрекеттермен салыстыруға болады. Мұнда компанияның өмір сүруі мен табысы ресурстарды алу жылдамдығына және сапалы ақпаратқа, яғни уақтылы және негізделген шешімдер қабылдауға байланысты (Рис. 7.17).

Деректерді визуализациялау «барлау» болса, деректер аналитикасы «оқ-дәрілер» болып табылады, олар әрекет ету үшін қажет. Ол «не істеу керек?» және «қалай істеу керек?» деген сұрақтарға жауап береді, нарықта бәсекелестік артықшылықтарды алу негізін қалыптастырады.

Аналитика деректерді бөлшектенген күйден құрылымдалған және мазмұнды ақпаратқа айналдырады, оның негізінде шешімдер қабылданады.

Аналитиктер мен менеджерлердің міндеті – тек ақпаратты интерпретациялау емес, негізделген шешімдер ұсыну, тенденцияларды анықтау, әртүрлі деректер типтері арасындағы өзара байланысты анықтау және оларды жобаның мақсаттары мен ерекшеліктеріне сәйкес жіктеу. Визуализация құралдары мен статистикалық талдау әдістерін пайдалана отырып, олар деректерді компанияның стратегиялық активіне айналдырады.



Рис. 7.17 Деректерді талдау, ақырында, жиналған ақпаратты шешім қабылдау көзіне айналдырады.

Аналитика процесінде шынайы негізделген шешімдер қабылдау үшін деректерге қойылатын сұрақтарды дұрыс формулировкалау қажет. Бұл сұрақтардың сапасы алынатын инсайттардың тәрнедігіне және, соның салдарынан, басқарушылық шешімдердің сапасына тікелей әсер етеді.

Өткен тек бүгінгі жазбаларда болғаны үшін ғана бар. Ал осы жазбалардың не екенін біз қандай сұрақтар қойсақ, соған байланысты. Басқа тарих жоқ [134].

Джон Арчибалд Уилер, физик 1982

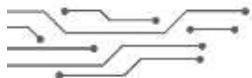
Терең сұрақтар қою және сыни ойлау өнері – деректермен жұмыс істеудегі ең маңызды дағды. Көптеген адамдар қарапайым, беткейлік сұрақтар қоюға бейім, олардың жауаптары айтартықтай күш-жігерді талап етпейді. Алайда, шынайы талдау мазмұнды және ойластырылған сұрақтардан басталады, олар ақпараттағы жасырын өзара байланыстар мен себеп-салдарлық тәуелділіктерді ашуға қабілетті, олар бірнеше деңгейдегі пайымдаулардың артында болуы мүмкін.

"Деректермен басқарылатын трансформация: Қазірдің өзінде масштабта жеделдету" (BCG, 2017) зерттеуіне сәйкес, сәтті цифрлық трансформация аналитикалық мүмкіндіктерге, өзгерістерді басқару бағдарламаларына инвестицияларды және бизнес мақсаттарын ИТ бастамаларымен келісуді талап етеді. Деректерге бағытталған мәдениетті қалыптастыратын компания аналитикалық деректерді пайдалану мүмкіндіктеріне инвестиция салып, жаңа ойлау, мінез-құлыштың және жұмыс тәсілдерін енгізу үшін өзгерістерді басқару бағдарламаларын іске қосуы тиіс.

Аналитикалық мәдениетті дамытуға, деректермен жұмыс істеу құралдарын жетілдіруге және маңандарды оқытуға инвестициясыз, компаниялар болашақта ескірген немесе толық емес ақпарат негізінде шешім қабылдау қаупіне ұшырайды – немесе HiPPO-менеджерлердің субъективті пікірлеріне сүйенеді.

Аналитиканы және ақпараттық панельдерді тұрақты жаңартудың өзектілігі мен қажеттілігін түсіну, басшылықты аналитикалық процестерді автоматтандырудың маңыздылығын түсінуге әкеледі. Автоматтандыру шешім қабылдау жылдамдығын арттырады, адам факторының әсерін азайтады және деректердің өзектілігін қамтамасыз етеді. Ақпарат көлемінің экспоненциалды өсу жағдайында жылдамдық тек бәсекелестік артықшылық емес, тұрақты табистың кілтті факторы болып табылады.

Деректерді талдау және өндөу процестерін автоматтандыру ETL (Extract, Transform, Load) тақырыбымен тығыз байланысты. Автоматтандыру процесінде деректерді түрлендіру қажет болғандай, ETL процесінде деректер әртүрлі көздерден алынып, қажетті талаптарға сәйкес түрлендіріліп, мақсатты жүйелерге жүктеледі.



ТАРАУ 7.2. ДЕРЕК АҒЫНЫ ҚОЛДАН БОЛМАЙ: ETL НИГЕ ҚАЖЕТ

ETL АВТОМАТТАНДЫРУ: ШЫҒЫНДАРДЫ ТӨМЕНДЕТУ ЖӘНЕ ДЕРЕКТЕРМЕН ЖҰМЫС ИСТЕУДІ ЖЫЛДАМДАТУ

Негізгі тиімділік көрсеткіштері (KPI) деректер көлемі мен команда санының артуына қарамастан өспей қалғанда, компания басшылығы автоматтандыру процестерінің қажеттілігін түсінуге келеді. Ерте ме, кеш пе, бұл түсінік кешенді автоматтандыруды бастауға ынталандыруға айналады, оның негізгі мақсаты процестердің күрделілігін азайту, өндөу жылдамдығын арттыру және адам фактологияның тәуелділігін төменидете.

McKinsey зерттеуіне сәйкес "Деректер архитектурасын инновацияларды ынталандыру үшін қалай құру керек - бүгін және ертең" (2022), ағынды деректер архитектураларын пайдаланатын компаниялар елеулі артықшылықтарға ие, себебі олар ақпаратты нақты уақыт режимінде талдай алады. Ағынды технологиялар нақты уақыт режимінде хабарламаларды тікелей талдауға және өндірісте деректерді талдау арқылы болжамды қызмет көрсетуге мүмкіндік береді.

Процесті жеңілдету – бұл автоматтандыру, мұнда дәстүрлі қолмен басқару функциялары алгоритмдер мен жүйелермен ауыстырылады.

Автоматтандыру мәселесі, яғни "деректерді өңдеудегі адамның рөлін минимизациялау", әр компания үшін қайтымсыз және өте сезімтал процесс болып табылады. Кәсіби саладағы мамандар, технологиялық ортада жұмыс орындарын жоғалту қаупін сезініп, әріптестеріне өз әдістері мен жұмыс ерекшеліктерін толық ашуға батылдық танытпайды.

Егер өзіңе дұшпан тапқың келсе, бірденені өзгертуге тырыс.

– Вудро Вильсон, сатушылар конгресінде сөйлеген сөз, Детройт, 1916 жыл.

Автоматизацияға айқын артықшылықтарға қарамастан, көптеген компаниялардың күнделікті тәжірибесінде, әсіресе инженерлік деректермен жұмыс істеу саласында, қолмен еңбек үлесі әлі де жоғары. Қазіргі жағдайды айқын көрсету үшін, осындаи процестер аясында деректерді ретті өңдеудің типтік мысалын қарастырайық.

Деректермен қолмен жұмыс істеу тәртібін CAD дерекқорларынан алынған ақпаратпен өзара әрекеттесу мысалында көрсетуге болады. CAD (BIM) бөлімдерінде атрибуттық кестелерді жасау немесе жобалық деректер негізінде құжаттама жасау үшін деректерді дәстүрлі өңдеу ("қолмен" ETL процесі) келесі тәртіpte жүзеге асырылады (Рис. 7.21):-

- Қолмен шығару (Extract): пайдаланушы жобаны қолмен ашады - CAD (BIM) қосымшасын іске қосу арқылы (Рис. 7.21 қадам 1). -
- Верификация: келесі кезеңде әдетте деректерді дайындау және олардың сапасын бағалау үшін бірнеше плагиндер немесе көмекші қосымшалар қолмен іске қосылады (Рис. 7.21 қадам 2-3).

- Қолмен түрлендіру (Transform): дайындықтан кейін деректерді өңдеу басталады, бұл деректерді экспорттауға дайындау үшін әртүрлі бағдарламалық құралдарды қолмен басқаруды талап етеді (Рис. 7.21 қадам 4). -
- Қолмен экспорттау (Load): түрлендірілген деректерді сыртқы жүйелерге, деректер форматтарына және құжаттарға қолмен экспорттау (Рис. 7.21 қадам 5). -

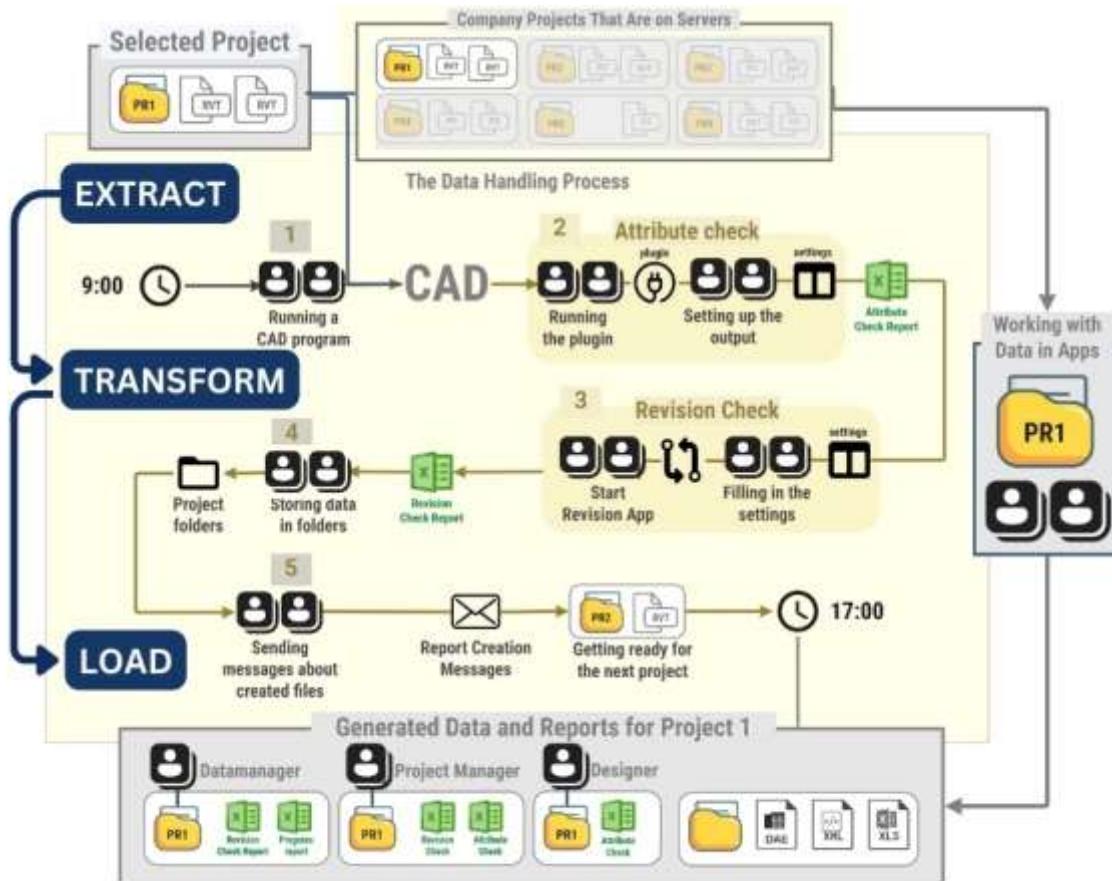


Рис. 7.21 Дәстүрлі қолмен ETL өңдеу жеке техникалық маманның тілектері мен физикалық мүмкіндіктерімен шектеледі.

Мұндай жұмыс процесі классикалық ETL процесінің мысалы болып табылады - деректерді шығару, түрлендіру және жүктегу (ETL). Басқа салаларда автоматтандырылған ETL құбырлары стандартқа айналғанымен, құрылышы саласында әлі де қолмен еңбек басым, бұл процестерді баяулатып, шығындарды арттырады.

ETL (Extract, Transform, Load) - бұл деректерді әртүрлі көздерден шығару, қажетті форматқа түрлендіру және мақсатты жүйеге жүктегу процесі, әрі қарай талдау және пайдалану үшін.

ETL - деректерді өңдеудің үш негізгі компонентін білдіретін процесс: Extract, Transform және Load (Рис. 7.22):-

- Extract - деректерді әртүрлі көздерден (файлдар, дерекқорлар, API) шығару.
- Transform - деректерді тазалау, агрегаттау, нормализациялау және логикалық өңдеу.
- Load - құрылымдалған ақпаратты деректер қоймасына, есепке немесе BI жүйесіне жүктеу.

Бұрын кітапта ETL концепциясы тек эпизодтық түрде қарастырылған: құрылымдалмаған сканер-ленген құжатты құрылымдалған кестелік форматқа түрлендіру кезінде (Рис. 4.11), талаптарды формализациялау контекстінде, өмірлік және іскерлік процестерді жүйелендіруге мүмкіндік беретін (Рис. 4.420), сондай-ақ CAD шешімдерінен деректерді тексеру және өңдеу автоматизациясы аясында. Енді ETL-ді типтік жұмыс процестері контекстінде тереңірек қарастырайық.-

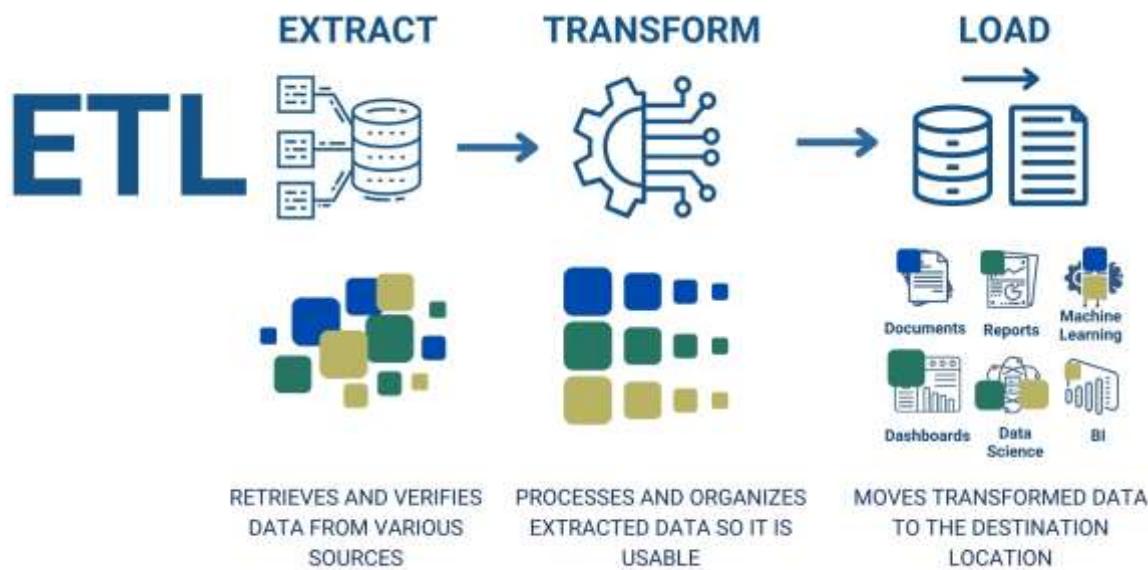


Рис. 7.22 ETL қайталанатын деректерді өңдеу тапсырмаларын автоматтандырады.

Қолмен немесе жартылай автоматтандырылған ETL процесі деректерді жинаудан есептерді генерациялауға дейін барлық қадамдарды қолмен басқаратын менеджер немесе техникалық маманың болуын талап етеді. Мұндай процесс, әсіресе шектеулі жұмыс күні (мысалы, 9:00-ден 17:00-ге дейін) жағдайында, едәуір уақытты алады.

Көптеген компаниялар тәмен тиімділік пен баяу жылдамдық мәселесін модульдік интеграцияланған шешімдерді (ERP, PMIS, CPM, CAFM және т.б.) сатып алу арқылы шешүге тырысады, олар кейіннен сыртқы вендорлар мен консультанттар тарарапынан жетілдіріледі. Алайда, мұндай вендорлар мен үшінші тарарап әзірлеушілері жиі критикалық тәуелділік нүктесіне айналады: олардың техникалық шектеулери бүкіл жүйенің және бизнесің өнімділігіне тікелей әсер етеді, бұл проприетарлық жүйелер мен форматтар туралы алдыңғы тарауларда егжей-тегжейлі баяндалған. Фрагментация мен тәуелділік туғызатын мәселелер туралы біз «Құрылыш бизнесі деректер хаосында қалай батады» тарауында кеңінен талқыладық.

Егер компания вендорлардың бірінің ірі модульдік платформасын енгізуге дайын болмаса, ол автоматизацияның баламалы жолдарын іздей бастайды. Олардың бірі – әр кезең (шығару, түрлендіру, валидация, жүктөу) скриптер түрінде жүзеге асырылатын өз модульдік ашық ETL-конвейерлерін әзірлеу.

ETL-дің автоматтандырылған нұсқасында (Рис. 7.21) жұмыс процесі модульдік код ретінде көрінеді, ол деректерді өңдеуден басталып, оларды ашық құрылымдалған формаға ауыстырады. Құрылымдалған деректер алынғаннан кейін, автоматты түрде, кесте бойынша, әртүрлі сценарийлер немесе модульдер өзгерістерді тексеру, түрлендіру және хабарламаларды жіберу үшін іске қосылады (Рис. 7.23).--

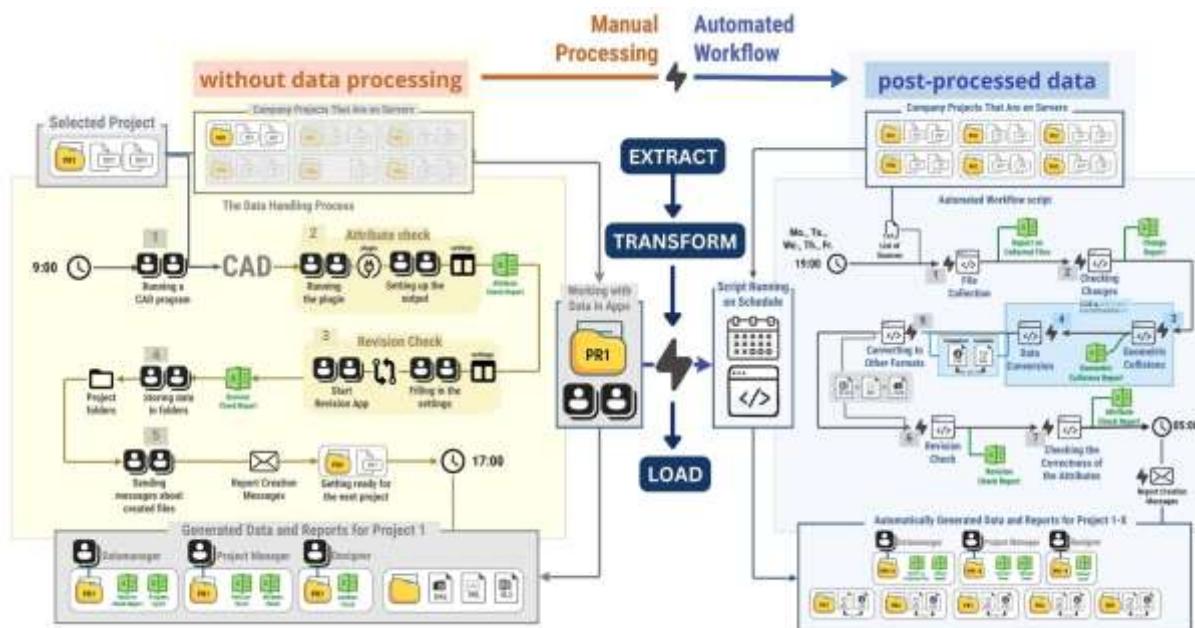


Рис. 7.23 Сол жақта қолмен өңдеу, оң жақта - автоматтандырылған процесс, ол дәстүрлі қолмен өңдеуден айырмашылығы, пайдаланушының мүмкіндітерімен шектелмейді.

Автоматтандырылған жұмыс процесінде деректерді өңдеу ETL(derekterin aldyndan ala oңdeu arkyly) женілдетіледі: құрылымдау және унификациялау.

Дәстүрлі өңдеу әдістерінде мамандар деректермен «олар бар» күйінде жұмыс істейді – олар жүйелерден немесе бағдарламалық қамтамасыз етуден алынған күйінде. Автоматтандырылған процестерде, керісінше, деректер жиі алдымен ETL-пайплин арқылы өтеді, онда олар келісілген құрылым мен форматқа келтіріледі, әрі қарай пайдалану мен талдау үшін жарамды.

ETL процесін қарастырайық, ол деректер кестелерін тексеру процесін көрсетеді, «Деректерді тексеру және тексеру нәтижелері» тарауында сипатталған (Рис. 4.413). Бұл үшін біз автоматтандырылған талдау және деректерді өңдеу процестері үшін Pandas кітапханасын LLM-мен бірге қолданамыз.-

ETL Extract: деректерді жинау

ETL процесінің бірінші кезеңі – шығару (Extract) – тексеруге және өңдеуге жататын деректер мас-сивтерін жинау үшін код жазудан басталады. Осы мақсатта жұмыс серверінің барлық папкаларын сканерлеп, белгілі бір формат пен мазмұндағы құжаттарды жинап, содан кейін оларды құрылымдалған формаға түрлендіреміз. Бұл процесс «Құрылымдалған және мәтіндік деректерді құрылымдалған формаға түрлендіру» және «CAD (BIM) деректерін құрылымдалған формаға түрлендіру» тарауларында ежей-тегжейлі қарастырылған (Рис. 4.11 - Рис. 4.112).-

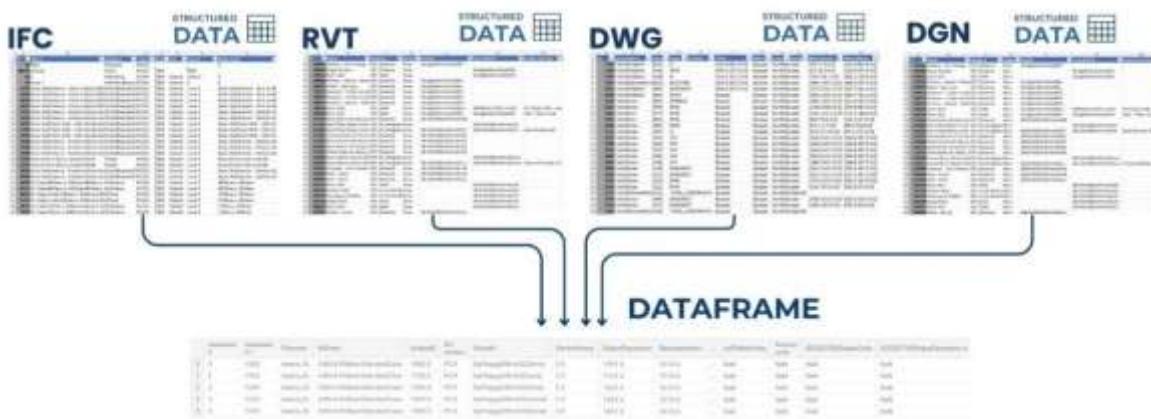


Рис. 7.24 CAD (BIM) деректерін бір үлкен деректер фрейміне түрлендіру, ол жобаның барлық бөлімдерін қамтитын болады.

Наглядный мысал ретінде деректерді жүктеу кезеңінде Extract және барлық CAD- (BIM-) жобаларының кестесін алу (Рис. 7.24) үшін RVT және IFC форматтарына кері инженерингті қолдайтын конвертерлер пайдаланылады, олар барлық жобалардан құрылымдалған кестелерді алғып, оларды бір үлкен DataFrame кестесіне біріктіреді.-

```

1 import os
2 import subprocess
3 import time
4 import pandas as pd
5
6 path_conv = r'C:\DDC_2023\\' # Where RvtExporter.exe|IfcExporter.exe is located
7 path = r'C:\IFCprojects\\' # Where Revit|IFC files are stored
8
9 def convert_and_wait(path_conv, exporter_name, file_path, extension):
10     # Start the conversion process
11     subprocess.Popen([os.path.join(path_conv, exporter_name),
12                      file_path], cwd=path_conv)
13     output_file = os.path.join(path,
14                             f"{os.path.splitext(os.path.basename(file_path))[0]}_{extension}.xlsx")
15
16 # Conversion process for RVT and IFC files
17 for file in os.listdir(path):
18     full_path = os.path.join(path, file)
19     if file.endswith('.ifc'):
20         convert_and_wait(path_conv, 'IfcExporter.exe', full_path, 'ifc')
21     elif file.endswith('.rvt'):
22         convert_and_wait(path_conv, 'RvtExporter.exe', full_path, 'rvt')
23
24 # Combine converted Excel files into one dataframe
25 df = pd.concat([pd.read_excel(os.path.join(path, f)) for f in os.listdir(path)
26 if f.endswith('.xlsx')], ignore_index=True)

```

[Run in IDE](#)

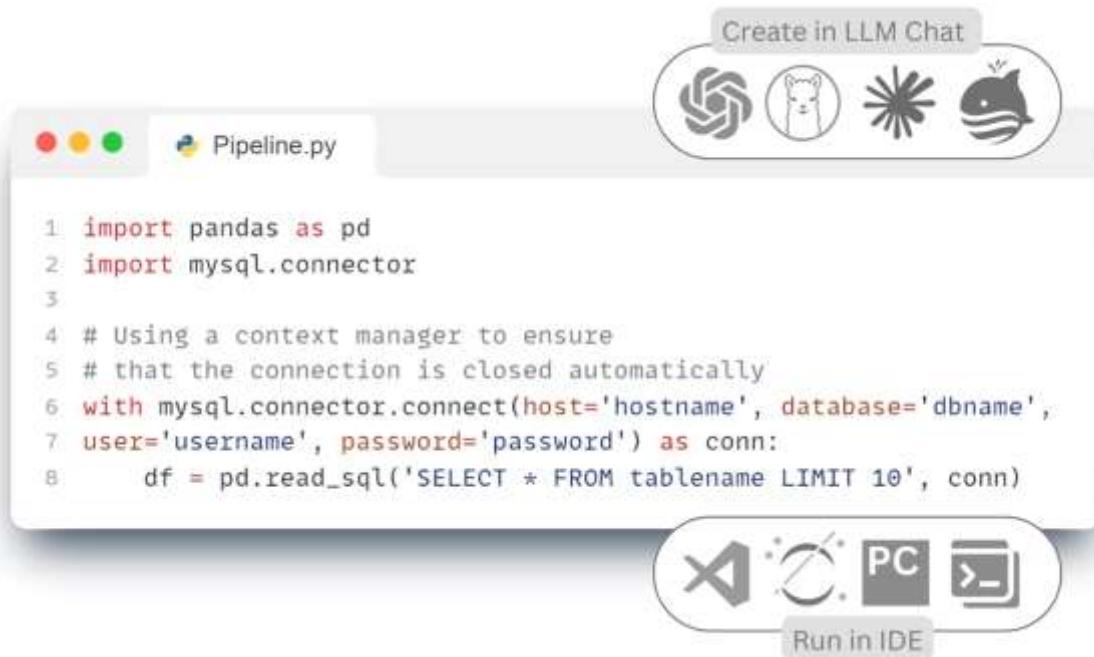
Рис. 7.25 Python коды мен RVT және IFC файлдарын көрі инженериялау құралдарының SDK-сы арқылы бір үлкен құрылымдалған (df) DataFrame-ге түрлендіру.

Pandas DataFrame-ге деректерді әртүрлі көздерден жүктеуге болады, соның ішінде CSV мәтіндік файлдары, Excel кестелері, JSON және XML файлдары, үлкен көлемдегі деректерді сақтау форматтары, мысалы, Parquet және HDF5, сондай-ақ MySQL, PostgreSQL, SQLite, Microsoft SQL Server, Oracle және басқа базалардан. Сонымен қатар, Pandas API, веб-беттер, Google BigQuery, Amazon Redshift және Snowflake сияқты бүлтты қызыметтер мен сақтау жүйелерінен деректерді жүктеуді қолдайды.

- ❷ Деректер базасынан ақпарат жинау және қосылу үшін код жазу үшін LLM чатында (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе кез келген басқа) осындаі мәтіндік сұраныс жіберіңіз:

Пожалуйста, напиши пример подключения к MySQL и преобразования данных в DataFrame ↴

■ LLM-нің жауабы:



```

1 import pandas as pd
2 import mysql.connector
3
4 # Using a context manager to ensure
5 # that the connection is closed automatically
6 with mysql.connector.connect(host='hostname', database='dbname',
7     user='username', password='password') as conn:
8     df = pd.read_sql('SELECT * FROM tablename LIMIT 10', conn)

```

The screenshot shows a code editor window titled "Pipeline.py". The code imports pandas and mysql.connector, then uses a context manager to connect to a MySQL database and read data into a DataFrame. Below the code editor is a "Run in IDE" button with icons for VS Code, CircleCI, PC, and a terminal.

Рис. 7.26 Python арқылы MySQL дереккөрүна қосылу және MySQL деректерін DataFrame-ге импорттау мысалы.

Алынған кодты (Рис. 7.25, Рис. 7.26) жоғарыда аталған танымал IDE-лерде оффлайн режимінде іске қосуға болады: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагинімен, Thonny, Wing IDE, IntelliJ IDEA Python плагинімен, JupyterLab немесе танымал онлайн құралдар: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.-

Мультиформатты деректерді "df" айнымалысына жүктегеннен кейін (Рис. 7.25 – 25-ші жол; Рис. 7.26 – 8-ші жол), деректерді Pandas DataFrame форматына конверттедік – деректерді өңдеудің ең танымал құрылымдарының бірі, ол жолдар мен бағандардан тұратын екі өлшемді кесте. ETL-Pipelines-те қолданылатын Parquet, Apache ORC, JSON, Feather, HDF5 сияқты сақтау форматтары және заманауи деректер қоймалары туралы "Құрылыш саласындағы деректерді сақтау және басқару" тарауында (Рис. 8.12) толығырақ айтып береміз.–

Деректерді алу және құрылымдау (Extract) кезеңінен кейін ақпараттың біртұтас массиві (Рис. 7.25, Рис. 7.26) қалыптасады, ол әрі қарай өңдеуге дайын. Алайда, бұл деректерді мақсатты жүйелерге жүктемес бұрын немесе талдау үшін пайдаланбас бұрын, олардың сапасы, тұтастығы және белгіленген талаптарға сәйкестігін тексеру қажет. Деректерді түрлендіру (Transform) – бұл кезеңде жүзеге асырылады, бұл келесі қорытындылар мен шешімдердің сенімділігін қамтамасыз ететін негізгі қадам.–

ETL Transform: тексеру ережелерін қолдану және түрлендіру

Transform кезеңінде деректерді өңдеу және түрлендіру жүзеге асырылады. Бұл процесс дұрыстылықты тексеруді, нормализацияны, жоғалған мәндерді толтыруды және автоматтандырылған

Құралдармен валидациялауды қамтуы мүмкін.

PwC-нің "Деректерге негізделген. Студенттерге бизнес әлеміндегі жылдам өзгерістерде табысқа жету үшін не қажет" (2015) [9] зерттеуіне сәйкес, қазіргі заманғы аудиттік компаниялар деректерді ірікеп текстеруден бас тартып, автоматтандырылған құралдарды пайдалана отырып, ақпарат массивін талдауға көшуде. Мұндай тәсіл тек есептіліктегі сәйкесіздіктерді анықтауға ғана емес, сонымен қатар бизнес процестерін оңтайландыру бойынша ұсыныстар жасауға мүмкіндік береді.

Құрылыс саласында үқсас әдістер, мысалы, жобалық деректерді автоматты түрде валидациялау, құрылыс сапасын бақылау және мердігерлердің жұмыс тиімділігін бағалау үшін қолданылуы мүмкін. Деректерді өңдеуді автоматтандыру және жеделдешу үшін қолданылатын құралдардың бірі - ETL процесіндегі деректерді түрлендіру кезеңінде регулярлы өрнектерді (RegEx) қолдану. RegEx деректер жолдарын тиімді текстеруге, сәйкесіздіктерді анықтауға және ақпараттың тұтастырын минималды ресурстармен қамтамасыз етуге мүмкіндік береді. RegEx туралы толығырақ біз "Талаптарды құрылымдық форматта аудару" бөлімінде айттық.-

Практикалық мысалды қарастырайық: жылжымайтын мүлік басқару жүйесінде (RPM) менеджер объектілердің негізгі атрибуттарына қойылатын талаптарды белгілейді. Түрлендіру кезеңінде келесі параметрлерді валидациялау қажет:-

- объектілердің идентификаторларының форматтарын текстеру (атрибут "ID")
- ауыстыру кепілдігінің мерзімін бақылау (атрибут "Гарантийный срок")
- элементтердің ауыстыру циклына текстеру (атрибут "Требования к обслуживанию")

ID	Element	Warranty Period	Replacement Cycle	Maintenance Requirements
W-NEW	Window	-	20 years	Annual Inspection
W-OLD1	Window	8 years	15 years	Biannual Inspection
W-OLD2	Window	8 years	15 years	Biannual Inspection
D-122	Door	15 years	25 years	Biennial Varnishing

Сурет 7.27 Сапаны текстеру атрибуттарға және олардың шекті мәндеріне талаптарды орнатудан басталады.

Параметрлерді текстеру үшін шекті мәндерді орнату үшін, мысалы, тәжірибеліден "ID" атрибуты үшін рұқсат етілген мәндер тек "W-NEW", "W-OLD1" немесе "D-122" сияқты жолдық мәндер болуы мүмкін екенін білеміз, мұнда бірінші символ әріп, одан кейін дефис, содан кейін 'NEW', 'OLD' немесе кез келген үш таңбалы сан (Сурет 7.27). Бұл идентификаторларды валидациялау үшін келесі регулярлы өрнек (RegEx) қолданылуы мүмкін:-

```
^W-NEW$| ^W-OLD[0-9]+$| ^D-1[0-9]{2}$
```

Бұл шаблон деректердегі барлық идентификаторлардың белгіленген критерийлерге сәйкес келетінін қамтамасыз етеді. Егер қандай да бір мән тексеруден өтпесе, жүйе қатені тіркейді. Деректерді түрлендіру үшін Python кодын жасау және алынған деректерді нәтижелер кестесін құру үшін пайдалану үшін, LLM чатында сұраныс формулоривкасын жасаймыз.

 **LLM-ге мәтіндік сұраныс:**

'W-NEW' немесе 'W-OLD' форматындағы идентификаторларды RegEx арқылы тексеретін, 'A' мен 'G' аралығындағы әріптермен энергия тиімділігін, кепілдік мерзімін және жылдардағы сандық мәндері бар ауыстыру циклын тексеру үшін DataFrame бағандарын тексеру коды жазыңыз ↴

 **LLM-нің жауабы:**



```
1 # Regex
2 import re
3
4 # Define regex patterns for each column
5 patterns = {
6     'Window_ID': r'^W-NEW|W-OLD[0-9]+|D-1[0-9]{2}$',
7     'Energy_Efficiency': r'^[A-G]\w*$',
8     'Warranty_Period': r'^\d+\w{3}\d+$',
9     'Replacement_Cycle': r'^\d+\w{3}\d+$'
10 }
11
12 # Define a verification function using regex
13 def verify_data(column, pattern):
14     # Apply the regex pattern to the series and return the boolean results
15     return column.apply(lambda x: bool(re.match(pattern, str(x))))
16
17 # Apply the verification function to the DataFrame for each column with its respective pattern
18 for column, pattern in patterns.items():
19     df[f'verified_{column}'] = verify_data(df[column], pattern)
20
```

Сурет 7.28 Процесті автоматтандыру коды параметрлер датафреймінің бағандарына шаблондар RegEx қолдану арқылы тексеру.

Автоматты түрде жасалған Python коды (Сурет 7.28) "re" кітапханасын (регулярлы өрнектер RegEx) пайдалана отырып, DataFrame элементінің әрбір атрибутын тексеретін функцияны анықтайды. Әрбір көрсетілген баған (атрибут) үшін функция RegEx шаблонын қолдана отырып, әр жазбаның күтілетін форматқа сәйкестігін тексереді және нәтижелерді жаңа атрибут-бағанға (False/True) қосады.-

Мұндағы автоматтандырылған тексеру деректердің белгіленген талаптарға формалды сәйкестігін қамтамасыз етеді және түрлендіру кезеңінде сапаны бақылау жүйесінің бір бөлігі ретінде пайдаланылуы мүмкін.

Трансформация кезеңі сәтті аяқталғаннан кейін және сапа тексерілгеннен кейін, деректер мақсатты жүйелерге жүктелуге дайын. Өзгертулған және тексерілген деректер CSV, JSON, Excel, дерекқорлар және басқа форматтарға экспортталуы мүмкін, кейінгі пайдалану үшін. Сондай-ақ, тапсырмаға байланысты нәтижелер есептер, графиктер немесе аналитикалық дэшбордтар түрінде ұсынылуы мүмкін.

ETL Load: НӘТИЖЕЛЕРДІ ГРАФИКТЕР МЕН ДИАГРАММАЛАР ТҮРІНДЕ КӨРСЕТУ

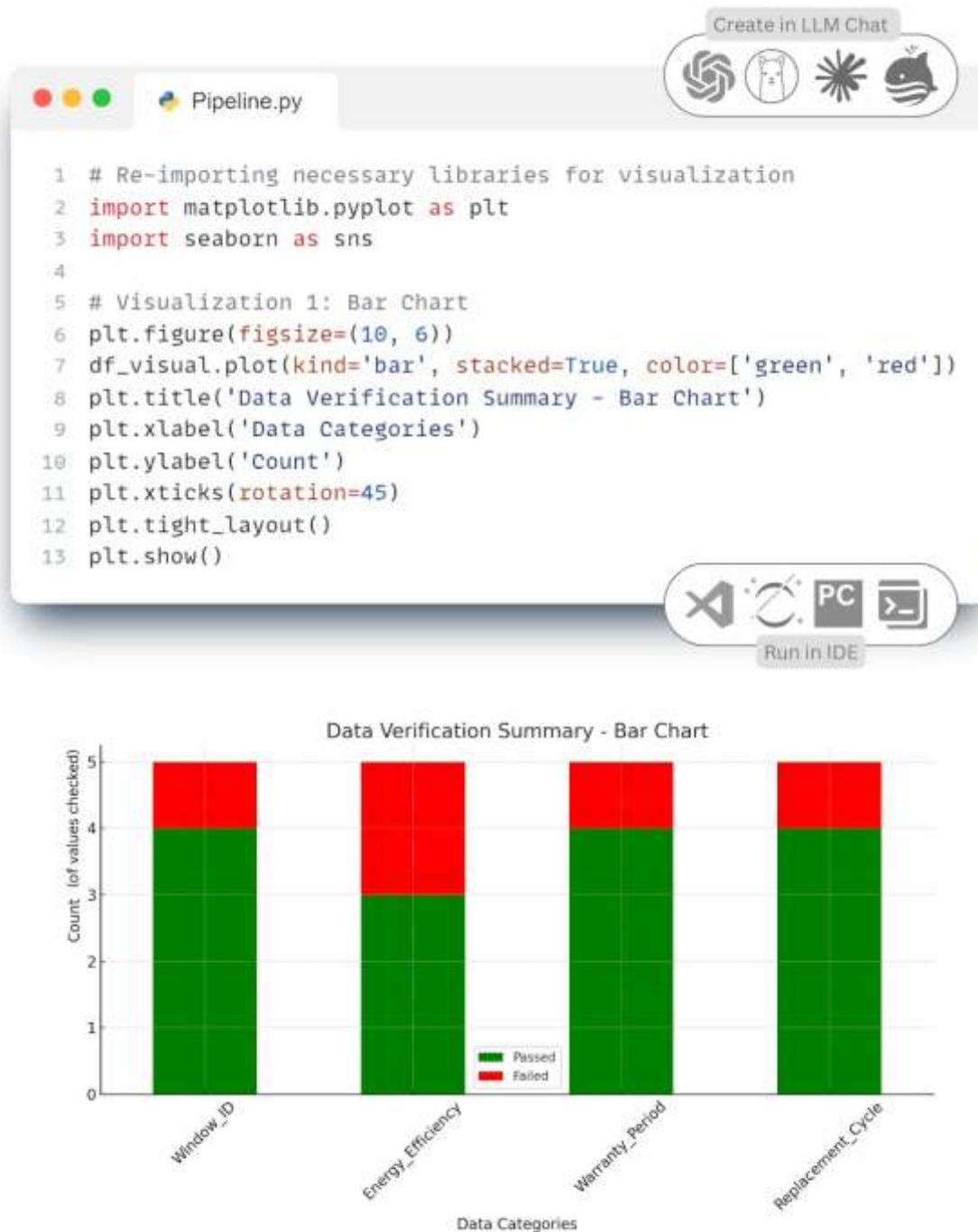
Трансформация кезеңі аяқталғаннан кейін, деректер құрылымдалған формаға келтіріліп, тексерілгеннен кейін, соңғы кезең – жүктеу кезеңі басталады, онда деректер мақсатты жүйеге жүктелуі немесе талдау үшін визуализациялануы мүмкін. Деректердің визуалды көрсетілімі ауытқуларды жедел анықтауға, таралымдарды талдауға және жобаның барлық қатысушыларына, соның ішінде техникалық дайындыққа ие емес адамдарға, негізгі қорытындыларды жеткізуға мүмкіндік береді.

Ақпаратты кестелер мен сандар түрінде ұсынудың орнына, біз инфографика, графиктер және приборлық панельдер (дашбордтар) қолдана аламыз. Python тілінде құрылымдалған деректерді визуализациялаудың ең кең таралған және икемді құралдарының бірі – Matplotlib кітапханасы (Рис. 7.29, Рис. 7.210). Ол статикалық, анимациялық және интерактивті графиктер жасауға мүмкіндік береді және диаграммалардың кең ауқымын қолдайды. –

- RPM жүйесінен атрибуттардың тексеру нәтижелерін визуализациялау үшін (Рис. 7.27) келесі тілдік модельге сұраныс жасауға болады:-

Жоғарыда келтірілген DataFrame деректерін визуализациялау үшін код жазыңыз, гистограмма арқылы нәтижелерді көрсету үшін, атрибуттардағы қателердің жиілігін көрсету.-

- ❸ LLM-нің код түріндегі жауабы және кодтың орындалу нәтижесінде дайын визуализация тікелей LLM чатында:



```

1 # Re-importing necessary libraries for visualization
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Visualization 1: Bar Chart
6 plt.figure(figsize=(10, 6))
7 df_visual.plot(kind='bar', stacked=True, color=['green', 'red'])
8 plt.title('Data Verification Summary - Bar Chart')
9 plt.xlabel('Data Categories')
10 plt.ylabel('Count')
11 plt.xticks(rotation=45)
12 plt.tight_layout()
13 plt.show()

```

The screenshot shows a Jupyter Notebook cell with the title "Pipeline.py". The cell contains Python code for generating a stacked bar chart. The chart is titled "Data Verification Summary - Bar Chart" and displays the count of values checked for four data categories: "Window_ID", "Energy_Efficiency", "Warranty_Period", and "Replacement_Cycle". The y-axis is labeled "Count (of values checked)" and ranges from 0 to 5. The x-axis is labeled "Data Categories". A legend indicates that green represents "Passed" and red represents "Failed". The chart shows that for "Window_ID", there are approximately 4 Passed and 1 Failed value. For "Energy_Efficiency", there are approximately 3 Passed and 2 Failed values. For "Warranty_Period", there are approximately 4 Passed and 1 Failed value. For "Replacement_Cycle", there are approximately 4 Passed and 1 Failed value.

Рис. 7.29 RPM жүйесінен атрибуттардың мәндерін тексеру нәтижелерін визуализациялау (Рис. 7.27) жүктеу кезеңінде гистограмма түрінде.-

- ❹ Структуранған деректерді әртүрлі форматтарда ұсынуға мүмкіндік беретін көптеген

ашық және тегін визуализация кітапханалары бар. Нәтижелерді визуализациялауды көлесі график түрімен жалғастырайық:

Сол деректерді сыйықтық график түрінде бейнелеңіз.

❷ LLM-нің жауабы:

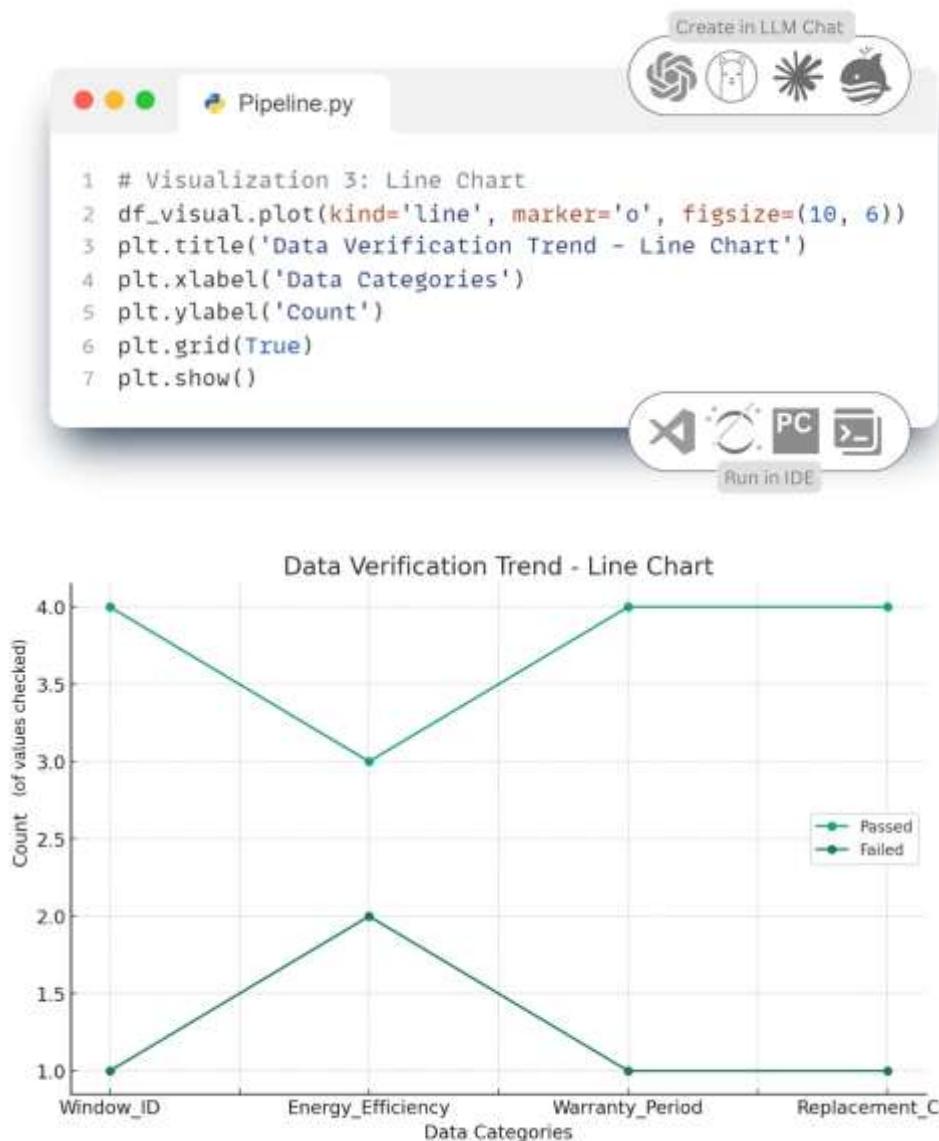


Рис. 7.210 Тексеру деректерін (Рис. 7.28) Matplotlib кітапханасы арқылы алынған сыйықтық диаграмма түрінде визуализациялау.-

Структураланған деректерді әртүрлі форматтарда ұсынуға мүмкіндік беретін көптеген ашық және тегін визуализация кітапханалары бар, мысалы:

- Seaborn – статистикалық графиктер үшін (Рис. 7.211)-
- Plotly – интерактивті веб-визуализациялар үшін (Рис. 7.212, Рис. 7.16)--
- Altair – декларативті визуализация үшін
- Dash немесе Streamlit – толыққанды дашбордтар жасау үшін

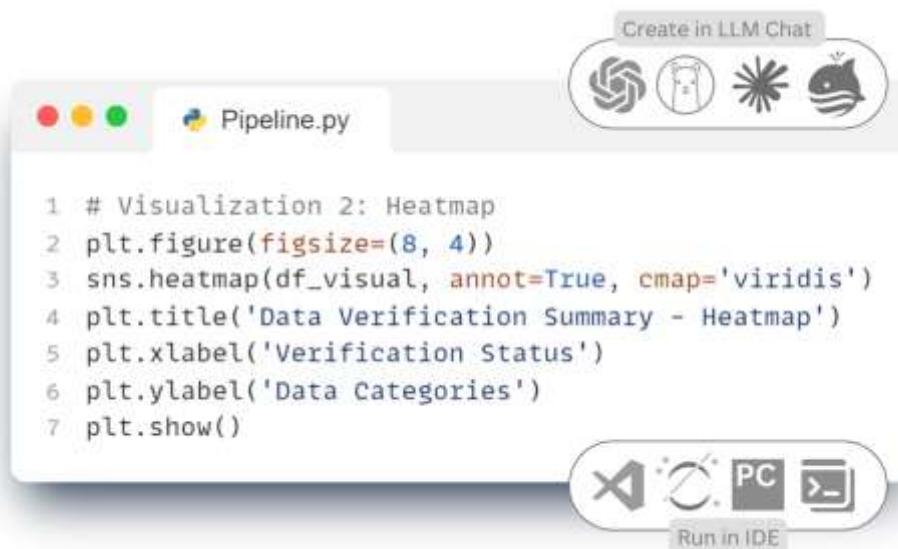
Сонымен қатар, визуализация үшін нақты кітапханаларды білу міндетті емес – заманауи құралдар, соның ішінде LLM, тапсырманың сипаттамасына негізделген графиктер мен толық қосымшаларды құру үшін автоматты түрде код генерациялауға мүмкіндік береді.

Құралды таңдау жобаның міндеттеріне байланысты: есеп, презентация немесе онлайн мониторинг панелі болсын. Мысалы, ашық кодты Seaborn кітапханасы категориялық деректермен жұмыс істеу үшін өте жақсы, заңдылықтар мен тенденцияларды анықтауға көмектеседі.

- Seaborn кітапханасын жұмыс барысында көру үшін, сіз LLM-ден тікелей қажетті кітапханы пайдалануды сұрай аласыз немесе LLM-пен жұмыс істеуді жалғастыру барысында осында мәтіндік сұраныс жібере аласыз:

Нәтижелер үшін жылу картасын көрсетіңіз.

- LLM-нің коды мен дайын график түріндегі жауабы, графикті IDE-ге көшіруге болатын код, ал графикті құжатқа қою үшін көшіруге немесе сақтауға болады:



The screenshot shows a software interface titled "Pipeline.py". At the top, there are three colored window control buttons (red, yellow, green) and a "Pipeline.py" logo. To the right of the logo is a "Create in LLM Chat" button and a row of four icons: a neural network, a document, a starburst, and a whale. Below this is a code editor window containing the following Python code:

```

1 # Visualization 2: Heatmap
2 plt.figure(figsize=(8, 4))
3 sns.heatmap(df_visual, annot=True, cmap='viridis')
4 plt.title('Data Verification Summary - Heatmap')
5 plt.xlabel('Verification Status')
6 plt.ylabel('Data Categories')
7 plt.show()

```

At the bottom of the interface is a toolbar with icons for "Run in IDE" (containing VS Code, PyCharm, and Jupyter Notebook symbols), "PC", and a "Run" button.



Рис. 7.211 Деректерді тексеру нәтижелерін визуализациялау (Рис. 7.28) Seaborn кітапханасы арқылы.-

Интерактивті тәсілді таңдағандар үшін динамикалық диаграммалар мен панельдер жасауға мүмкіндік беретін құралдар бар. Plotly кітапханасы (Рис. 7.16, Рис. 7.212) жоғары интерактивті диаграммалар мен панельдер жасау мүмкіндігін ұсынады, оларды веб-беттерге енгізуге және пайдаланушыға деректермен нақты уақыт режимінде өзара әрекеттесуге мүмкіндік береді.-

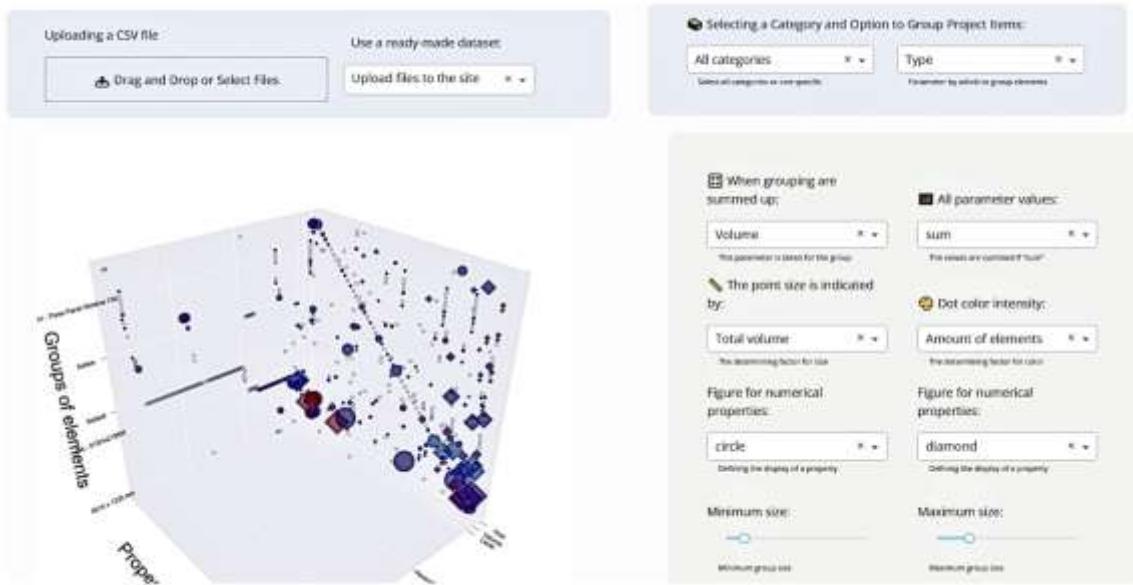


Рис. 7.212 CAD- (BIM-) жобасындағы элементтердің атрибуттарын Plotly кітапханасы арқылы интерактивті 3D-визуализациялау.

Bokeh, Dash және Streamlit сияқты мамандандырылған ашық кітапханалар деректерді ұсынудың ыңғайлы тәсілін қамтамасыз етеді, терен веб-дамыту білімдерін талап етпейді. Bokeh күрделі интерактивті графиктер үшін, Dash толыққанды аналитикалық дашбордтар құру үшін, ал Streamlit деректерді талдау үшін веб-қосымшаларды жылдам жасауға мүмкіндік береді.

Мұндай визуализация құралдарының арқасында әзірлеушілер мен аналитиктер нәтижелерді

әріптестер мен мұдделі тараптар арасында тиімді таратат алады, деректермен интуитивті өзара әрекеттесуді қамтамасыз етіп, шешім қабылдау процесін жеңілдетеді.

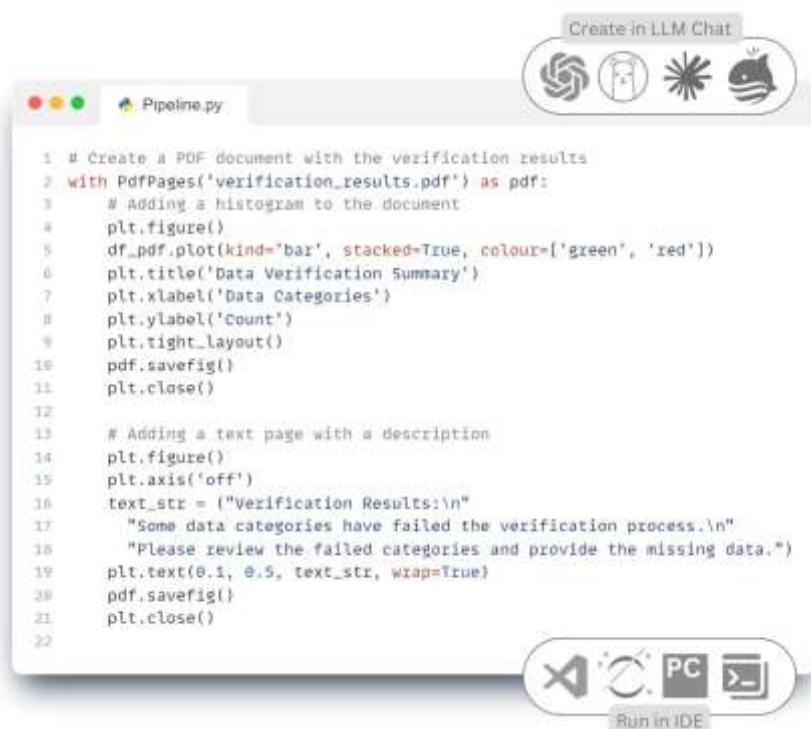
ETL Load: PDF ҚҰЖАТТАРЫН АВТОМАТТЫ ТҮРДЕ ЖАСАУ

Жүктеу кезеңінде деректерді визуализациялаумен қатар, оларды кестелерге немесе деректер базаларына экспорттауға, сондай-ақ қажетті графиктер, диаграммалар мен негізгі аналитикалық көрсеткіштерді қамтитын автоматтандырылған есептерді қалыптастыруға болады, оларды менеджер немесе нәтижелерді күтіп отырған маман алады. Автоматтандырылған есептер түсініктемелер мен деректердің мәтіндік интерпретациясын, сондай-ақ визуализацияның көрнекі элементтерін – кестелер, графиктерді қамтуы мүмкін.

- Алдыңғы тарауларда жүргізілген тексеру нәтижелері бойынша гистограмма (Рис. 7.29) және талдау сипаттамасымен PDF-есепті жасау үшін, LLM-мен диалогты жалғастыру барысында сұраныс формулировкасы жеткілікті, мысалы:-

Жоғарыда алынған деректердің нәтижелерімен гистограмма және сипаттамасы бар PDF-файл жасау үшін код жазыңыз, және кейбір категориялардың тексеруден өтпегені туралы мәтін түрінде ескерту жазыңыз, сондай-ақ жетіспейтін деректерді толтыру қажет ↴

- LLM-нің коды мен нәтижелермен дайын PDF-тің жауабы:



```

1 # Create a PDF document with the verification results
2 with PdfPages('verification_results.pdf') as pdf:
3     # Adding a histogram to the document
4     plt.figure()
5     df_pdf.plot(kind='bar', stacked=True, colour=['green', 'red'])
6     plt.title('Data Verification Summary')
7     plt.xlabel('Data Categories')
8     plt.ylabel('Count')
9     plt.tight_layout()
10    pdf.savefig()
11    plt.close()
12
13    # Adding a text page with a description
14    plt.figure()
15    plt.axis('off')
16    text_str = ("Verification Results:\n"
17                "Some data categories have failed the verification process.\n"
18                "Please review the failed categories and provide the missing data.")
19    plt.text(0.1, 0.5, text_str, wrap=True)
20    pdf.savefig()
21    plt.close()
22

```

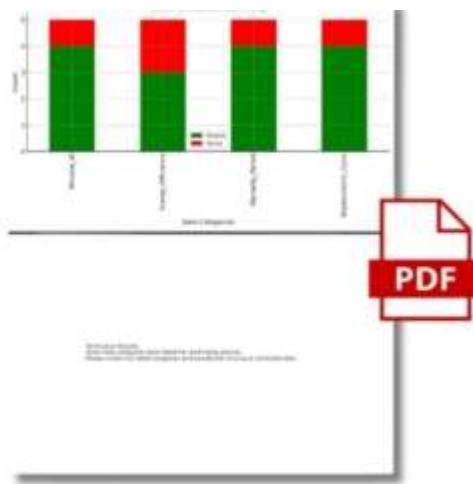


Рис. 7.213 Автоматты код тестілік деректермен гистограмманы және тексеру нәтижелерімен мәтінді қамтитын PDF құжатын жасайды.

LLM көмегімен 20 жолдан тұратын автоматтандырылған шешім қажетті PDF (немесе DOC) құжатын гистограмма түрінде визуализациямен, тексеруден өткен және өтпеген деректердің санын көрсеттін атрибуттармен, сондай-ақ нәтижелердің қысқаша сипаттамасы мен әрі қарайғы әрекеттер бойынша ұсыныстарды қамтитын мәтін блогын қосу арқылы бірден жасайды.-

Автоматтандырылған құжаттарды генерациялау – жүктеу кезеңінің негізгі элементі, әсіресе жобалық қызмет жағдайында, мұнда есептілікті дайындаудың жылдамдығы мен дәлдігі өте маңызды.

ETL Load: FPDF ПЕН ҚҰЖАТТАРДЫ АВТОМАТТЫ ГЕНЕРАЦИЯЛАУ

ETL жүктеу кезеңінде есептілікті автоматтандыру – деректерді өңдеудің маңызды кезеңі, әсіресе талдау нәтижелері беруге және қабылдауға ыңғайлы форматта ұсынылуы қажет болғанда. Құрылыш саласында бұл жиі жұмыс барысы, жобалық деректер статистикасы, сапа тексеру есептері немесе қаржылық құжаттама туралы есептер үшін өзекті.

FPDF кітапханасы, Python және PHP үшін қолжетімді, осындағы тапсырмаларды орындау үшін ең ыңғайлы құралдардың бірі болып табылады.

Ашиқ FPDF кітапханасы код арқылы құжаттарды қалыптастырудың икемді тәсілін ұсынады, бұл заголовкаларды, мәтінді, кестелерді және суреттерді қосуға мүмкіндік береді. Кодты қолдану қолмен өңдеуге қарағанда қателер санын азайтып, PDF форматында есептерді дайындау процесін жеделдетеді.

PDF құжатын жасаудың негізгі кезеңдерінің бірі – тақырыптар мен негізгі мәтінді комментарийлер немесе сипаттама түрінде қосу. Алайда есепті қалыптастыру кезінде тек мәтінді қосу ғана емес, сонымен қатар оны дұрыс құрылымдау да маңызды. Тақырыптар, шегіністер, жолдар арасындағы интервалдар – бұлардың барлығы құжаттың оқылуына әсер етеді. FPDF пайдалану арқылы форматтау параметрлерін орнатуға, элементтердің орналасуын басқаруға және құжаттың стилін баптауға болады.

FPDF HTML-мен жұмыс істеу принципі бойынша өте үқсас. HTML-мен таныс адамдар FPDF-ті пайдалана отырып, кез келген күрделі PDF құжаттарын оңай жасай алады, себебі код құрылымы HTML белгілеуіне үқсас: тақырыптарды, мәтінді, суреттерді және кестелерді қосу үқсас тәсілмен жүзеге асырылады. HTML-мен таныс емес адамдар алаңдамауы керек – қажетті құжаттың безендірілуі үшін кодты тез арада жасауға көмектесетін LLM-ді пайдалануға болады.

- Келесі мысал есепті тақырыппен және негізгі мәтінмен қалай қалыптастыруды көрсетеді. Бұл кодты Python-ды қолдайтын кез келген IDE-де орындау қажетті тақырып пен мәтінді қамтитын PDF-файлды жасайды.

```
FPDF кітапханасын импорттау # PDF құжатын жасау үшін FPDF(pdf = FPDF())
pdf.add_page() # Бет қосу

pdf.set_font("Arial", style='B', size=16) # Шрифт орнату: Arial, қалың, өлшемі
16pdf.cell(200, 10, "Жоба бойынша есеп", ln=True, align='C') # Заголовокты жасай-
мыз және оны орталықтандырамызpdf.set_font("Arial",size=12) # Шрифтті Arial-ға, өлшемі 12-ге өзгер-
теміз.Бұл құжат жобалық файлдардың тексеру нәтижелері бойынша деректерді
қамтиды...pdf.output(r"C:\reports\report.pdf") # PDF-файлды сақтаймыз
```



Рис. 7.214 Бірнеше жол кодты пайдалана отырып, Python бағдарламасында қажетті PDF құжатын автоматты түрде генерациялай аламыз.

Есептерді дайындау барысында деректердің, олардан қалыптасатын құжаттардың, сирек статикалық күйде қалатынын ескеру маңызды. Заголовкалар мен мәтін блоктары (Сурет 7.214) жиі динамикалық түрде қалыптасып, ETL процесінің Transform кезеңінде мәндерді алады.

Кодты пайдалану жобаның атауын, есептің қалыптасу күнін, сондай-ақ қатысушылар немесе ағымдағы мәртебе туралы ақпаратты қамтитын құжаттарды жасауға мүмкіндік береді. Кодтағы айнымалыларды пайдалану осы деректерді есептің қажетті жерлеріне автоматты түрде енгізуге мүмкіндік береді, бұл жіберу алдында қолмен өңдеуді то-лығымен жояды.

Тек мәтін мен тақырыптардан бөлек, жобалық құжаттамада кестелер ерекше орын алады. Эр құжатта құрылымдалған деректер бар: объектілердің сипаттамасынан бастап тексеру нәтижелеріне дейін. Transform кезеңінен алынған деректер негізінде автоматты түрде кестелерді қалыптастыру құжаттарды дайындау процесін жылдамдатуға ғана емес, ақпаратты көшіру кезінде қателіктерді минимизациялауға да мүмкіндік береді. FPDF PDF-файлдарға кестелерді (мәтін немесе сурет түрінде) енгізуге, ұяшықтардың шекараларын, бағаналардың өлшемдерін және шрифттерді

анықтауға мүмкіндік береді. Бұл динамикалық деректермен жұмыс істегендеге, жолдар мен бағаналардың саны құжаттың міндеттеріне байланысты өзгеруі мүмкін болғанда, әсіресе ыңғайлышты.

■ Келесі мысал материалдар тізімін, сметалық есептеулерді немесе параметрлерді тексеру нәтижелерін автоматты түрде кесте жасау жолын көрсетеді:

```
data = [ ["Элемент", "Количество", "Цена"], # Бағаналардың тақырыптары      ["Бетон", "10 м³", "$500."], # Бірінші жолдың деректері
         ["Арматура", "2 т.", "$600"], # Екінші жолдың деректері      ["Кирпич", "5000 шт.", "$750"] # Үшінші жолдың деректері]

pdf = FPDF() # PDF-документті жасаймызды pdf.add_page() # Бет қоса-  
мызды pdf.set_font("Arial", size=12) # Шрифт орнатамыз

for row in data: # Кесте жолдарын айналып өтеміз      for item in row: # Жолдағы  
ұяшықтарды айналып өтеміз
    pdf.cell(60, 10, item, border=1) # Шекарасы бар, ені 60 және биіктігі 10 ұяшық  
жасаймызды pdf.ln() # Келесі жолға өтемізді pdf.output(r"C:\reports\table.pdf") # PDF-файлды  
сақтаймыз
```

Item	Quantity	Price
Concrete	10 м³	\$500
Rebar	2 т.	\$600
Brick	5000 pcs.	\$750



Кестелік ақпаратты PDF-те автоматты түрде генерациялауға болады, тек мәтін ғана
емес, Transform кезеңінде алынған кез келген кестелік ақпарат.

Нақты есеп беру сценарийлерінде кестелер, әдетте, деректерді трансформациялау кезеңінде алынған динамикалық түрде қалыптасқан ақпаратты білдіреді. Келтірілген мысалда кесте PDF-документке статикалық түрде енгізіледі: data сөздігіне (кодтың бірінші жолы) мысал үшін деректер енгізілген, нақты жағдайларда мүндай data айнымалысы, мысалы, деректер фреймін топтастырғаннан кейін автоматты түрде толтырылады.

Практикада мүндай кестелер, әдетте, әртүрлі динамикалық көздерден алынған құрылымдалған деректер негізінде құрылады: деректер базалары, Excel файлдары, API интерфейстері немесе аналитикалық есептеулердің нәтижелері. Көбінесе Transform (ETL) кезеңінде деректер агрегатталады, топтастырылады немесе сұзгіленеді – және тек содан кейін графиктер немесе екі өлшемді кестелер түрінде қорытындыға айналады, есептерде көрсетіледі. Бұл кестенің мазмұны таңдалған параметрлерге, талдау кезеңіне, жобалық сұзгілерге немесе пайдаланушы параметрлеріне байланысты.

өзгеруі мүмкін дегенді білдіреді.

Динамикалық деректер фреймдері мен деректер жиынтықтарын Transform кезеңінде пайдалану есептерді қалыптастыру процесін Load кезеңінде максималды икемді, масштабталатын және оңай қайталанатын етеді, қолмен араласуды қажет етпейді.

Кестелер мен мәтіндерден бөлек, FPDF сонымен қатар кестелік деректер графиктерін қосуға қолдау көрсетеді, бұл есепке Matplotlib немесе жоғарыда қарастырылған басқа визуализация кітапханалары арқылы жасалған суреттерді енгізуге мүмкіндік береді. Код арқылы құжатты кез келген графиктермен, диаграммалармен және схемалармен толықтыруға болады.

- Python FPDF кітапханасын пайдаланып, PDF құжатына Matplotlib арқылы алдын ала жасалған графикті қосамыз:

```
import matplotlib.pyplot as plt # Графиктер жасау үшін Matplotlib-ті импорттаймыз
fig, ax = plt.subplots() # График үшін фигура мен осытерді құрыс categories = ["Бетон", "Арматура", "Кирпич"] #
Санаттардың атаулары values = [50000, 60000, 75000] # Санаттар бойынша мәндер

ax.bar(categories, values) # Бағандық диаграмманы құралыт.уlabel("Баға, $.") # Y осін белгілеу
plt.title("Шығындардың таралуы") # Тақырып
plt.savefig(r"C:\reports\chart\chart.png") # Графикті сурет ретінде сақтау
pdf = FPDF() # PDF-документті жасайды pdf.add_page() # Бет қосамыз pdf.set_font("Arial", size=12) # Шрифт орнатау
мыз

pdf.cell(200, 10, "Шығындар графигі", ln=True,
align='C') # Тақырып қосу
pdf.image(r"C:\reports\chart\chart.png", x=10, y=30, w=100) # PDF-ке суретті енгізу (x, y - координаттар, w - ені)
pdf.output(r"C:\reports\chart_report.pdf") # PDF-файлды сақтау
```

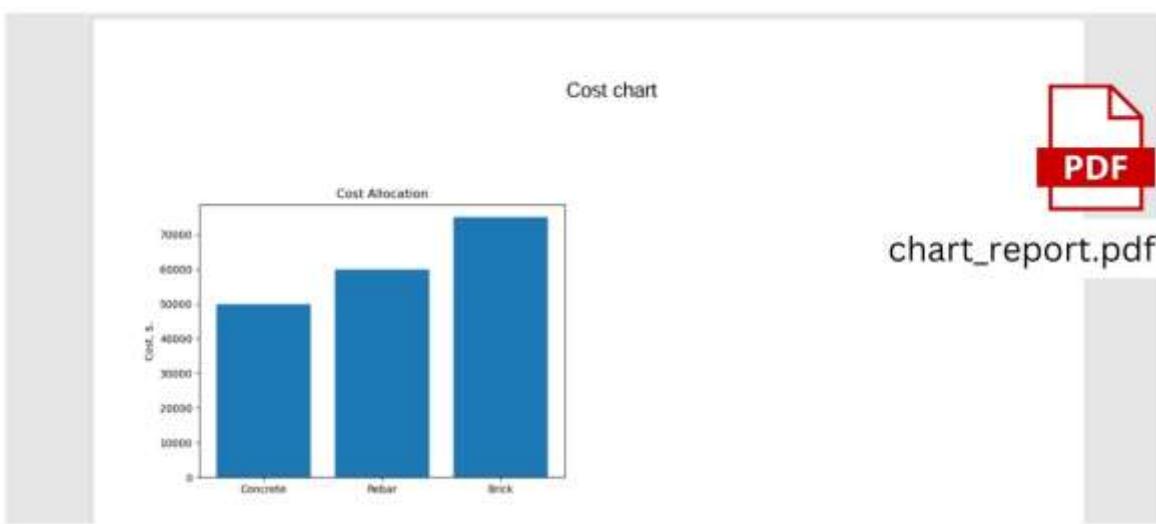


Рис. 7.216 Он шақты код жолымен графикті генерациялап, оны сақтап, кейін PDF құжатына қажетті жерге енгізуге болады.

FPDF арқасында құжаттарды дайындау процесі мен логикасы айқын, жылдам және ыңғайлышты болады. Кодқа енгізілген шаблондар құжаттарды өзекті деректермен генерациялауға мүмкіндік береді, қолмен толтыру қажеттілігін жояды.

ETL автоматизациясын пайдалану арқылы – қолмен есептерді дайындаудың ауыртпалығын жою, мамандар деректерді талдауға және шешім қабылдауға назар аудара алады, ал қажетті құралды таңдау мәселе сімен айналыспайды.

Осылайша, FPDF кітапханасы кез келген күрделілік деңгейіндегі құжаттарды автоматтандырылған түрде жасау үшін икемді құралды ұсынады – қысқа техникалық есептерден бастап, кестелер мен графиктері бар кешенді аналитикалық шолуларға дейін, бұл құжат айналымын жылдамдатуға ғана емес, сонымен қатар деректерді қолмен енгізу мен форматтаумен байланысты қателіктердің ықтималдығын едәуір төмендетуге мүмкіндік береді.

ETL Load: ЕСЕПТЕРДІ ҚҰРАСТАЫРУ ЖӘНЕ БАСҚА ЖҮЙЕЛЕРГЕ ЖҮКТЕУ

Load кезеңінде кестелер, графиктер және белгіленген талаптарға сәйкес дайындалған PDF есептері түрінде нәтижелер қалыптастырылады. Одан әрі, бұл деректерді машинамен оқылатын форматтарға (мысалы, CSV) экспорттау мүмкіндігі бар, бұл сыртқы жүйелермен интеграция үшін қажет – ERP, CAFM, CPM, BI платформалары және басқа корпоративтік немесе салалық шешімдер. CSV-дан басқа, экспорт XLSX, JSON, XML форматтарына немесе автоматты ақпарат алмасуды қолдайтын дерекқорларға тікелей жүргізілуі мүмкін.

- Load кезеңін автоматтандыру үшін сәйкес кодты қалыптастыру үшін LLM интерфейсіне сұраныс беру жеткілікті, мысалы: ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude немесе QWEN:

Деректерді тексеру нәтижелері туралы есеп жасау үшін код жазыңыз, онда 'verified_' префиксі бар бағандар 'Passed' және 'Failed' деп қайта аталады, жіберілген мәндер 0-мен ауыстырылады, содан кейін барлық тексерулерден өткен жолдар ғана CSV файлына экспортталады.

- LLM-нің жауабы:

```

1 # Create a report on the results of data validation
2 report = df.filter(like='verified_').apply(lambda x: x.value_counts()); T
3 report.columns = ['Passed', 'Failed']
4 report.fillna(0, inplace=True)
5
6 # Export the cleaned and verified data to a CSV file
7 verified_df = df[df.filter(like='verified_').all(axis=1)]
8 verified_df.to_csv('verified_data.csv', index=False)

```

Рис. 7.217 Трансформация кезеңінде алынған тексерілген деректер соғы датафреймнен CSV файлына экспортталауды, бұл басқа жүйелермен интеграция үшін қажет.

Берілген кодта (Рис. 7.217) ETL-процесінің соғы кезеңі – Load жүзеге асырылады, оның барысында тексерілген деректер CSV форматында сақталады, бұл көптеген сыртқы жүйелер мен дәрекқорлармен үйлесімді. Осылайша, біз ETL-процесінің толық циклын аяқтадық, оған деректерді алу, түрлендіру, визуализациялау, құжаттау және қажетті жүйелер мен форматтарға экспорттау кіреді, бұл ақпаратпен жұмыс істеудің қайта өндірілуін, ашықтығын және автоматтандырылуын қамтамасыз етеді.

Дамытылған ETL-конвейерін (pipeline) бір жобаларды өңдеу үшін де, жүздеген және мындаған кіріс деректерін, құжаттар, суреттер, сканерлер, CAD-жобалар, нұктелер бұлттары, PDF-файлдар немесе басқа да таралған жүйелерден келетін деректерді талдау үшін де қолдануға болады. Процесті толық автоматтандыру мүмкіндігі ETL-ді тек техникалық өңдеу құралына айналдырмай, цифрлық құрылыш ақпараттық инфрақұрылымының негізіне айналдырады.

LLM АРҚЫЛЫ ETL: PDF ҚҰЖАТТАРЫНАН ДЕРЕКТЕРДІ КӨРСЕТУ

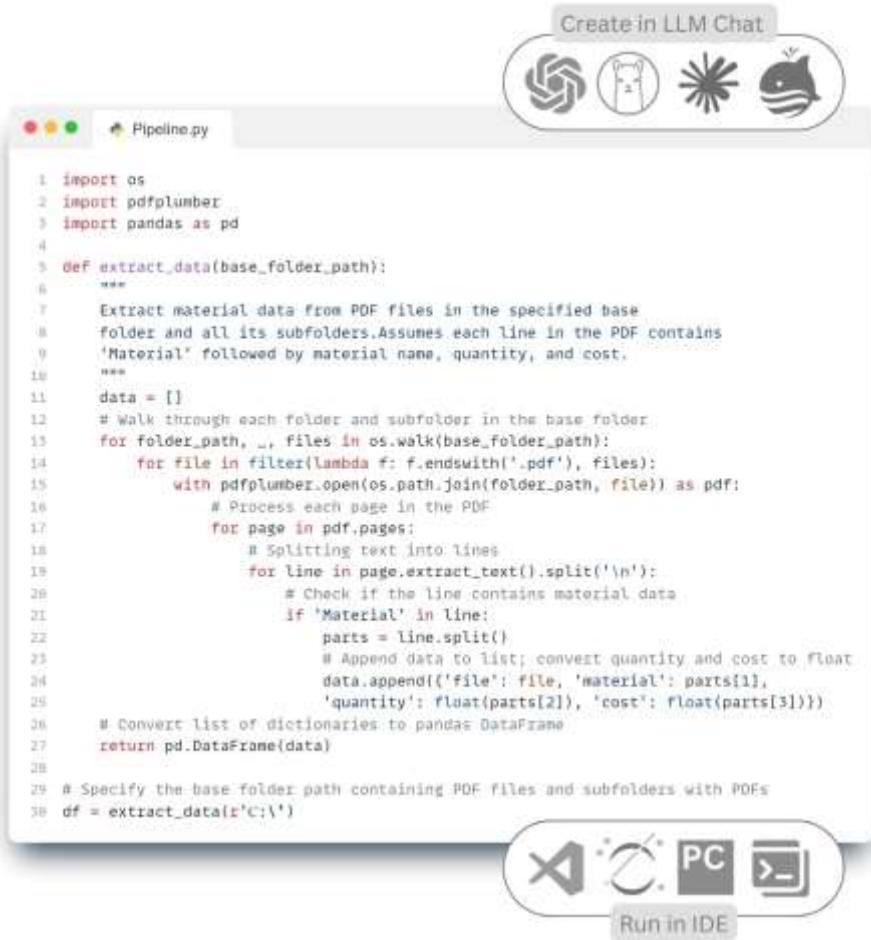
Енді деректермен жұмыс істеудің барлық негізгі кезеңдерін қамтитын толыққанды ETL-процесін жинақтау уақыты келді – алу, түрлендіру және жүктеу. PDF-документтерді қолмен жұмыссыз өңдеуге мүмкіндік беретін автоматтты ETL-Pipeline құрастырамыз – құжаттардан деректерді алу, визуализациялау, талдау және басқа жүйелерге беру.

Біздің мысалымызда ETL процесі промпттар арқылы сипатталады, олар тілдік модельге (LLM) ETL процестерін түсіндіруі және алынуы тиіс соғы нәтижені сипаттауы керек. Бұл жағдайда тапсырма – көрсетілген папка мен оның ішкі папкаларында барлық PDF-файлдарын табу, олардан тиісті ақпаратты – мысалы, материалдардың атаулары, саны және құны – алу және нәтижені құрылымдалған кесте (DataFrame) түрінде ұсыну.

- ❷ LLM-ге көптеген PDF құжаттарынан деректерді автоматтты түрде алу және Extract кезеңі үшін деректер датафреймін жасау туралы алғашқы мәтіндік сұраныс:

Берілген папка мен оның ішкі папкаларындағы PDF-файлдардан материалдар туралы ақпаратты алу үшін код жазыңыз. PDF-дегі деректер материалдың атауын, санын және құнын қамтиды. Нәтижені DataFrame-ге сақтау қажет ↴

2 LLM-нің жауабы:



```

1 import os
2 import pdfplumber
3 import pandas as pd
4
5 def extract_data(base_folder_path):
6     """
7         Extract material data from PDF files in the specified base
8         folder and all its subfolders. Assumes each line in the PDF contains
9         'Material' followed by material name, quantity, and cost.
10    """
11    data = []
12    # Walk through each folder and subfolder in the base folder
13    for folder_path, _, files in os.walk(base_folder_path):
14        for file in filter(lambda f: f.endswith('.pdf'), files):
15            with pdfplumber.open(os.path.join(folder_path, file)) as pdf:
16                # Process each page in the PDF
17                for page in pdf.pages:
18                    # Splitting text into lines
19                    for line in page.extract_text().split('\n'):
20                        # Check if the line contains material data
21                        if 'Material' in line:
22                            parts = line.split()
23                            # Append data to list; convert quantity and cost to float
24                            data.append({'file': file, 'material': parts[1],
25                                         'quantity': float(parts[2]), 'cost': float(parts[3])})
26    # Convert list of dictionaries to pandas DataFrame
27    return pd.DataFrame(data)
28
29 # Specify the base folder path containing PDF files and subfolders with PDFs
30 df = extract_data(r'C:\')

```

Рис. 7.218 LLM PDF-файлдардан деректерді алу үшін берілген папка мен оның ішкі папкаларында Python-кодын жасайды.

LLM-нің жауабы (Рис. 7.218) – барлық папкаларды автоматты түрде айналып өтіп, табылған PDF-файлдарды ашып, олардан мәтіндік ақпаратты алып, оны кестеге түрлендіретін дайын Python-скрипт. Чатта алынған кодты PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагинімен, Thonny, Wing IDE, IntelliJ IDEA Python плагинімен, JupyterLab немесе танымал онлайн құралдар Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker платформаларында іске қосуға болады.-

Процесті жеңілдету үшін, LLM-ден кодты көшіріп алып, IDE-де қолданудың орнына, PDF-файлдардың оншақтысын тікелей LLM чатына жүктеп, кодты көрү немесе іске қосу қажеттілігі жоқ, кестені алуға болады (Рис. 7.219). Бұл кодтың орындалуының нәтижесі – таңдалған атрибуттармен құрылымдалған кесте.-

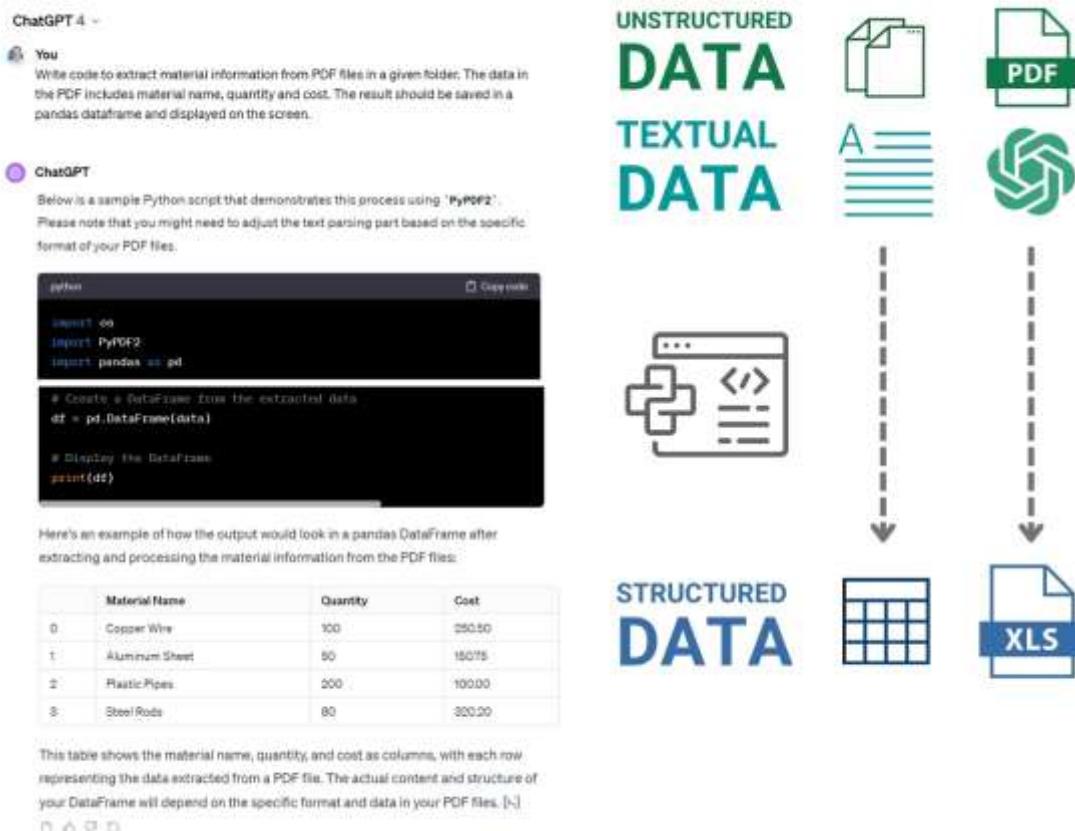


Рис. 7.219 LLM-нің PDF-файлдардан деректерді құрылымдалған DataFrame түрінде алу нәтижесі, таңдалған атрибуттармен.

Келесі кезеңде алынған деректер бойынша тілдік модельден – мысалы, материалдардың құнын және пайдалану көлемдерін салыстыруды және кейінгі талдау үшін негіз болатын бірнеше визуализация мысалдарын жасауды сұраймыз.

- Чатта LLM-мен жалғастыра отырып, Transform кезеңінде алынған кестелерден бірнеше графиктер құруды сұраңыз (Сурет 7.218):

DataFrame-ден әр материалдың жалпы құнын және санын визуализациялаңыз (Сурет 7.218) ↵-

```

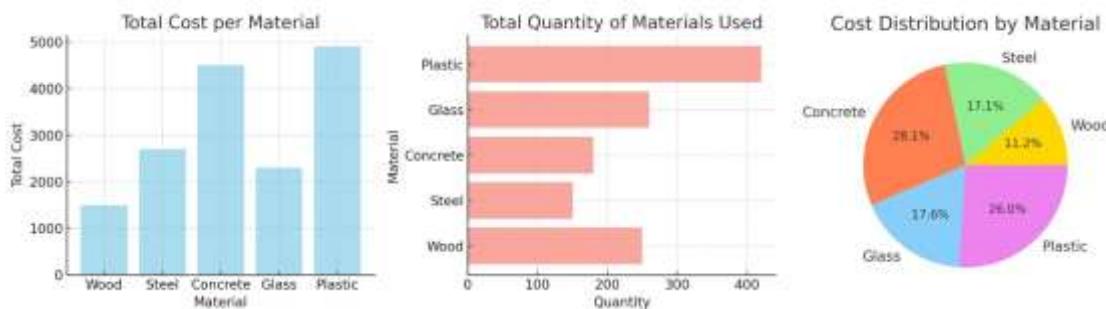
1 import matplotlib.pyplot as plt
2
3 # Plotting total cost per material
4 # Group data by 'material' and sum up the 'cost' for each material
5 df.groupby('material')['cost'].sum().plot(kind='bar')
6 plt.title('Total Cost per Material')
7 plt.xlabel('Material')
8 plt.ylabel('Total Cost')
9 plt.show()
10
11 # Plotting total quantity of materials used
12 # Group data by 'material' and sum up the 'quantity' for each material
13 df.groupby('material')['quantity'].sum().plot(kind='barh')
14 plt.title('Total Quantity of Materials Used')
15 plt.xlabel('Quantity')
16 plt.ylabel('Material')
17 plt.show()
18

```

Run in IDE

Сурет 7.220 LLM-моделінің Python кодындағы жауаптары, matplotlib кітапханасы арқылы деректерді визуализациялау үшін.

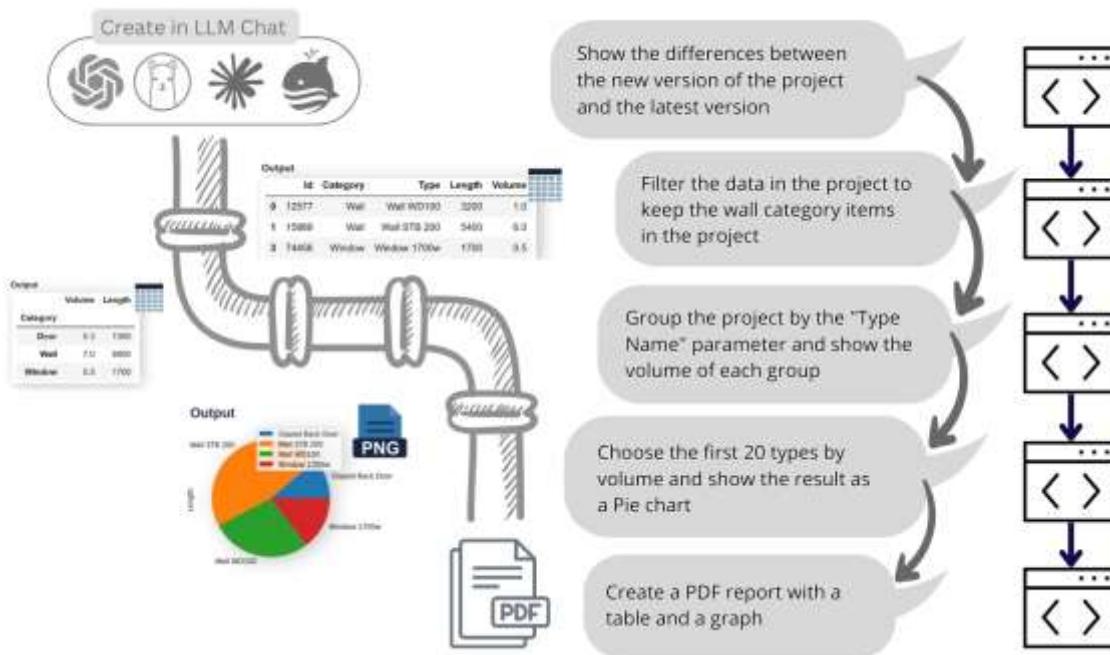
LLM автоматты түрде Python-кодын (Сурет 7.220) генерациялап, орындалп, matplotlib кітапханасын пайдаланады. Бұл кодты орындағаннан кейін, біз құрылыш жобаларындағы материалдардың шығындары мен көлемдерін графиктер түрінде тікелей чатта аламыз (Сурет 7.221), бұл аналитикалық жұмыстың айтарлықтай жеңілдетеді.-



Сурет 7.221 DataFrame-ден жиналған деректер негізінде графиктер түрінде LLM-нің жауаптарын визуализациялау.

ETL кодын жазу, талдау және кодты орындау, нәтижелерді визуализациялау бойынша қолдау LLM арқылы қарапайым мәтіндік сұраулар арқылы қолжетімді, бағдарламалаудың негіздерін үйреноудің қажеті жоқ. LLM сияқты ИИ құралдарының пайда болуы бағдарламалада мен деректерді өндеді, ал ол автомотивтандыруға деген көзқарасты анық өзгертеуді (Сурет 7.222).-

RwC-нің «Сіздің бизнесіңіз үшін жасанды интеллектінің нақты құны неде және одан қалай пайда табуға болады?» (2017) [139] есебіне сәйкес, процестерді автоматтандыру және өнімділікті арттыру экономикалық өсудің негізгі драйверлері болады. Және еңбек өнімділігінің артуы 2017-2030 жылдар аралығында ИИ арқылы ЖІӨ-нің 55%-дан астамын қамтамасыз етеді деп күтілуде.



Сурет 7.222 ИИ LLM болашақ жобаларда қолданылатын кодтың черновигін қалыптастыруға көмектеседі, LLM-ді пайдаланудың қажеті жоқ.

ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok сияқты құралдарды, сондай-ақ ашық деректер мен ашық бастапқы кодты бағдарламалық қамтамасыз етуді пайдалана отырып, біз бұрын тек мамандандырылған, жоғары шығынды және қызмет көрсетуі қын модульдік патенттеген жүйелер арқылы жүзеге асырылған процестерді автоматтандыра аламыз.

Құрылыс контекстінде бұл автоматтандырылған Pipeline деректерді өңдеу процестерін бірінші болып енгізген компаниялар үшін елеулі артықшылықтар береді – жобаларды басқаруда тиімділікті арттырудан бастап, қаржылық шығындарды азайтуға және фрагменттеген қосымшалар мен оқшауланған деректер қоймаларын жоюға дейін.

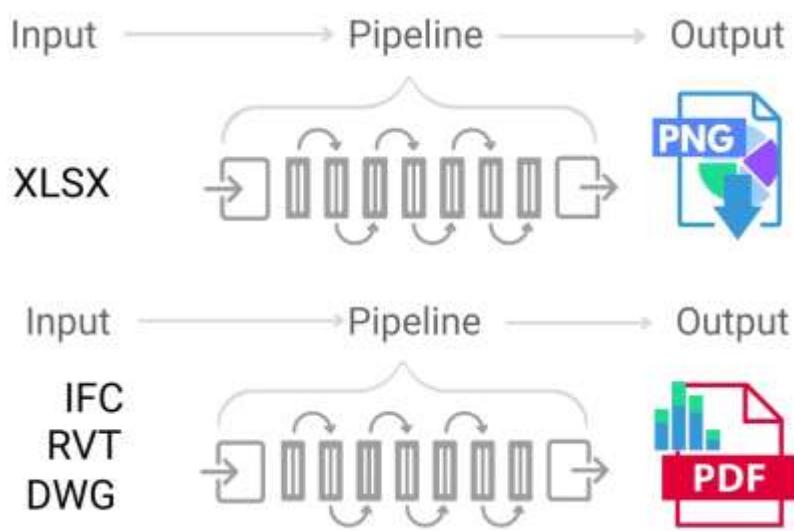
ETL процесінде бизнес міндеттерін орындау логикасы – аналитика мен деректерді өңдеу процестерін автоматтандырудың маңызды бөлігі, бұл кеңірек түсінікті – конвейерлердің (Pipelines) ерекше түрі.

ТАРАУ 7.3. АВТОМАТТЫ ETL ДЕРЕКТЕР КОНВЕЙЕРІ

Pipeline: Автоматты ETL деректер конвейері

ETL процесі дәстүрлі түрде аналитикалық жүйелерде деректерді өңдеу үшін қолданылады, құрылымдаған және құрылымдаған дереккөздерді қамтиды. Алайда, қазіргі цифрлық ортада Pipeline (конвейер) термині жиі қолданылады, ол кез келген өңдеу тізбегін сипаттайды, мұнда бір кезеңнің шығысы келесі кезеңнің кірісі болады.

Мұндай тәсіл тек деректерге ғана емес, сонымен қатар басқа автоматтандыру түрлеріне де қолданылады: тапсырмаларды өңдеу, есептілікті құру, бағдарламалық қамтамасыз етумен интеграция және цифрлық құжат айналымы (Сурет 7.31).-



Сурет 7.31 Pipeline – бұл өңдеу тізбегі, мұнда бір кезеңнің шығысы келесі кезеңнің кірісі болады.

Пайплайнды қолдану автоматтандырудың негізгі элементтерінің бірі болып табылады, әсіресе әртүрлі типтегі деректермен жұмыс істегендеге. Конвейерлік архитектура күрделі өңдеу кезеңдерін модульдік, ретті және басқарылатын форматта ұйымдастыруға мүмкіндік береді, бұл кодтың оқылуын арттырады, қолдауды жеңілдетеді және кезең-кезеңімен отладка мен масштабталатын тестілеуді жүзеге асыруға мүмкіндік береді.

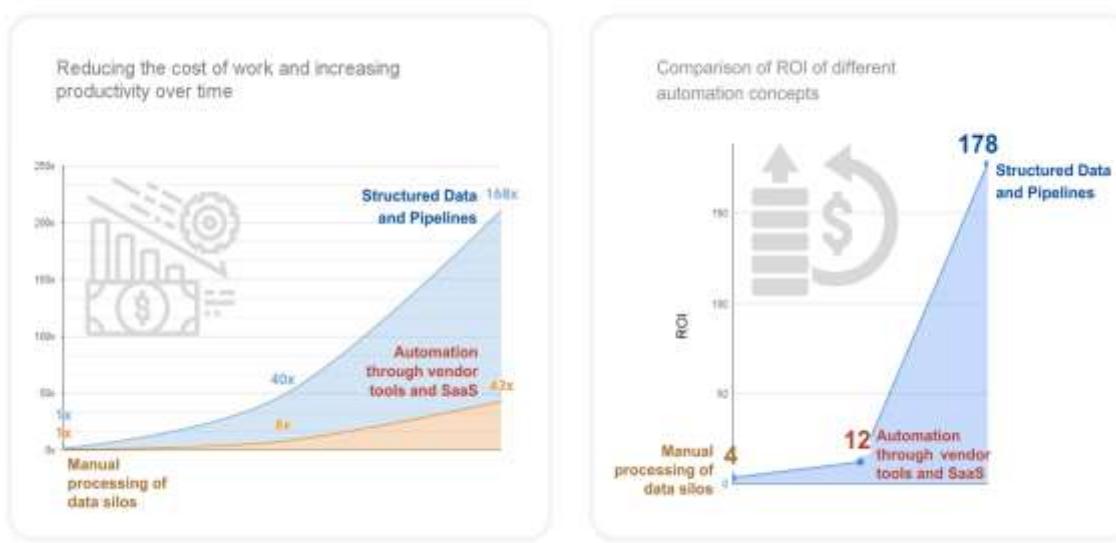


Рис. 7.32 ROI Пайплайн деректерді тексеру процесінің уақыты классикалық құралдармен өңдеуге қарағанда он есе және жүзделген есе қысқарады [74].

Проприетарлық жүйелерде (ERP, PMIS, CAD және т.б.) қолмен жұмыс істеуден айырмашылығы, конвейерлік деректерді өңдеу (Рис. 7.32) тапсырмаларды орындау жылдамдығын едәуір арттыруға, қайталараптың жұмыстардан аулақ болуға және процестерді қажетті уақытта автоматты түрде іске қосуға мүмкіндік береді (Рис. 7.33).--

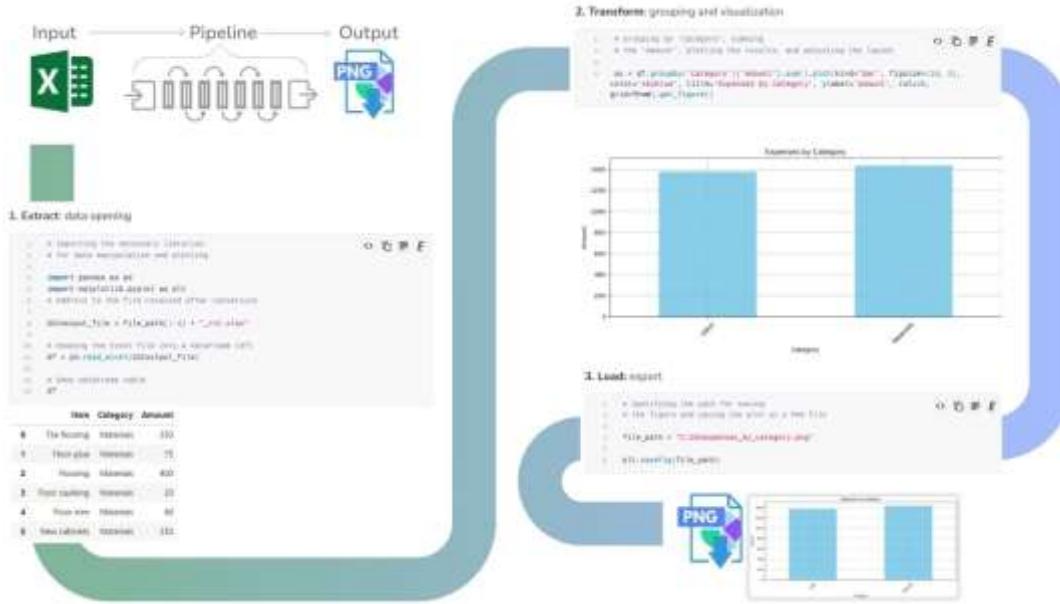


Рис. 7.33 Excel-ді ашпай-ақ XLSX файлының кестелік деректерінен графикті автоматты түрде алу үшін ETL Пайплайнның мысалы.

Ағынды деректерді өңдеу және автоматтандырылған Пайплайн құру үшін, ETL процесіне үқсас, дегеккөздерді алдын ала анықтау және олардың жинау мерзімдерін белгілеу қажет – нақты бизнес-процесс үшін немесе компанияның барлық аясында.

Құрылымың жобаларында деректер әртүрлі жаңарту жиілігімен көптеген гетерогенді дереккөздерден келеді. Сенімді деректер витринасын қалыптастыру үшін ақпаратты алу және жаңарту сәтін фиксируе критикалық маңызды. Бұл уақытылы шешім қабылдауды қамтамасыз етеді және жобаны басқарудың жеделдігін арттырады.

Бір нұсқа – жинау процесін фиксиленген уақытта іске қосу – мысалы, жұмыс күні аяқталғаннан кейін 19:00-де. Осы сәтте әртүрлі жүйелер мен қоймалардан деректерді агрегаттауға жауап беретін бірінші скрипт іске қосылады (Рис. 7.34 қадам 1). Одан кейін деректерді автоматтандырылған өндіріс және аналитикаға жарамды құрылымдық форматқа түрлендіру жүзеге асырылады (Рис. 7.34 қадам 2-4). Соңғы кезеңде, дайындалған деректерді пайдалана отырып, автоматты түрде есептер, дашбордтар және алдыңғы тарауларда сипатталған басқа өнімдер қалыптастырылады (Рис. 7.34 қадам 6-7). Нәтижесінде, таңертен 05:00-ге қарай менеджерлерде қажетті форматта жобаның ағымдағы есептері болады (Рис. 7.35).--

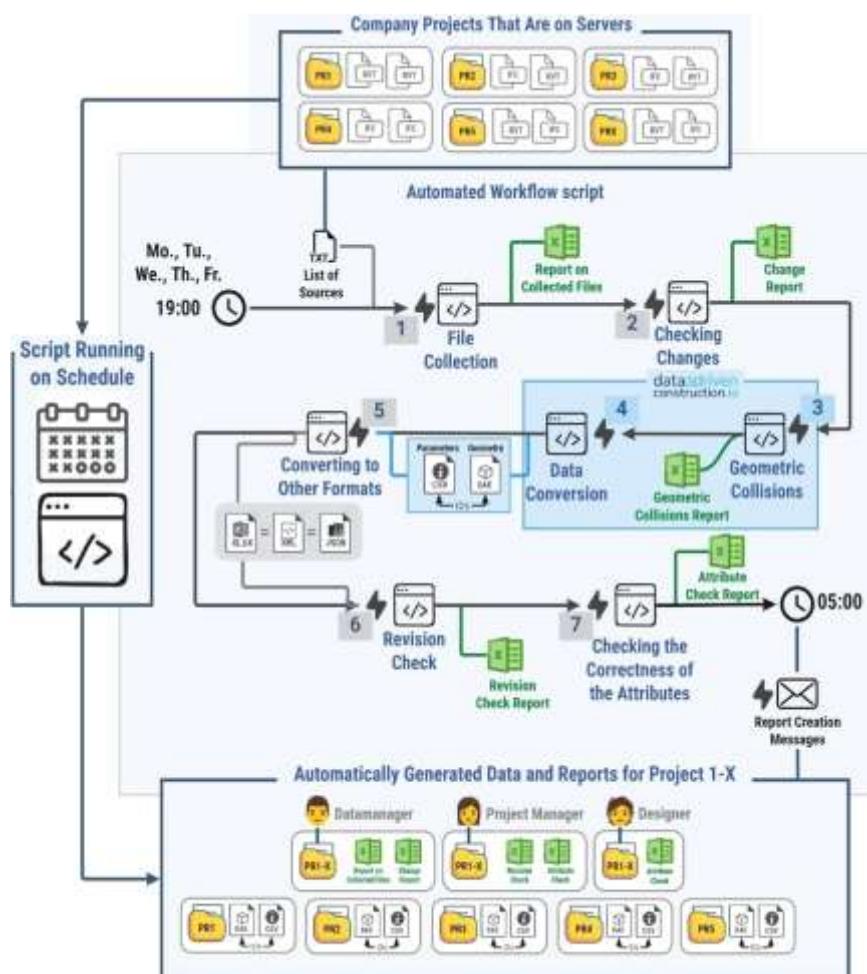


Рис. 7.34 Пайплайндағы деректер, кешке автоматты түрде жиналған, түнде өнделеді, осылайша таңертен менеджерлерде ағымдағы есептер мен жаңа есептер болады.

Деректерді уақтылы жинау, KPI анықтау, түрлендіру процестерін автоматтандыру және ақпараттық панельдер арқылы визуализация – деректер негізінде шешім қабылдаудың табысты элементтері.

Мұндай автоматтандырылған процестер (Рис. 7.34) толық автономиямен орындалуы мүмкін: олар кесте бойынша іске қосылады, оператордың қатысусының деректерді өндейді және бұлтта немесе компанияның жеке серверінде орналастырылуы мүмкін (Рис. 7.35). Бұл ETL конвейерлерін бар ИТ-инфрақұрылымға интеграциялауға мүмкіндік береді, деректерді бақылауды сақтай отырып, масштабтау кезінде икемділікті қамтамасыз етеді.

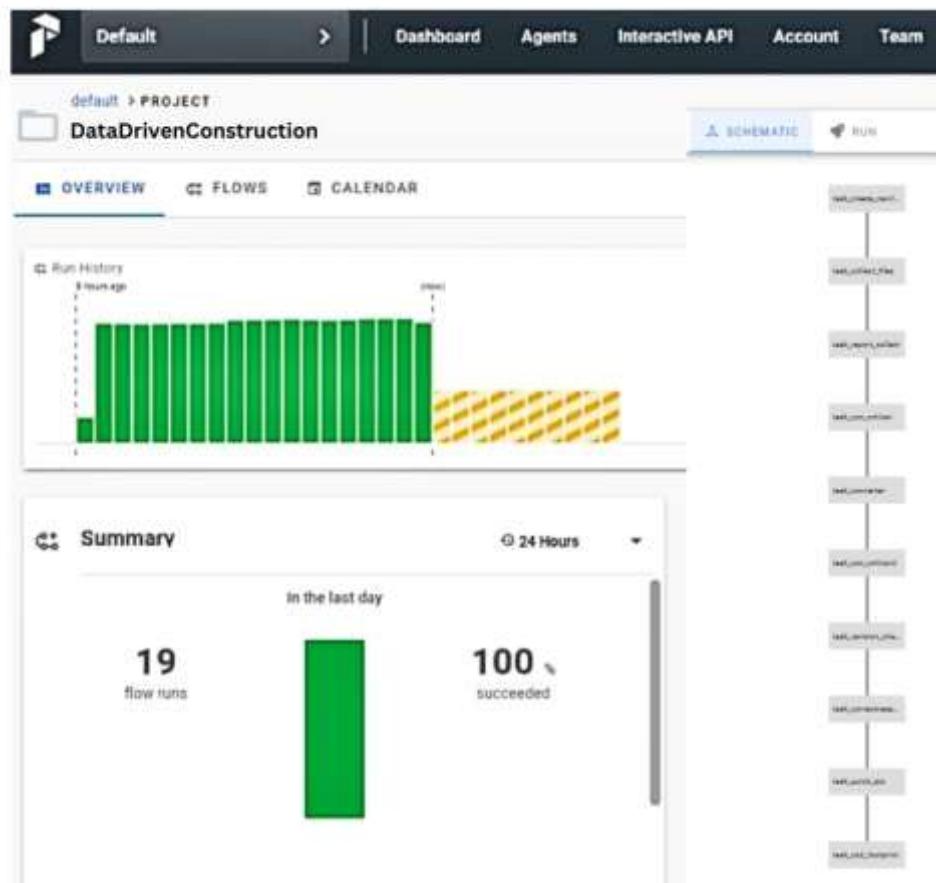


Рис. 7.35 Prefect платформасында автоматты ETL конвейері (Рис. 7.34), онда 10 python скрипті әр жұмыс күні 19:00-ден кейін кезекпен іске қосылады.-

Жұмыс процестерін автоматтандыру команда өнімділігін арттырып қана қоймай, маңызды және рутинді емес тапсырмаларға уақыт босатады, сонымен қатар бизнес-процестерге жасанды интеллект (ЖИ) технологияларын енгізуін алғашқы маңызды қадамы болып табылады, бұл туралы біз "Болжамдар мен машиналық оқыту" бөлімінде толығырақ айтып өтеміз.

Pipeline-ETL ДЕРЕКТЕРДІ ТЕКСЕРУ ПРОЦЕСІ LLM АРҚЫЛЫ

Деректерге талаптарды жасау және ETL автоматтандыруына арналған алдыңғы тарауларда біз деректерді дайындау, түрлендіру, тексеру және визуализациялау процесін кезең-кезеңімен қарастырық. Бұл әрекеттер нақты тапсырманы орындайтын код блоктары түрінде жүзеге асырылды (Рис. 7.218 - Рис. 7.220).-

Енді біздің алдымызда осы элементтерді біртұтас, байланысты және автоматтандырылған деректер өңдеу конвейеріне - ETL-Pipeline - біріктіру мақсаты түр, онда барлық кезеңдер (жүктеу, тексеру, визуализация, экспорт) бір автоорындалатын скрипtte тізбектей орындалады.

Келесі мысалда деректерді өңдеудің толық циклі жүзеге асырылады: бастапқы CSV файлын жүктеуден → құрылым мен мәндерді тексеру үшін регулярлы өрнектерді қолдануға → нәтижелерді есептеуге → PDF форматында визуалды есепті генерациялауға дейін.

- ❶ Қажетті кодты алу үшін келесі мәтіндік сұранысты LLM-ге қолдануға болады:

CSV-ден деректерді жүктеуді, DataFrame деректерін регулярлы өрнектермен тексеруді, 'W-NEW' немесе 'W-OLD' форматындағы идентификаторларды, 'A' мен 'G' ара-лығындағы әріптермен энергия тиімділігін, жылдармен сандық мәндермен кепілдік мерзімін және ауыстыру циклін тексеруді, сонымен қатар тексеруден өткен және өтпеген мәндерді есептеуді, нәтижелердің гистограммасын генерациялауды және мәтіндік сипаттаманы қосатын есепті жасауды сұрай.

 LLM-нің жауабы:



```

1 # Importing necessary libraries
2 import pandas as pd
3 import re
4 import matplotlib.pyplot as plt
5 from matplotlib.backends.backend_pdf import PdfPages
6
7
8 # Load data
9 df = pd.read_csv('path_to_csv_file.csv')
10
11 # Define patterns and verify data
12 patterns = {
13     'Window_ID': r'^W-NEW|W-OLD[0-9]+|D-1[0-9]{2})$',
14     'Energy_Efficiency': r'^[A-G]\+*$', 
15     'Warranty_Period': r'^\d+\syears$', 
16     'Replacement_Cycle': r'^\d+\syears$'
17 }
18
19 for column, pattern in patterns.items():
20     df[f'verified_{column}'] = df[column].apply(lambda x: bool(re.match(pattern, str(x))))
21
22 # Create report
23 report = df.filter(like='verified_').apply(lambda x: x.value_counts()).T
24 report.columns = ['Passed', 'Failed']
25
26 # Generate PDF
27 pdf_path = '/mnt/data/verification_results_full_pipeline.pdf'
28 with PdfPages(pdf_path) as pdf:
29     # Plot and save bar chart
30     report.plot(kind='bar', stacked=True, color=['green', 'red'])
31     plt.title('Data Verification Summary')
32     plt.xlabel('Data Categories')
33     plt.ylabel('Count')
34     plt.tight_layout()
35     pdf.savefig()
36     plt.close()
37
38     # Add text page
39     plt.figure()
40     plt.axis('off')
41     text_str = "Verification Results Summary"
42     plt.text(0.1, 0.5, text_str, wrap=True)
43     pdf.savefig()
44     plt.close()
45
46 pdf_path
47

```

Рис. 7.36 Pipeline (ETL) деректерді өндеудің толық циклын автоматтандырады: жүктеуден және тексеруден бастап PDF форматында құрылымдық есепті жасаумен аяқталады.

LLM чаты немесе DfE ішіндегі автоматтандырылған код (Рис. 7.36) CSV файлының деректерін берілген регулярлы өрнектерді пайдалана отырып тексереді, тексеруден өткен және өтпеген жазбалар саны туралы есеп жасайды, содан кейін тексеру нәтижелерін PDF файлы ретінде сақтайды.

ETL-конвейерінің мұндай құрылымы, мұнда әр кезең - деректерді жүктеуден есепті қалыптастыруға дейін - жеке модуль ретінде жүзеге асырылған, ашықтықты, масштабталуды және қайталануды қамтамасыз етеді. Тексеру логикасын Python тілінде оңай оқылатын код түрінде көрсету процессін тек әзірлеушілерге ғана емес, деректерді басқару, сапа және аналитика саласындағы мамандарға да түсінікті етеді.

Деректерді өңдеуді автоматтандыру үшін Pipeline тәсілін қолдану процестерді стандарттауға, олардың қайталанғыштығын арттыруға және жаңа жобаларға бейімделуді жеңілдетуге мүмкіндік береді. Осылайша, деректерді талдау үшін біртұтас методология қалыптасады, дереккөзге немесе тапсырма түріне қарамастан - стандарттарға сәйкестікті тексеру, есепті генерациялау немесе деректерді сыртқы жүйелерге беру.

Мұндай автоматтандыру адам факторының әсерін азайтады, меншікті шешімдерге тәуелділікті төмендетеді және нәтижелердің дәлдігі мен сенімділігін арттырады, оларды жобалар деңгейінде жедел аналитика үшін де, компаниялар деңгейінде стратегиялық аналитика үшін де жарамды етеді.

Pipeline-ETL: CAD (BIM) ЖОБА ЭЛЕМЕНТТЕРІНІҢ ДЕРЕКТЕРІН ЖӘНЕ АҚПАРАТЫН ТЕКСЕРУ

CAD (BIM) жүйелерінен және деректер базаларынан алынған мәліметтер – құрылыш компанияларының бизнесіндегі ең күрделі және динамикалық жаңартылатын дереккөздердің бірі. Бұл қосымшалар жобаны геометрия арқылы сипаттап қана қоймай, сонымен қатар оны көпқабатты мәтіндік ақпаратпен толықтырады: көлемдер, материалдардың қасиеттері, бөлмелердің тағайындалуы, энергия тиімділігі деңгейлері, рұқсаттар, пайдалану мерзімдері және басқа атрибуттар.

CAD модельдеріндегі объектілерге берілетін атрибуттар жобалау кезеңінде қалыптасады және одан әрі бизнес-процестердің негізін құрайды, соның ішінде шығындарды есептеу, графиктерді құру, өмірлік циклды бағалау және ERP және CAFM жүйелерімен интеграциялау, мұнда процестердің тиімділігі көп жағдайда жобалау бөлімдерінен келетін деректердің сапасына байланысты.

CAD (BIM) модельдеріндегі атрибуттарды тексерудің дәстүрлі тәсілі қолмен валидациялауды (Рис. 7.21) қамтиды, бұл үлкен модельдер көлемінде ұзақ және қыимбат процесс болып табылады. Қазіргі құрылыш жобаларының көлемі мен санына және олардың тұрақты жаңартуларына байланысты, деректерді тексеру және трансформациялау процесі тұрақсыз және ауыртпалықты болып отыр.-

Генеральді мердігерлер мен жобаларды басқарушылар көптеген жобалық деректерді, соның ішінде бірдей модельдердің бірнеше нұсқалары мен фрагменттерін өндедеу қажеттілігімен бетпе-бет келеді. Деректер жобалау үйымдарынан RVT, DWG, DGN, IFC, NWD және басқа форматтарда (Рис. 3.114) келеді және салалық және корпоративтік стандарттарға сәйкестігін тұрақты тексеруді талап етеді.

Қолмен әрекеттерге тәуелділік пен мамандандырылған бағдарламаларға байланысты деректерді валидациялау процесі компанияның модельдермен байланысты жұмыс процестерінде тар орын болып табылады. Автоматтандыру және құрылымдық талаптарды қолдану бұл тәуелділікті жоюға мүмкіндік береді, деректерді тексеру жылдамдығы мен сенімділігін бірнеше есе арттырады (Рис. 7.37).

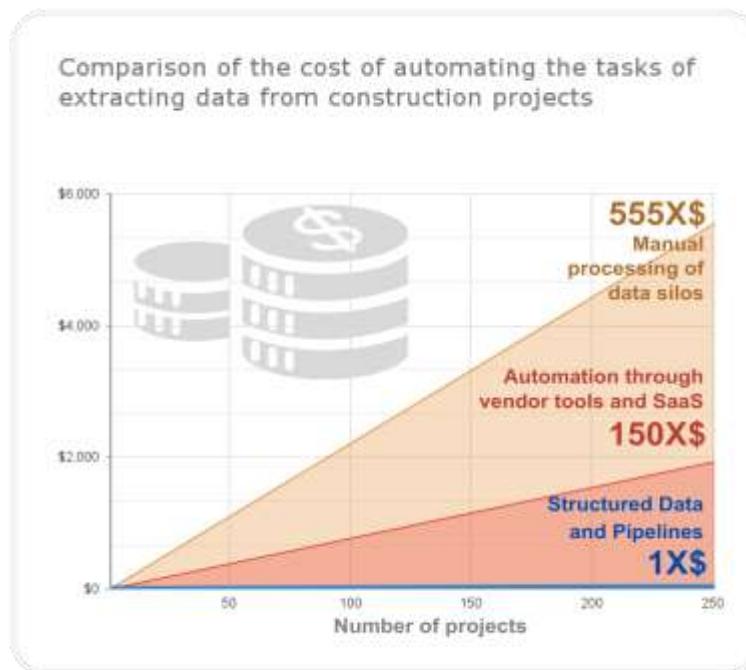


Рис. 7.37 Автоматтандыру деректерді тексеру және өндедеу жылдамдығын арттырады, бұл жұмыстардың құнын бірнеше есе төмендетеді [140].

CAD деректерін тексеру процесі әртүрлі жабық (RVT, DWG, DGN, NWS және т.б.) немесе ашық жартылай құрылымдық және параметрлік форматтардан (IFC, CPXML, USD) деректерді жүктеуді (ETL кезеңі Extract) қамтиды, мұнда әр атрибутқа және оның мәндеріне ережелер кестелері (Transform кезеңі) қолданылуы мүмкін, бұл үшін регулярлы өрнектер (RegEx) пайдаланылады (Рис. 7.38), бұл процесс кітаптың төртінші бөлімінде егжей-тегжейлі қарастырылған. -

Қателер туралы есепті PDF форматында және сәтті тексерілген жазбаларды шығару (Load кезеңі) тексерілген объектілерді ғана қамтитын құрылымдық форматтарға шығумен аяқталуы тиіс, оларды одан әрі процестер үшін пайдалануға болады.

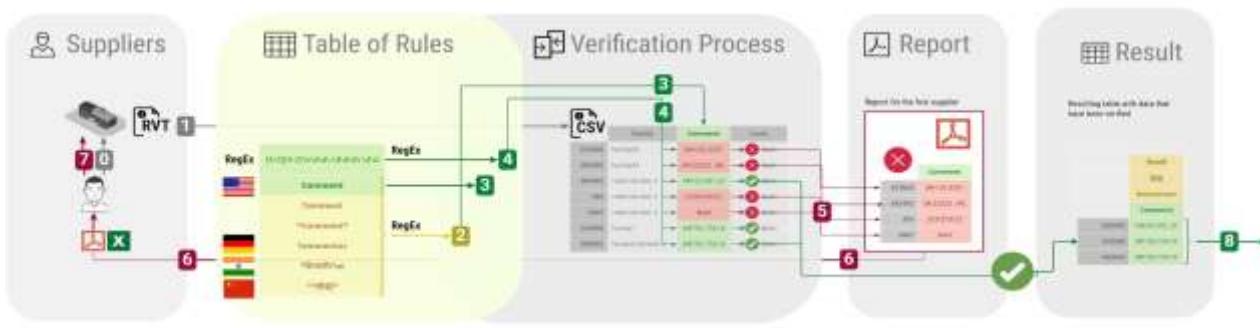


Рис. 7.38 Жобалық деректерді жеткізушілерден соңғы есепке дейін деректерді тексеру процесі, регулярлары өрнектер арқылы тексерілген.

CAD (BIM) жүйелерінен алынған деректерді тексеруді автоматтандыру, құрылымдық талаптар болған жағдайда және жаңа деректердің ағынды түрде келуі кезінде, ETL-Pipelines арқылы өндеді (Рис. 7.39), бұл валидация процесінде қолмен араласуды азайтады (әрбір тексеру және деректерге талаптарды құру процесі алдыңғы тарауларда қарастырылған).

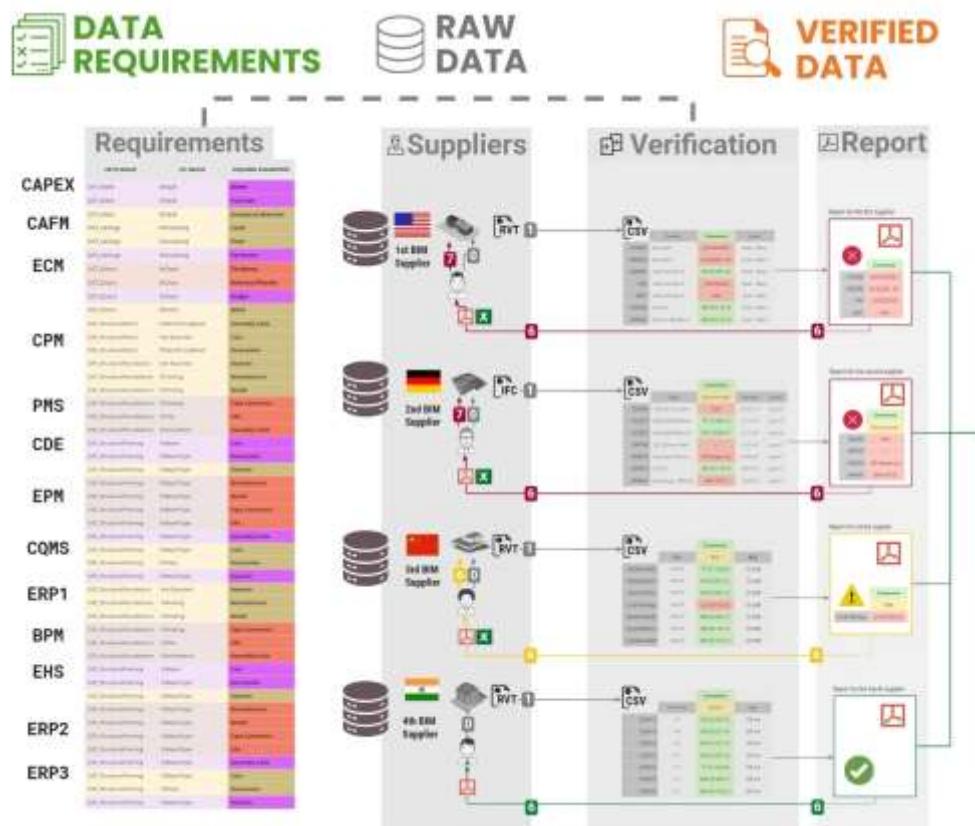


Рис. 7.39 ETL арқылы деректерді тексеруді автоматтандыру құрылымынан жобаларын басқаруды жеңілдетеді, процестерді жеделдету арқасында.

Дәстүрлі түрде мердігерлер мен CAD (BIM) мамандары ұсынған модельдерді тексеру бірнеше күннен бірнеше аптаға дейін созылуы мүмкін. Алайда автоматтандырылған ETL процестерін енгізу

арқылы бұл мерзімді бірнеше минутқа қысқартуға болады. Стандартты жағдайда мердігер: «Модель тексерілді және талаптарға сәйкес келеді» деп мәлімдейді. Мұндай мәлімдеме мердігердің деректер сапасы туралы мәлімдемесін тексеру тізбегін іске қосады:

- Жоба менеджері – «Мердігер: «Модель тексерілді, бәрі дұрыс» деп мәлімдейді».
- Деректер менеджері – «Валидацияны жүктейміз»:
- Pandas-те қарапайым скрипт секундына бұзушылықтарды анықтайды. Автоматтандыру дауларды жояды:
- Санат: OST_StructuralColumns, Параметр: FireRating IS NULL.
- Бұзушылық ID тізімін генерациялаймыз → Excel/PDF-ке экспорттаймыз.

Pandas-те қарапайым скрипт секундына бұзушылықтарды анықтайды:

```
df = model_data[model_data["Category"] == "OST_StructuralColumns"] # Фильтрация
issues = df[df["FireRating"].isnull()] # Бос мәндер
issues[["ElementID"]].to_excel("fire_rating_issues.xlsx") # ID экспорттау
```

- Деректер менеджері жоба менеджеріне – «Тексеру 18 бағанның FireRating параметрі толтырылмағанын көрсетті».
- Жоба менеджері мердігерге – «Модель қайта өңдеуге қайтарылады: FireRating параметрі міндетті, оның болмауы қабылдауды мүмкін емес етеді»

Нәтижесінде CAD моделі сапа тексерісінен өтпейді, автоматтандыру дауларды жояды, ал мердігер дерлік бірден проблемалық элементтердің ID тізімімен құрылымдалған есеп алады. Осылайша валидация процесі ашық, қайталанатын және адам факторының әсерінен қорғалған болады (Рис. 7.310).

Мұндай тәсіл деректерді тексеру процесін инженерлік функцияға айналдырады, ал қолмен сапаны бақылау емес. Бұл тек өнімділікті арттырып қана қоймай, компанияның барлық жобаларына бірдей логиканы қолдануға мүмкіндік береді, жобалаудан бастап пайдалану кезеңіне дейінгі процестердің цифрлық трансформациясын қамтамасыз етеді.

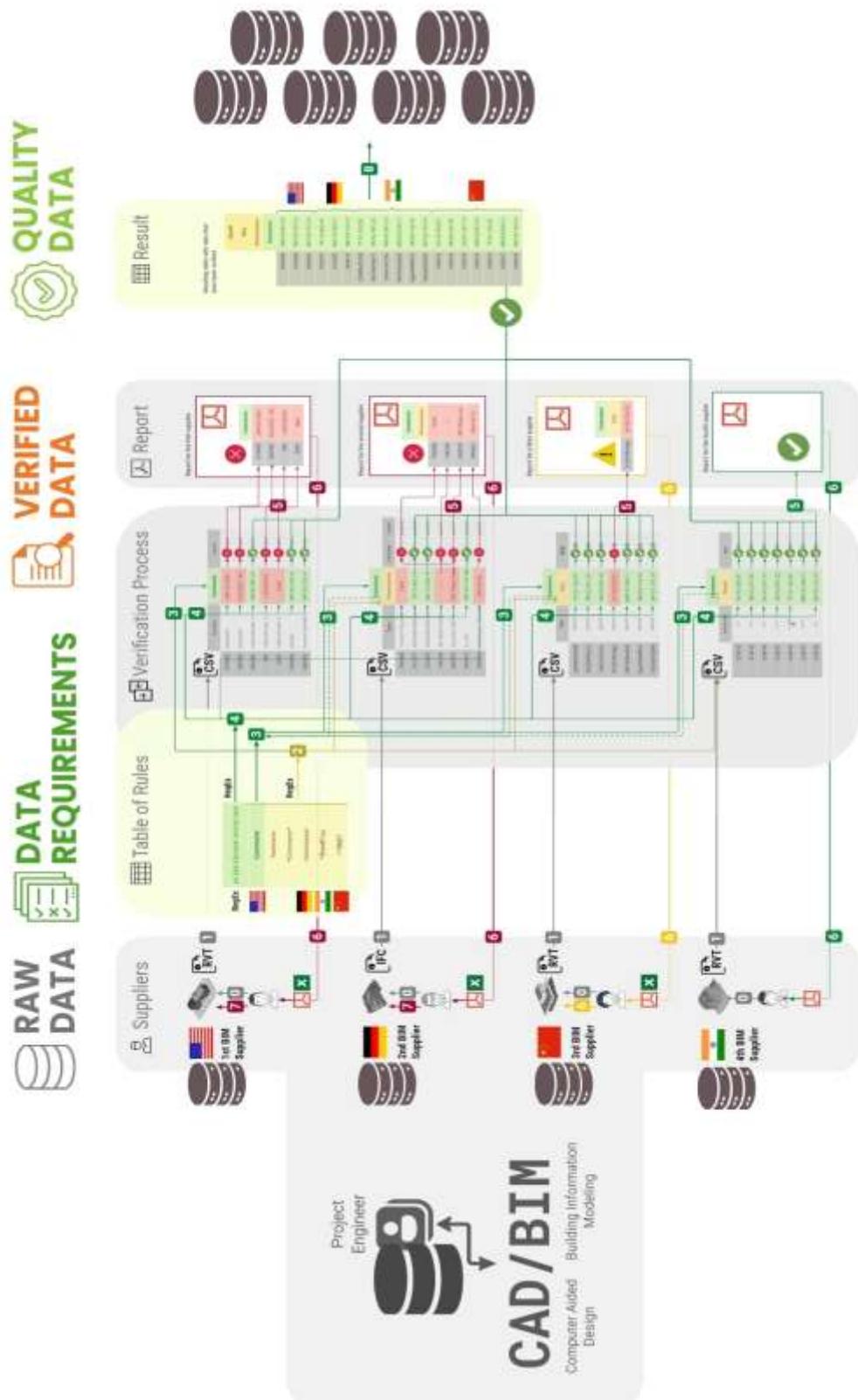


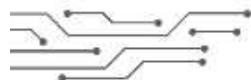
Рис. 7.310 Элементтердің атрибуттарын автоматтандырылған тексеру адам факторының әсерін жояды және қателіктер ықтималдығын төмендетеді.

Автоматтандырылған конвейерлерді қолдану арқылы (Рис. 7.310) CAD- (BIM-) жүйелерінен сапалы деректерді күтетін жүйе пайдаланушылары қажетті шығу деректерін - кестелер, құжаттар, суреттер - бірден ала алады және оларды жұмыс тапсырмаларына тез интеграциялай алады.

Бақылау, өңдеу және талдауды автоматтандыру құрылымын басқару тәсілдеріне, әсіресе әртүрлі жүйелердің өзара әрекеттесуіне, құрделі және қымбат модульдік меншікті жүйелерді немесе жабдықтаушылардың жабық шешімдерін пайдаланбай, өзгерістер енгізеді.

Концепциялар мен маркетингтік аббревиатураалар келіп-кетсе де, деректер талаптарын тексеру процестері бизнес-процестердің ажырамас бөлігі болып қала береді. Жаңа арнайы форматтар мен стандарттарды жасауға тырысадының орнына, құрылымы саласы басқа экономикалық салаларда тиімділігін дәлелдеген құралдарға назар аударуы керек. Бұғаңға күні деректерді өңдеуді автоматтандыру және процестерді интеграциялау үшін қуатты платформалар бар, олар компанияларға рутиндік операцияларға жұмсалатын уақытты едәуір қысқартуға және Extract, Transform және Load процестеріндегі қателіктерді минимизациялауға мүмкіндік береді.

ETL процестерін автоматтандыру және оркестрациялау бойынша танымал шешімдердің бірі Apache Airflow болып табылады, ол құрделі есептеу процестерін үйімдастыруға және ETL конвейерлерін басқаруға мүмкіндік береді. Airflow-мен қатар, деректерді маршрутизациялау және ағынды өңдеу үшін Apache NiFi, сондай-ақ бизнес процестерін автоматтандыру үшін n8n сияқты басқа да шешімдер белсенді түрде қолданылады.



ТАРАУ 7.4. ETL ЖӘНЕ ЖҰМЫС ПРОЦЕСТЕРІН ОРКЕСТРАЦИЯЛАУ: ПРАКТИКАЛЫҚ ШЕШІМДЕР

DAG ЖӘНЕ Apache Airflow: ЖҰМЫС ПРОЦЕСТЕРІН АВТОМАТТАНДЫРУ ЖӘНЕ ОРКЕСТРАЦИЯЛАУ

Apache Airflow – ETL конвейерлерін автоматтандыру, оркестрациялау және мониторинг жүргізу үшін арналған тегін ашық кодты платформа.

Күн сайын үлкен деректермен жұмыс істегендегі:

- Эртүрлі көздерден файлдарды жүктеу - Extract (мысалы, жеткізуашілерден немесе клиенттерден).
- Бұл деректерді қажетті форматқа түрлендіру - Transform (құрылымдау, тазалау және тексеру).
- Нәтижелерді тексеруге жіберу және есептер жасау - Load (қажетті жүйелерге, құжаттарға, деректер базаларына немесе дашбордтарға шығару).

Мұндай ETL процесстерін қолмен орындау едәуір уақытты алады және адам факторы байланысты қателіктерге әкеледі. Деректер көзіне өзгеріс енгізу немесе бір кезеңдегі ақау кідірістер мен дұрыс емес нәтижелерге себеп болуы мүмкін.

Apache Airflow сияқты автоматтандыру құралдары сенімді ETL конвейерін құруға, қателіктерді минимизациялауға, деректерді өңдеу уақытын қысқартуға және олардың әр кезеңдегі дұрыстығын қамтамасыз етуге мүмкіндік береді. Apache Airflow негізінде DAG (Directed Acyclic Graph) – бағытталған ацикличикалық граф концепциясы жатыр, онда әрбір тапсырма (оператор) басқа тәуелділіктермен байланысты және нақты тәртіпте орындалады. DAG циклдарды жояды, бұл тапсырмаларды орындаудың логикалық және болжамды құрылымын қамтамасыз етеді.

Airflow тәуелділіктерді басқару, орындау кестесін бақылау, күйді қадағалау және ақауларға автоматты жауап беру сияқты оркестрацияны өзіне алады. Мұндай тәсіл қолмен араласуды минимизациялайды және бүкіл процестің сенімділігін қамтамасыз етеді.

Тапсырмаларды оркестрациялаушы – құрделі есептей және ақпараттық ортада тапсырмаларды орындауды басқару және бақылау үшін арналған құрал немесе жүйе. Ол тапсырмаларды орналастыру, автоматтандыру және орындауды басқару процесін жеңілдетеді, бұл жұмыс тиімділігін арттыруға және ресурстарды оңтайландауға мүмкіндік береді.

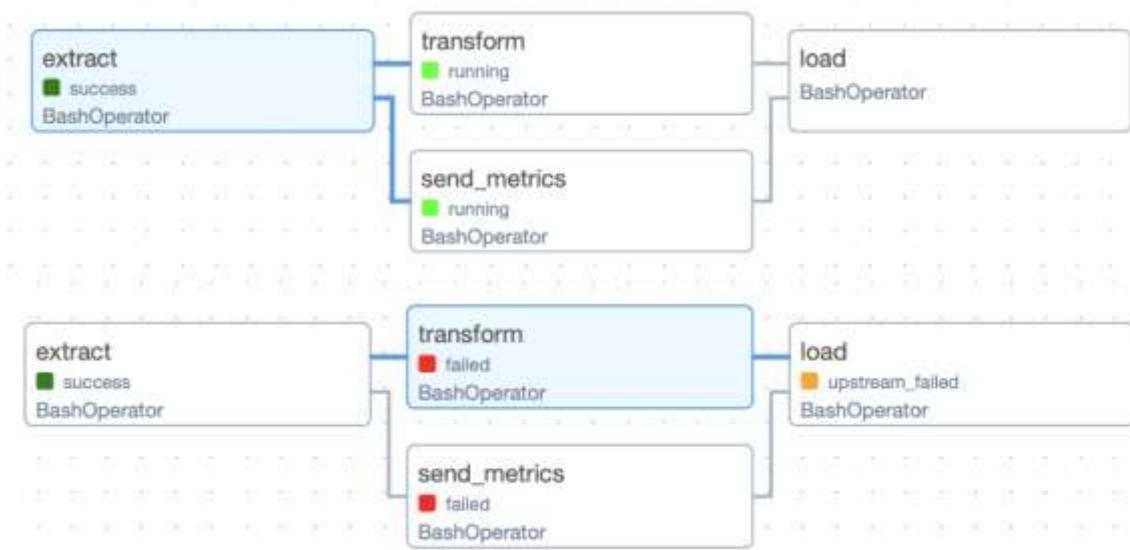


Рис. 7.41 Apache Airflow пайдалануышы интерфейсінде DAG-ETL визуализациялауға, орындау журналдарын қараяға, тапсырманың іске қосылу күйін көрге және тағы басқа мүмкіндіктерді ұсынады.

Airflow дегенде Distributed computing, деректерді өндеде, ETL (Extract, Transform, Load) процесстерін басқару, тапсырмаларды жоспарлау және деректермен жұмыс істеудің басқа сценарийлері үшін оркестрациялау және автоматтандыру үшін кеңінен қолданылады. Әдеткі бойынша Apache Airflow SQLite-ті деректер базасы ретінде пайдаланады.

ETL-ге үқсас қаралапайым DAG мысалы тапсырмаларды – Extract, Transform және Load-ты қамтиды. Пайдалануушы интерфейсі арқылы басқарылатын графта (Рис. 7.41) тапсырмалардың (код фрагменттерінің) орындалу тәртібі анықталған: мысалы, алдымен extract орындалады, содан кейін transform (және sending_metrics), ал жұмыс аяқталады load тапсырмасымен. Барлық тапсырмалар орындалғаннан кейін, деректерді жүктөу процесі сәтті деп есептеледі.-

Apache Airflow: ETL АВТОМАТТАНДЫРУ БОЙЫНША ПРАКТИКАЛЫҚ ҚОЛДАНУ

Apache Airflow деректерді өндеде процесстерін үйімдастыру үшін кеңінен қолданылады, бұл икемді ETL-конвейерлерін құруға мүмкіндік береді. Apache Airflow веб-интерфейс арқылы да, Python-код арқылы да іске қосылуы мүмкін. Веб-интерфейсте әкімшілер мен әзірлеушілер DAG-тарды визуалды түрде бақылап, тапсырмаларды іске қосып, орындалу нәтижелерін талдай алады.--

DAG пайдалану арқылы тапсырмалардың нақты орындалу ретін анықтауға, олардың арасындағы тәуелділіктерді басқаруға және бастапқы деректердегі өзгерістерге автоматты түрде жауап беруге болады. Airflow-ты есептілікті автоматтандыру үшін пайдалану мысалын қарастырайық.-



Рис. 7.42 Apache Airflow пайдалану арқылы деректерді өңдеу үшін ETL-конвейерінің концепциясы.

Бұл мысалда (Рис. 7.42) ETL-конвейерінің шеңберінде негізгі тапсырмаларды орындайтын DAG

қарастырылған: -

- Excel-файлдарын оқу (Extract):• Белгіленген директориядағы барлық файлдарды ретті түрде қарай. • Эр файлдан деректерді pandas кітапханасын пайдаланып оқу. • Барлық деректерді бір DataFrame-ге біріктіру.
- PDF-документ жасау (Transform):• Біріктірілген DataFrame-ді HTML-кестеге түрлендіру. • Кестені PDF форматында сақтау (демонстрациялық нұсқада – HTML арқылы).
- Есепті электрондық пошта арқылы жіберу (Load):• PDF-документті электрондық пошта арқылы жіберу үшін EmailOperator-ды қолдану.
- DAG-ты баптау:• Тапсырмалардың орындалу ретін анықтау: деректерді шығару → есепті қалыптастыру → жіберу.
• Жоспарланған іске қосу кестесін тағайындау (@monthly – әр айдың бірінші күні).
- Автоматты ETL мысалында (Рис. 7.42) Excel-файлдарынан деректер жинау, PDF-документ жасау және оны электрондық пошта арқылы жіберу көрсетілген. Бұл Airflow-ты пайдаланудың көптеген мүмкін сценарийлерінің бірі ғана. Бұл мысалды нақты тапсырмаларға бейімдеуге болады, деректерді өндеду процестерін жеңілдету және автоматтандыру үшін.-

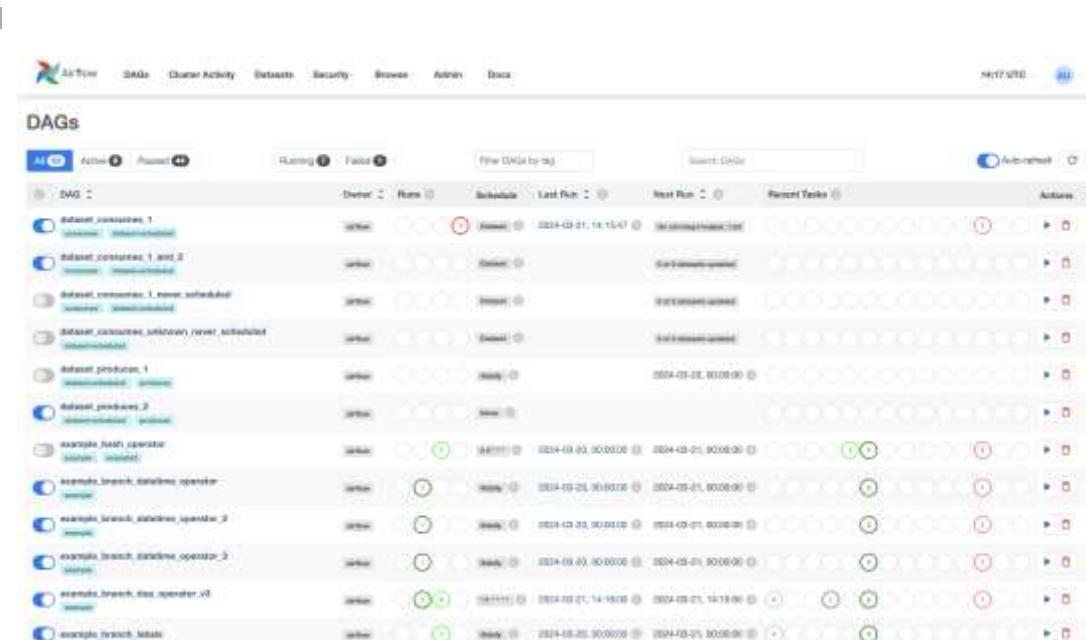


Рис. 7.43 Соңғы іске қосу туралы ақпаратпен DAG топтарының шолуы.

- Apache Airflow веб-интерфейсі (Рис. 7.43) деректер жұмыс процестерін басқару үшін кешенді визуалды орта ұсынады. Ол DAG-тарды интерактивті графтар түрінде көрсетеді, мұнда түйіндер тапсырмаларды, ал қабырғалар олардың арасындағы тәуелділіктерді

білдіреді, бұл курделі деректер өңдеу процестерін оңай бақылауға мүмкіндік береді. Интерфейс тапсырмалардың орындалу күйі, іске қосу тарихы, егжей-тегжейлі журналдар және өнімділік метрикалары туралы ақпаратты қамтитын мониторинг панелін қамтиды. Әкімшілер тапсырмаларды қолмен іске қосып, сәтсіз операцияларды қайта іске қосып, DAG-тарды тоқтатып, орта айнымалыларын баптай алады – мұның бәрі интуитивті түсінікті пайдаланушы интерфейсі арқылы.

Мұндай архитектура деректерді тексеру, орындалу күйі туралы хабарламалар, сыртқы API немесе деректер базаларымен интеграциямен толықтырылуы мүмкін. Airflow DAG-ты икемді түрде бейімдеуге мүмкіндік береді: жаңа тапсырмаларды қосу, олардың ретін өзгерту, тізбектерді біріктіру – бұл оны курделі деректер өңдеу процестерін автоматтандыру үшін тиімді құрал етеді. Airflow веб-интерфейсінде DAG іске қосылған кезде (Рис. 7.43, Рис. 7.44) тапсырмалардың орындалу күйін бақылауға болады. Жүйе түс кодын пайдаланады:–

- Жасыл – тапсырма сәтті орындалды.
- Сары – процесс орындалып жатыр.
- Қызыл – тапсырманы орындау кезінде қате.

Қателіктер болған жағдайда (мысалы, файл жоқ немесе деректер құрылымы бұзылған) жүйе автоматты түрде хабарлама жіберуді бастайды.

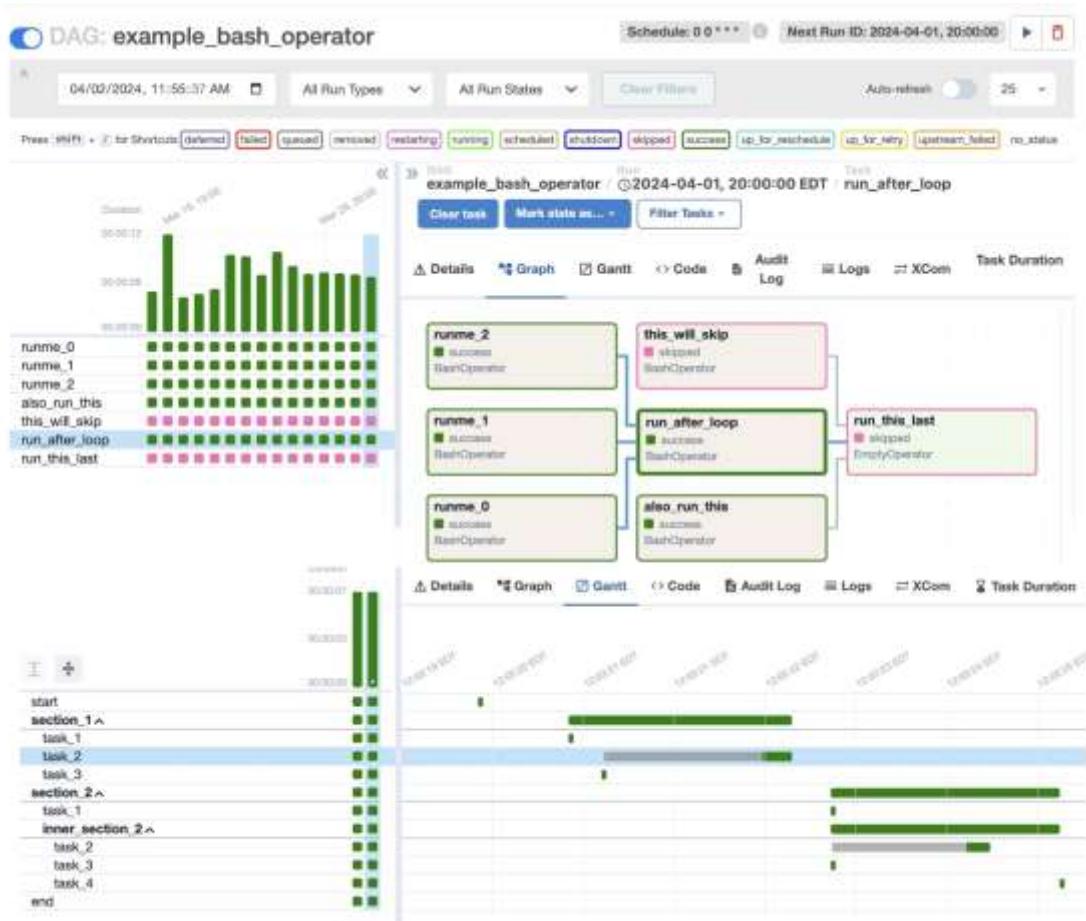


Рис. 7.44 Apache Airflow проблемаларды диагностикалауды, процестерді оңтайландыруды және командалардың күрделі деректер өндеу пайплайндарымен бірлесіп жұмыс істеуін айтартықтай жөнілдетеді.

Apache Airflow рутиналық тапсырмаларды автоматтандыру арқылы оларды қолмен орындау қажеттілігін жояды. Ол процестердің орындалуын мониторингтеу және қателер туралы жедел хабарлау арқылы сенімділікті қамтамасыз етеді. Жүйенің икемділігі жаңа тапсырмаларды оңай қосуға немесе бар тапсырмаларды өзгертуге мүмкіндік береді, жұмыс процестерін өзгөрмелі талаптарға бейімдейді.

Apache Airflow-қа қосымша, жұмыс процестерін оркестрациялауға арналған үқсас құралдар бар. Мысалы, ашық және тегін Prefect (Рис. 7.35) қарапайым синтаксис ұсынады және Python-мен жақсы интеграцияланады, ал Spotify әзірлеген Luigi үқсас функционалдылықты қамтамасыз етеді және үлкен деректермен жақсы жұмыс істейді. Сондай-ақ, модульдік және масштабталуға бағытталған заманауи тәсілдерді ұсыннатын Kronos және Dagster-ді атап өткен жөн. Тапсырмаларды оркестрациялау құралын таңдау жобаның нақты қажеттіліктеріне байланысты, бірақ олардың барлығы күрделі ETL деректер өндеу процестерін автоматтандыруға көмектеседі.-

Ерекше атап өтуді қажет ететін Apache NiFi – әртүрлі жүйелер арасында деректерді ағынды (streaming) өндеу және маршрутизациялау үшін арналған ашық кодты платформа. Airflow-тан айырмашылығы, ол пакетпен өндеуге және тәуелділіктерді басқаруға бағытталған, NiFi нақты

уақыт режимінде, деректерді өңдеуді және жүйелер арасында икемді маршрутизацияны қамтамасыз етеді.

Apache NiFi ДЕРЕКТЕРДІ ЖОЛДАНДЫРУ ЖӘНЕ ТҮРЛЕНДІРУ ҮШІН

Apache NiFi – әртүрлі жүйелер арасында деректер ағындарын автоматтандыруға арналған қуатты ашық кодты платформа. 2006 жылы АҚШ Үлттық қауіпсіздік агенттігі (NSA) үшін "Niagara Files" деген атпен ішкі қажеттіліктер үшін әзірленген. 2014 жылы жоба ашылып, Apache Software Foundation-ға берілді, олардың технологияларды беру бастамасының бір бөлігі болды [141].

Apache NiFi нақты уақыт режимінде деректерді жинау, өңдеу және беру үшін жобаланған. Airflowтан айырмашылығы, ол пакетпен тапсырмалармен жұмыс істейді және нақты кестелерді талап етеді, NiFi ағынды өңдеу режимінде жұмыс істейді, әртүрлі қызметтер арасында деректерді үздіксіз беруге мүмкіндік береді.

Apache NiFi IoT құрылғыларымен, құрылыш объектілерінің сенсорларымен, мониторинг жүйелерімен интеграциялау үшін тамаша. Мысалы, серверде CAD форматтарын ағынды тексеру кезінде деректердегі өзгерістерге дереу жауап беру қажет болуы мүмкін.

NiFi-дің кіріктірілген фильтрация, трансформация және маршрутизация құралдарының арқасында деректерді стандарттау (Transform кезеңі) және оларды сақтау немесе аналитикалық жүйелерге беру (Load) мүмкіндігі бар. Оның басты артықшылықтарының бірі – кіріктірілген қауіпсіздік және қол жеткізу бақылауы, бұл оны конфиденциалды ақпаратты өңдеу үшін сенімді шешім етеді.

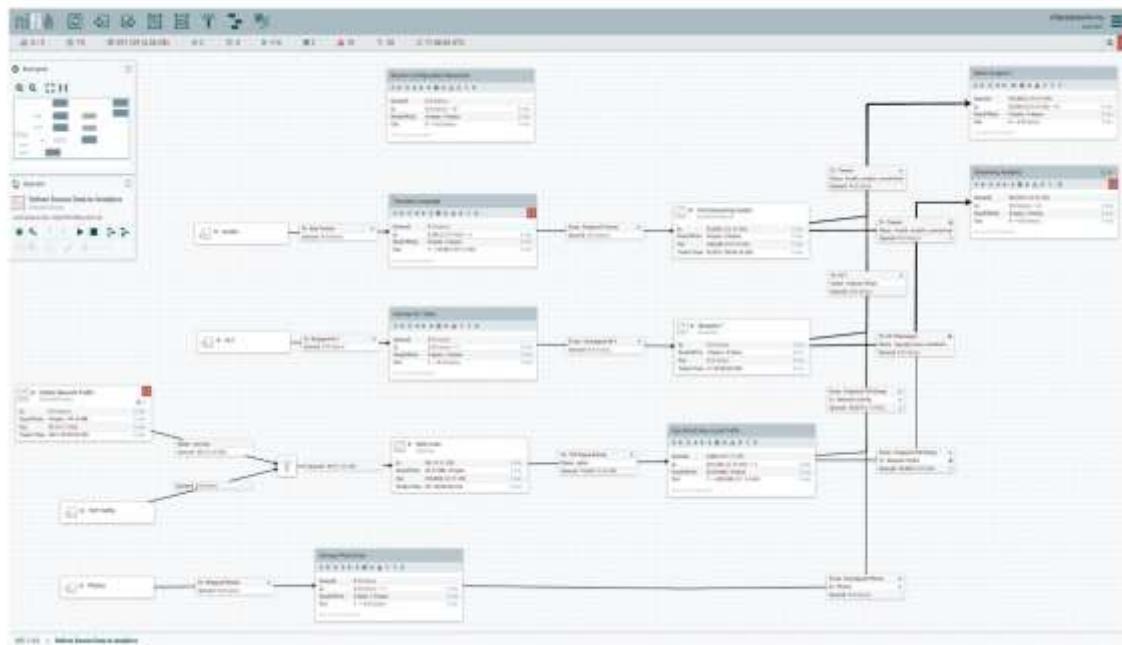


Рис. 7.45 Apache NiFi интерфейсіндегі деректер ағынның графикалық көрінісі.

Apache NiFi нақты уақыт режимінде деректерді ағынды беру, фильтрациялау және маршрутизациялау міндеттерін тиімді түрде шешеді. Ол жүйелер арасында ақпараттың тұрақты берілуі мен жоғары өткізу қабілеті маңызды техникалық сценарийлер үшін тамаша.

Дегенмен, негізгі мақсат әртүрлі қызметтерді интеграциялау, рутиналық операцияларды автоматтандыру және терең бағдарламалау білімінсіз жұмыс процесстерін жылдам баптау болған жағдайда, кіру шегі тәмен және максималды икемділікке ие шешімдер сұранысқа ие. Мұндай құралдардың бірі n8n – бизнес-автоматтандыру мен визуалды оркестрацияга бағытталған Low-Code/No-Code платформасы.

n8n ТӨМЕН КОД, НОЛЬ КОД ПРОЦЕСТЕРДІ ОРКЕСТРАЦИЯЛАУ

n8n – автоматтандырылған жұмыс процесстерін құруға арналған Open Source Low-Code / No-Code платформа, қолданудың қарапайымдылығымен, икемділігімен және сыртқы қызметтермен жылдам интеграциялау мүмкіндігімен ерекшеленеді.

No-Code – код жазбай цифрлық өнімдерді жасау әдісі. Процестің барлық элементтері – логикадан бастап интерфейске дейін – тек визуалды құралдар арқылы жүзеге асырылады. No-Code платформалары техникалық дайындықсыз пайдаланушыларға бағытталған және автоматтандыру, формалар, интеграциялар және веб-қосымшаларды жылдам жасауға мүмкіндік береді. Мысал: пайдаланущы автоматты түрде хабарламаларды жіберуді немесе Google Sheets-пен интеграцияны drag-and-drop интерфейсі арқылы бағдарламалау білімінсіз баптайды.

Ашық бастапқы код пен жергілікті орналастыру мүмкіндігі арқасында n8n автоматтандыру және ETL Pipelines құру процесстерінде компанияларға өз деректеріне толық бақылау береді, қауіпсіздікте қамтамасыз етеді және бұлттық провайдерлерден тәуелсіздік береді.

Python білімін талап ететін қатал оркестрациямен есептеу тапсырмаларына бағытталған Apache Airflow-тан айырмашылығы, n8n бағдарламалау тілдерін білмей сценарийлерді құруға мүмкіндік беретін визуалды редакторды ұсынады. Оның интерфейсі код жазбай автоматтандырылған процесстерді жасауға мүмкіндік берсе де, қурделі сценарийлерде пайдаланушылар өздерінің JavaScript және Python функцияларын қосып, мүмкіндіктерін кеңейте алады.-

Low-Code – бағдарламалық қамтамасыз етуді әзірлеу тәсілі, мұнда қосымша немесе процестің негізгі логикасы графикалық интерфейс пен визуалды элементтерді пайдалана отырып жасалады, ал бағдарламалық код тек функционалдылықты баптау немесе кеңейту үшін қолданылады. Low-Code платформалары шешімдерді әзірлеуді едәуір жылдамдатуға мүмкіндік береді, бұл процесс тек бағдарламашыларды ғана емес, базалық техникалық дағдылары бар бизнес-пайдаланушыларды да қамтиды. Мысал: пайдаланущы дайын блоктардан бизнес-процесті жинап, қажет болған жағдайда JavaScript немесе Python-да өз скриптін қосуға болады.

n8n кіру шегі тәмен платформа ретінде позицияланса да, құрделі автоматтандыру сценарийлерін жасау үшін бағдарламалау негіздері, веб-технологияларды түсіну және API-мен жұмыс істеу дағдылары пайдалы. Жүйенің икемділігі оны автоматтандырылған деректерді өңдеуден бастап,

мессенджерлермен, IoT құрылғыларымен және бұлттық қызметтермен интеграциялауға дейінгі кең ауқымды тапсырмаларға бейімдеуге мүмкіндік береді.

n8n-ді пайдаланудың негізгі ерекшеліктері мен артықшылықтары:

- Ашық бастапқы код пен жергілікті орналастыру мүмкіндігі деректерге толық бақылауды, қауіпсіздік талаптарына сәйкестікті және бұлттық провайдерлерден тәуелсіздікті қамта-масыз етеді.
- CRM, ERP, электрондық коммерция, бұлттық платформалар, мессенджерлер және дерек-тер базалары сияқты 330-дан астам қызметпен интеграция.
- Сценарийлердің икемділігі: қарапайым хабарламалардан бастап API сұрауларын өңдеу, шешім қабылдау логикасы және ИИ қызметтерін қосу арқылы күрделі тізбектерге дейін.
- JavaScript және Python қолдауы: қажет болған жағдайда пайдаланушылар автоматизация мүмкіндіктерін кеңейту үшін пайдаланушы кодын енгізе алады.
- Интуитивті визуалды интерфейс: процестің барлық кезеңдерін жылдам баптауға және ви-зуализациялауға мүмкіндік береді.

Low-Code класындағы платформалар минималды код көлемімен цифрлық шешімдер жасау құралдарын ұсынады, бұл терең техникалық сараптамасы жоқ, бірақ процестерді автоматизациялауға мүқтаж командалар үшін ете қолайлы.

Құрылыш саласында n8n әртүрлі процестерді автоматизациялау үшін пайдаланылуы мүмкін, мысалы, жобаларды басқару жүйелерімен интеграция, ағынды тексеру, дайын есептер мен хаттарды жазу, материалдар қорын автоматтый түрде жаңарту, командаларға тапсырмалардың мәртебесі туралы хабарламаларды жіберу және тағы басқалар. n8n-де бапталған Pipeline қолмен операциялардың санын бірнеше есе азайтуға, қателіктер ықтималдығын төмендетуге және жобаларды орындау үшін шешім қабылдауды жеделдетуге мүмкіндік береді.

Сіз n8n.io/workflows сайтында қолжетімді, тегін және ашық екі мыңдан астам дайын n8n Pipeline-ды таңдай аласыз, бұл құрылыш саласындағы жұмыс процестерін және жеке тапсырмаларды автоматизациялауға, рутиналық операцияларды азайтуға көмектеседі.

n8n.io сайтында қолжетімді тегін Pipeline шаблондарының бірін алайық, ол Gmail-де жауаптардың, черновиктерін автоматтый түрде жасайды, бұл көп көлемдегі хат алатын немесе жауап жазуда қындық қөретін пайдаланушыларға көмектеседі.-

n8n шаблоны "Gmail AI Авто жауап беруші: Kіріс хаттарға жауап черновиктерін жаса" (Рис. 7.46) кіріс хабарламаларды ChatGPT-тің LLM көмегімен талдайды, жауап беру қажеттілігін анықтайды, ChatGPT көмегімен черновикті қалыптастырады және мәтінді HTML форматына түрлендіріп, Gmail-де хабарламалар тізбегіне қосады. Бұл жағдайда хат автоматтый түрде жіберілмейді, бұл жауапты қолмен өңдеуге және бекітуге мүмкіндік береді. Баптау шамамен 10 минутты алады және Gmail API-дің OAuth конфигурациясын және OpenAI API интеграциясын қамтиды. Нәтижесінде, хаттардың мазмұнын бақылаусыз автоматизациялауға арналған ыңғайлы және тегін шешім алынады.

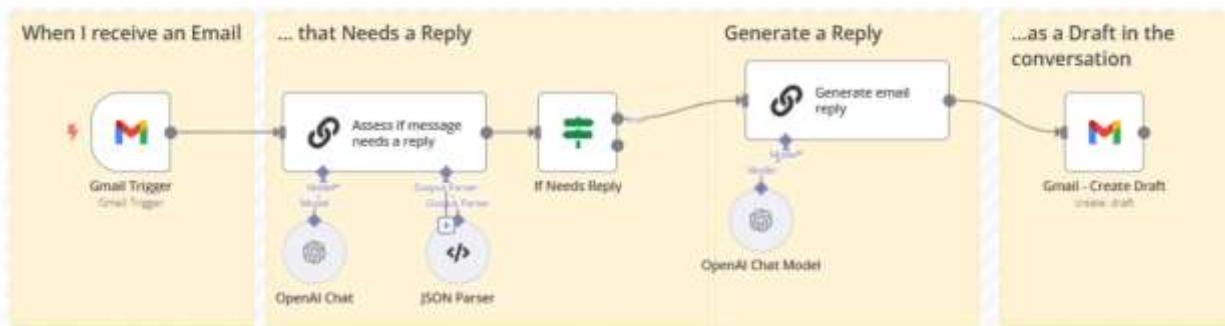


Рис. 7.46 n8n көмегімен электрондық хаттарға жауаптарды автоматты түрде генерациялау процесі.

n8n-мен автоматизацияның тағы бір мысалы – жылжымайтын мүлік нарығындағы тиімді мәмілелерді іздеу. N8n Pipeline “Zillow API, Google Sheets және Gmail көмегімен жылжымайтын мүлік мәмілелерін автоматизациялау” күн сайын берілген критерийлерге сәйкес келетін ағымдағы ұсыныстарды жинайды, Zillow API-ды пайдалана отырып. Ол автоматты түрде негізгі инвестициялық метрикаларды (Cash on Cash ROI, Monthly Cash Flow, Down Payment) есептейді, Google Sheets-ті жаңартады және электрондық поштаға қорытынды есеп жібереді (Рис. 7.47), бұл инвесторларға уақытты үнемдеуге және ең жақсы ұсыныстарға тез жауап беруге мүмкіндік береді.-

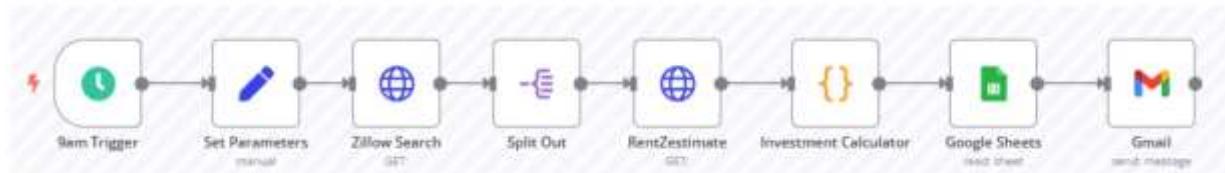


Рис. 7.47 Жылжымайтын мүліктің инвестициялық тартымдылығын бағалаудың автоматтандырылған процесі.

n8n-нің икемділігі мен кеңейтілуі арқасында, n8n цифрлық трансформацияға және нарықтағы бәсекеге қабілеттілікті арттыруға ұмытылатын компаниялар үшін құнды құралға айналады, салыстырмалы түрде қарапайым және тегін ашық кодты құралдармен.

Apache NiFi, Airflow және n8n құралдарын деректерді өңдеудің үш деңгейі ретінде қарастыруға болады (Рис. 7.48). NiFi деректер ағынын басқара отырып, олардың жеткізілуі мен трансформациясын қамтамасыз етеді, Airflow тапсырмалардың орындалуын оркестрлейді, деректерді өңдеу конвейерлеріне біркітіреді, ал n8n сыртқы қызметтермен интеграцияны автоматтандырады және бизнес-логиканы басқарады. -




	The main task	Approach
Apache NiFi	Streaming and data transformation	Real-time stream processing
Apache Airflow	Task orchestration, ETL pipelines	Batch planning, DAG processes
n8n	Integration, automation of business logic	Low-code visual orchestration

Рис. 7.48 Apache Airflow, Apache NiFi және n8n қазіргі заманғы деректерді басқару архитектурасының үш өзара толықтыратын деңгейі ретінде қарастырылуы мүмкін.

Бірлесіп, бұл тегін және ашық құралдар құрылыс индустриясында деректер мен процестерді басқарудың тиімді экожүйесін қалыптастыруы мүмкін, компанияларға ақпаратты тиімді пайдалану арқылы шешім қабылдауға және процестерді автоматтандыруға мүмкіндік береді.

КЕЛЕСІ ҚАДАМДАР: ҚОЛДАН ОПЕРАЦИЯЛАРДАН АНАЛИТИКАЛЫҚ ШЕШІМДЕРГЕ ӨТҮ

Қазіргі заманғы құрылыс компаниялары жоғары белгісіздік жағдайында жұмыс істейді: материалдар бағасының өзгеруі, жеткізілімдердің кешігі, жұмыс күшінің жетіспеушілігі және жобалардың қатаң мерзімдері. Аналитикалық дэшбордтар, ETL-конвейерлері және BI-жүйелерін пайдалану компанияларға проблемалық аймақтарды тез анықтауға, ресурстардың тиімділігін бағалауға және өзгерістерді қаржылық шығындарға әкелмей тұрып болжауға көмектеседі.

Осы бөлімнің қорытындысында, қарастырылған технологияларды құнделікті тапсырмаларда қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

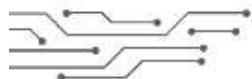
- Деректерді визуализациялау және аналитикалық панельдерді енгізіңіз
- Негізгі көрсеткіштерді (KPI) мониторингтеу үшін ақпараттық панельдер жасау процесін менгеріңіз
- Деректеріңіз үшін визуализация құралдарын (Power BI, Tableau, Matplotlib, Plotly) пайдаланыңыз
- Деректерді ETL процестері арқылы өңдеуді автоматтандырыңыз
- ETL процестері арқылы әртүрлі көздерден (құжаттар, кестелер, CAD) деректерді автоматтарты түрде жинауды үйімдастырыңыз
- Деректерді трансформациялауды (мысалы, регулярлы өрнектер арқылы тексеру немесе есептеу) Python скриптерін қолдана отырып үйімдастырыңыз
- Excel файлдарынан деректерді пайдаланып немесе басқа PDF құжаттарынан ақпаратты шығарып, FPDF кітапханасы арқылы PDF (немесе DOC) форматында автоматтты есептерді

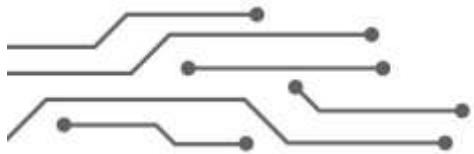
қалыптастыруды орнатуға тырысының

- Автоматтандыру үшін тілдік модельдерді (LLM) пайдаланыңыз
- Неструктуриленген құжаттардан деректерді шығарып, талдауға көмектесетін кодты генерациялау үшін үлкен тілдік модельдерді (LLM) пайдаланыңыз
- n8n автоматтандыру құралымен танысының және олардың сайтында дайын шаблондар мен кейстерді зерттеңіз. Сіздің жұмысыныңыздан қандай процесстерді No-Code/Low-Code тәсілімен толық автоматтандыруға болатынын анықтаңыз

Деректерге аналитикалық көзқарас пен процесстерді автоматтандыру тек рутиндік операцияларға жұмсалатын уақытты қысқартпайды, сонымен қатар қабылданатын шешімдердің сапасын арттырады. Визуалды аналитика құралдары мен ETL-конвейерлерін енгізетін компаниялар өзгерістерге жедел жауап беру мүмкіндігіне ие болады.

Бизнес-процесстерді автоматтандыру n8n, Airflow және NiFi сияқты құралдарды пайдалану арқылы цифрлық жетілуге алғашқы қадам болып табылады. Келесі кезең – автоматтандырудың негізінде жатқан деректерді сапалы сақтау және басқару. Сегізінші бөлімде құрылыш компоненттерінде құжаттар мен әртүрлі форматтағы файлдардың хаосынан орталықтандырылған қоймалар мен аналитикалық платформаларға қалай тұрақты деректер сақтау архитектурасын құра алатыны турали егжей-тегжейлі қарастырамыз.





VIII БӨЛІМ

ДЕРЕКТЕРДІ САҚТАУ ЖӘНЕ БАСҚАРУ САЛАСЫНДА

Сегізінші бөлім құрылым саласындағы заманауи деректерді сақтау және басқару технологияларын зерттейді. Мұнда үлкен көлемдегі ақпаратпен жұмыс істеуге арналған тиімді форматтар – қарапайым CSV және XLSX-тан бастап, өнімділігі жоғары Apache Parquet және ORC-ке дейін олардың мүмкіндіктері мен шектеулерін егжей-тегжейлі салыстыру жүргізіледі. Деректер қоймалары (DWH), деректер көлдері (Data Lakes) және олардың гибридтік шешімдері (Data Lakehouse) концепциялары, сондай-ақ деректерді басқару (Data Governance) және ақпаратпен жұмыс істеудегі минимализм (Data Minimalism) принциптері қарастырылады. "Деректер батпағы" (Data Swamp) мәселелері мен ақпараттық жүйелердегі хаостың алдын алу стратегиялары жан-жақты талқыланады. Деректермен жұмыс істеудің жаңа тәсілдері, соның ішінде векторлық деректер базалары және олардың құрылым саласындағы Bounding Box концепциясы арқылы қолданылуы ұсынылады. Бұл бөлімде DataOps және VectorOps методологиялары деректермен жұмыс процестерін ұйымдастырудың жаңа стандарттары ретінде де қарастырылады.

ТАРАУ 8.1. ДЕРЕКТЕР ИНФРАСТРУКТУРАСЫ: САҚТАУ ФОРМАТТАРЫНАН ЦИФРЛЫҚ ҚОЙМАЛАРҒА ДЕЙІН

ДЕРЕК АТОМДАРЫ: АҚПАРАТТЫ ТИІМДІ БАСҚАРУДЫҢ НЕГІЗІ

Әлемдегі барлық нәрсе ең кішкентай құрылыш блоктарынан - атомдар мен молекулалардан тұрады, және уақыт өте келе барлық тірі және өлі заттар міндепті түрде осы бастапқы күйіне оралады. Табиғатта бұл процесс таңғажайып жылдамдықпен жүреді, біз оны адам басқаруындағы процестерге көшіруге тырысамыз.

Орманда кез келген тірі организмдер уақыт өте келе жаңа өсімдіктердің негізі болып табылатын қоректік затқа айналады. Бұл өсімдіктер, өз кезегінде, миллиондаған жылдар бұрын ғаламды жасаған атомдардан тұратын жаңа тірі жаратылыстардың азығына айналады.

Бизнесте де күрделі көпқабатты құрылымдарды ең негізгі, минималды өндөлетін бірліктерге бөлу маңызды – табиғаттағы атомдар мен молекулаларға ұқсас. Бұл деректер атомдарын тиімді сақтау мен басқаруға мүмкіндік береді, оларды бай, құнарлы негізге айналдырып, аналитика мен шешім қабылдау сапасының өсуіне негізгі ресурсқа айналдырады.



Рис. 8.11 Талдау және шешім қабылдау бұрын өндөліп, сақталған қайта пайдаланылатын деректерге негізделеді.

Музыкалық композициялар ноталардан тұрады, олар бірігіп, күрделі музыкалық шығармаларды жасайды, ал сөздер примитивті бірліктен - дыбыстық әріптен құралады. Табиғат, ғылым, экономика, өнер немесе технология болсын, әлем өздігінен жою, құрылым, циклділік және құрылышқа

ұмтылуда таңғажайып бірлік пен гармонияны көрсетеді. Дәл сол сияқты, өзіндік құнды есептеу жүйелеріндегі процестер ең кішкентай құрылымдалған бірліктерге - ресурстардың баптарына - есептеулер мен кестелер деңгейінде бөлінеді. Содан кейін бұл бірліктер, ноталар сияқты, күрделі есептеулер мен графиктерді қалыптастыру үшін пайдаланылады. Автоматтандырылған жобалау жүйелері де дәл осындаи принциппен жұмыс істейді, онда күрделі архитектуралық және инженерлік жобалар базалық элементтерден - жеке элементтер мен кітапханалық компоненттерден - құралады, олардан күрделі ғимарат немесе құрылыс жобасының толық 3D моделі жасалады.

Табиғат пен ғылымға тән циклділік және құрылымдылық концепциясы қазіргі деректер әлемінде де көрініс табады. Табиғатта барлық тірі жәндіктер атомдар мен молекулаларға қайта оралатын болса, қазіргі деректер өңдеу құралдарында ақпарат ең прimitivtі формаға өтуді көздейді.

Ең кішкентай элементтер, олардың шектеулі бөлінбеушілігімен, бизнес-процестердің негізгі құрылышы блоктары болып табылады. Бастапқыдан-ақ, осы ең кішкентай құрылышы блоктарын әртүрлі көздерден жинау, құрылымдау (атомдарға бөлу) және сақтау тәсілдерін мұқият ойластыру маңызды. Деректерді үйімдастыру және сақтау - бұл тек олардың құрамдас бөліктерге бөлінуі мәселесі емес. Деректердің интеграциясы мен құрылымдалған сақталуын қамтамасыз ету де маңызды, соңдықтан деректер қажет болған кезде оңай алынып, талданып, шешім қабылдау үшін пайдаланылуы мүмкін.

Ақпаратты тиімді өңдеу үшін деректерді сақтау форматтары мен әдістерін мұқият таңдау қажет - бұл ағаштардың өсуі үшін топырақты дайындау сияқты. Деректер қоймалары ақпараттың жоғары сапасы мен өзектілігін қамтамасыз ету үшін үйімдастырылуы тиіс, артық немесе сәйкес келмейтін деректерді жою керек. Бұл «ақпараттық топырақ» неғұрлым жақсы құрылымдалған болса, пайдаланушылар қажетті деректерді тезірек және дәлірек таба алады және аналитикалық тапсырмаларды шешеді.

АҚПАРАТ САҚТАУ: ФАЙЛДАР НЕМЕСЕ ДЕРЕКТЕР

Деректер қоймалары компанияларға әртүрлі жүйелерден ақпарат жинап, біріктіруге мүмкіндік береді, осылайша кейінгі аналитика үшін біртұтас орталық құрады. Жиналған тарихи деректер процестерді тереңірек талдауға ғана емес, сонымен қатар бизнестің тиімділігіне әсер етуі мүмкін занұлыштарды анықтауға мүмкіндік береді.

Мысалы, компания бірнеше обьектілерді бір уақытта жүргізеді. Инженер қанша бетон құйылғанын және қанша көлемді сатып алу қажет екенін түсінгісі келеді. Дәстүрлі тәсілмен оған серверде бірнеше сметалық кестелерді ашып, оларды орындалған жұмыстар актілерімен салыстыру және ағымдағы қойма қалдықтарын тексеру үшін қолмен іздеу қажет болады. Бұл бірнеше сағат, тіпті күндер алуы мүмкін. ETL процестері мен автоматты скрипттер болған жағдайда да, тапсырма жартылай қолмен қалады: инженерге әлі де сервердегі папкаларға немесе нақты файлдарға жолды қолмен көрсету қажет. Бұл автоматтандырудың жалпы әсерін төмендетеді, себебі ол құнды жұмыс уақытын алып отырады.

Деректерді басқаруға көшу кезінде инженер сервердің файлдық жүйесімен жұмыс істеудің орнына, ақпаратты нақты уақыт режимінде жақартылатын біртұтас сақтау құрылымына қол жеткізеді. Бір сұрау – код, SQL-сұрау немесе тіпті LLM-агентіне жүгіну түрінде – алдын ала қайта дайындалған

және деректер қоймасында біріктірілген жағдайда, ағымдағы қалдықтар, орындалған жұмыстардың көлемі және алдағы жеткізілімдер туралы дәл деректерді бірден алуға мүмкіндік береді, мұнда папкалар бойынша шарлаудың, ондаған файлдарды ашудың және мәндерді қолмен сәйкестендірудің қажеті жоқ.

Ұзақ уақыт бойы құрылыш компаниялары PDF құжаттарын, DWG сызбаларын, RVT модельдерін және жүздеген, мындаған Excel кестелері мен басқа да шашыраңқы форматтарды пайдаланды, олар компания серверлерінде белгілі бір папкаларда сақталды, бұл ақпаратты іздеуді, тексеруді және талдауды қынданатты. Нәтижесінде, жобалар аяқталғаннан кейін қалған файлдар көбінесе серверге архивтік папка-сақтағыштарға қайта көшіріледі, олар кейіннен іс жүзінде пайдаланылмайды. Мұндай дәстүрлі файлдық деректерді сақтау, деректер ағынының артуымен, адам фактоның қателіктеріне осал болуына байланысты өзектілігін жоғалтады.

Файл – бұл деректерді сақтайтын оқшауланған контейнер. Файлдар адамдар үшін жасалады, жүйелер үшін емес, сондықтан оларды қолмен ашу, оқу және интерпретациялау қажет. Мысал ретінде Excel кестесі, PDF құжаты немесе CAD сызбасы келтірілуі мүмкін, оларды қажетті ақпаратқа қол жеткізу үшін арнайы құралда ашу қажет. Структуризацияланған шығару мен өңдеу болмаса, ақпаратта пайдаланылмай қалады.

Деректер, өз кезегінде, автоматты түрде байланысты, жаңартылатын және талданатын машина оқитын ақпарат. Біртұтас деректер қоймасында (мысалы, деректер базасы, DWH немесе Data Lake) ақпарат кестелер, жазбалар және байланыстар түрінде ұсынылады. Бұл біртекті сақтау, автоматты сұрауларды орындау, мәндерді талдау және нақты уақыт режимінде есептілікті құру мүмкіндігін қамтамасыз етеді.

Файлдардың орнына деректерді пайдалану (Рис. 8.11) қолмен іздеу процесінен бас тартуға және өңдеу процестерін унификациялауға мүмкіндік береді. Мұндай тәсілді енгізген компаниялар ақпаратқа қол жеткізуідің жылдамдығы мен оны бизнес-процестерге жылдам интеграциялау мүмкіндігі арқасында бәсекелестік артықшылыққа ие болады.

Файлдарды пайдаланудан деректерге көшу – құрылыш саласының болашағын анықтайтын 不可避免 ное өзгеріс.

Әрбір құрылыш компаниясы маңызды таңдау алдында тұрады: ақпаратты адамдар арнайы бағдарламалар арқылы оқу үшін оқшауланған файлдар мен силостарда сақтауды жалғастыру немесе оны алғашқы өңдеу кезеңдерінде құрылымдық деректерге түрлендіріп, автоматтандырылған жобаларды басқару үшін біртұтас интеграцияланған цифрлық негіз құру.

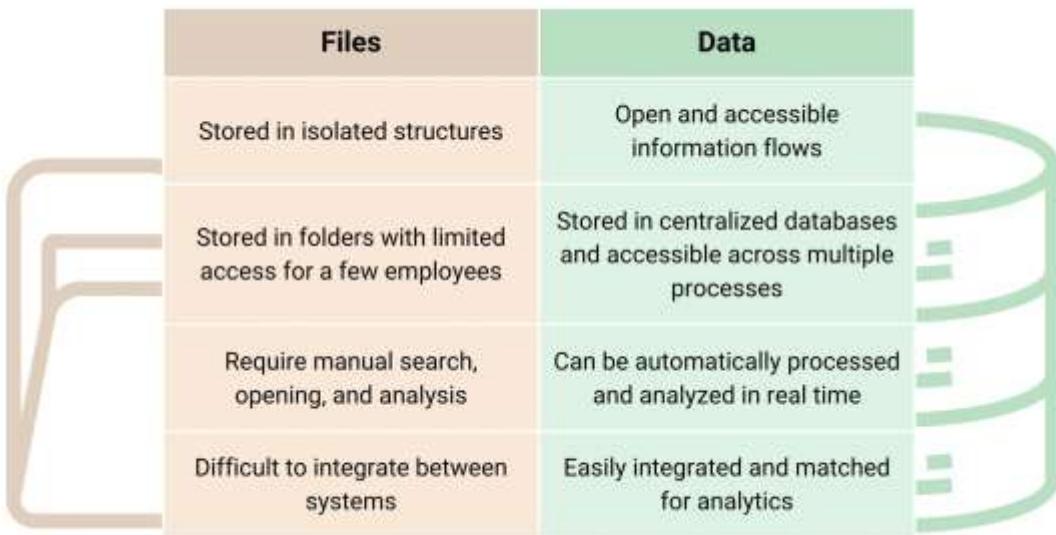


Рис. 8.11 Ақпарат ағынының эволюциясы: оқшауланған файлдардан интеграцияланған деректерге.

Ақпарат көлемінің жедел өсіу жағдайында дәстүрлі файлдарды сақтау және өндеу әдістері барған сайын тиімділігін жоғалтуда. Құрылыш саласында, басқа секторларда сияқты, әртүрлі форматтағы файлдармен немесе бір-бірімен байланыссыз деректер базаларымен жұмыс істеу жеткіліксіз.

Цифрлық технологиялар дәуірінде бәсекелестік қабілетін сақтағысы келетін компаниялар интеграцияланған цифрлық платформаларға, үлкен деректер технологияларын пайдалануға және автоматандырылған аналитикалық жүйелерге көшүге мәжбүр болады.

Файлдық сақтау жүйесінен деректермен жұмыс істеуге көшу ақпаратты басқару тәсілдерін қайта қарауды және орталықтандырылған қоймаларға интеграциялау үшін қолайлы форматтарды санауды түрде таңдауды талап етеді. Бұл таңдау деректерді өндеу тиімділігіне, оларға қол жеткізу жылдамдығына және компанияның цифрлық процестеріне интеграциялау жеңілдігіне әсер етеді.

ҮЛКЕН ДЕРЕКТЕРДІ САҚТАУ: ТАНЫМАЛ ФОРМАТТАРДЫ ТАЛДАУ ЖӘНЕ ОЛАРДЫҢ ТИІМДІЛІГІ

Сақтау форматтары аналитикалық инфрақұрылымның масштабталуын, сенімділігін және өнімділігін қамтамасыз етуде маңызды рөл атқарады. Деректерді талдау және өндеу үшін – мысалы, сұзу, топтау және агрегаттау – біздің мысалдарымызда Pandas DataFrame қолданылды – оперативті жадтағы деректермен жұмыс істеуге арналған танымал құрылым.

Алайда, Pandas DataFrame өз сақтау форматына ие емес, сондықтан өндеу аяқталғаннан кейін деректер сыртқы форматтардың біріне экспортталауды – көбінесе CSV немесе XLSX. Бұл кестелік форматтар алмасу үшін ыңғайлы және көптеген сыртқы жүйелермен үйлесімді, бірақ сақтау тиімділігінің төмендігі, сығудың болмауы және нұсқалаудың әлсіз қолдауы сияқты бірқатар шектесулерге ие.

- CSV (Comma-Separated Values): әртүрлі платформалар мен құралдармен кеңінен қолдау

көрсетілетін қарапайым мәтіндік формат. Ол қолдануға оңай, бірақ курделі деректер типтерін және сығуды қолдамайды.

- XLSX (Excel Open XML Spreadsheet): Microsoft Excel файлдарының форматы, формулалар, диаграммалар және стильдеу сияқты күрделі функцияларды қолдайды. Ол деректерді қолмен талдау және визуализациялау үшін ыңғайлы, бірақ ауқымды деректерді өңдеу үшін оңтайландырылмаған.

Танымал кестелік XLSX және CSV форматтарынан басқа, құрылымдық деректерді тиімді сақтау үшін бірнеше танымал форматтар бар (Рис. 8.12), олардың әрқайсысы деректерді сақтау және талдау бойынша нақты талаптарға байланысты ерекше артықшылықтарға ие.-

- Apache Parquet: деректерді бағандық сақтау үшін арналған файл форматы, деректерді талдау жүйелерінде қолдануға оңтайландырылған. Ол деректерді сығу және кодтау бойынша тиімді схемаларды ұсынады, бұл оны күрделі құрылымдық деректер мен үлкен деректерді өңдеу үшін тамаша етеді.
- Apache ORC (Optimized Row Columnar - оңтайландырылған қатарлық баған): Parquet-ке ұқсас, ORC жоғары сығу дәрежесін және деректерді тиімді сақтауды қамтамасыз етеді. Ол ауыр оқу операцияларына оңтайландырылған және деректер көлдерінде сақтау үшін жақсы келеді.
- JSON (JavaScript Object Notation): JSON деректерді сақтау түрғысынан Parquet немесе ORC сияқты екілік форматтарға қарағанда тиімді емес, бірақ ол өте қолжетімді және жұмыс істеуге оңай, бұл оны оқылымдылық пен веб-технологиялармен үйлесімділіктің маңызды болған сценарийлері үшін тамаша етеді.
- Feather: аналитикаға бағытталған жылдам, жеңіл және қолдануға оңай бинарлық бағандық деректерді сақтау форматы. Ол Python (Pandas) және R арасында деректерді тиімді беру үшін әзірленген, бұл оны осы бағдарламалар орталарын қамтитын жобалар үшін тамаша таңдау етеді.
- HDF5 (Hierarchical Data Format version 5): үлкен деректер көлемдерін сақтау және ұйымдастыру үшін арналған. Ол деректердің кең ауқымын қолдайды және күрделі деректер жиһақтарымен жұмыс істеуге өте қолайлы. HDF5 ғылыми есептеулерде үлкен деректер жиһынтықтарын тиімді сақтау және оларға қол жеткізу қабілеті арқасында ерекше танымал.



		XLSX	CSV	Apache Parquet	HDF5	Pandas DataFrame
Storage	Tabular	Tabular	Columnar	Hierarchical	Tabular	
Usage	Office tasks, data presentation	Simple data exchange	Big data, analytics	Scientific data, large volumes	Data analysis, manipulation	
Compression	Built-in	None	High	Built-in	None (in-memory)	
Performance	Low	Medium	High	High	High (memory dependent)	
Complexity	High (formatting, styles)	Low	Medium	Medium	Low	
Data Type Support	Limited	Very limited	Extended	Extended	Extended	
Scalability	Low	Low	High	High	Medium (memory limited)	

Рис. 8.12 Деректер форматтарын сақтау және өңдеу аспекттері бойынша негізгі айырмашылықтарды көрсететін салыстыру.

ETL процесіндегі Load кезеңінде қолданылатын форматтарды салыстырмалы талдау үшін файлдардың өлшемдері мен оқу уақытын көрсететін кесте жасалды (Рис. 8.13). Зерттеу барысында бірдей деректермен файлдар қолданылды: кесте 10 000 жолдан және 10 бағаннан тұрады, кездейсоқ мәндермен толтырылған.

Зерттеуге келесі сақтау форматтары енгізілді: CSV, Parquet, XLSX және HDF5, сондай-ақ олардың ZIP архивтеріндегі сығылған нұсқалары. Бастапқы деректер NumPy кітапханасын пайдалана отырып генерацияланды және Pandas DataFrame құрылымында ұсынылды. Тестілеу процесі келесі кезеңдерден тұрды:

- Файлдарды сақтау: деректер фреймі төрт түрлі форматта сақталды: CSV, Parquet, XLSX және HDF5. Эр формат деректерді сақтау тәсілі бойынша ерекше ерекшеліктерге ие, бұл файлдың өлшеміне және оқу жылдамдығына әсер етеді.
- Файлдарды ZIP форматында сығу: стандартты сығудың тиімділігін талдау үшін, әр файл ZIP архивіне қосымша сығылды.
- Файлдарды оқу (ETL – Load): ZIP-тан шығарылғаннан кейін әр файлдың оқу уақыты өлшенді. Бұл архивтен алынған деректерге қол жеткізу жылдамдығын бағалауға мүмкіндік береді.

Pandas DataFrame өлшемдері немесе оқу уақыты талдауы кезінде тікелей қолданылмағанын атап өткен жөн, себебі ол деректерді сақтау үшін дербес форматты білдірмейді. Ол тек деректерді генерациялау және әртүрлі форматтарға сақтау үшін аралық құрылым ретінде қызмет етті.



Рис. 8.13 Деректерді сақтау форматтарын өлшемі мен оқу жылдамдығы бойынша салыстыру.

CSV және HDF5 файлдары (Рис. 8.13) сығу тиімділігін жоғары көрсетіп, ZIP-ке орау кезінде өлшемдерін едәуір азайтады, бұл сақтау оптимизациясын талап ететін сценарийлерде әсіресе пайдалы болуы мүмкін. XLSX файлдары, керісінше, сығуға мүлдем көнбейді, және олардың ZIP-дегі өлшемі бастапқы өлшеммен салыстырмалы болып қалады, бұл оларды үлкен деректер көлемінде немесе деректерге қол жеткізу жылдамдығы маңызды жағдайларда пайдалану үшін кем тиімді етеді. Сонымен қатар, XLSX үшін оқу уақыты басқа форматтармен салыстырғанда айтарлықтай жоғары, бұл оны деректерді жылдам оқу операциялары үшін кем қолайлы етеді. Apache Parquet өзінің колонкалық құрылымы арқасында аналитикалық тапсырмалар мен үлкен деректер көлемі үшін жоғары тиімділікті көрсетті.-

Apache Parquet ПЕН ДЕРЕКТЕРДІ САҚТАУДЫ ОПТИМАЛДАНДЫРУ

Үлкен деректерді сақтау және өндөу үшін танымал форматтардың бірі Apache Parquet болып табылады. Бұл формат колонкалық сақтау үшін арнайы әзірленген (Pandas-қа үқсас), бұл сақтау көлемін едәуір азайтуға және аналитикалық сұраулардың жылдамдығын арттыруға мүмкіндік береді. Дәстүрлі форматтар, мысалы, CSV және XLSX, Parquet форматында кіріктілген сұғуды қолдайды және үлкен деректерді өндөу жүйелерімен, соның ішінде Spark, Hadoop және бұлтты сақтау жүйелерімен жұмыс істеуге оңтайландырылған.

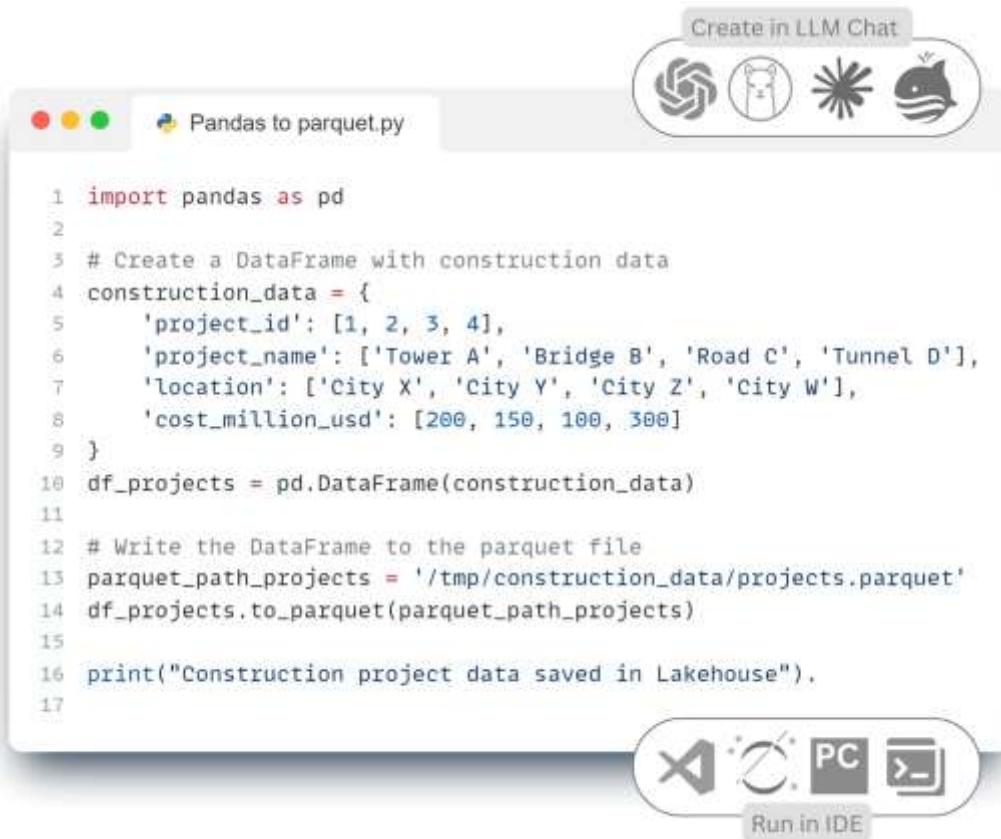
Parquet-тің негізгі ерекшеліктері сығу мен деректер кодтауды қамтиды, бұл сақтау көлемін едәуір азайтады және деректерді оқу операцияларын жылдамдатады, себебі тек қажетті бағандармен жұмыс істейді, барлық деректер жолдарымен емес.

Apache Parquet форматына деректерді түрлендіру үшін қажетті кодты алу қаншалықты оңай екенін көрсету үшін LLM-ді пайдаланамыз.

- ❶ LLM чатында мәтіндік сұраныс жіберіңіз (ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN):

Pandas DataFrame-ден Apache Parquet-ке деректерді сақтау үшін код жазыңыз. ↵

- ❷ LLM-нің жауабы:



The screenshot shows a user interface for generating code. At the top, there's a button labeled "Create in LLM Chat" and a row of icons for different AI models: ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude, and QWEN. Below this is a code editor window with the title "Pandas to parquet.py". The code itself is as follows:

```

1 import pandas as pd
2
3 # Create a DataFrame with construction data
4 construction_data = {
5     'project_id': [1, 2, 3, 4],
6     'project_name': ['Tower A', 'Bridge B', 'Road C', 'Tunnel D'],
7     'location': ['City X', 'City Y', 'City Z', 'City W'],
8     'cost_million_usd': [200, 150, 100, 300]
9 }
10 df_projects = pd.DataFrame(construction_data)
11
12 # Write the DataFrame to the parquet file
13 parquet_path_projects = '/tmp/construction_data/projects.parquet'
14 df_projects.to_parquet(parquet_path_projects)
15
16 print("Construction project data saved in Lakehouse").
17

```

At the bottom of the code editor, there's a "Run in IDE" button with icons for VS Code, PC, and a terminal.

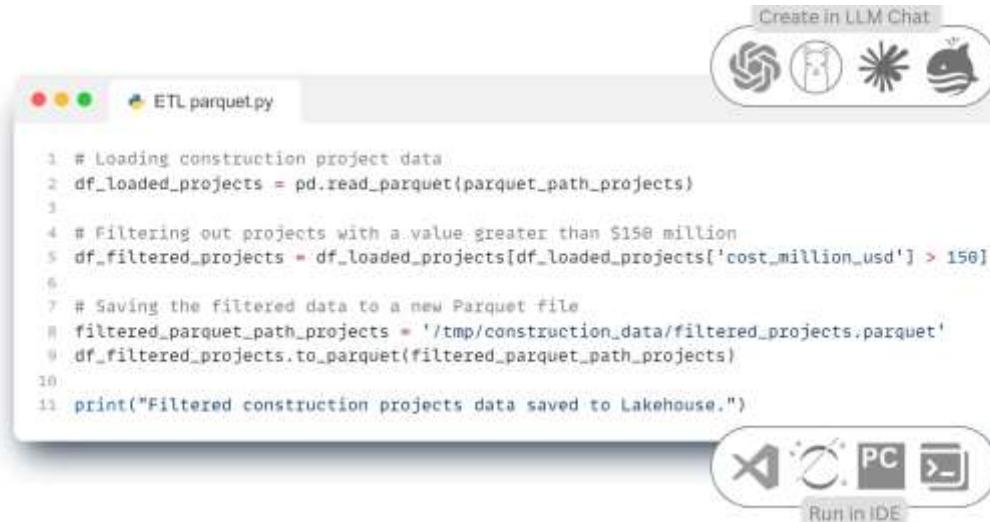
Рис. 8.14 Деректер фреймін жедел жадтан Apache Parquet сақтау форматына тиімді түрде бірнеше жол Python арқылы беру.

Келесі мысал: ETL процесін Parquet форматында сақталған деректермен модельдеу, „cost_million_usd“ атрибутының белгілі бір мәні бойынша жобаларды сұзу (Сурет 8.14).-

❷ Чаттың жалғасында, LLM-ге мәтіндік сұраныс жіберіңіз:

Код жазыңыз, онда біз кестедегі деректерді сүзгілеп, құны 150 миллион доллардан асатын жобаларды (кесте жолдарын) Apache Parquet деректерінен сақтағыңыз келеді. ↵

❸ LLM-нің жауабы:



```

1 # Loading construction project data
2 df_loaded_projects = pd.read_parquet(parquet_path_projects)
3
4 # Filtering out projects with a value greater than $150 million
5 df_filtered_projects = df_loaded_projects[df_loaded_projects['cost_million_usd'] > 150]
6
7 # Saving the filtered data to a new Parquet file
8 filtered_parquet_path_projects = '/tmp/construction_data/filtered_projects.parquet'
9 df_filtered_projects.to_parquet(filtered_parquet_path_projects)
10
11 print("Filtered construction projects data saved to Lakehouse.")

```



Сурет 8.15 Apache Parquet форматындағы деректермен жұмыс істегендеге ETL процесі басқа құрылымдық форматтармен бірдей көрінеді.

Parquet форматында (XLSX, CSV және т.б. қатысты) сақтау көлемін едәуір азайтып, іздеу операцияларын жылдамдатады. Осы себепті ол деректерді сақтау мен талдау үшін тамаша болып табылады. Parquet әртүрлі өндеу жүйелерімен интеграциялана отырып, гибридті архитектураларда тиімді қолжетімділікті қамтамасыз етеді.

Дегенмен, тиімді сақтау форматы – бұл деректермен толыққанды жұмыс істеудің тек бір элементі. Тұрақты және масштабталатын орта құры үшін деректерді басқарудың нақты жобаланған архитектурасы қажет. Бұл функцияны DWH (Data Warehouse) сыныбының жүйелері атқарады. Олар әртүрлі көздерден деректерді агрегациялауды, бизнес-процесстердің айқындылығын және BI құралдары мен машиналық оқыту алгоритмдерін пайдалана отырып кешенді талдауды қамтамасыз етеді.

DWH: Деректер қоймасының (Data Warehouse) САҚТАУ ОРНЫ

Parquet форматы үлкен көлемдегі ақпаратты тиімді сақтау үшін оңтайландырылғандай, Data Warehouse аналитиканы, болжауды және басқарушылық шешімдерді қолдау мақсатында деректерді интеграциялау мен құрылымдауға арналған.

Заманауи компанияларда деректер көптеген әртүрлі көздерден келеді: ERP, CAFM, CPM, CRM жүйелері, бухгалтерлік және қойма есебі, цифрлық CAD ғимарат модельдері, IoT сенсорлары және басқа шешімдер. Толық көрініс алу үшін деректерді жинау жеткіліксіз – оларды ұйымдастыру,

стандарттау және бір орталықта жинақтау қажет. Бұл функцияны DWH атқарады – әртүрлі көздерден ақпаратты агрегациялауға, құрылымдауға және аналитика мен стратегиялық басқару үшін қолжетімді етуге мүмкіндік беретін орталықтандырылған сақтау жүйесі.

DWH (Data Warehouse) – бұл көптеген көздерден ақпаратты агрегациялап, құрылымдап, аналитика мен есептілік үшін қолжетімді ететін орталықтандырылған деректерді сақтау жүйесі.

Көптеген компанияларда деректер әртүрлі жүйелерде шашырап жатыр, оларды кітаптың алғашқы бөлімдерінде қарастырдық (Сурет 1.24). DWH осы көздерді интеграциялап, ақпараттың толық айқындылығын және сенімділігін қамтамасыз етеді. DWH деректер қоймасы – бұл көптеген көздерден деректерді жинап, өндеп, сақтайтын мамандандырылған база (үлкен деректер базасы). DWH-ның негізгі сипаттамалары:-

- ETL процестерін (Extract, Transform, Load) пайдалану – деректерді көздерден шығару, оларды тазалау, түрлендіру, қоймаға жүктеу және осы процестерді автоматтандыру, бұл туралы кітаптың жетінші бөлімінде айтылған.
- Деректердің гранулярлығы – DWH-да деректер агрегатталған түрде (қорытынды есептер) және детальды түрде (шикі деректер) сақталуы мүмкін. 2024 жылдан бастап CAD-вендорлары гранулярланған деректер туралы айтуда [125], бұл, мүмкін, салада цифрлық ғимарат модельдерімен жұмыс істеу үшін мамандандырылған бүлттық қоймаларды пайдалану үшін дайындықтың белгісі болып табылады.
- Аналитика мен болжауды қолдау – деректер қоймалары BI құралдары, Big Data талдауы және машиналық оқыту үшін негізді қамтамасыз етеді.
- DWH бизнес-аналитиканың негізі болып табылады, негізгі тиімділік көрсеткіштерін талдауға, сатылымдарды, сатып алуларды және шығындарды болжауға, сондай-ақ есептілікті автоматтандыруға және деректерді визуализациялауға мүмкіндік береді (Рис. 8.16).-

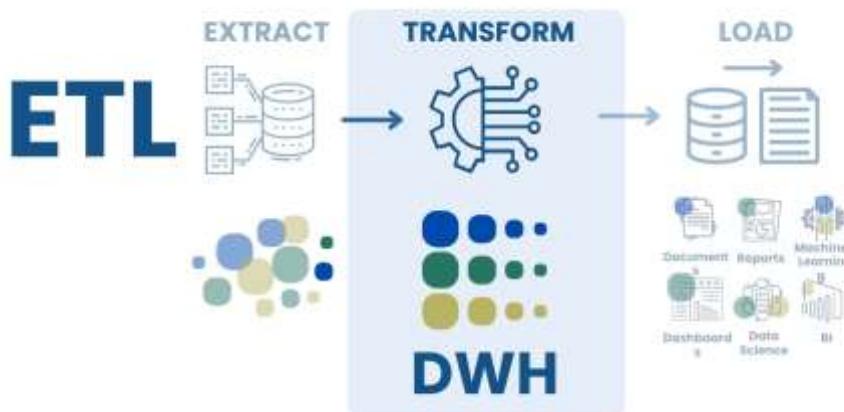


Рис. 8.16 DWH ETL процесінде орталық қойма рөлін атқара алады, мұнда әртүрлі жүйелерден алынған деректер трансформация және жүктеу кезеңдерінен өтеді.

DWH ақпаратты интеграциялау, тазарту және құрылымдау бойынша маңызды рөл атқарады, биzenes-аналитика мен шешім қабылдау процестері үшін берік негіз құрады. Алайда, қазіргі жағдайда деректер көлемі тез өсіп, олардың көздері барған сайын әртүрлі бола бастаған кезде, дәстүрлі ақпарат сақтау тәсілі DWH көбінесе ELT және Data Lake тәсілдерімен кеңейтуді талап етеді.

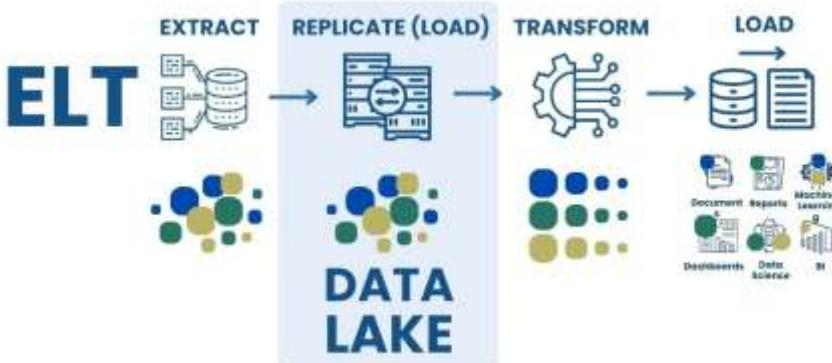
Data Lake - ETL-ден ELT-ге ЭВОЛЮЦИЯ: ДӘСТҮРЛІ ТАЗАЛАУДАН ИНДИВИДУАЛДЫ ӨНДІРІСКЕ

Классикалық DWH деректер қоймалары аналитикалық сұраулар үшін оңтайланырылған форматта құрылымдалған деректерді сақтау үшін арналған, бірақ олар құрылымдалмаған деректермен және масштабталу мәселелерімен бетпе-бет келді. Осы мәселелерге жауап ретінде деректер көлдері (Data Lakes) пайда болды, олар әртүрлі деректердің үлкен көлемдерін икемді сақтау мүмкіндігін ұсынады.

Data Lake дәстүрлі DWH тәсіліне балама ұсынып, құрылымдалмаған, жартылай құрылымдалған және шикі деректермен алдын ала қатты схемасыз жұмыс істеуге мүмкіндік береді. Бұл сақтау әдісі деректерді нақты уақыт режимінде өңдеу, машиналық оқыту және жетілдірілген аналитика үшін жиі өзекті. DWH-дан айырмашылығы, деректерді жүктемес бұрын құрылымдайтын және агрегаттайтын Data Lake ақпаратты бастапқы күйінде сақтауға мүмкіндік береді, осылайша икемділік пен масштабталуды қамтамасыз етеді.

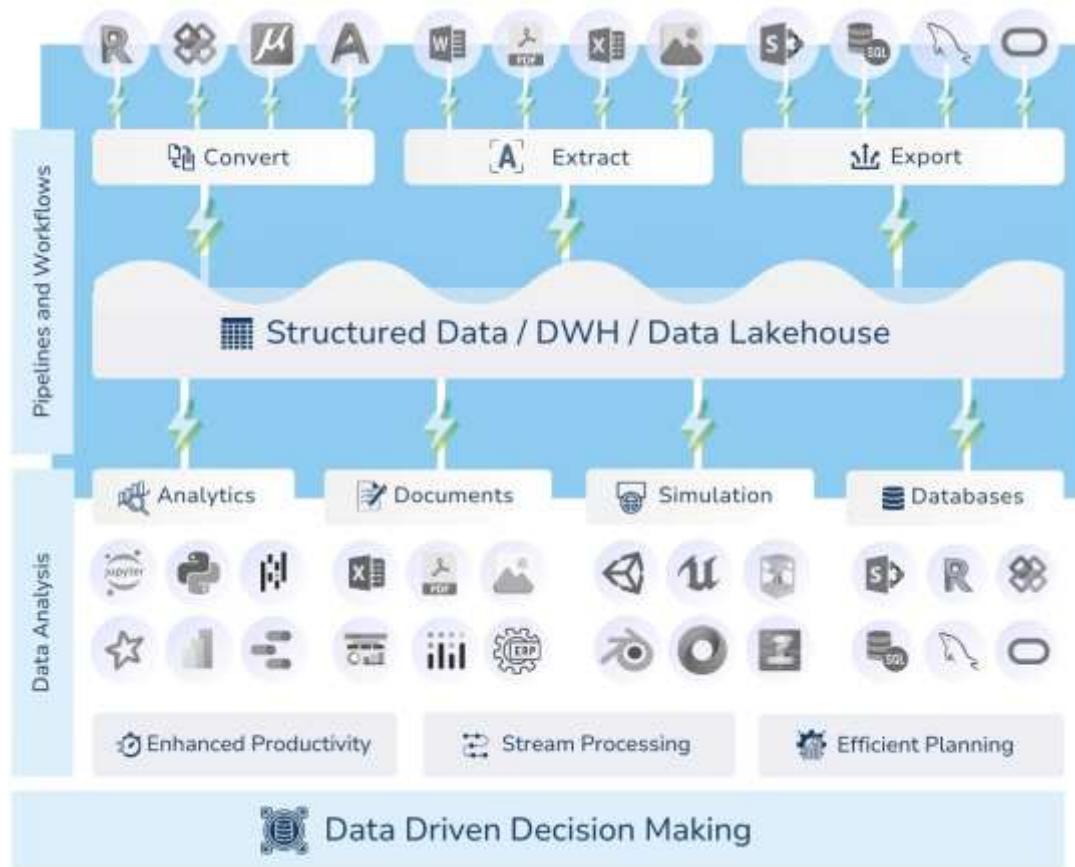
Дәстүрлі деректер қоймалары (RDBMS, DWH) туралы разочарование және «үлкен деректерге» қызығушылық деректер көлдерінің пайда болуына әкелді, мұнда күрделі ETL орнына деректер жайғана әлсіз құрылымдалған қоймаға жүктеледі, ал олардың өнделуі талдау кезеңінде жүзеге асырылады:

- Дәстүрлі деректер қоймаларында деректер әдетте жүктелмес бұрын алдын ала өңдеуден, трансформациядан және тазартудан өтеді (ETL - Extract, Transform, Load) (Рис. 8.16). Бұл деректерді нақты болашақ аналитика және есептілік міндеттерін шешу үшін құрылымдауға және оңтайланыруға мүмкіндік береді. Негізгі назар сұраулардың жоғары өнімділігін және деректердің тұтастығын сақтауға аударылады. Алайда, мұндай тәсіл жаңа деректер түрлерін интеграциялау және тез өзгеретін деректер схемалары түрфысынан қымбат және икемсіз болуы мүмкін.-
- Деректер көлдері, екінші жағынан, үлкен көлемдегі өнделмеген деректерді олардың бастапқы форматында сақтау үшін арналған. ETL (Extract, Transform, Load) процесінің орнына ELT (Extract, Load, Transform) келеді, мұнда деректер алдымен "осылай" жүктеледі және тек содан кейін қажет болған жағдайда түрлендіріледі және талданады. Бұл әртүрлі деректерді, соның ішінде құрылымдалмаған деректерді, мысалы, мәтін, суреттер және журналдар сияқты, сақтау үшін үлкен икемділік пен мүмкіндіктерді қамтамасыз етеді.
-



Деректер көлінде ETL-ге қарағанда ELT қолданылады, мұнда ақпарат алдымен "шикі" күйінде жүктеледі, ал түрлендіру жүктеу кезеңінде жүзеге асырылады.

Дәстүрлі деректер қоймалары деректерді алдын ала өңдеуге бағытталған, бұл сұраулардың жоғары өнімділігін қамтамасыз етеді, ал деректер көлдерінде икемділікке басымдық беріледі: олар өндемеген деректерді сақтайтында және қажет болған жағдайда оларды түрлендіреді.-



Заманауи қойма концепциялары барлық деректер түрлерін сақтау және өңдеу мақсатында бағытталған.

Дегенмен, деректер көлдерінің барлық артықшылықтарына қарамастан, кемшіліктері де бар. Қатаң құрылымның болмауы және ақпаратты басқарудың қындығы хаосқа әкелуі мүмкін, мұнда деректер қайталаңады, бір-біріне қарсы келеді немесе өзектілігін жоғалтады. Сонымен қатар,

мұндағы қоймада деректерді іздеу және талдау едәуір күш-жігерді талап етеді, әсіресе әртүрлі ақпаратпен жұмыс істегендегі. Осы шектеулерді жеңу және дәстүрлі қоймалардың және деректер көлдерінің ең жақсы қасиеттерін біріктіру үшін Data Lakehouse архитектурасы әзірленді.

Data Lakehouse АРХИТЕКТУРАСЫ: САҚТАУ ОРЫНДАРЫ МЕН ДЕРЕК ОЗЕРІНІҢ СИНЕРГИЯСЫ

DWH (құрылымдылық, басқару, аналитикадағы жоғары өнімділік) және Data Lake (масштабталушылық, әртүрлі деректермен жұмыс) ең жақсы сипаттамаларын біріктіру үшін Data Lakehouse тәсілі әзірленді. Бұл архитектура деректер көлдерінің икемділігін дәстүрлі қоймаларға тән қуатты өңдеу және басқару құралдарымен біріктіреді, сақтау, аналитика және машиналық оқыту арасында теңгерімді қамтамасыз етеді. Data Lakehouse - деректер көлдері мен деректер қоймаларының синтезі, бірінші тараптың икемділігі мен масштабталушылығын екіншісінің басқарылуымен және сұрауларды оңтайландырумен біріктіреді.

Data Lakehouse - деректер көлдерінің икемділігі мен масштабталушылығын деректер қоймаларындағы басқару мен сұраулардың өнімділігімен біріктіруге үмтүлательн архитектуралық тәсіл.-

Data Lakehouse-тың негізгі ерекшеліктері мыналарды қамтиды:

- Деректерді сақтау үшін ашық формат: деректерді сақтау үшін Apache Parquet сияқты ашық форматтарды пайдалану тиімділік пен сұрауларды оңтайландыруды қамтамасыз етеді.
- Тек оқу схемасы: дәстүрлі DWH-дағы тек жазу схемасынан айырмашылығы, Lakehouse тек оқу схемасын қолдайды, бұл деректер құрылымын басқаруды икемді етеді.
- Икемділік және масштабталушылық: құрылымдалған және құрылымдалмаған деректерді сақтау және талдауды қолдайды, сақтау деңгейінде оңтайландыру арқылы сұраулардың жоғары өнімділігін қамтамасыз етеді.

Data Lakehouse екі тәсілдің артықшылықтарын біріктіретін компромисс шешімін ұсынады, бұл оны деректерді өңдеуде икемділікті талап ететін заманауи аналитикалық жүктемелер үшін идеалды етеді.



Data Lakehouse - күрделі және үнемі өзгеріп отыратын талаптарды қанағаттандыру үшін жасалған деректерді сақтау жүйелерінің келесі буыны.

Заманауи деректер қоймаларының идеясы қарапайым көрінеді: егер барлық деректер бір жерде болса, оларды талдау оқайырақ. Алайда, тәжірибеде бәрі де оқай емес. Мысалы, компания дәстүрлі есептеу және басқару жүйелерінен (ERP, PMIS, CAFM және т.б.) толықтай бас тартуға шешім қабылдаса, оларды барлығына қолжетімді бір үлкен деректер көлімен алмастырады. Не болады? Мүмкін, хаос басталады: деректер қайталанатын, бір-біріне қарама-қайшы келетін, ал маңызды ақпарат жоғалып немесе бұрмаланған болады. Деректер көлі тек аналитика үшін пайдаланылса да, оны дұрыс басқармаса, елеулі қыындықтар туындаиды:

- Деректерді түсіну қыын: дәстүрлі жүйелерде деректердің нақты құрылымы бар, ал деректер көлінде тек үлкен файлдар мен кестелердің жинағы бар. Бір нәрсені табу үшін маман әр жол мен бағанның не үшін жауап беретінін түсініү керек.
- Деректер дәл емес болуы мүмкін: егер бір жерде бірдей ақпараттың бірнеше нұсқасы сақталса, қайсысы өзекті екенін түсінү қыын. Нәтижесінде, ескірғен немесе қате деректер негізінде шешімдер қабылданады.
- Деректерді жұмысқа дайындау қыын: деректерді тек сақтау ғана емес, сонымен қатар оларды есептер, графиктер, кестелер түрінде ыңғайлы түрде ұсыну қажет. Дәстүрлі жүйелерде бұл автоматты түрде жүзеге асырылады, ал деректер көлдерінде қосынша өңдеуді талап етеді.

Нәтижесінде, әрбір деректерді сақтау концепциясының өз ерекшеліктері, өңдеу тәсілдері және бизнесте қолдану әдістері бар. Дәстүрлі деректер базалары транзакциялық операцияларға бағытталған, деректер қоймалары (DWH) аналитика үшін құрылым ұсынады, деректер көлдері (Data Lake) ақпаратты шикі күйінде сақтайды, ал гибридті қоймалар (Data Lakehouse) DWH мен Data Lake-тің артықшылықтарын біріктіреді.-

	Traditional Approach	Data Warehouse	Data Lake	Data Lakehouse
Data Types	Relational Databases	Structured, ready for analytics	Raw, semi-structured, or unstructured	Mix of structured and unstructured
Use Cases	Transactional Systems	Reporting, dashboards, BI	Big data storage, AI, advanced analytics	Hybrid analytics, AI, real-time data
Processing	OLTP – real-time transactions	ETL – clean and structure before analysis	ELT – store raw data, transform later	ELT with optimized storage and real-time processing
Storage	On-premise servers	Centralized, SQL-based	Decentralized, flexible formats	Combines advantages of DWH and DL
Common Tools	MySQL, PostgreSQL	Snowflake, Redshift, BigQuery	Hadoop, AWS S3, Azure Data Lake	Databricks, Snowflake, Google BigLake

DWH, Data Lake және Data Lakehouse: деректер түрлери, пайдалану сценарийлері, өңдеу әдістері және сақтау тәсілдері бойынша негізгі айырмашылықтар.

Деректерді сақтау архитектурасын таңдау – бизнес қажеттіліктеріне, ақпарат көлеміне және аналитика талаптарына байланысты күрделі процесс. Әрбір шешімнің өз артықшылықтары мен кемшіліктері бар: DWH құрылымдылықты қамтамасыз етеді, Data Lake икемділікті, ал Lakehouse екеуінің арасындағы тепе-тендікті ұсынады. Ұйымдар сирек жағдайда бір ғана деректер архитектурасымен шектеледі.

Таңдалған архитектураға қарамастан, автоматтандырылған деректерді басқару жүйелері қолмен жүргізілетін әдістерден айтарлықтай артық. Олар адам қателіктерін минимизациялауға, ақпаратты өңдеуді жеделдетуге, бизнес-процестердің барлық кезеңдерінде деректердің ашықтығы мен бақылауын қамтамасыз етуге мүмкіндік береді.

Егер орталықтандырылған деректер қоймалары көптеген экономикалық салаларда салалық стандартқа айналса, құрылыш саласында жағдай әлі де фрагменттелген. Мұнда деректер әртүрлі платформалар арасында (CDE, PMIS, ERP және т.б.) бөлінген, бұл болып жатқан жағдайдың біртұтас көрінісін жасауға кедегі келтіреді және осы көздерді тұтас, аналитикалық түрғыдан пайдалы цифрлық ортаға біріктіруге қабілетті архитектураларды талап етеді.

CDE, PMIS, ERP НЕМЕСЕ DWH ЖӘНЕ Data Lake

Құрылыш және жобалау саласында жұмыс істейтін кейір компанияларда ISO 19650 стандартына сәйкес жалпы деректер ортасы (Common Data Environment, CDE) концепциясы қолданылууда. Негізінде, CDE басқа салалардағы деректер қоймасының (DWH) функцияларын орындаиды: ақпаратты орталықтандырады, нұсқаларды бақылауды қамтамасыз етеді, тексерілген ақпаратқа қолжетімділікті ұсынады.

Жалпы деректер ортасы (CDE) – бұл объектінің өмірлік циклі кезеңдерінде жобалық ақпаратты басқару, сақтау, алмасу және бірлесіп жұмыс істеу үшін пайдаланылатын орталықтандырылған цифрлық кеңістік. CDE көбінесе бұлттық технологияларды пайдалану арқылы жүзеге асырылады және CAD (BIM) жүйелерімен интеграцияланады.

Қаржы секторы, бөлшек сауда, логистика және өнеркәсіп орталықтандырылған деректерді басқару жүйелерін онжылдықтар бойы пайдаланып келеді, олар әртүрлі көздерден ақпаратты біріктіріп, оның өзектілігін бақылап, аналитиканы қамтамасыз етеді. CDE осы принциптерді дамыта отырып, оларды ғимараттарды жобалау және басқару міндеттеріне бейімдейді.

DWH сияқты, CDE деректерді құрылымдайды, өзгерістерді тіркейді және тексерілген ақпаратқа бірегей қолжетімділікті қамтамасыз етеді. Бұлттық технологияларға көшу және аналитикалық құралдармен интеграциялану барысында олардың арасындағы айырмашылықтар барған сайын азаяды. CDE-ге 2023 жылдан бастап CAD жеткізушилерімен талқыланатын гранулирленген деректерді қосу арқылы класикалық DWH-мен тағы да көп параллельдер көргө болады.

Бұрын «Құрылыш ERP және PMIS жүйелері» бөлімінде біз PMIS (Жобаларды Басқару Ақпарат Жүйесі) және ERP (Кәсіпорынды Ресурстарды Жоспарлау) қарастырыдық. Құрылыш жобаларында CDE мен PMIS бірге жұмыс істейді: CDE деректер қоймасы ретінде қызмет етеді, оған сыйбалар, модельдер және жобалық құжаттама кіреді, ал PMIS мерзімдер, тапсырмалар, ресурстар және бюджет сияқты процестерді басқарады.

ERP, бизнес басқару үшін жауап беріп (қаржы, сатып алу, персонал, өндіріс), PMIS-пен интеграциялануы мүмкін, бұл компания деңгейінде шығындар мен бюджеттерді бақылауды қамтамасыз етеді. Аналитика мен есептілік үшін DWH қолданылуы мүмкін, ол CDE, PMIS және ERP-ден деректерді жинап, құрылымдап, агрегаттауға көмектеседі, бұл қаржылық көрсеткіштер KPI (ROI) бағалауға және заңдылықтарды анықтауға мүмкіндік береді. Өз кезегінде, Data Lake (DL) DWH-ны толықтыра алады, ол шикі және құрылымдалмаған деректерді (мысалы, логтар, сенсорлық деректер, суреттер) сақтайды. Бұл деректер өндөліп, DWH-ға жүктеліп, әрі қарай талдау үшін пайдаланылады.

Осылайша, CDE мен PMIS жобаларды басқаруға, ERP бизнес процестеріне, ал DWH мен Data Lake аналитика мен деректермен жұмыс істеуге бағытталған.

CDE, PMIS және ERP жүйелерін DWH және Data Lake-пен салыстырғанда, вендорларға тәуелсіздік, құны, интеграцияның икемділігі, деректердің тәуелсіздігі, өзгерістерге бейімделу жылдамдығы, сондай-ақ аналитикалық мүмкіндіктер түрфысынан елеулі айырмашылықтарды байқауға болады. Дәстүрлі жүйелер, мысалы, CDE, PMIS және ERP, көбінесе нақты шешімдер мен вендор стандарттарымен байланысты, бұл оларды аз икемді етеді және лицензиялар мен қолдау арқылы құнын артырады. Сонымен қатар, мұндай жүйелердегі деректер көбінесе меншікті жабық форматтарда болады, бұл олардың пайдаланылуын және талдауын шектейді.-

		CDE, PMIS, ERP	DWH, Data Lake
	Vendor Dependency	High (tied to specific solutions and standards of vendors)	Low (flexibility in tool and platform choice)
	Integration Flexibility	Limited (integration depends on vendor solutions)	High (easily integrates with various data sources)
	Cost	High (licensing and support costs)	Relatively lower (use of open technologies and platforms)
	Data Independence	Low (data often locked in proprietary formats)	High (data stored in open and accessible formats)
	Adaptability to Changes	Slow (changes require vendor approval and integration)	Fast (adaptation and data structure modification without intermediaries)
	Analytical Capabilities	Limited (dependent on vendor-provided solutions)	Extensive (support for a wide range of analytical tools)

DWH және Data Lake CDE, PMIS және ERP жүйелеріне қарағанда деректердің үлкен икемділігі мен тәуелсіздігін ұсынады.

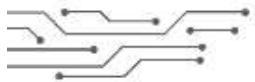
Оларға қарағанда, DWH және Data Lake әртүрлі дереккөздермен интеграциялануда үлкен икемділік ұсынады, ал ашық технологиялар мен платформаларды пайдалану жалпы меншік құнын төмендетьге көмектеседі. Сонымен қатар, DWH және Data Lake кең ауқымды аналитикалық құралдарды қолдайды, бұл талдау мен басқару мүмкіндіктерін кеңейтеді.

CAD форматтарына арналған көрініс инженерия құралдарының дамуы және CAD қосымшаларының деректер базаларына қол жеткізудің болуы жобалық деректердің көптеген мердігерлік және жобалық үйімдарда жұмыс істейтін мамандарға қолжетімді болуы тиіс екендігін ескере отырып, жабық, оқшауланған платформаларды пайдалануды жалғастырудың қаншалықты негізді екендігі туралы мәселе өткір түр.

Белгілі бір вендорға тәуелділік деректерді басқарудағы икемділікті едәуір шектеп, жобадағы өзгерістерге реакцияны баяулатып, қатысуышылар арасындағы тиімді ынтымақтастыққа кедерігі келтіруі мүмкін.

Деректерді басқарудың дәстүрлі тәсілдері – DWH, Data Lake, CDE және PMIS – негізінен ақпаратты сақтау, құрылымдау және өндөуге бағытталған. Алайда, жасанды интеллект пен машиналық оқытудың дамуы жаңа деректерді үйімдастыру тәсілдеріне қажеттілікті арттырады, олар тек агрегаттаумен шектелмей, күрделі өзара байланыстарды анықтауға, жасырын зандылықтарды табуға және ең релевантты ақпаратқа жедел қол жеткізуді қамтамасыз етуге мүмкіндік береді.

Бұл бағытта векторлық деректер базалары ерекше рөл атқара бастайды – жоғары өлшемді әмбебдингтермен жұмыс істеуге оңтайландырылған жаңа типтегі сақтау орындары.



ТАРАУ 8.2.

ДЕРЕКТЕРДІ САҚТАУ ЖӘНЕ ХАОС ТУДЫРУДЫ БОЛДЫРМАУ

ВЕКТОРЛЫҚ ДЕРЕК БАЗАЛАРЫ ЖӘНЕ БОКС ШЕГІ

Векторлық деректер базалары – деректерді тек сақтап қана қоймай, мағына бойынша іздеуді, объектілерді семантикалық жақындық бойынша салыстыруды және интеллектуалды жүйелерді құруды қамтамасыз ететін жаңа класс сақтау орындары. Дәстүрлі деректер базаларынан айырмашылығы, олар дәл сәйкестіктерге бағытталған, векторлық БД белгілер негізінде ұқсас объектілерді табады – тіпті дәл сәйкестік болмаған жағдайда да.

Векторлық деректер базасы – деректерді көп өлшемді векторлар түрінде сақтайтын мамандандырылған деректер базасы, әрбір вектор белгілі бір сипаттамаларды немесе қасиеттерді білдіреді. Бұл векторлардың өлшемдері деректердің күрделілігіне байланысты әртүрлі болуы мүмкін (бір жағдайда бірнеше өлшем, ал басқа жағдайда – мындаған).

Векторлық базалардың негізгі артықшылығы – семантикалық маңыздылық бойынша іздеу, дәл мәндердің сәйкестігі бойынша емес. SQL және Pandas сұрауларындағы «тең» немесе «құрамында» фильтрлерінің орнына, белгілер кеңістігінде жақын көршілерді іздеу (k -NN) қолданылады (k -NN туралы келесі бөлімде толығырақ айтылады).

LLM (Large Language Models) және генеративті модельдердің дамуы деректер базаларымен өзара әрекеттесуді өзгертуде. Енді деректерді табиғи тілде сұрау, құжаттар бойынша семантикалық іздеу, кілттік терминдерді автоматты түрде шығару және объектілер арасындағы контексттік байланыстарды құру мүмкіндігі бар – мұның барлығы SQL-ды менгеру немесе кесте құрылымын білу қажеттілігінсіз. Бұл туралы «LLM және олардың деректерді өңдеудегі және бизнес-процесстердегі рөлі» бөлімінде айтылған.

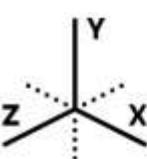
Алайда, LLM ақпаратты автоматты түрде құрылымдамайтынын және тәртіпке келтірмейтінін түсіну маңызды. Модель тек деректер массивінде "жүзеді" және сұрау контекстіне негізделген ең қолайлы фрагментті табады. Егер деректер алдын ала тазартылmasa немесе түрлендірілмесе, терең іздеу (deep search) цифрлық "қоқыста" жауап табуға ұқсас болады – жұмыс істеуі мүмкін, бірақ нәтижелердің сапасы төмен болады. Идеал жағдай – деректерді алдын ала құрылымдау (мысалы, құжаттарды Markdown форматына аудыстыру) және векторлық базаға жүктеу. Бұл шығару дәлдігі мен релеванттылығын едәуір арттырады.

Бастапқыда векторлық деректер базалары машиналық оқуда қолданылды, бірақ бүгінгі таңда олар іздеу жүйелерінде, контентті персонализациялауда, интеллектуалдық аналитикада кеңінен қолданыс табуда.

Векторлық тәсілдің құрылышы саласындағы ең айқын мысалдарының бірі – шектейтін параллелепипед (Bounding Box). Бұл геометриялық құрылым үш өлшемді кеңістіктегі объектінің шекараларын

сипаттайды. Шектейтін параллелепипед X, Y және Z осьтері бойынша минималды және максималды координаталармен анықталады, объектінің айналасында «қорап» құра отырып. Бұл әдіс элементтің өлшемдері мен орналасуын бүкіл геометрияны талдаусыз бағалауға мүмкіндік береді.

Әрбір Bounding Box-ты көп өлшемді кеңістікте вектор ретінде көрсетуге болады: мысалы, [x, y, z, ені, биіктігі, теренідігі] – бұл 6 өлшем (Сурет 8.21).-



Bounding Box

	minX	maxX	minY	maxY	minZ	maxZ	Width	Height	Depth
Column	-15	-5	-25	-15	0	10	10	10	20
Stairs	-5	5	-15	-5	0	10	10	10	10
Door	5	15	5	15	0	10	10	10	10
Window	25	35	-35	-25	10	30	10	20	20
Balcony	15	25	-5	5	20	40	10	20	20

Рис. 8.21 Bounding Box элементтерінің координаттары мен олардың жобадағы орналасуы векторлық деректер базасының аналогиясы болып табылады.

Мұндай деректерді ұсыну объектілер арасындағы қылыштарды тексеру, ғимарат элементтерінің кеңістіктік таралуын жоспарлау және автоматтандырылған есептеулерді орындау сияқты көптеген міндеттерді жеңілдетеді. Bounding Box күрделі үш өлшемді модельдер мен дәстүрлі векторлық деректер базалары арасында көпір қызметтің атқара алады, архитектуралық және инженерлік мөдельдеуде екі тәсілдің артықшылықтарын тиімді пайдалануға мүмкіндік береді.

Bounding Box – бұл "геометрияның векторизациясы", ал әмбеддинг (абстрактіні түрлендіру тәсілі) – "мағынаның векторизациясы". Екі тәсіл де қолмен іздеуден интеллектуалды іздеуге көшуге мүмкіндік береді, бұл 3D-объектілер жобаның моделінде немесе мәтіндең үғымдарда болуы мүмкін.

Жобада объектілерді іздеу (мысалы, "барлық ені 1,5 м-ден асатын терезелерді табу") векторлық деректер базасындағы жақын көршілерді іздеуге (k-NN) үқсас, мұнда критерийлер "аймақты" белгілер кеңістігінде анықтайды. (жақын k-NN көршілерін іздеу туралы толығырақ машиналық оқытуға арналған келесі бөлімде талқылаймыз) (Сурет 8.22). Егер шекаралық қорапқа атриуттарға қосылыш параметрлер (материал, салмақ, өндіріс мерзімі) қосылса, онда кесте жоғары өлшемді векторға айналады, мұнда әрбір атриут - жаңа өлшем. Бұл қазіргі заманғы векторлық деректер базаларына жақын, мұнда өлшемдер жүздеген немесе мындаған болып есептеледі (мысалы, нейрондық желілерден алынған әмбеддинг).-

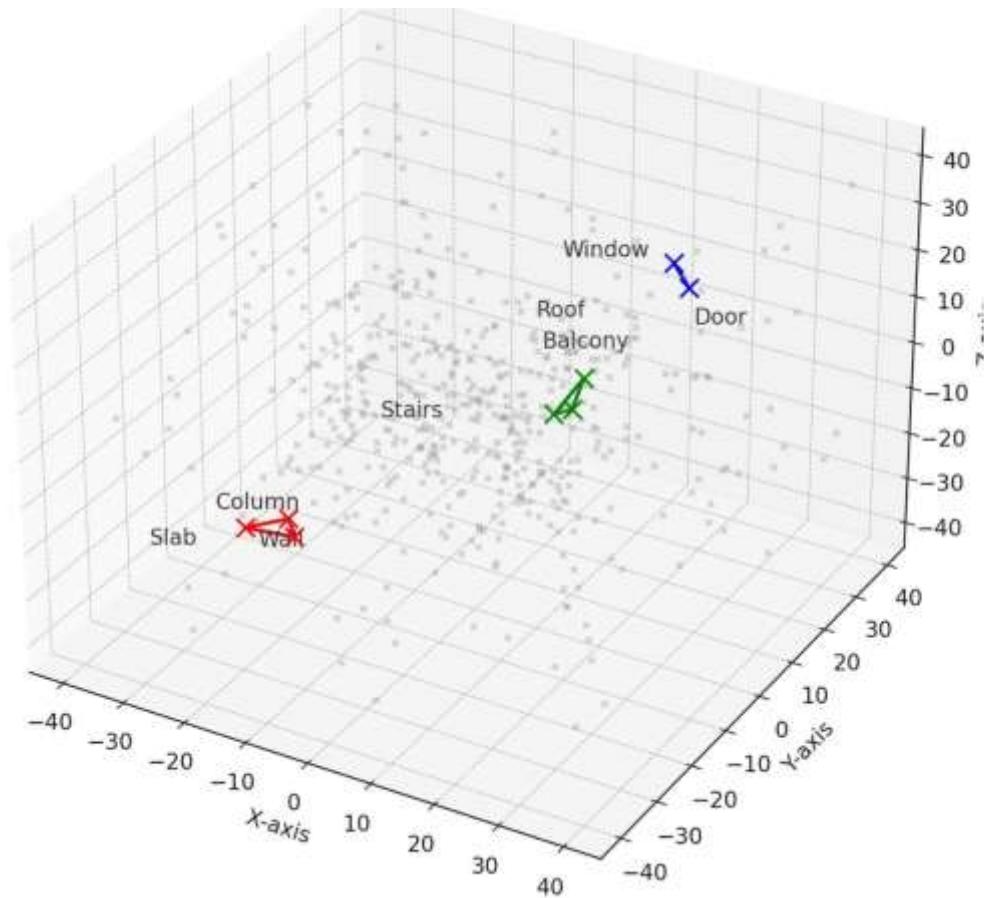


Рис. 8.22 Жобада объектілерді векторлық деректер базаларын пайдалана отырып іздеу.

Bounding Box-те қолданылатын тәсіл тек геометриялық объектілерге ғана емес, сонымен қатар мәтіндер мен тілдерді талдауға да қолданылады. Деректердің векторлық ұсыныстары табиғи тілдерді өңдеуде (NLP) белсенді түрде пайдаланылады. Құрылыш жобасындағы объектілерді кеңістіктік жақындық бойынша топтастыру сияқты, мәтіндегі сөздерді олардың мағыналық және контекстік жақындықтарына қарай талдауға болады.-

Мысалы, "архитектор", "құрылыш", "жобалау" сөздері векторлық кеңістікте жақын орналасады, себебі олар үқсас мағынаға ие. LLM-да бұл механизм автоматты түрде, қолмен классификациялаудың қажеті жоқ, мүмкіндік береді.

- Мәтіннің тақырыбын анықтау.
- Мазмұнды құжаттар бойынша семантикалық іздеу жүргізу.
- Автоматты аннотациялар мен мәтіннің резюмесін жасау.
- Синонимдер мен өзара байланысты терминдерді табу.

Векторлық деректер базалары мәтінді талдауға және оған байланысты терминдерді табуға мүмкіндік береді, дәл Bounding Box 3D модельдеріндегі кеңістіктік объектілерді талдауға көмектеседі.

седі. Жобаның элементтерінің Bounding Box мысалы векторлық ұсыныстың тек "жасанды" ML концепциясы емес, деректерді құрылымдаудың табиғи тәсілі екенін түсінуге көмектеседі, бұл қолданбалы міндеттерді шешу үшін, CAD жобасындағы колонналарды іздеу немесе дерекқордағы семантикалық жағынан жақын суреттерді табу болсын.

Деректер базаларымен жұмыс істейтін мамандар векторлық қоймаларды назарға алуы тиіс. Олардың таралуы классикалық реляциялық жүйелер мен AI-бағдарланған технологиялардың бірігін көрсетеді, бұл болашақтың гибридтік шешімдерін қалыптастырады.

Күрделі және ауқымды AI-қосымшаларды өзірлейтін пайдаланушылар векторлық іздеу үшін мамандандырылған деректер базаларын қолданады. Сонымен қатар, тек белгілі AI-функцияларын интеграциялау қажет болғандар, әдетте, өздері пайдаланатын деректер базаларында (PostgreSQL, Redis) векторлық іздеудің кіріктілген мүмкіндіктерін таңдайды.

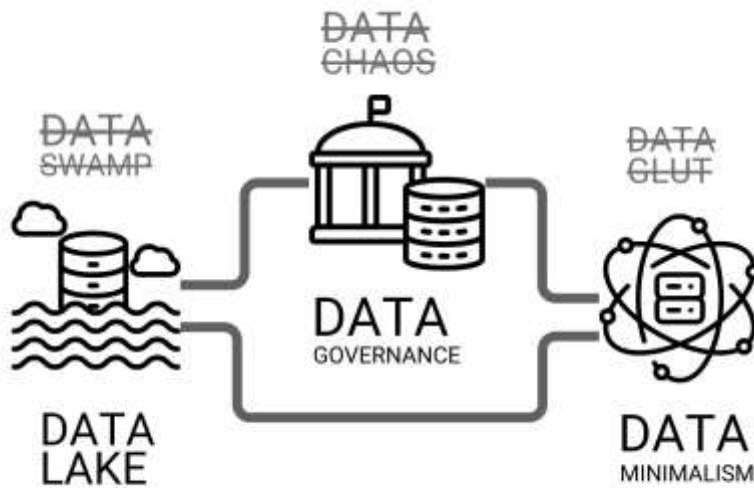
DWH, Data Lake, CDE, PMIS сияқты жүйелер, векторлық деректер базалары және басқалары деректерді сақтау мен басқарудың әртүрлі тәсілдерін ұсынса да, олардың тиімділігі тек архитектурамен емес, сонымен қатар деректердің қаншалықты дұрыс ұйымдастырылғаны мен басқарылуымен анықталады. Заманауи шешімдерді – векторлық базалар, классикалық реляциялық СУБД немесе Data Lake типті қоймаларды қолданғаның өзінде, деректерді басқару, құрылымдау және жаңарту бойынша нақты ережелердің болмауы, пайдаланушыларды шашыраңқы файлдар мен әртүрлі форматтағы деректермен жұмыс істеген кездегі қиындықтарға әкелуі мүмкін.

Деректерді тиімді басқару (Data Governance) жоқ болса, ең қуатты шешімдер де хаосты және құрылымда маған ақпарат массивтеріне айналуы мүмкін, деректер көлдерін батпақтарға (Data Swamp) айналдырады. Мұны болдырмай үшін компаниялар тек қойма архитектурасын дұрыс таңдаумен шектелмей, деректер минимализмі (Data Minimalism), қолжетімділікті басқару және сапа бақылау стратегияларын енгізу тиіс, бұл деректерді шешім қабылдаудың тиімді құралына айналдыруға мүмкіндік береді.

ДЕРЕКТЕРДІ БАСҚАРУ (Data Governance), ДЕРЕКТЕРДІҢ МИНИМАЛИЗМІ (Data Minimalism) ЖӘНЕ ДЕРЕКТЕРДІҢ БАТПАҒЫ (Data Swamp)

Деректерді басқару (Data Governance), деректер минимализмі (Data Minimalism) және деректер батпағын (Data Swamp) болдырмай концепцияларын түсіну мен енгізу – деректер қоймаларын тиімді басқарудың және олардың бизнес үшін құндылығын қамтамасыз етудің негізгі факторлары.-

Gartner зерттеуінде (2017) сәйкес, үлкен деректер саласындағы жобалардың 85%-ы сәтсіздікке ұшырайды, және бұл сәтсіздіктің негізгі себептерінің бірі – деректердің сапасын және оларды басқаруды жеткіліксіз басқару.



Деректерді басқарудың негізгі аспектілерінің бірі – Data Governance және Data Minimalism.

Деректерді басқару (Data Governance) – деректерді барлық бизнес-процестерде дұрыс және тиімді пайдалануды қамтамасыз ететін деректерді басқарудың негізгі компоненті. Бұл тек ережелер мен процедураларды орнатумен ғана емес, сонымен қатар деректердің қолжетімділігін, сенімділігін және қауіпсіздігін қамтамасыз етумен байланысты.

- Деректерді анықтау және класификациялау: ұйымдарға қажетті сущностерді түсінуге және оларды қалай пайдалану керектігін анықтауға мүмкіндік беретін сущностердің нақты анықтамасы мен класификациясы.
- Қол жеткізу құқықтары мен басқару: деректерге қол жеткізу және оларды басқару саясатын және процедураларын әзірлеу, тек авторизацияланған пайдаланушылардың белгілі деректерге қол жеткізуін қамтамасыз етеді.
- Сыртқы қауіптерден деректерді қорғау: сыртқы қауіптерден деректерді қорғау - деректерді басқарудың негізгі аспектілерінің бірі. Бұл тек техникалық шараларды ғана емес, сонымен қатар қызметкерлерді ақпараттық қауіпсіздік негіздеріне оқытуды да қамтиды.

Деректер минимализмі (Data Minimalism) - деректерді ең құнды және маңызды элементтер мен атрибуттарға дейін қысқартуға бағытталған тәсіл, бұл шығындарды азайтуға және деректерді пайдаланудың тиімділігін арттыруға мүмкіндік береді. -

- Шешім қабылдау процесін жеңілдету: обьектілер мен олардың атрибуттарын ең маңыздысына дейін қысқарту шешім қабылдау процесін жеңілдетеді, деректерді талдау мен өңдеуге қажетті уақыт пен ресурстарды азайтады.
- Маңыздыға назар аудару: ең релевантты элементтер мен атрибуттарды таңдау бизнестің шын мәнінде маңызды ақпаратқа назар аударуына мүмкіндік береді, шу мен қажетсіз деректерді жояды.
- Ресурстарды тиімді бөлу: деректерді минимизациялау ресурстарды тиімдірек бөлуге мүмкіндік береді, деректерді сақтау мен өңдеуге кететін шығындарды азайтып, олардың сапасы мен қауіпсіздігін арттырады.

Деректермен жұмыс істеу логикасы олардың құрылуынан емес (Рис. 8.24), деректерді генерациялау процесі басталмай тұрып, осы деректерді болашақта пайдалану сценарийлерін түсінуден басталуы тиіс. Мұндай тәсіл атрибуттарға, олардың типтеріне және шекті мәндеріне минималды қажетті талаптарды алдын ала анықтауға мүмкіндік береді. Дәл осы талаптар ақпараттық модельде дұрыс және тұрақты элементтерді құрудың негізін қалыптастырады. Деректерді қолдану мақсаттары мен тәсілдерін алдын ала ойластыру талдауға жарамды құрылымды қалыптастыруға ықпал етеді. Деректерді модельдеудің концептуалдық, логикалық және физикалық деңгейлеріндегі тәсілдер туралы біз «Деректерді модельдеу: концептуалдық, логикалық және физикалық модель» бөлімінде айттық.

Дәстүрлі бизнес-процестерде құрылыш компанияларының деректерді өңдеуі көбінесе деректерді батпаққа тастауға ұқсайды, мұнда алдымен деректер жасалады, содан кейін мамандар оларды басқа жүйелер мен құралдарға интеграциялауға тырысады.

Деректер батпағы (Data Swamp) - деректерді бақылаусыз жинау мен сақтау нәтижесінде пайда болатын, дұрыс үйімдастырылмаған, құрылымдалмаған және басқарылмаған деректер, нәтижесінде деректер құрылымсыз, пайдалану қыын және құнсыз болады.

Ақпарат ағынын деректер батпағына айналдыруды қалай болдырмауға болады:

- Деректер құрылымын басқару: деректердің құрылымдылығын және класификациясын қамтамасыз ету «деректер батпағын» болдырмауға көмектеседі, оларды тәртіпке келтіріп, қолжетімді етеді.
- Деректерді түсіну және интерпретациялау: деректердің шығу тегі, модификациялары мен мағыналарының нақты сипаттамасы деректердің дұрыс түсінілуі мен интерпретациялануын қамтамасыз етеді.
- Деректер сапасын сақтау: деректерді тұрақты түрде күтіп ұстау және тазалау олардың сапасын, өзектілігін және аналитикалық және бизнес-процестер үшін құндылығын сақтауға көмектеседі.

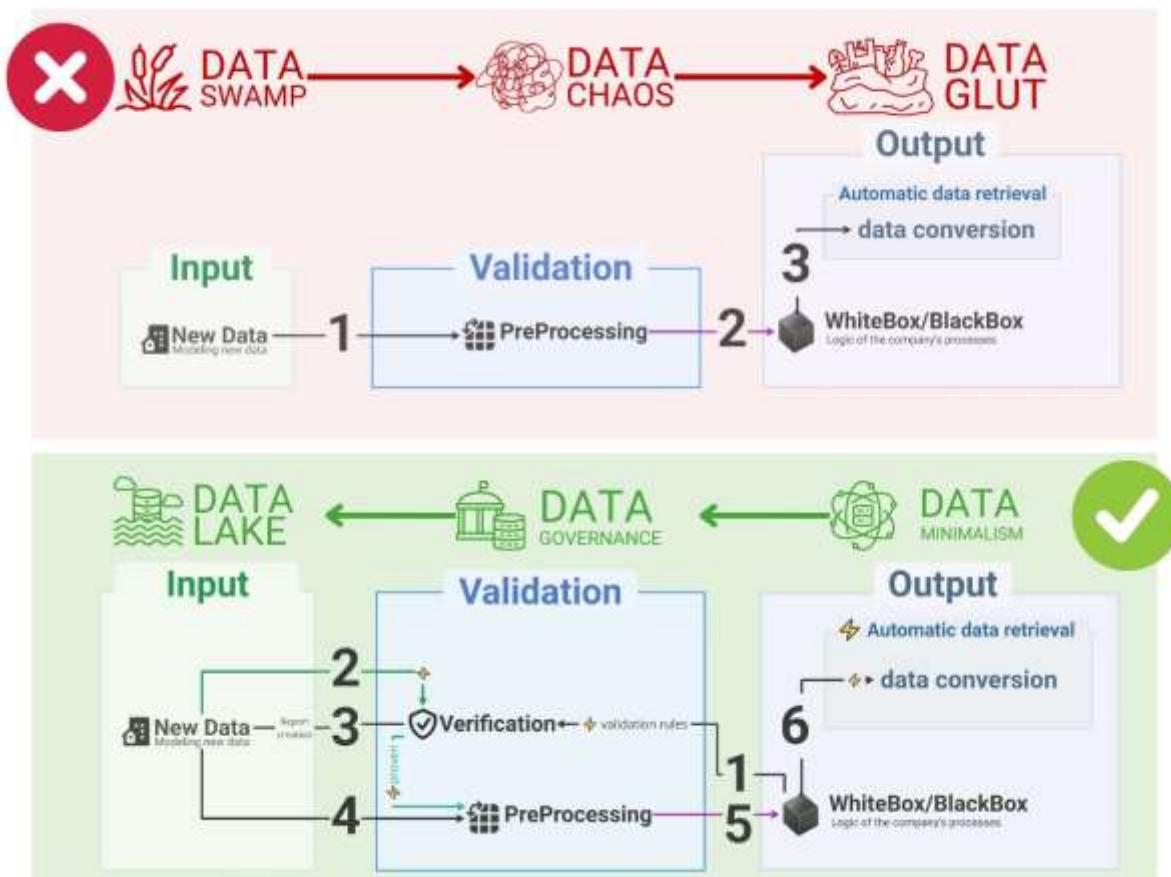


Рис. 8.24 Деректер қоймасында тәртіпсіздікті болдырмау үшін деректерді құру процесін атрибуттарға талаптарды жинаудан бастау қажет.

Деректерді басқару принциптерін және деректер минимализмін деректерді басқару процестеріне интеграциялау, сондай-ақ деректер қоймаларын деректер батпағына айналдыруды белсенді түрде болдырмау арқылы үйімдар өз деректерінің әлеуетін барынша пайдалана алады.

Деректермен жұмыс істеудің әволюциясындағы келесі кезең – басқару және минимализм мәселелерін шешудан кейін автоматты өңдеуді стандарттау, сапаны қамтамасыз ету және деректерді талдау, түрлендіру және шешім қабылдау үшін ыңғайлы ететін әдістерді енгізу. Дәл осы мәселелермен DataOps және VectorOps әдістемелері айналысады, олар үлкен көлемдегі ақпаратпен және машиналық оқытумен жұмыс істейтін компаниялар үшін маңызды құралдарға айналуда.

DataOps ЖӘНЕ VectorOps: ДЕРЕКТЕРМЕН ЖҰМЫС ИСТЕУДІҢ ЖАҢА СТАНДАРТТАРЫ

Data Governance деректерді бақылау және үйімдастырумен айналысса, DataOps олардың дәлдігін,

келісімділігін және компания ішінде үздіксіз қозғалысын қамтамасыз етеді. Бұл құрылым саласындағы бірқатар бизнес-кейстер үшін әсіресе маңызды, өйткені деректер үздіксіз генерациялаңып, уақытында өндөлуді талап етеді. Мысалы, ақпараттық модельдер, жобалық талаптар және аналитикалық есептер әртүрлі жүйелер арасында бір жұмыс күні ішінде синхрондалуы қажет болған жағдайларда, DataOps рөлі шешуші болуы мүмкін. Ол деректерді өңдеу үшін түрақты және қайталанатын процестерді құруға мүмкіндік береді, кешіктірuler мен ақпараттың өзектілігін жоғалту қаупін азайтады.

Data Governance деңгейінде басқару жеткіліксіз – деректер тек сақталмай, күнделікті операцияларда белсенді түрде пайдаланылуы маңызды. Дәл осы жерде DataOps алға шығады – автоматтандыруға, интеграцияға және деректердің үздіксіз ағынын қамтамасыз етуге бағытталған әдістеме.

DataOps ұйымдардағы деректер ағындарын ынтымақтастықты, интеграцияны және автоматтандыруды жақсартуға бағытталған. DataOps практикаларын енгізу деректердің дәлдігі, келісімділігі және қолжетімділігін арттыруға ықпал етеді, бұл деректерге бағытталған қосымшалар үшін өте маңызды.

DataOps әкожүйесіндегі негізгі құралдар – Apache Airflow (Сурет 7.44) – жұмыс процестерін оркестрациялау үшін, және Apache NiFi (Сурет 7.45) – деректер ағындарын маршрутизациялау және түрлендіру үшін. Бұл технологиялар бірге деректерді автоматты өңдеу, бақылау және жүйелер арасында интеграцияны қамтамасыз ететін икемді, сенімді және масштабалатын деректер конвейерлерін құруға мүмкіндік береді (толығырақ «Автоматты ETL-конвейер» бөлімінде). DataOps тәсілін құрылым процестеріне енгізгенде төрт негізгі аспекттің ескеру маңызды:–

- Адамдар мен құралдар деректерден маңызды: деректердің бөлшектенген қоймалары басты проблема ретінде қарастырылуы мүмкін, алайда шын мәнінде жағдай қүрделірек. Деректердің фрагментациясынан басқа, командалардың оқшаулануы және олардың пайдаланатын құралдарының әртүрлілігі маңызды рөл атқарады. Құрылым саласында деректермен жұмыс істейтін мамандар әртүрлі бағыттардан: деректер инженерлері мен аналитиктер, BI және визуализация командалары, сондай-ақ жобаларды басқару және сапа бойынша сарапшылар. Әрқайсысының өз жұмыс әдістері бар, сондықтан деректердің қатысушылар арасында еркін берілуін қамтамасыз ететін әкожүйені құру маңызды факторға айналады, бұл ақпараттың біртұтас, келісілген нұсқасын қамтамасыз етеді.
- Тестілеуді автоматтандыру және қателерді анықтау: құрылым деректері әрқашан қателіктерді қамтиды, бұл модельдердегі дәлсіздіктер, есептеулердегі қателер немесе ескірген спецификациялар болуы мүмкін. Деректерді түрақты тестілеу және қайталанатын қателерді жою олардың сапасын айтарлықтай арттырады. DataOps аясында деректердің дұрыстығын бақылау және тексеру үшін автоматтандырылған механизмдерді енгізу қажет, олар деректердің дұрыстығын қадағалап, қателерді талдайды және зандылықтарды анықтайды, сондай-ақ әр жұмыс цикліне жүйелік ақауларды тіркеп, жояды. Тексерудің автоматтандыру деңгейі негұрлым жоғары болса, деректердің жалпы сапасы соғұрлым жоғары болады және соңғы кезеңдердегі қателердің ықтималдығы соғұрлым төмен болады.
- Деректер бағдарламалық код сияқты тестіленуі тиіс: құрылым қосымшаларының көпшілігі деректерді өңдеуге негізделген, алайда олардың бақылауы жиі екінші рөлде қалады. Егер машиналық оқыту модельдері дәлсіз деректерге оқытылса, бұл қате болжамдарға және қаржылық шығындарға әкеледі. DataOps аясында деректер бағдарламалық код сияқты

мүқият тексеруден өтуі тиіс: логикалық тексерулер, стресс-тесттер, кіріс мәндері өзгерген кезде модельдердің мінез-құлқын бағалау. Тек тексерілген және сенімді деректер басқарушылық шешімдер қабылдау үшін негіз ретінде пайдаланылуы мүмкін.

- Деректердің бақылауы өнімділікке зиян келтірмей: деректерді мониторингтеу тек метрикаларды жинау емес, сапаны басқарудың стратегиялық құралы. DataOps тиімді жұмыс істеуі үшін бақылау деректермен жұмыс істеудің барлық кезеңдерінде, жобалаудан бастап пайдалану кезеңіне дейін енгізілуі тиіс. Сонымен қатар, мониторинг жүйені баяулатпау маңызды. Құрылыштың жобаларында деректерді жинау ғана емес, сонымен қатар бұл жұмысты (мысалы, жобалаушылар) жүргізетін мамандардың жұмысын ешқандай бұзбайтындей етіп жасау да маңызды. Бұл баланс деректердің сапасын өнімділікке зиян келтірмей бақылауға мүмкіндік береді.

DataOps – деректер мамандары үшін қосымша жүктеме емес, олардың жұмысының негізі. DataOps енгізу арқылы құрылыштық компаниялары деректерді басқаруда хаотикалық басқарудан тиімді экожүйеге өтуді қамтамасыз ете алады, онда деректер бизнес үшін жұмыс істейді.

VectorOps, өз кезегінде, DataOps әволюциясының келесі кезеңі болып табылады, ол көпөлшемді векторлық деректерді өңдеуге, сақтауға және талдауға бағытталған (бұл алдыңғы тарауда талқыланды). Бұл цифрлық егіздер, нейрондық желі модельдері және семантикалық іздеу сияқты салаларда, олар құрылыштық саласына енуде, әсіресе өзекті. VectorOps векторлық деректер базаларына сүйенеді, олар көпөлшемді объектілердің тиімді сақталуын, индекстеуін және іздеуін қамтамасыз етеді.

VectorOps – DataOps-тан кейінгі қадам, құрылыштық саласында векторлық деректерді өңдеуге, талдауға және пайдалануға бағытталған. DataOps-тан айырмашылығы, ол деректердің ағынына, келісімділігіне және сапасына назар аударады, VectorOps көпөлшемді объектілердің көріністерін басқаруға назар аударады, олар машиналық оқыту үшін қажет.

Дәстүрлі тәсілдерден айырмашылығы, VectorOps объектілердің дәл сипаттамасын алуға мүмкіндік береді, бұл цифрлық егіздер, генеративті жобалау жүйелері және CAD деректеріндегі қателерді автоматты түрде анықтау үшін критикалық маңызды. DataOps және VectorOps-тың бірлескен енгізілүі класикалық кестелерден семантикалық жағынан бай кеңістіктік модельдерге дейін үлкен көлемдегі ақпаратпен масштабталатын, автоматтандырылған жұмыстың берік негізін қалыптастырады.

КЕЛЕСІ ҚАДАМДАР: ХАОТТЫ САҚТАУДАН ҚҰРЫЛЫМДЫ САҚТАУ ОРНЫНА ӨТУ

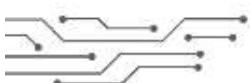
Дәстүрлі құрылыштық деректерін сақтау тәсілдері жиі маңызды ақпараттың талдау мен шешім қабылдауға қолжетімсіз болуына әкелетін "ақпарат силостарының" пайда болуына себеп болады. Заманауи деректерді сақтау концепциялары, мысалы, Data Warehouse, Data Lake және олардың гибридтері, шашыраңқы ақпаратты біріктіруге және оны деректерді ағынды өңдеу мен бизнес-аналитика үшін орталықтандырылған түрде қолжетімді етуге мүмкіндік береді. Деректерді сақтау үшін қолайлы архитектуралы таңдау ғана емес, сонымен қатар деректерді басқару (Data Governance) және деректер минимализмі (Data Minimalism) принциптерін енгізу маңызды, бұл хранилищалардың бақылаусыз "деректер батпақтарына" (Data Swamp) айналуын болдырмауға көмектеседі.

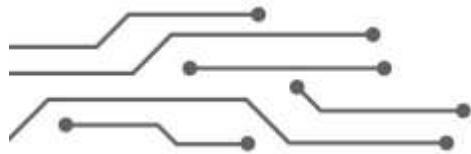
Осы бөлімнің қорытындысында, қарастырылған концепцияларды құнделікті міндеттеріңізде қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

- Деректерді сақтау үшін тиімді форматтарды таңдаңыз
- CSV және XLSX форматтарынан үлкен деректер көлемін сақтау үшін тиімдірек форматтарға (Apache Parquet, ORC) өтіңіз
- Өзгерістерді бақылау үшін деректердің нұсқаларын енгізіңіз
- Ақпараттың құрылымы мен шығу тегін сипаттау үшін метадеректерді пайдаланыңыз
- Компанияның біртұтас деректер архитектурасын жасаңыз
- Деректерді сақтау архитектураларын салыстырыңыз: RDBMS, DWH және Data Lake. Масштабталу, дереккөздерді интеграциялау және аналитикалық өндеу бойынша міндеттеріңізге ең жақсы жауап беретінін таңдаңыз
- Әртүрлі дереккөздерден деректерді алу, жүктеу және түрлендіру (ETL) процестерінің картасын жобалаңыз. Кілттік қадамдар мен интеграция нұктелерін визуализациялау үшін Miro, Lucidchart немесе Draw.io сияқты визуализация құралдарын пайдаланыңыз
- Data Governance және Data Minimalism практикаларын енгізіңіз
- Data Minimalism тәсілін ұстаныңыз – тек шын мәнінде құнды ақпаратты сақтаңыз және өндешіңіз
- Data Governance принциптерін енгізіңіз – деректер үшін жауапкершілікті анықтаңыз, сапаны және ашықтықты қамтамасыз етіңіз
- Деректерді басқару саясаты мен DataOps, VectorOps концепциялары туралы көбірек біліңіз
- DataOps аясында деректер сапасының критерийлерін және тексеру процедуralарын анықтаңыз

Жақсы үйімдастырылған деректерді сақтау компанияның аналитикалық процестерін орталықтандыру үшін негіз жасайды. Файлдарды хаотикалық жинақтаудан құрылымдық хранилишаларға көшу ақпаратты стратегиялық активке айналдырып, негізделген шешімдер қабылдауға және бизнес-процессердің тиімділігін арттыруға көмектеседі.

Деректерді жинау, түрлендіру, талдау және құрылымдық сақтау процестері автоматтандырылып, стандартталғаннан кейін, цифрлық трансформацияның келесі кезеңі үлкен деректермен (Big Data) толыққанды жұмыс істеу болады.





ІХ БӨЛІМ

ҮЛКЕН ДЕРЕКТЕР, МАШИНАЛЫҚ ОҚУ ЖӘНЕ БОЛЖАУ

Тоғызыншы бөлім үлкен деректерге, машиналық оқытуға және құрылыш саласындағы болжаяуға арналған. Мұнда интуитивті шешім қабылдаудан тарихи деректерге негізделген объективті талдауға көшу қарастырылады. Практикалық мысалдар арқылы құрылыш саласындағы үлкен деректерді талдау көрсетіледі – Сан-Францискодың құрылыш рұқсаттары деректер жиынтығын талдаудан CAD жобаларындағы миллиондаған элементтерді өндөуге дейін. Машиналық оқыту әдістеріне, құрылыш жобаларының құны мен мерзімдерін болжаяуға ерекше назар аударылады, сыйықтық регрессия және к-жаңашыл көршілер алгоритмдерінің егжей-тегжейлі талдауы беріледі. Құрылымдалған деректердің болжамдық модельдер үшін негіз болып, тәуекелдерді бағалауға, ресурстарды оңтайландыруға және жобаларды басқарудың тиімділігін арттыруға мүмкіндік беретіні көрсетіледі. Бөлім сондай-ақ деректердің репрезентативті үлгілерін таңдау бойынша ұсыныстарды қамтиды және тиімді талдау үшін әрдайым үлкен ақпарат массивтері қажет емес екенін түсіндіреді.

ТАРАУ 9.1. УЛКЕН ДЕРЕКТЕР ЖӘНЕ ОЛАРДЫҢ ТАЛДАНУЫ

Улкен деректер құрылыс саласында: интуициядан болжамдылыққа.

"Улкен деректер" термині нақты анықтамасы жок. Бастапқыда бұл концепция ақпарат көлемі дәстүрлі өңдеу әдістерінің мүмкіндіктерінен асып кеткен кезде пайда болды. Бүгінгі күні көптеген салаларда, соның ішінде құрылыс саласында, деректердің көлемі мен күрделілігі соншалықты артты, олар компьютерлердің локалды жадында сақталмайды және оларды өңдеу үшін жаңа технологияларды қолдануды талап етеді.

Улкен деректермен жұмыс істеудің мәні тек сақтау мен өңдеуде ғана емес, сонымен қатар болжау мүмкіндігінде. Құрылыс саласында Big Data интуитивті шешімдерден, кестелер мен визуализациялардың субъективті интерпретациясына негізделген, нақты бақылаулар мен статистикамен расталған негізделген болжамдарға жол ашады.

Жиі кездесетін пікірге қарамастан, үлкен деректермен жұмыс істеудің мақсаты "машиналы адам сияқты ойлауға мәжбүрлеу" емес, деректер массивтерін талдау үшін математикалық модельдер мен алгоритмдерді қолдану, заңдылықтарды анықтау, оқығаларды болжау және процестерді оңтайландыру болып табылады.

Улкен деректер (Big Data) – бұл алгоритмдерден тұратын сұық әлем емес, адам әсерінен айырылған. Керісінше, үлкен деректер біздің инстинктімізben, қателіктерімізben және шығармашылығымызben тығыз байланысты. Адам ойлауының кемшілігі ерекше шешімдер табуға және серпілістер жасауға мүмкіндік береді.

Цифрлық технологиялардың дамуы арқасында құрылыс саласында IT саласынан келген деректерді өңдеу әдістері белсенді түрде қолданылуда. Pandas және Apache Parquet сияқты құралдардың арқасында құрылымдалған және құрылымдалмаған деректер біріктіріліп, ақпаратқа қолжетімділікті жеңілдетеді және талдаудағы шығындарды азайтады, ал құжаттар немесе CAD жобаларынан алынған үлкен деректер жиынтықтары (Сурет 9.210 - Сурет 9.212) жобаның өмірлік циклі кезеңдерінің барлығында деректерді жинауға, талдауға және болжауға мүмкіндік береді.–

Улкен деректер құрылыс саласына трансформациялық әсер етеді, оны әртүрлі аспектілерде ықпал етеді. Big Data технологияларын қолдану бірнеше негізгі бағыттарда нәтижелер береді, соның ішінде:

- **Инвестициялық әлеуетті талдау – алдыңғы обьектілердің деректері негізінде жобалардың кірістілігі мен қайтарым мерзімдерін болжау.**
- **Болжамдық қызмет көрсету – жабдықтың ықтимал ақауларын олардың нақты пайда болуынан бұрын анықтау, бұл тоқтап қалу уақытын азайтады.**
- **Жеткізу тізбегін оңтайландыру – ақауларды болжау және логистиканың тиімділігін арттыру.**
- **Энергия тиімділігін талдау – төмен энергия тұтыну деңгейі бар ғимараттарды жобалауға көмек.**

- Қауіпсіздікті мониторингтеу – құрылым алаңындағы жағдайларды бақылау үшін датчики мен киуге болатын құрылғыларды пайдалану.
- Сапаны бақылау – технологиялық нормаларға сәйкестікті нақты уақыт режимінде бақылау.
- Еңбек ресурстарын басқару – өнімділікті талдау және кадрлық қажеттіліктерді болжау.

Құрылым саласында деректер аналитикасы мен болжау қажеттілігі жоқ сала табу қыын. Болжау алгоритмдерінің басты артықшылығы – олардың деректер жинақталған сайын өзін-өзі үйрету және үнемі жетілдіру қабілеті.

Алдағы уақытта жасанды интеллект құрылымшыларға тек көмек көрсетіп қана қоймай, жобалау процестерінен бастап ғимараттарды пайдалану мәселелеріне дейінгі маңызды шешімдерді қабылдайтын болады.

Болжаулар қалай қалыптасатыны және оқытылатын модельдер қалай пайдаланылатыны туралы кітаптың келесі бөлімінде: «Машиналық оқыту және болжаулар» баяндалады.

Улken деректермен толыққанды жұмыс істеуге көшу үшін аналитикаға деген көзқарасты өзгерту қажет. Егер бұрынғы классикалық жүйелерде себеп-салдарлық байланыстарға басты назар аударылса, онда үлken деректер аналитикасында статистикалық заңдылықтар мен корреляцияларды іздеуге акцент жасалады, бұл жасырын өзара байланыстарды анықтауға және обьектілердің мінез-құлқын толық түсінбей-ақ болжауға мүмкіндік береді.

Улken деректердің тиімділігі туралы мәселе: корреляция, статистика және деректерді іріктеу.

Дәстүрлі түрде құрылым субъективті гипотезалар мен жеке тәжірибеге сүйенді. Инженерлер материалдың қалай әрекет ететінін, конструкцияның қандай жүктемелерді көтеретінін және жобаның қанша уақытқа созылатынын белгілі бір ықтималдықпен болжады. Бұл болжамдар практикада тексерілді, көбінесе уақыт, ресурстар және болашақтағы қауіптер бағасына.

Улken деректердің пайда болуымен көзқарас түбегейлі өзгеруде: шешімдер енді интуитивті болжамдарға емес, ауқымды ақпарат массивтерін талдауға негізделеді. Құрылым интуиция өнерінен болжамдар дәлдігіне айналуда.

Улken деректерді пайдалану идеясына көшу, сөзсіз, деректер көлемінің қаншалықты маңызды екендігі және сенімді болжау аналитикасы үшін қанша ақпараттың шын мәнінде қажет екендігі туралы маңызды сұрақты көтереді. «Деректер неғұрлым көп болса, дәлдік соғұрлым жоғары» деген кең таралған пікір статистика түрғысынан әрдайым негізделген емес.

1934 жылды статистик Ежи Нейман дәлелдегендей, статистикалық қорытындылардың дәлдігі кілті деректер көлемінде емес, олардың репрезентативтілігі мен кездейсоқтығында.

Бұл құрылым саласында өсіреле өзекті, мұнда IoT-сенсорлары, сканерлер, бақылау камeralары, дрондар және тіпті әртүрлі CAD-модельдері арқылы үлken көлемдегі деректер жиналады, бұл

«соқыр жерлердің», деректердің шығуларын және бүрмалануын туынданту қаупін арттырады.

Жол жабыны жағдайын мониторингтеу мысалын қарастырайық. Барлық жол участкелері туралы толық деректер жиынтығы X Гб көлемді алғып, өңдеуге шамамен бір күн уақыт қажет етеді. Сонымен қатар, тек әр 50-ші участкені қамтитын кездейсоқ таңдау тек X/50 Гб көлемді алады және жарты сағатта өндөледі, бұл белгілі бір есептеулер үшін бағалаудың аналогиялық дәлдігін қамтамасыз етеді (Рис. 9.11).

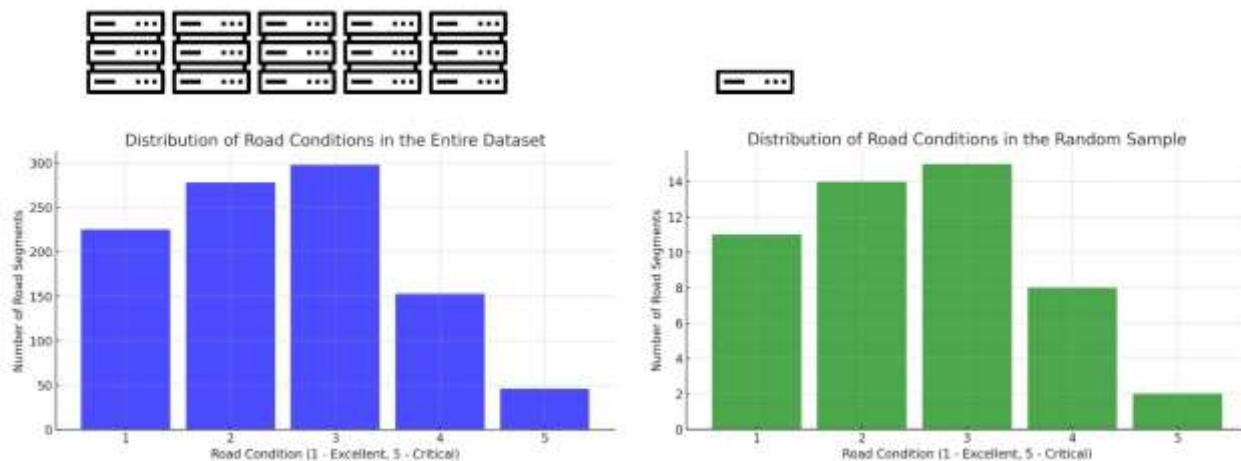


Рис. 9.11 Жол жабыны жағдайының гистограммалары: толық деректер жиынтығы мен кездейсоқ таңдау бірдей нәтижелерді көрсетеді.

Осылайша, деректерді сәтті талдаудың негізгі факторы көбінесе олардың көлемі емес, таңдау ре-презентативтілігі мен қолданылатын өңдеу әдістерінің сапасы болуы мүмкін. Кездейсоқ таңдау мен таңдаулы тәсілге көшу құрылыш индустрисында ойлау тәсілін өзгертуге талап етеді. Тарихи тұрғыдан алғанда, компаниялар «деректер неғұрлым көп болса, соғұрлым жақсы» деген логикаға сүйенди, барлық мүмкін көрсеткіштерді қамту максималды дәлдікті қамтамасыз етеді деп есептеді.

Бұл тәсіл жобаларды басқарудағы танымал жаңсақ пікірге үқсас: «мен неғұрлым көп мамандарды тартсам, жұмыс соғұрлым тиімді болады». Алайда, кадрлармен болғандай, сапа мен құралдар салына қарағанда маңыздырақ. Деректер немесе жобаның қатысуышылары арасындағы өзара байланыстарды (корреляцияларды) ескермей, көлемнің өсуі тек шу, бүрмалаулар, дублирлеу және негізсіз шығуларға әкелуі мүмкін.

Нәтижесінде, көбінесе сапалы дайындалған, тұрақты және негізделген болжамдар беруге қабілетті, бірақ көлемі аз деректер жиынтығын ұстау әлдеқайда өнімдірек болады, хаотикалық, бірақ мас-сивті ақпаратқа сүйенуден гөрі, онда көптеген қарама-қайшы сигналдар бар.

Артық деректер көлемі тек үлкен дәлдікті қамтамасыз етпейді, сонымен қатар шу, артық белгілер, жасырын корреляциялар және релятивті емес ақпараттың болуы салдарынан бүрмаланған қорытындыларға әкелуі мүмкін. Мұндай жағдайда модельдердің артық оқытылу қаупі артады және аналитикалық нәтижелердің сенімділігі төмендейді.

Құрылыш саласында үлкен деректермен жұмыс істеудегі басты қыындық - деректердің оңтайлы саны мен сапасын анықтау. Мысалы, бетон конструкцияларының жағдайын мониторингтеу

кезінде мыңдаған датчиктерді пайдалану және ақпаратты әр минут сайын жинау сақтау және талдау жүйесін жүктеу мүмкін. Алайда, корреляцияны талдау жүргізіп, 10% ең ақпаратты датчиктерді таңдау арқылы болжамдардың дәлдігін іс жүзінде бірдей деңгейде алуға болады, ресурстарды бірнеше есе, кейде ондаған және жұздеген есе аз жұмысай отырып.

Кішірек деректер жиынтығын пайдалану қажетті сақтау көлемін және өндедеу уақытын қысқартады, бұл деректерді сақтау мен талдауға жұмысалатын шығындарды едәуір азайтады және көбінесе кездейсоқ таңдау болжамдық аналитика үшін, әсіресе үлкен инфрақұрылымдық жобаларда немесе нақты уақыт режимінде жұмыс істегендегі, тамаша шешім болып табылады. Соңында, құрылыш процестерінің тиімділігі жиналған деректердің көлемімен емес, олардың талдау сапасымен анықтады. Сындарлы көзқарас пен мұқият талдау болмаса, деректер қате қорытындыларға әкелуі мүмкін.

Белгілі бір деректер көлемінен кейін әр жаңа ақпарат бірлік пайдалы нәтижені береді. Шексіз ақпарат жинаудың орнына, оның репрезентативтілігі мен талдау әдістеріне назар аудару маңызды (Рис. 9.22).

Бұл феномен Аллен Уоллис тарапынан жақсы сипатталған, ол АҚШ Әскери-теңіз күштерінің екі баллистикалық снаряд конструкцияларын тестілеу мысалында статистикалық әдістерді қолдануды көрсетеді.

Әскери-теңіз күштері екі баллистикалық снаряд конструкциясын (A және B) парлы ату сериясын өткізу арқылы тестіледі. Әр раундта A 1 немесе 0 алады, B-мен салыстырғанда оның сипаттамалары жақсы немесе нашар болғанына байланысты, және көрініше. Стандартты статистикалық тәсіл фиксиленген сынақтар санын (мысалы, 1000) өткізіп, жеңімпазды пайыздық үлестіру негізінде анықтауды ұсынады (мысалы, егер A 53%-дан астам жағдайда 1 алса, оны ең жақсы деп санайды). Аллен Уоллис капитан Гаррет Л. Шайлдермен осындағы мәселені талқылағанда, капитан мұндай тесттің, Алленнің әңгімесін келтіре отырып, пайдасты болуы мүмкін екенін айтты. Егер Шайлдер сияқты тәжірибелі және ақылды артиллериялық офицер болса, ол алғашқы бірнеше жүз атудан кейін экспериментті аяқтаудың қажеті жоқ екенін көрер еді, не жаңа әдіс айқын нашар, не ол күтілгеннен айқын асып түседі.— АҚШ-тың Колумбия университетіндегі статистикалық зерттеулер үкіметтік тобы, екінші дүниежүзілік соғыс кезеңі.

Бұл принцип әртүрлі салаларда кеңінен қолданылады. Медицинада, мысалы, жаңа дәрілердің клиникалық зерттеулері кездейсоқ пациенттер тобына жүргізіледі, бұл бүкіл әлемдегі адамдар популляциясында дәріні тестілеудің қажеті жоқ статистикалық маңызды нәтижелер алуға мүмкіндік береді. Экономика мен социологияда қоғамның пікірін көрсету үшін репрезентативті сауалнамалар жүргізіледі, бұл елдің әрбір тұрғынын сұраудың қажеті жоқ.

Мемлекеттер мен зерттеу үйімдары әлеуметтік тенденцияларды түсіну үшін халықтың шағын топтарына сауалнамалар жүргізгендей, құрылыш саласындағы компаниялар жобаларды басқару үшін тиімді мониторинг және болжау жасау үшін кездейсоқ деректерді пайдалана алады.

Үлкен деректер әлеуметтік ғылымдарға қөзқарасты өзгерте алады, бірақ статистикалық сауаттылықты алмастырмайды.

– Томас Ландсалл-Велфэр, «Қазіргі уақытта ұлттың көңіл-күйін болжай», *Significance v. 9(4)*, 2012 жыл.

Ресурстарды үнемдеу түрғысынан, болашақ болжамдар мен шешімдер қабылдау үшін деректер жинау кезінде маңызды сұрақ туындаиды: үлкен деректер массивтерін жинау мен өндөуге едәуір қаражат жұмысаудың мағынасы бар ма, әлде біртіндеп масштабтауға болатын әлдеқайда аз және арзан тесттік деректер жиынтығын пайдалану тиімдірек пе? Кездейсоқ таңдамалардың тиімділігі компанияларға деректерді жинау мен модельдерді оқытуға жұмсалатын шығындарды он есе немесе тіпті мың есе қысқартуға мүмкіндік береді, бұл кешенді қамтуды талап етпейтін, бірақ жеткілікті дәлдік пен репрезентативтілікті қамтамасыз ететін деректер жинау әдістерін таңдау арқылы жүзеге асырылады. Мұндай тәсіл тіпті шағын компанияларға үлкен корпорациялар деңгейінде нәтижелерге қол жеткізуге мүмкіндік береді, бұл шығындарды оңтайландыруға және ресурстары шектеулі болған жағдайда негізделген шешімдер қабылдау процесін жеделдетуге үмтүлатын компаниялар үшін маңызды. Келесі тарауларда қоғамдық деректер жиынтықтарын пайдалана отырып, аналитика мен болжамның мысалдары қарастырылады.

Үлкен деректер: Сан-Франциско дағы миллион құрылыс рұқсатын талдау.

Ашық деректермен жұмыс істеу алдыңғы тарауларда қарастырылған принциптерді тәжірибеде қолдануға ерекше мүмкіндік береді: тиімді белгілерді таңдау, репрезентативті іріктеу, визуализация және сыни талдау. Бұл тарауда біз Сан-Франциско дағы құрылыс белсенділігін зерттеу үшін ашық деректерді, атап айтқанда, құрылыс рұқсаттары бойынша миллионнан астам жазбаларды қалай пайдалануға болатынын қарастырамыз.

Сан-Францисконың "Гимараттар департаменті" [148] ұсынған миллионнан астам рұқсаттар туралы ашық деректер (Рис. 9.12) бізге CSV форматындағы екі деректер жинағын пайдалана отырып, қаладағы құрылыс белсенділігін ғана емес, сонымен қатар 1980 жылдан 2019 жылға дейінгі Сан-Франциско құрылыс саласындағы соңғы тенденциялар мен тарихты сыни талдауға мүмкіндік береді.

Деректер жинағынан визуализациялар жасау үшін қолданылған код мысалдары (Рис. 9.13- Рис. 9.18), сондай-ақ код, түсіндірмелер мен комментарийлермен бірге визуалды графиктерді "Сан-Франциско. Құрылыс секторы 1980–2019 жж." [149] сұрауымен Kaggle платформасынан табуға болады.

Building Permits on or after January 1, 2013

Building Permits before January 1, 2013

June 13, 2020 2,237 views June 13, 2020 880 views

permit_creation_date	description	current_status	current_status_date	filed_date	issued_date	completed_date
07/01/1998	repair stucco	complete	07/07/1998	07/01/1998	07/01/1998	07/07/1998
permit_number	permit_expiration_date	estimated_cost	revised_cost	existing_use	Zipcode	Location
362780	9812394	11/01/1998	780.0	NaN	1 family dwelling	(37.7963468760490, -122.4322641443574)
570817	200412131233	06/13/2005	9000.0	9000.0	apartments	(37.72925858009388, -122.4644245687462)
198411	9202396	09/18/1992	9000.0	NaN	apartments	(37.79506002552974, -122.39593224461805)

Рис. 9.12 Деректер жинақтары әртүрлі объектілердің атрибуттари бар құрылымсұраштырылған туралы ақпараттың қамтиды.

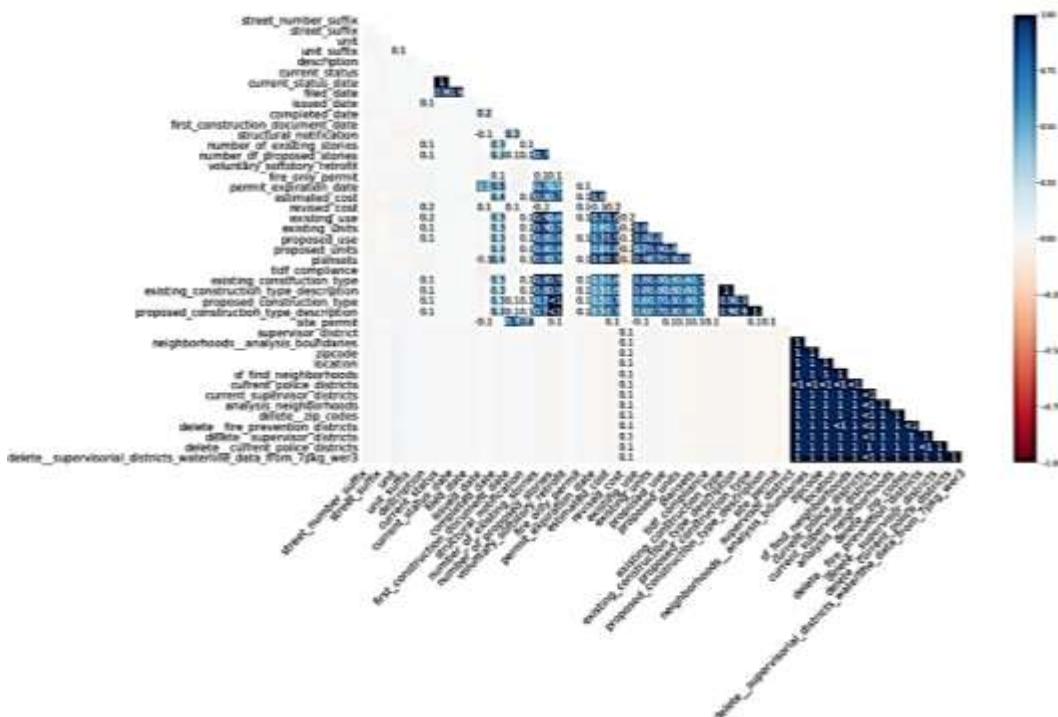


Рис. 9.13 Тепловая карта (Pandas және Seaborn) деректер жинағының барлық атрибуттарын визуализациялад, атрибуттар жүргізуары арасындағы өзара байланысты анықтауда көмектеседі.

Сан-Францисконың ғимараттар департаменті ұсынған кестеден (Рис. 9.12) ешқандай тенденциялар немесе қорытындылар көрінбейді. Кестелік түрде берілген құрғақ сандар шешім қабылдау үшін негіз бола алмайды. Деректерді визуалды түрде түсінікті ету үшін, деректерді визуализациялау туралы тарауларда егжей-тегжейлі талқыланғандай, оларды кітаптың жетінші бөлімінде қарастырылған әртүрлі кітапханалар арқылы визуализациялау қажет.

Pandas DataFrame және Python визуализациялау кітапханаларын пайдалана отырып, 1 137 695 рұқсаттың [148] деректерін талдай отырып, Сан-Францискодағы құрылыш белсенділігінің Кремний алқабындағы технологиялық сектордың экономикалық циклдарымен тығыз байланысты деген қорытынды жасауға болады (Рис. 9.14). -

Экономикалық бүмдер мен құлдыраулар құрылыш жобаларының саны мен құнына елеулі әсер етеді. Мысалы, құрылыш белсенділігінің алғашқы шыңы 1980-ші жылдардың ортасындағы электроника бұмымен сәйкес келді (Pandas және Matplotlib қолданылды), ал кейінгі шыңдар мен құлдыраулар дотком бұмымен және соңғы жылдардағы технологиялық бүммен байланысты болды.

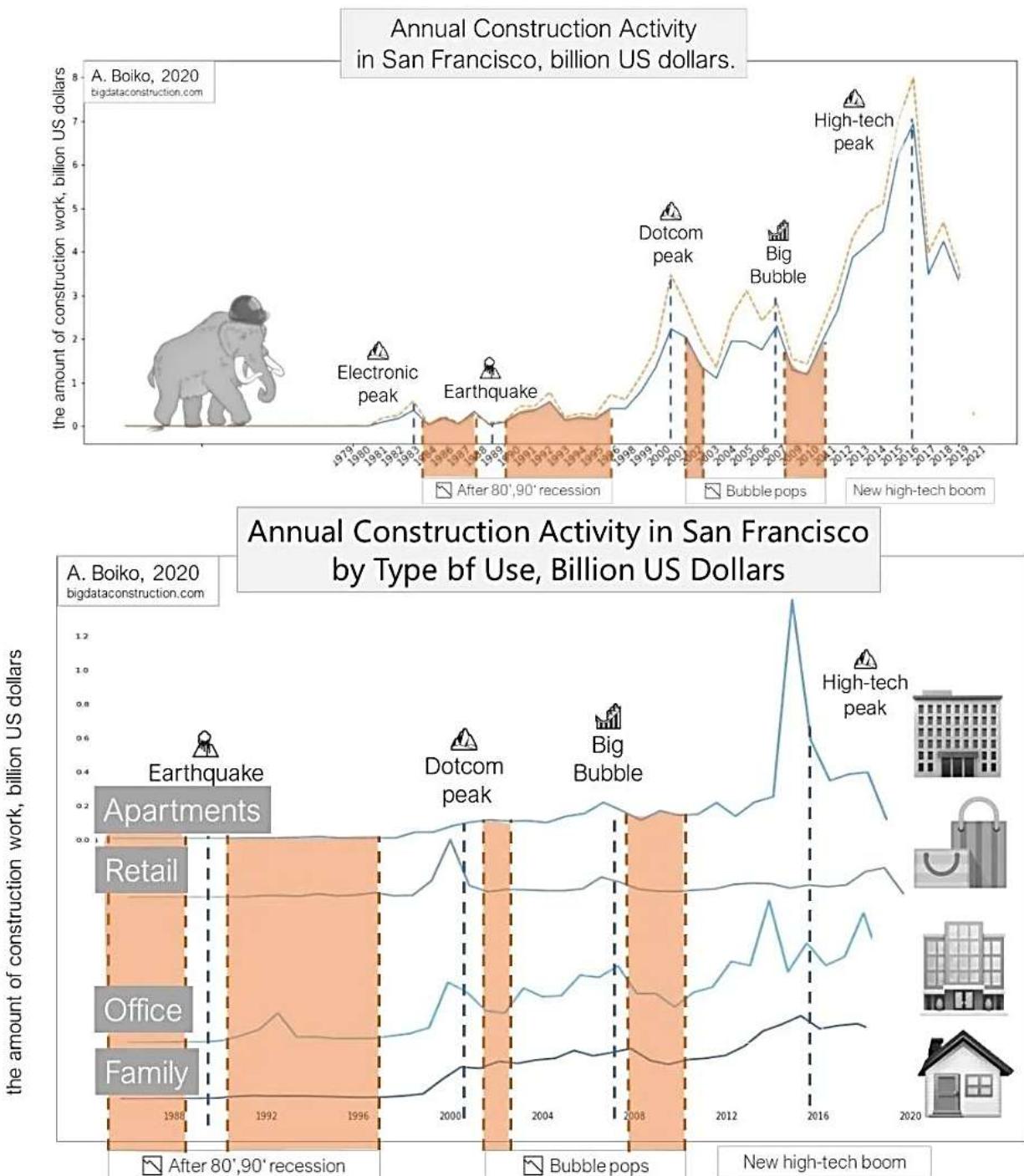


Рис. 9.14 Сан-Францискодағы жылжымайтын мүлік саласында инвестициялар Кремний алқабындағы технологиялық дамуымен корреляцияланады.

Деректер аналитикасы Сан-Францискодағы соңғы онжылдықта құрылыш және қайта құруға салынған 91,5 миллиард доллардың 75%-дан астамы - яғни, 2 км радиустағы орталық қалада шоғырланғанын бағалауға мүмкіндік береді (Рис. 9.15 - Pandas және Folium визуализациялау кітапханасы қолданылды), бұл осы орталық аймақтардағы инвестициялардың жоғары тығыздығын көрсетеді.

Құрылыш рұқсаттарының орташа құны ауданға байланысты айтартылғанда өзгеріп отырады, соның

ішінде қала орталығындағы өтінімдер қала шетіне қарағанда үш есе қымбат, бұл жердің, еңбек күшінің, материалдардың жоғары құнымен және энергия тиімділігін арттыру үшін қымбат материалдарды қолдануды талап ететін қатаң құрылыш нормаларымен байланысты.

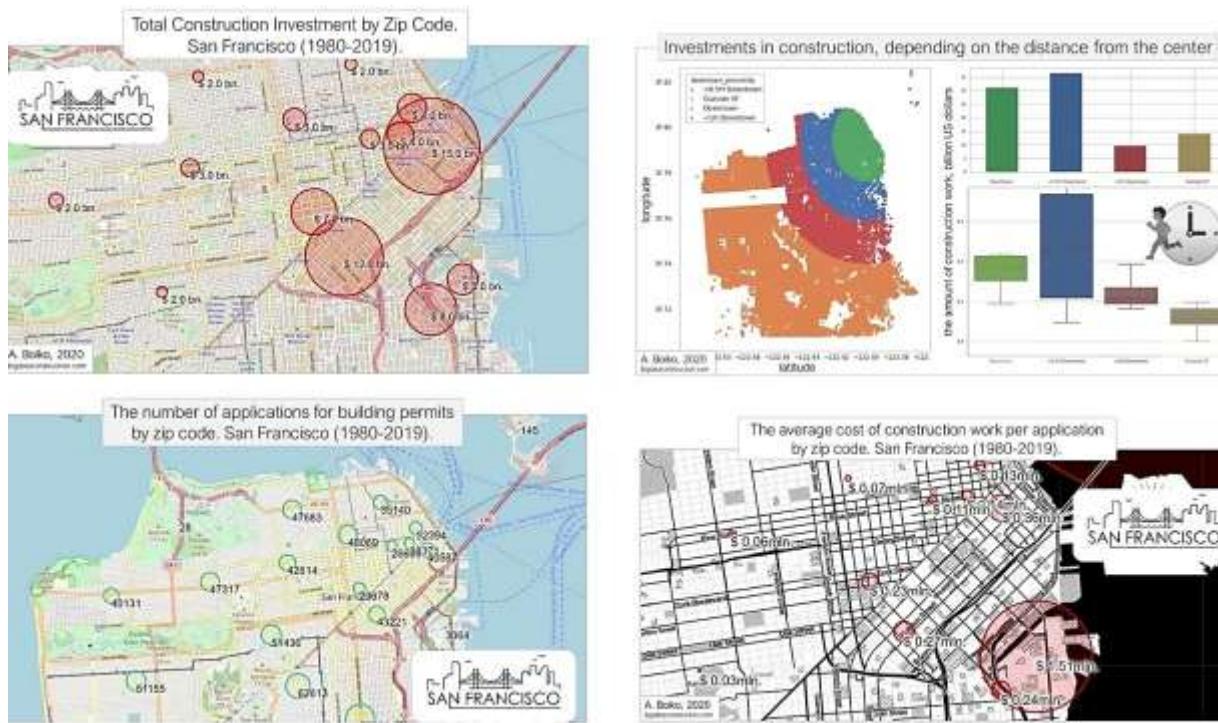


Рис. 9.15 Сан-Францискода құрылышқа салынған инвестициялардың 75 пайызы (91,5 миллиард доллар) қала орталығында шоғырланған.

Деректер жиынтығы тұрғын үйлердің түрлері бойынша ғана емес, сонымен қатар қала аудандары мен жеке мекенжайлар (пошталық индекстер) бойынша орташа жөндеу бағаларын есептеуге мүмкіндік береді. Сан-Францискода тұрғын үй жөндеуінің құны әртүрлі жөндеу түрлері мен тұрғын үйлер үшін айқын тенденцияларды көрсетеді (Рис. 9.16 - Pandas және Matplotlib қолданылды). Асүйді жөндеу ванна бөлмесін жөндеуден айтарлықтай қымбат: бір отбасылық үйдегі асүй жөндеудің орташа құны шамамен 28 000 долларды құрайды, ал екі отбасылық үйде 25 000 доллар.

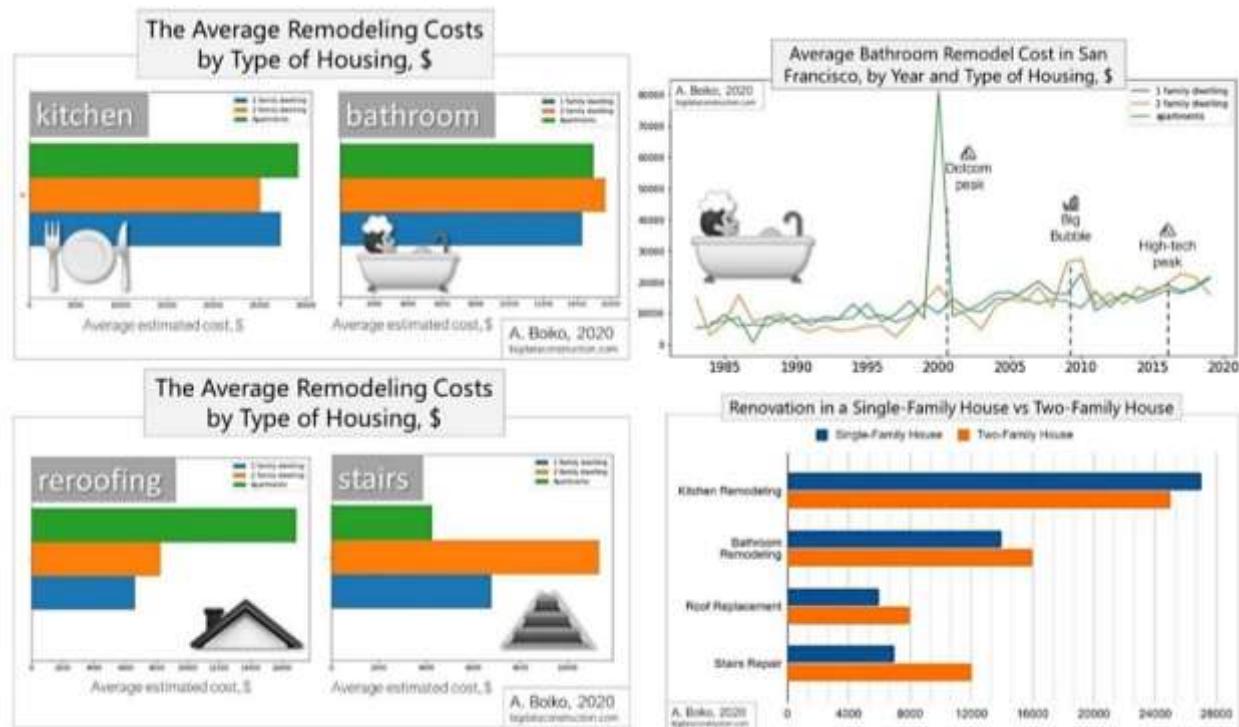


Рис. 9.16 Сан-Францискода асүйлерді жөндеу ванна бөлмелерін жөндеуден шамамен екі есе қымбат, және үй иелері негізгі тұрғын үй жөндеу шығындарын жабу үшін 15 жыл бойы айына 350 доллар жинақтауға мәжбүр.

Сан-Францискода құрылыш құнының инфляциясын жылдар бойы тұрғын үй түрлері мен жылдар бойынша топтастырылған деректерді талдау арқылы бақылауға болады (Рис. 9.17 - Pandas және Seaborn қолданылды), бұл 1990 жылдан бастап жөндеу құнының тұрақты өсуін көрсетеді және көпқабатты үйлердің жөндеу құны бойынша қысқа мерзімді үш жылдық циклдарды анықтайды.

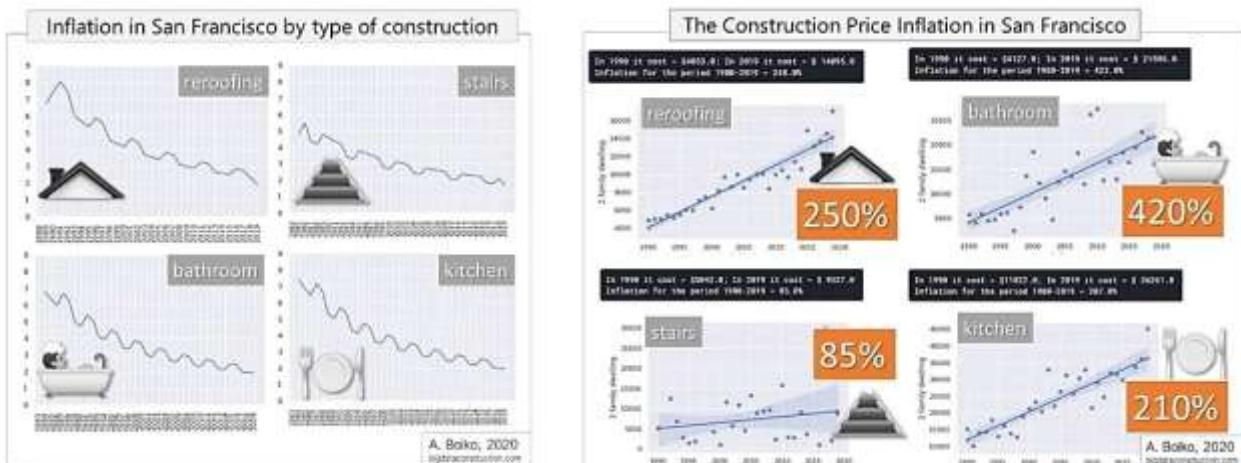


Рис. 9.17 1980 жылдан 2019 жылға дейін Сан-Францискода ванна бөлмелерін жөндеу құны бес есе өсті, ал шатыр мен асүйді жөндеу үш есе қымбаттады, ал баспалдақтарды жөндеу тек 85 пайызға қымбаттады.

Сан-Франциско құрылыш департаментінің ашық деректерін зерттеу (Рис. 9.13) қаладағы құрылыш

Құнының өте өзгермелі және жиі болжанбайтынын, көптеген факторлардың әсеріне үшірайтынын анықтауға мүмкіндік береді. Бұл факторлардың арасында экономикалық өсу, технологиялық инновациялар және әртүрлі түрғын үй түрлерінің ерекше талаптары бар.

Бұрын мұндай талдау жүргізу үшін бағдарламалау мен аналитикада терең білім қажет болатын. Алайда LLM құралдарының пайда болуымен, процесс құрылымы саласындағы мамандардың кең ауқымына, жобалау бөлімдеріндегі инженерлерден бастап компаниялардың жоғары басшылығына дейін қолжетімді және түсінікті болды.

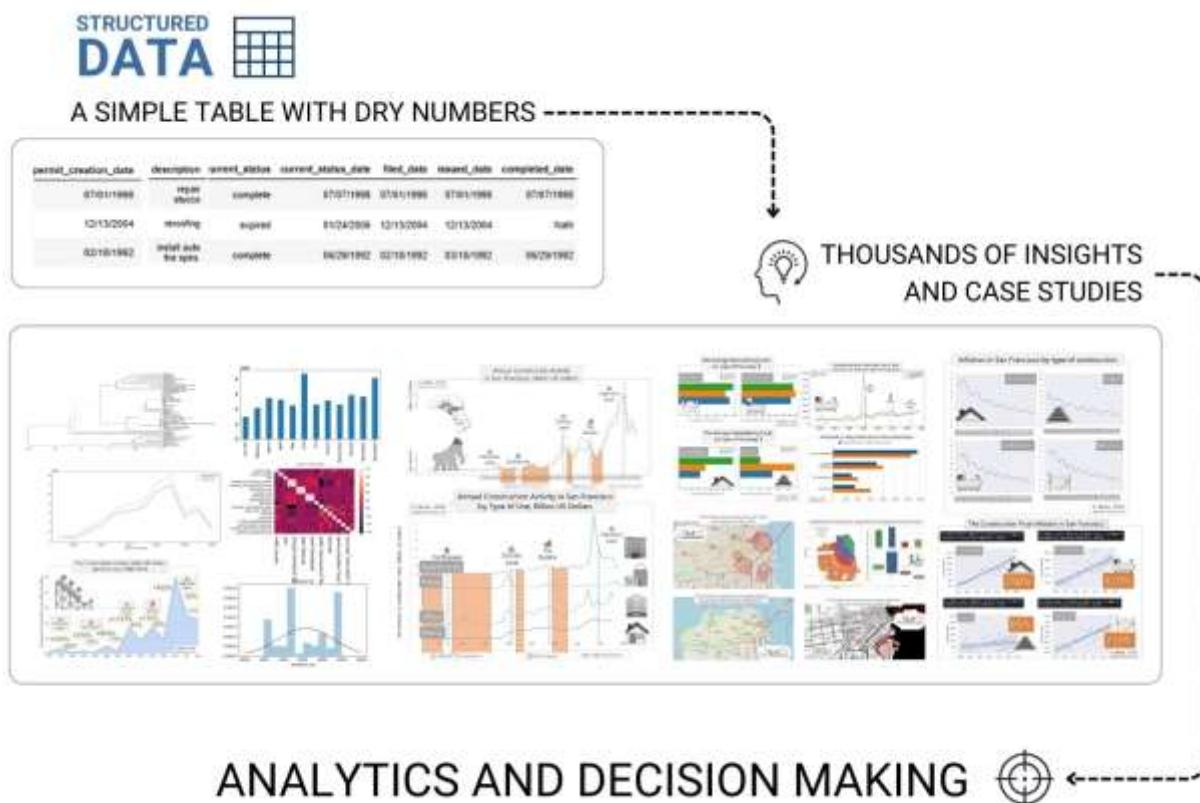


Рис. 9.18 Көрнекі деректерге көшу, жасырын үлгілерді тану арқылы шешім қабылдауды автоматтандыруға мүмкіндік береді.

Сан-Франциско құрылымы басқармасының кестелік деректер жиынтығынан алынған деректерді талдағанымыздай, біз кез келген деректер жиынтықтарын - суреттер мен құжаттардан IoT деректеріне немесе CAD базаларынан алынған деректерге дейін визуализациялап, талдай аламыз.

CAD (BIM) деректеріне негізделген үлкен деректердің мысалы

Келесі мысалда CAD (BIM) құралдарынан алынған деректерді пайдалана отырып, үлкен деректер жиынтығын талдаймыз. Үлкен деректер жиынтығын жинау және құру үшін арнайы автоматтандырылған веб-сканер (скрипт) қолданылды, ол тегін архитектуралық модельдер ұсынатын сайттардан жобалық файлдарды автоматты түрде іздеу және жинау үшін бапталды. Бірнеше күн ішінде сканер 4 596 IFC файлын және 6 471 RVT файлын, сондай-ақ 156 024 DWG файлын сәтті тауып, жүктеп алды[149].

RVT және IFC форматтарындағы жобаларды жинап, оларды тегін SDK-лар арқылы құрылымдық CSV форматына конвертациялағаннан кейін, шамамен 10 мың RVT және IFC жобасы бір үлкен Apache Parquet кестесіне жинақталып, Pandas DataFrame-ге талдау үшін жүктелді (Рис. 9.19).-

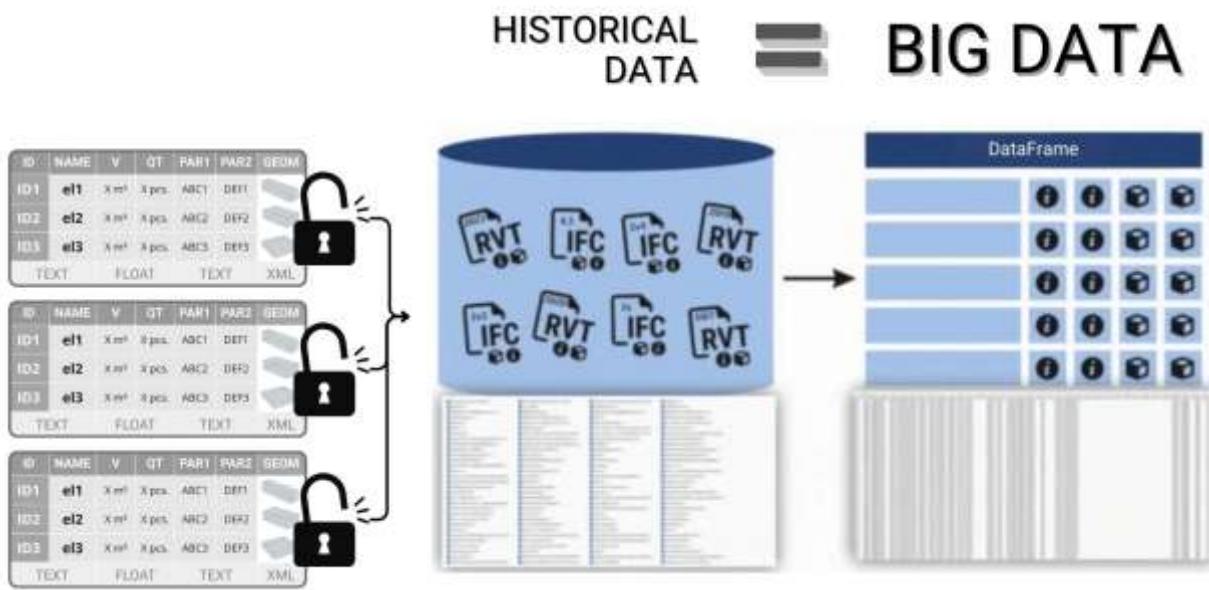


Рис. 9.19 Структурирленген жобалық деректер кез келген жобаларды бір екі өлшемді кестеге біріктіруге мүмкіндік береді.

Бұл ауқымды жинақтағы деректер келесі ақпараттарды қамтиды: IFC файлдарының жиынтығы шамамен 4 миллион элементті (жол) және 24 962 атрибутты (баған) қамтиды, ал RVT файлдарының жиынтығы шамамен 6 миллион элементтен (жол) тұрады және 27 025 түрлі атрибутты (баған) қамтиды.

Бұл ақпараттық жиынтықтар (Рис. 9.110) миллиондаған элементтерді қамтиды, олардың әрқайсысы үшін геометрияның Bounding Box (жобадағы объектінің шекараларын анықтайтын тікбұрыш) координаттары қосынша алынып, жалпы кестеге енгізілді – әр элементтің PNG форматындағы суреттері және DAE (Collada) ашық форматындағы геометриясы жасалды.

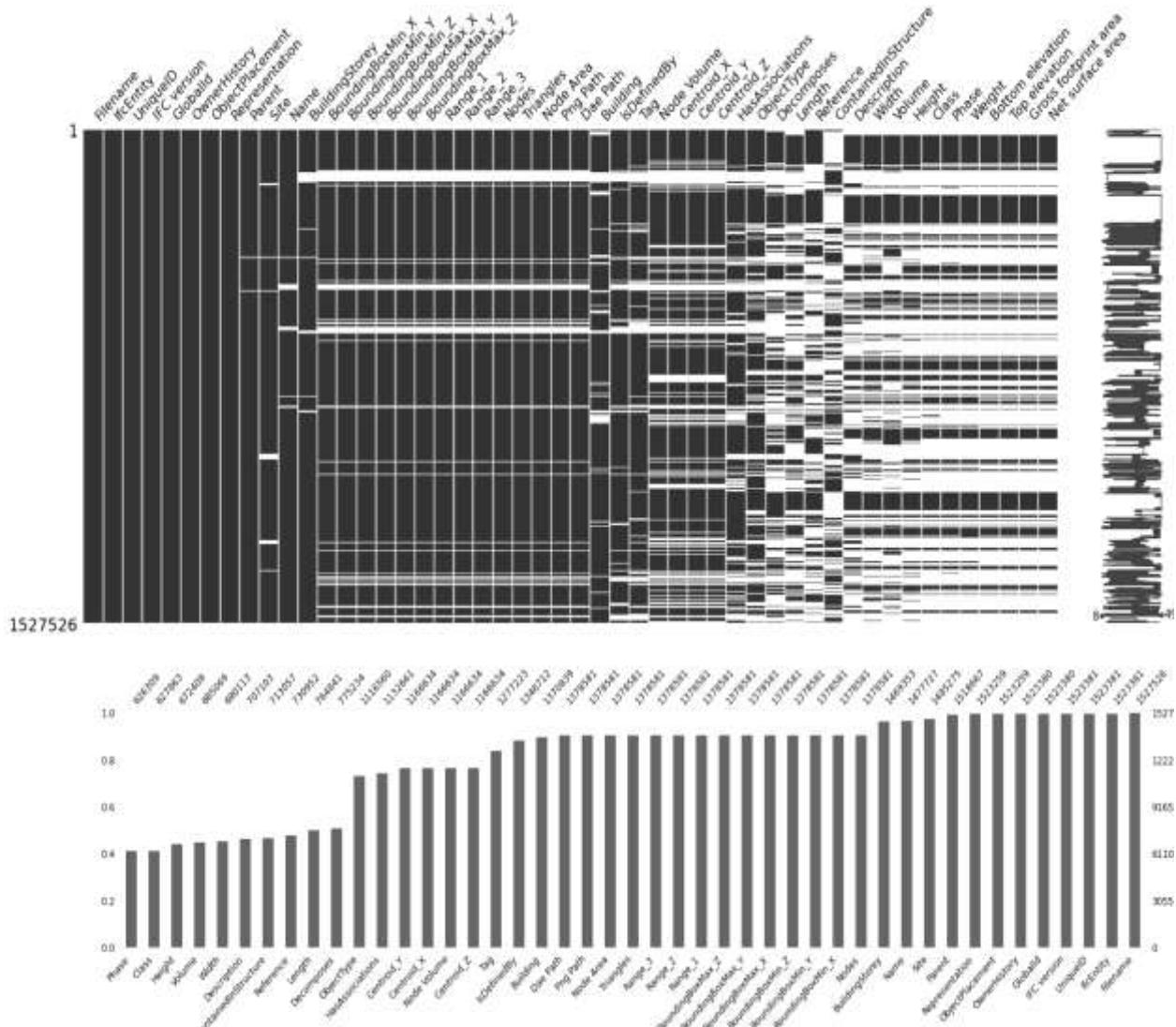


Рис. 9.110 1,5 миллион элементтен тұратын сабсет және алғашқы 100 атрибуттың толтырылу визуализациясы (missingno кітапханасы) гистограмма түрінде.

Осылайша, біз 4 596 IFC жобасынан және 6 471 RVT жобасынан он миллиондаған элементтер туралы барлық ақпаратты алдық, мұнда барлық элементтердің атрибуттары мен олардың геометриясы (Bounding Box) бір кесте (DataFrame) құрылымдық формасына айналдырылды (Рис. 9.110 - да-тағреимнің толтырылу деректері гистограммалар түрінде көрінеді).-

Талдау барысында құрылған гистограммалар (Рис. 9.110, Рис. 9.26, Рис. 9.27) деректердің тығыздығын және бағандардағы мәндердің жиілігін жылдам бағалауға мүмкіндік береді. Бұл атрибуттардың таралуы, бос орындардың болуы және машиналық оқыту модельдерін құру кезінде жеке атрибуттардың әлеуетті пайдалығын зерттеуге алғашқы түсінік береді.-

Бұл деректер массивінің практикалық қолдану мысалдарының бірі (Рис. 9.110) "5000 IFC және RVT жобасы" жобасы болып табылады [149], ол Kaggle платформасында қолжетімді. Оnda деректерді алдын ала өндөу мен талдаудан бастап, Python кітапханаларын – pandas, matplotlib, seaborn, folium

және басқаларын пайдалана отырып, нәтижелерді визуализациялауға дейінгі толық Pipeline шешімі бар Jupyter Notebook ұсынылған (Рис. 9.111).-



Рис. 9.111 CAD (BIM) форматтарынан алынған деректерді Python визуализация кітапханалары мен pandas кітапханасы арқылы талдау мысалдары.

Метаинформация негізінде жобалардың қай қалаларда әзірленгенін анықтауға болады және бұл деректерді картада көрсетуге болады (мысалы, folium кітапханасын пайдалана отырып). Сонымен қатар, деректердегі уақыт белгілері файлдардың сақталу немесе өндеду заңдылықтарын зерттеуге мүмкіндік береді: аптаның күндері, тәулік уақыты және айлар бойынша.

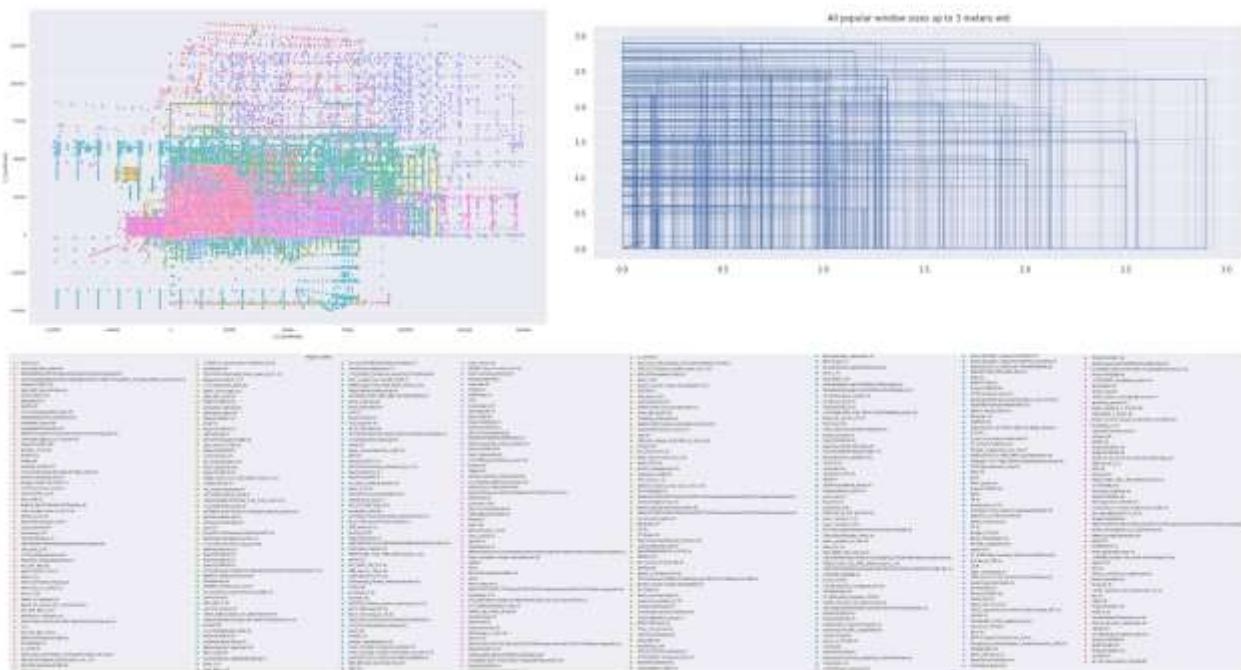


Рис. 9.112 Барлық колонналардың геометриялық орналасуын және 3 метрге дейінгі барлық терезелердің өлшемдерін төменгі графиктегі жобалар тізіміне сәйкес визуализациялау.

Модельдерден алынған Bounding Box түріндегі геометриялық параметрлер агрегаттық талдауға

жатады. Мысалы, Рис. 9.112-де екі график көрсетілген: сол жақтағы график барлық жобалар бойынша колонналар арасындағы қашықтықтардың таралуын нөлдік нүктеге қатысты көрсетеді, ал оң жақтағы график – «Category» параметрі «OST_Windows», «IfcWindows» мәні бойынша топтастырылған он мыңдаған терезе элементтерінің (3 метрге дейінгі) өлшемдерін көрсетеді.

Бұл мысал үшін аналитикалық Pipeline коды мен деректер жиынтығы Kaggle сайтында "5000 жобалар IFC және RVT | DataDrivenConstruction.io" атауымен қолжетімді. Бұл дайын Pipeline мен деректер жиынтығын Kaggle сайтында онлайн режимінде немесе PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse PyDev плагині, Thonny, Wing IDE, IntelliJ IDEA Python плагині, JupyterLab немесе танымал онлайн құралдарда Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker сияқты оффлайн режимінде көшіріп, тегін іске қосуға болады.

Улкен құрылымдалған деректер массивтерін өңдеу және зерттеу нәтижесінде алынған аналитикалық деректер құрылыш саласындағы шешім қабылдау процестерінде шешуші рөл атқарады.

Мұндай ақпаратты талдау арқылы өткен жобалардың деректері негізінде мамандар материалдар мен жұмыс күшіне қажеттіліктерді тиімді болжай алады және құрылыш басталмай тұрып жобалық шешімдерді оңтайландыра алады.

Алайда, егер жобалық деректер немесе құрылыш рұқсаттары салыстырмалы түрде статикалық ақпарат болса, баяу өзгеретін болса, құрылыш процесі әртүрлі датчиктер мен IoT құрылғыларымен қарқынды турде толығады: камeralар, автоматтандырылған мониторинг жүйелері, олар деректерді нақты уақыт режимінде жеткізеді – бұл құрылыш алаңын динамикалық цифрлық ортаға айналдырады, мұнда деректерді нақты уақыт режимінде талдау қажет.

IoT (Заттар Интернеті) және смарт-келісімшарттар

IoT (интернет заттары) цифрлық трансформацияның жаңа толқынын білдіреді, мұнда әрбір құрылғы өз IP-мекенжайына ие болып, жаһандық желінің бөлігіне айналады. IoT – физикалық обьектілерді интернетке қосу арқылы деректерді жинау, өңдеу және беру концепциясы. Құрылыш саласында бұл құрылыш процестерін нақты уақыт режимінде бақылау, материалдардың шығындарын минимизациялау, жабдықтың тозуын болжау және шешім қабылдауды автоматтандыру мүмкіндігін білдіреді.

CFMA мақаласына сәйкес «Қосылған құрылыш арқылы болашаққа дайындалу» [150], келесі онжылдықта құрылыш саласы ауқымды цифрлық трансформациядан өтеді, оның кульминациясы Connected Construction концепциясы – толық интеграцияланған және автоматтандырылған құрылыш алаңы болады.

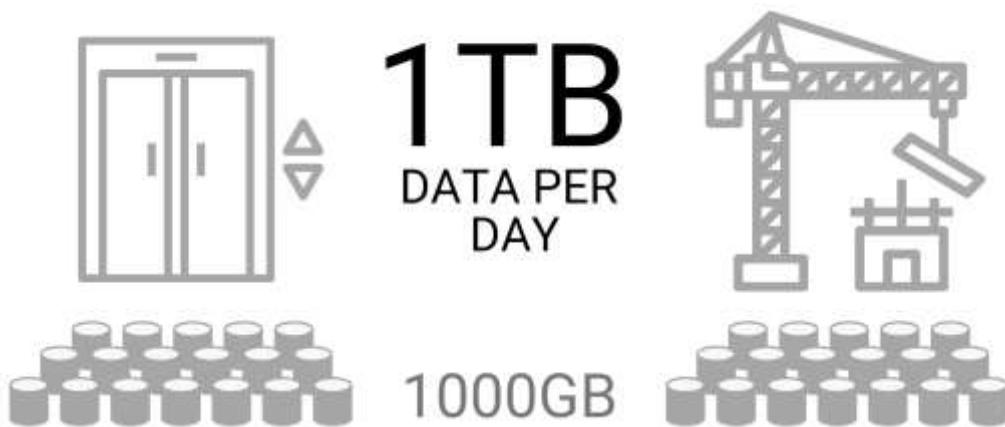


Рис. 9.113 IoT құрылғылары немесе құрылым алаңындағы деректерді беру құрылғылары күніне терабайт деректерді өндіру және беруі мүмкін.

Цифрлық құрылым алаңы барлық құрылым элементтерін – жоспарлаудан, логистикадан, жұмыстарды орындаудан және құрылым алаңындағы сапанды бақылаудан бастап стационарлық камералар мен квадрокоптерлер арқылы – біртұтас динамикалық цифрлық экожүйеге біріктіруді көздейді. Бұрын, кітаптың жетінші бөлімінде, біз Apache NiFi (Сурет 7.45) тегін және ашық құралын қарастырдық, ол деректерді нақты уақыт режимінде ағынды өңдеуді үйімдастыруға мүмкіндік береді – әртүрлі көздерден жинаудан бастап сақтау орындарына немесе аналитикалық платформаларға жеткізуге дейін.

Құрылым барысы, материалдардың шығыны, жабдықтың жағдайы және қауіпсіздік туралы деректер нақты уақыт режимінде аналитикалық жүйелерге беріледі (Сурет 9.113). Бұл әлеуетті тәуекелдерді болжауға, ауытқуларға жедел жауап беруге және алаңдағы процестерді оңтайландыруға мүмкіндік береді. Цифрлық құрылым алаңының негізгі компоненттеріне мыналар жатады:-

- IoT-сенсорлары – қоршаған орта параметрлерін бақылау, құрылым техникасын мониторингтеу және еңбек жағдайларын бақылау.
- Цифрлық егіздер – ғимараттар мен инфрақұрылымның виртуалды модельдері, мүмкін болатын ауытқуларды болжауға және қателіктерді болдырмауға мүмкіндік береді.
- Автоматтандырылған логистикалық жүйелер – үзілістер мен шығындарды азайту үшін жеткізу тізбектерін нақты уақыт режимінде басқару.
- Роботтандырылған құрылым кешендері – рутинді және қауіпті тапсырмаларды орындау үшін автономды машиналарды пайдалану.

Роботтандыру, IoT-тің кеңінен қолданылуы және "Connected Site (Construction)" цифрлық құрылым алаңы концепциясы тек тиімділікті арттырып, шығындарды азайтумен шектелмей, қауіпсіздік, тұрақты құрылым және болжамды жобаларды басқарудың жаңа дәүірін ашады.

IoT компоненттерінің маңызды құрамдас белгілі болып табылады. Олар құрылым алаңында материалдарды, техниканы және тіпті персоналды идентификациялау және бақылау үшін қолданылады, жобаның ресурстарын басқаруды және ашықтығын арттырады.

RFID технологиясы радиосигналдар арқылы объектілерді автоматты түрде тану үшін қолданылады. Ол үш негізгі элементтен тұрады:

- RFID-белгілері (пассивті немесе активті) – уникалды идентификаторды қамтиды және материалдарға, құралдарға немесе техникаға бекітіледі.
- Сканерлер – белгілерден ақпаратты оқытын және оны жүйеге беретін құрылғылар.
- Орталықтандырылған деректер базасы – объектілердің орналасуы, жағдайы және қозғалысы туралы ақпаратты сақтайды.

Құрылыш саласындағы RFID қолдану:

- Материалдардың автоматты есебі – дайын бетон өнімдерінде, арматурада немесе сэндвич-панельдер пакеттерінде белгілер қойып, қорларды бақылауға және үрлықтың алдын алуға мүмкіндік береді.
- Персоналдың жұмысын бақылау – қызметкерлердің RFID-бейдждері жұмыс уақытының басталуы мен аяқталуын тіркейді, жұмыс уақытын есепке алуды қамтамасыз етеді.
- Жабдықты мониторингтеу – RFID жүйесі техниканың қозғалысын бақылап, үзілістерді болдырмауға және логистиканың тиімділігін арттыруға мүмкіндік береді.

Бұл технологиялық жинақты блокчейн технологияларына негізделген смарт-келісімшарттар то-лықтырады, олар төлемдерді, жеткізулерді бақылауды және келісім-шарт шарттарын орындауды автоматтандыруға мүмкіндік береді, делдалдардың қажеттілігін жоя отырып, алайқтық пен кешіктіру тәуекелдерін азайтады.

Бұгінгі таңда, біртұтас деректер моделінің болмауынан, смарт-контракттар тек қатысуышылардың келісімін білдіретін код болып табылады. Алайда, Data-Centric тәсілі арақылы келісімшарт параметрлерінің жалпы моделін құру, оны блокчейнге кодтау және шарттардың орындалуын автоматтандыру мүмкіндігі бар.

Мысалы, жеткізу басқару жүйесінде смарт-контракт IoT датчиктері мен RFID белгілерінен жүк жеткізуді бақылап, оның келуі кезінде автоматты түрде төлемді аудара алады. Сол сияқты, құрылыш алаңында смарт-контракт дрондар немесе құрылыш сенсорларының деректері негізінде жұмыстардың кезеңінің аяқталғанын – мысалы, арматураны монтаждау немесе негізді құю – тіркеп, мердігерге келесі төлемді автоматты түрде бастауға мүмкіндік береді, бұл қолмен тексеру мен қағаз актілерін қажет етпейді.

Алайда, жаңа технологиялар мен халықаралық ұйымдардың стандарттау бойынша күш-жігеріне қарамастан, көптеген бәсекелес стандарттар IoT ландшафтын құрделендіреді.

Cisco компаниясының 2017 жылы жарияланған зерттеуі бойынша, IoT саласындағы базалардың шамамен 60%-ы концепцияны дәлелдеу кезеңінде тоқтап қалады, ал тек 26% компаниялар өз IoT жобаларын толықтай сәтті деп санайды. Сонымен қатар, аяқталған жобалардың үштен бірі белгіленген мақсаттарға жетпейді және енгізілгеннен кейін де сәтті деп танылмайды.

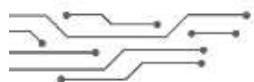
Негізгі себептердің бірі – әртүрлі сенсорлардан деректерді өндейтін платформалар арасындағы үй-лесімділіктің болмауы. Нәтижесінде, деректер жеке шешімдер шеңберінде оқшауланып қалады. Бұл тәсілге балама, біз осы кітапта қарастырған басқа ұқсас жағдайларда, деректерді негізгі актив ретінде қарастыратын архитектура болып табылады.

IoT сенсорлары тек жабдықтың техникалық жағдайын мониторингте ғана емес, сонымен қатар предиктивті аналитикада да маңызды рөл атқарады, құрылым алаңындағы қауіптерді төмендетуге және процестердің жалпы өнімділігін арттыруға мүмкіндік береді, ақаулар мен ауытқуларды болжау арқылы.

IoT сенсорлары мен RFID белгілері арқылы жиналған деректер машиналық оқыту алгоритмдерімен нақты уақыт режимінде өңделіп, аномалияларды анықтап, инженерлерді әлеуетті ақаулар туралы алдын ала хабардар етуге мүмкіндік береді. Бұл бетон конструкцияларындағы микротрециналардың пайда болуы немесе кранның жұмысындағы ерекше үзілістер сияқты техникалық ақаулар немесе регламенттің бұзылуын көрсетуі мүмкін. Сонымен қатар, жетілдірілген мінез-құлышты талдау алгоритмдері, мысалы, персоналдың физикалық шаршауын көрсететін мінез-құлыш паттерндерін тіркеуге мүмкіндік береді, бұл алаңдағы қызметкерлердің қауіпсіздігі мен әл-ауқатын проактивті басқару деңгейін арттырады.

Құрылым саласында апаттар мен ақаулар – техника немесе адамдар – сирек күтпеген жерден болады. Әдетте, оларға кішкентай ауытқулар алдын ала болады, олар байқалмай қалады. Предиктивті аналитика мен машиналық оқыту осы сигналдарды ерте кезеңдерде, критикалық салдарлар орын алmas бўрын, анықтауға мүмкіндік береді.

Егер құжаттар, жобалық файлдар және IoT құрылғылары мен RFID белгілерінен алынған деректер құрылым объектілерінің цифрлық ізі болса, машиналық оқыту одан пайдалы білімдерді алуға көмектеседі. Деректер көлемінің артуы мен деректерге қол жеткізу дің демократиялануы құрылым саласына аналитика, болжау және жасанды интеллекттің қолдану саласында жаңа мүмкіндіктер береді.



ТАРАУ 9.2. Машиналық оқыту және болжамдар

Машиналық оқыту мен жасанды интеллект біздің құрылымызды өзгертеді.

Құрылымы бизнесіндегі әртүрлі жүйелердің дерекқорлары – олардың міндетті түрде ескіріп, күрделеніп жатқан инфрақұрылымымен – болашақ шешімдер үшін құнарлы ортаға айналуда. Компания серверлері, орман сияқты, маңызды ақпараттың биомассасымен бай, көбінесе жер астында, папкалар мен серверлердің тереңдігінде жасырылған. Бұғынгі күні жасалған әртүрлі жүйелерден алынған деректер массалары - пайдаланылғаннан кейін, сервердің түбіне құлап, жылдар бойы тасқа айналғаннан кейін - болашақта машиналық оқыту және тілдік модельдер үшін отынға айналады. Орталықтандырылған қоймаларды пайдалана отырып, осы ішкі корпоративтік модельдер негізінде компаниялардың ішкі чаттары (мысалы, жергілікті конфигурацияланған ChatGPT, LlaMa, Mistral, DeepSeek) құрылып, ақпаратты жылдам және ыңғайлы түрде алуға, қажетті графиктер, панельдер мен құжаттарды қалыптастыруға мүмкіндік береді.

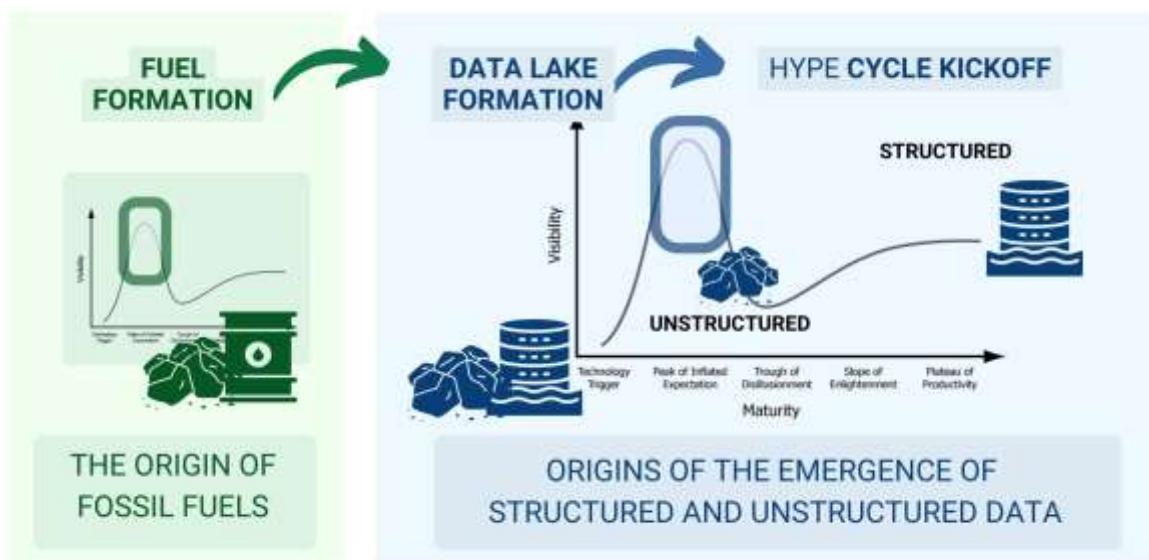


Рис. 9.21 Ағаштар көмірге айналғандай, ақпарат та уақыттың және аналитиктердің қысымымен бизнес үшін құнды энергия көздеріне айналады.

Өсімдік массасының тасқа айналуы, қысым мен температуралың әсерімен, әртүрлі уақытта өмір сүрген әртүрлі ағаш түрлерінің біртекті және ерекше құрылымдалған гомогенді массасын құрайды - көмір [152]. Дәл сол сияқты, әртүрлі форматтарда және әртүрлі уақытта қатты дискілерде жазылған ақпарат, аналитика бөлімдерінің қысымымен және сапа менеджментінің температурасының әсерімен, ақырында, құнды ақпараттың біртекті құрылымдалған массасын қалыптастырады (Рис. 9.21).

Бұл ақпараттың қабаттары (немесе жиі оқшауланған самородкалар) тәжірибелі аналитиктердің деңгектерді үйімдастыру бойынша мүқият жұмысы арқылы қалыптасады, олар, бір қарағанда, ескіріген деректерден құнды ақпаратты біртінде шығара бастайды.

Бұл пісіп жетілген деректер қабаттары тек «жанбай» қалғанда, бизнес-процестерде айнала бағтайды, шешімдерді байытып, процестерді жақсартады, компания машиналық оқыту мен жасанды интеллектке (Рис. 9.22) көшуге дайын болады.

Машиналық оқыту (ML - Machine learning) – жасанды интеллект мәселелерін шешу үшін қолданылатын әдістер класы. Машиналық оқыту алгоритмдері үлкен деректер масивтерінде зандылықтарды таниды және оларды өзін-өзі оқыту үшін пайдаланады. Әрбір жаңа деректер жиыны математикалық алгоритмдардың жетілдірілуіне және алынған ақпаратқа сәйкес бейімделуіне мүмкіндік береді, бұл ұсыныстар мен болжамдардың дәлдігін үнемі арттыруға мүмкіндік береді.

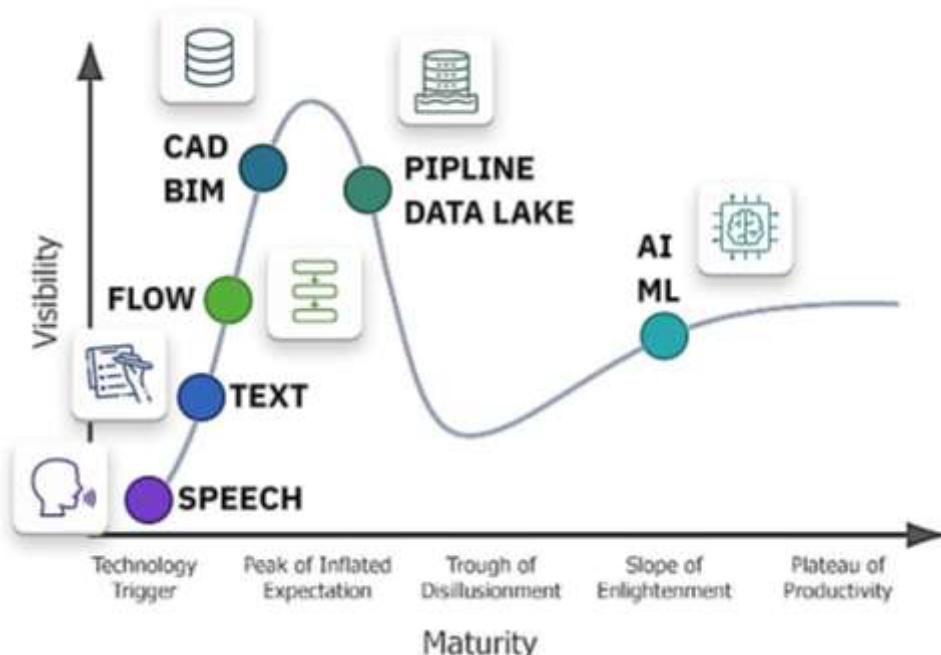


Рис. 9.22 Деректерді жасау технологияларының әлсіреуі және аналитикалық құралдарды қолдану машиналық оқыту тақырыбына жол ашады.

2023 жылы әлемдегі ең ірі инвестициялық қордың ықпалды бас директорының сұхбатында (қорқынышы бағдарламалық қамтамасыз ету өндіретін ең ірі компаниялардың негізгі акциялар пакеттеріне ие, сондай-ақ әлемдегі ең көп жылжымайтын мүлікке ие компанияларға иелік етеді [55]) - машиналық оқыту қорқынышы әлемін өзгертерді.

Жасанды интеллекттің үлкен әлеуеті бар. Ол біздің жұмысымызды, өмір сүруімізді өзгертерді. Жасанды интеллект пен робототехника біздің жұмысымызды және қорқынышымызды өзгертерді, және біз жасанды интеллект пен робототехниканы әлдеқайда жоғары өнімділікке қол жеткізу үшін құрал ретінде пайдалана аламыз [153]. – Әлемдегі ең ірі инвестициялық қордың бас директоры, сұхбат, 2023 жылдың қыркүйегі.

Машиналық оқыту (ML) үлкен деректер көлемдерін өңдеу арқылы жұмыс істейді, адамның ойлау аспектілерін имитациялау үшін статистикалық әдістерді пайдаланады. Алайда, көптеген компаниялар мұндай деректер жиынтығына ие емес, ал егер олар болса, жиі жеткіліксіз түрде белгіленген. Мұнда семантикалық технологиялар мен трансферлік оқыту көмектесе алады – ML-дің шағын деректер көлемдерімен жұмыс істеу тиімділігін арттыратын әдіс, оның тиімділігі осы бөлімнің алдыңғы тарауларында талқыланды.

Трансферлік оқытудың мәні – әрбір тапсырманы нөлден бастап өңдеу орнына, жақын салаларда алынған білімдерді пайдалану. Басқа экономикалық салалардан алынған паттерндер мен ашуулар құрылымындағы саласында бейімделіп, қолданылуы мүмкін екенін түсіну қажет. Мысалы, бөлшек саудада әзірленген логистикалық процестерді оңтайландыру әдістері құрылымын жеткізу тізбектерін басқарудың тиімділігін арттыруға көмектеседі. Қаржы саласында белсенді қолданылатын үлкен деректерді талдау құрылымындағы шығындарды болжау және тәуекелдерді басқару үшін пайдаланылуы мүмкін. Ал өнеркәсіпте дамып келе жатқан компьютерлік көру және робототехника технологиялары автоматтандырылған сапа бақылауында, қауіпсіздікті мониторингте және құрылымындағы объектілерді басқаруда қолданылуда.

Трансферлік оқыту инновацияларды енгізуді жеделдетуге ғана емес, сонымен қатар басқа салалардың жинақталған тәжірибелі пайдалана отырып, олардың дамуына көтөтін шығындарды азайтуға мүмкіндік береді.

$$\text{labor productivity in construction} = f(\text{AI})$$

Рис. 9.23 Жасанды интеллект және робототехника технологиялары құрылымындағы өнімділікті арттыру үшін болашақтың негізгі қозғаушы құші болады.

Адамның ойлауы да үқсас принцип бойынша жұмыс істейді: біз жаңа тапсырмаларды шешу үшін бұрын алынған білімдерге сүйенеміз (Рис. 4.419, Рис. 4.420, Рис. 4.421). Машиналық оқытудың бұл тәсілі де жұмыс істейді – деректер моделін жеңілдету және оны әлегантты ету алгоритмдер үшін тапсырманың күрделілігін төмендетуге мүмкіндік береді. Бұл, өз кезегінде, үлкен деректер көлемдеріне деген қажеттілікті азайтады және есептеу шығындарын қысқартады.–

Субъективті бағалаудан статистикалық болжамға.

Стратегиялық шешімдер жеке басшылардың интуициясына тәуелді болған дәүір (Рис. 9.24) өткен шаққа айналуда. Бәсекелестік пен күрделі экономикалық жағдайлардың артуымен субъективті тәсіл тым қауіпті және тиімді емес болып табылады. Жеке пікірлерге сүйенетін компаниялар обьективті деректерді талдауға негізделген өзгерістерге жедел жауап беру мүмкіндігін жоғалтады.

Бәсекелестік орта деректерге, статистикалық заңдылықтарға және есептелетін ықтималдықта негізделген дәлдік пен қайталанушылықты талап етеді. Шешімдер енді сезімге емес, аналитика мен машиналық оқыту арқылы алынған корреляцияларға, трендтерге және болжамдық модельдерге негізделуі тиіс. Бұл тек құралдардың өзгерісі емес — бұл ойлау логикасының өзгерісі: болжамдардан — дәлелдерге, субъективті ықтималдықтардан — статистикалық есептелген ауытқуларға, сезімнен — фактілерге.



Рис. 9.24 HiPPO (ең жоғары жалақы алатын қызметкердің пікірі) қабылдаған шешімдер дәүірі үлкен деректер мен машиналық оқытудың келуімен өткен шаққа кетеді.

Басшылар, тек өз сезімдеріне ғана сүйенуге дағыланған, жаңа шындықпен бетпе-бет келеді: бедел енді таңдауын анықтамайды. Басқарудың орталығында миллиондаған параметрлер мен векторларды талдайтын жүйелер орналасқан, олар жасырын заңдылықтарды анықтап, оңтайлы стратегияларды ұсынады.

Бұғаңғі күні компаниялардың ML енгізуден әлі де бас тартуының басты себебі — оның түсініксіздігі. Көптеген модельдер менеджерлер үшін "қара жәшік" ретінде жұмыс істейді, олар өздерінің қорытындыларына қалай келгенін түсіндірмейді. Бұл проблемаларға әкеледі: алгоритмдер стереотиптерді нығайтуы мүмкін және тіпті Microsoft чат-боты сияқты күлкілі жағдайларды тудыруы мүмкін.

"Deep Thinking" кітабында бұрынғы шахматтан әлем чемпионы Гарри Каспаров IBM Big Blue компьютерінен жеңілуі туралы ойлады. Ол ИИ-дің шынайы құндылығы адамның интеллектісін көшірмелеуде емес, біздің қабілеттерімізді толықтыруда екенін айтады. ИИ адамдар әлсіз болатын тапсырмаларды орындауы тиіс, ал адамдар шығармашылықты енгізеді. Компьютерлер шахматты талдаудың дәстүрлі тәсілін өзгертті. Шахмат партиялары туралы қызықты әңгімелер құрудың орнына, компьютерлік шахмат бағдарламалары әр жүрісті объективті бағалайды, тек оның нақты күшіне немесе әлсіздігіне негізделді. Каспаров адамның оқиғаларды байланыстыра қабылдауға бейімділігі, жеке әрекеттер ретінде емес, жиынтық ретінде, жи қате қорытындыларға әкелетінін атап өтеді – тек шахматта ғана емес, жалпы өмірде.

Сондықтан, егер сіз машиналық оқытуды болжау және талдау үшін пайдалануды жоспарласаңыз, алгоритмдердің қалай жұмыс істейтінін және деректердің қалай өндөлетінін түсіну маңызды, машиналық оқыту және ИИ құралдарын жұмысыныңда қолданар алдында. Ең жақсы бастама – практикалық тәжірибе.

Машиналық оқыту және болжау тақырыбымен танысадың ең ыңғайлы құралдарының бірі - Jupyter Notebook және танымал класикалық Titanic деректер жинағы, ол деректерді талдау және ML модельдерін құрудың негізгі әдістерін визуалды түрде мәнгеруге мүмкіндік береді.

Титаник деректер жиыны: Деректер аналитикасы мен үлкен деректер әлеміндегі Сәлем Әлем.

Деректер аналитикасында ML-ді қолданудың ең танымал мысалдарының бірі - "Титаник" деректер жинағын талдау, ол жолаушылардың тірі қалу ықтималдығын зерттеу үшін жи қолданылады. Бұл кестені зерттеу бағдарламалу тілдерін үйренудегі "Hello World" бағдарламасына аналогия болып табылады.

RMS Titanic-тің 1912 жылды суға батуынан 2224 адамның 1502-і қаза тапты. Titanic деректер жинағы тек жолаушының тірі қалғаны туралы ақпаратты ғана емес, сонымен қатар: жасы, жынысы, билет класы және басқа параметрлер сияқты атрибуттарды да қамтиды. Бұл деректер жинағы тегін қолжетімді және оны оффлайн және онлайн платформаларда ашып, талдауға болады.

Titanic деректер жинағына сілтеме:

<https://raw.githubusercontent.com/datasets/master/titanic.csv>

"LLM қолдайтын IDE және бағдарламалаудағы болашақ өзгерістер" бөлімінде Jupyter Notebook - деректерді талдау және машиналық оқыту үшін ең танымал даму ортасы туралы айтылған. Jupyter Notebook-тың тегін бүлттық аналогтары Kaggle және Google Collab платформалары болып табылады, олар Python кодын бағдарламалық қамтамасыз етусіз іске қосуға мүмкіндік береді және есептеу ресурстарына тегін қолжетімділікті ұсынады.

Kaggle – деректерді талдау, машиналық оқыту бойынша жарыстар үшін ең ірі платформа, кодты орындау үшін интеграцияланған орта. 2023 жылдың қазан айындағы жағдай бойынша Kaggle-да 194 елден 15 миллионнан астам пайдалануышы бар.

Kaggle платформасында Titanic деректер жинағын жүктеп, пайдаланыңыз (Рис. 9.25), деректер жинағын (оның көшірмесін) сақтау және Python кодын алдын ала орнатылған кітапханалармен тікелей браузерде іске қосу үшін, арнайы IDE орнатудың қажеті жоқ.-



Рис. 9.25 Titanic кестесінің статистикасы – деректерді талдауды және машиналық оқытуды үйрену үшін ең танымал оқу деректер жинағы.

Titanic деректер жинағы 1912 жылы RMS Titanic апартында болған 2224 жолаушы туралы мәліметтерді қамтиды. Набор екі бөлек кесте түрінде ұсынылған – оқу (train.csv) және тестілеу (test.csv) жиынтығы, бұл оны модельдерді оқыту және жаңа деректерде олардың дәлдігін бағалау үшін пайдалануға мүмкіндік береді.

Оқу деректер жинағы жолаушылардың атрибуттарын (жасы, жынысы, билет класы және басқалар) және кімнің аман қалғаны туралы ақпаратты (бинарлық мәндермен «Аман қалды» бағаны) қамтиды. Оқу деректер жинағы (Рис. 9.26 – train.csv файлы) модельді оқыту үшін пайдаланылады. Тестілеу деректер жинағы (Рис. 9.27 – test.csv файлы) тек жолаушылардың атрибуттарын қамтиды, аман қалғандар туралы ақпаратсыз (бір ғана «Аман қалды» бағаны жоқ). Тестілеу деректер жинағы модельдің жаңа деректердегі жұмысын тексеру және оның дәлдігін бағалау үшін арналған.-

Осылайша, бізде оқу және тестілеу деректер жинақтарында жолаушылардың дерлік бірдей атрибуттары бар. Бір ғана негізгі айырмашылық – тестілеу деректер жинағында «Аман қалды» – біз әртүрлі математикалық алгоритмдер арқылы болжауды үйренетін мақсатты айнымалы бағаны жоқ. Модельді құрғаннан кейін, біз модельдің нәтижесін тестілеу деректер жинағындағы нақты «Аман қалды» параметрімен салыстыра аламыз, бұл нәтижелерді бағалау үшін ескеріледі.

Оқу және тестілеу деректер жинақтарындағы негізгі кесте бағандары, жолаушылар параметрлері:

- PassengerId – жолаушының бірегей идентификаторы
- Survived – 1, егер жолаушы аман қалса, 0 – егер қаза тапса (тестілеу жиынтығында жоқ)
- Pclass – билет класы (1, 2 немесе 3)
- Name – жолаушының аты

- Sex – жолаушының жынысы (ер/әйел)
- Age – жасы
- SibSp – борттағы ағайын/сіңлілер немесе жұбайлар саны
- Parch – борттағы ата-аналар немесе балалар саны
- Ticket – билет нөмірі
- Fare – билет құны
- Cabin – каюта нөмірі (көптеген деректер жоқ)
- Embarked – отырғызу порты (C = Шербур, Q = Куинстаун, S = Саутгемптон)

Екі кестедегі жоғалған деректерді визуализациялау үшін missingno кітапханасын пайдалануға болады (Рис. 9.26, Рис. 9.27), ол жоғалған мәндерді гистограмма түрінде көрсетеді, мұнда ақ аландар жоқ деректерді көрсетеді. Мұндай визуализация деректерді өңдеуден бұрын олардың сапасын жылдам бағалауға мүмкіндік береді.-

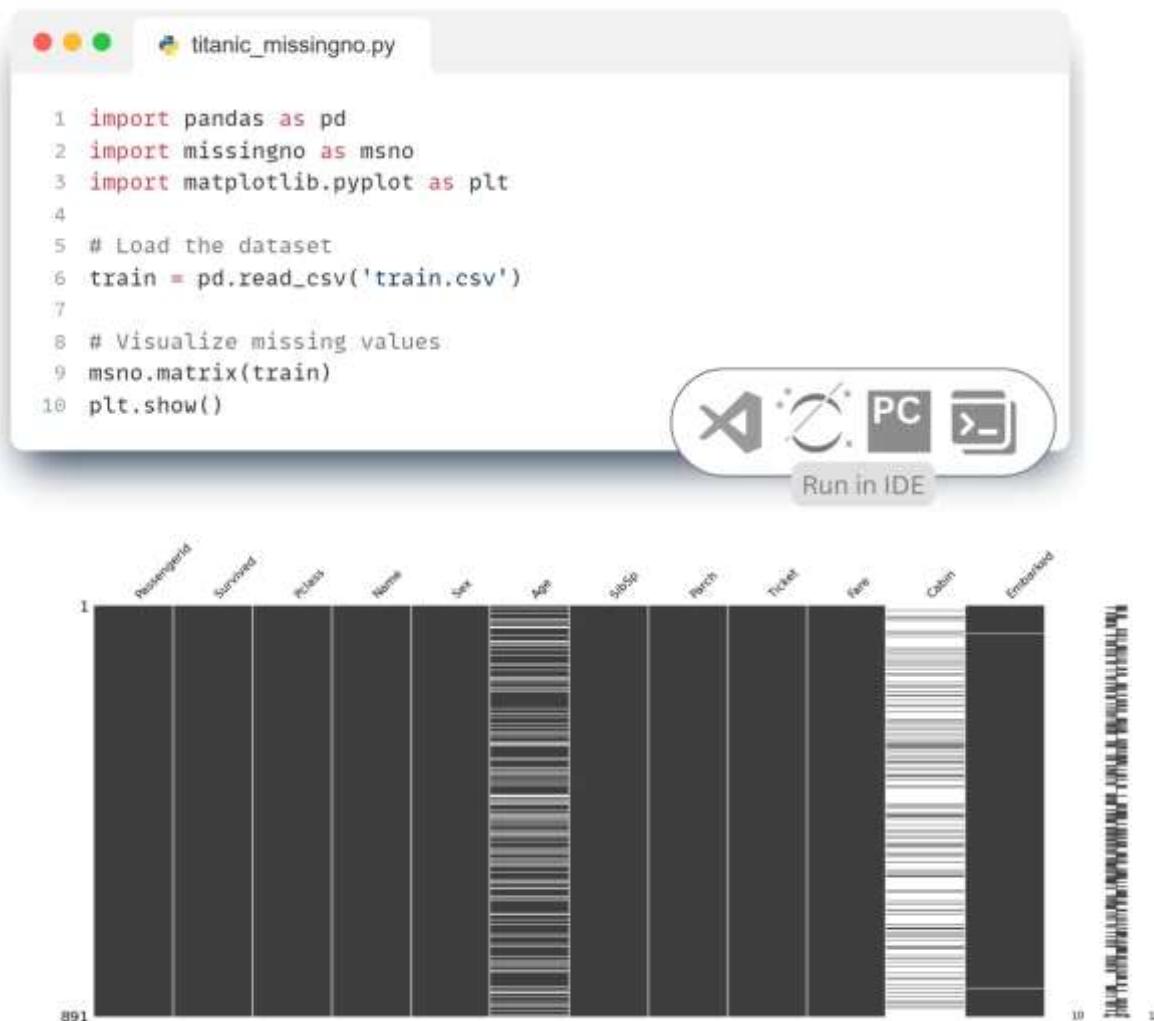


Рис. 9.26 Titanic оқу деректер жинағындағы жоғалған деректерді визуализациялау үшін бірнеше код жолдары пайдаланылады, мұнда оқыту үшін негізгі параметр «Аман қалды» параметрі болып табылады.

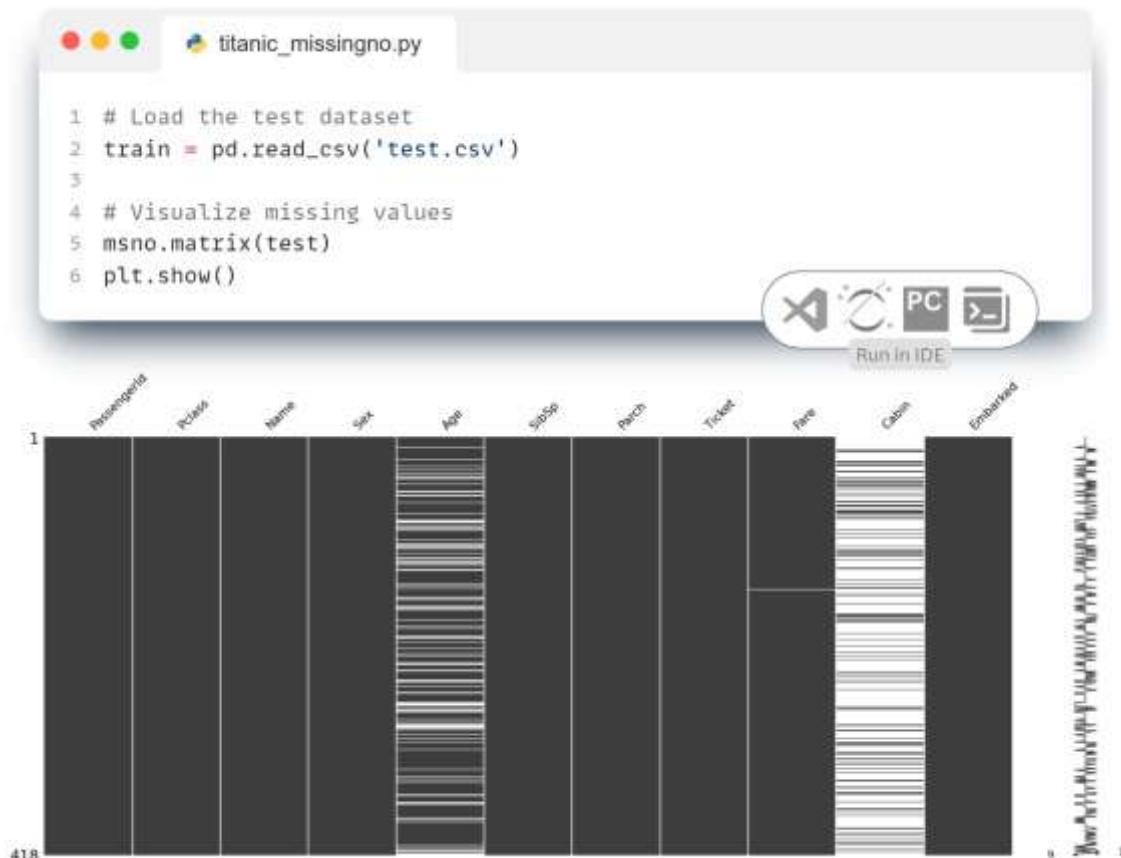


Рис. 9.27 Титаник тестілеу деректер жиынтығындағы жоғалған деректерді визуализациялау, ол тек жолаушылардың сипаттамаларын қамтиды, ақпаратсыз.

Гипотезаларды қалыптастырмас бұрын және деректер жиынтығы негізінде болжамдар жасағанға дейін, визуалды талдау деректердегі негізгі заңдылықтарды анықтауға, олардың сапасын бағалауға және мүмкін тәуелділіктерді анықтауға көмектеседі. Титаник деректер жиынтығын жақсы түсінуге көмектесетін көптеген визуализация әдістері бар. Жолаушылардың жас топтарын талдау үшін тарапу графиктерін, жыныс пен классқа байланысты тіршілік диаграммаларын, сондай-ақ ақпараттың сапасын бағалау және деректерді түсіну үшін жоғалған деректер матрикаларын пайдалануға болады.

- LLM-ден Титаник деректер жиынтығындағы деректерді визуализациялауға көмек сұрайық, үшін біз кез келген LLM моделіне (ChatGPT, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN немесе басқа) келесі мәтіндік сұранысты жібереміз:

Титаник деректер жиынтығы үшін бірнеше қарапайым графиктерді көрсетіңіз. Деректер жиынтығын өзініз жүктеп алыңыз және графиктерді көрсетіңіз ↗

- 2 LLM-нің жауаптары дайын код және деректер жиынтығының параметрлерін визуализациялайтын графиктер түрінде:

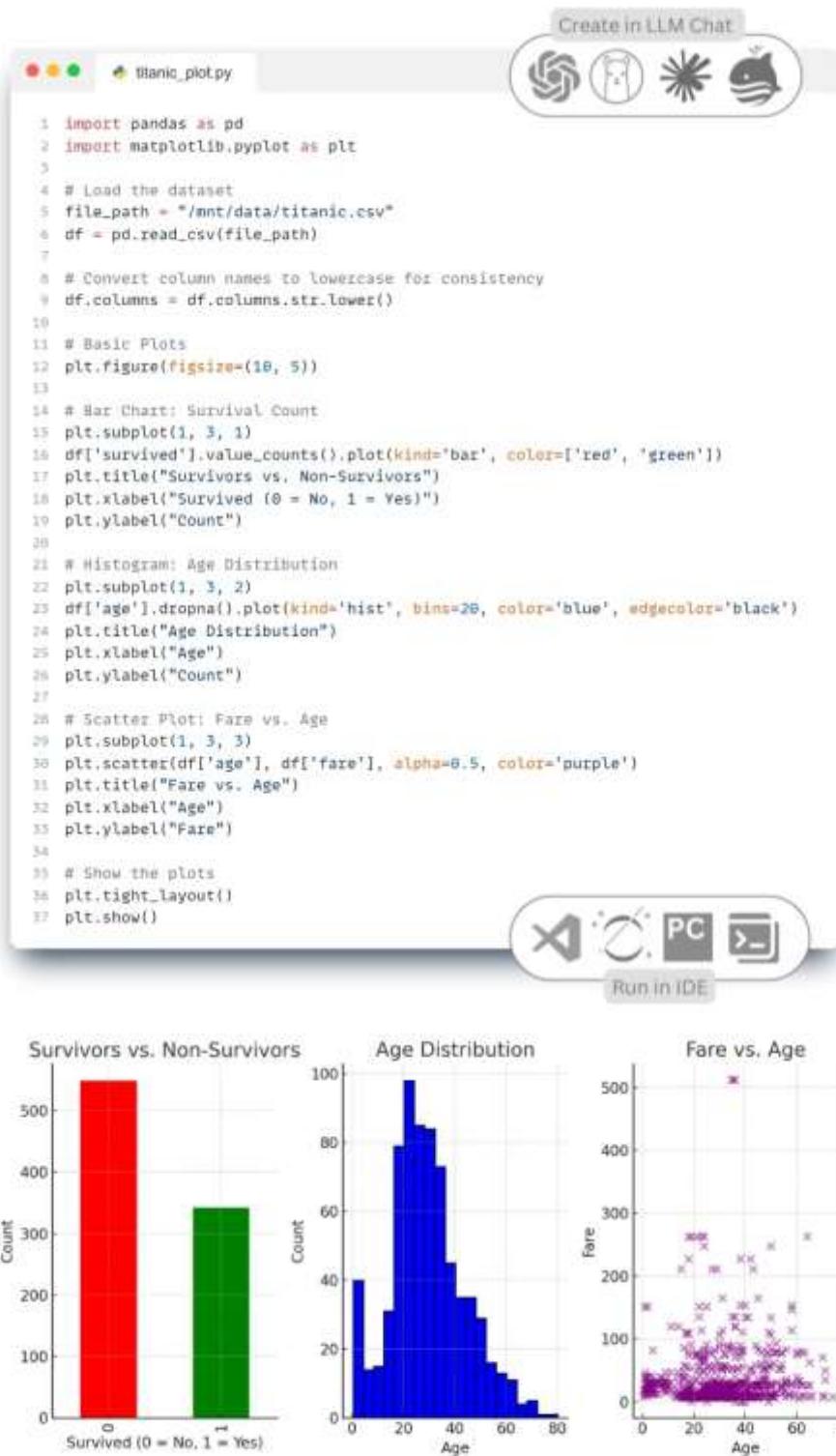


Рис. 9.28 LLM деректер жиынтығының визуализациясын бірден алуға көмектеседі.

Деректерді визуализациялау – бұл деректер жиынтығын машиналық оқыту моделін құруға дайындаудың маңызды кезеңі, оған тек деректерді түсінгеннен кейін ғана көшуге болады.

Машиналық оқыту іс-әрекетте: "Титаник" жолаушыларынан жобаларды басқаруға.

Титаник деректер жиынтығын зерттеуге арналған негізгі гипотеза – кейбір жолаушылар топтарының тірі қалу мүмкіндіктері жоғары болды.

Титаник жолаушыларының шағын кестесі әлем бойынша танымал болды, және миллиондаған адамдар оны оқыту, эксперименттер жүргізу және модельдерді тестілеу үшін пайдаланады, қандай алгоритмдер мен гипотезалар Титаник жолаушыларының жаттығу деректер жиынтығы негізінде максималды дәлдікпен тірі қалу модельнің құруға мүмкіндік береді.

Титаник деректер жиынтығының тартымдылығы оның компакттылығымен түсіндіріледі: бірнеше жұз жол мен он екі баған (Рис. 9.26) деректерді талдауға кең мүмкіндіктер береді. Деректер жиынтығы салыстырмалы түрде қарапайым, классикалық бинарлық классификация шешімінің мысалы болып табылады, мұнда тапсырманың мақсаты – тірі қалу – 0 немесе 1 форматында көрсетілген.

Джон Уилер «It from Bit» [7] еңбегінде әлемнің негізінде бинарлық шешімдер жатқанын айтады. Адамдармен басқарылатын бизнес те молекулалардан тұрады, шын мәнінде, бинарлық таңдау тізбегіне негізделген.

Сонымен қатар, деректер нақты тарихи оқиғаға негізделген, бұл оларды зерттеу үшін құнды етеді, жасанды мысалдардан айырмашылығы. Kaggle платформасында, Data Pipeline және ETL бойынша жұмыс істейтін ең ірі алаңдардың бірі, Титаник деректер жиынтығына негізделген тапсырмаларды шешуге 1 355 998 адам қатысып, 53 963 бірегей Data Pipeline шешімін әзірледі [157] (Рис. 9.29).-

Титаник жолаушылары туралы 1000 жол деректер мен 12 параметрдің болуы таңқаларлық, бірақ бұл миллиондаған гипотезалар, логикалық тізбектер мен бірегей Data-Pipelines үшін алаңға айналды. Кішігірім деректер жиынтығынан шексіз инсайттар, гипотезалар мен интерпретациялар туындаиды – қарапайым тірі қалу модельдерінен бастап, жасырын заңдылықтар мен күрделі ойлау лабиринттерін ескеретін күрделі ансамбльдерге дейін.

Machine Learning from Disaster

Submit Prediction

Data **Code** **Models** **Discussion** **Leaderboard** **Rules**

Titanic Tutorial
Updated 3y ago
29858 comments · Titanic - Machine Learning from Disaster

Titanic competition w/ TensorFlow Decision Forests
Updated 2y ago
Score: 0.80143 · 318 comments · Titanic - Machine Learning from Disaster

Titanic Data Science Solutions
Updated 6y ago
2590 comments · Titanic - Machine Learning from Disaster

Exploring Survival on the Titanic
Updated 7y ago
Score: 0.80382 · 1072 comments · Titanic - Machine Learning from Disaster

Рис. 9.29 Барлық 53 963 дайын және ашық Pipeline-дан алғашқы бес шешім. Kaggle платформасында бұл тапсырманы шешүге 1,5 миллионнан астам адам қатысқан [157].

Мұндай шағын кесте миллиондаған бірегей шешімдерді тудыратын болса (Рис. 9.29), онда нақты өндірістік құрылымдар деректер жиынтықтары туралы не айтуға болады, мұнда параметрлер он мыңдағанмен өлшенеді? -

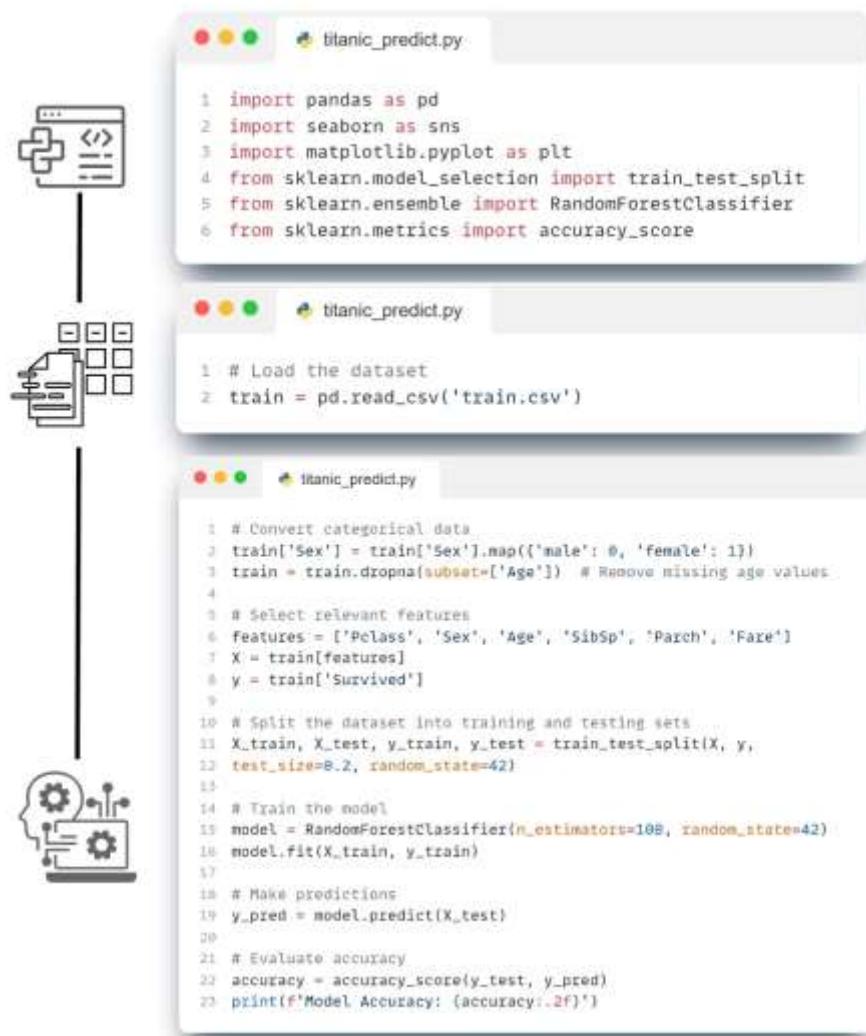
Стандартты CAD-жобасы салыстырмалы түрде шағын ғимарат үшін ондаған мың объектілерді қамтиды, олардың әрқайсысында геометриялық сипаттамалардан бастап, құндық және уақыттық атрибуттарға дейін мыңдаған параметрлер бар. Сіздің компанияңыздың соңғы жылдардағы барлық жобаларының деректерінде қанша әлеуетті инсайт, өзара байланыс, болжамдар мен басқару гипотезалары бар екенін елестетіңіз. Тарихи жобалық деректер — бұл тек архив емес, үйімнің тірі жадысы, оның цифрлық ізи, оны талдау арқылы көптеген бірегей гипотезаларды құруға болады.

Ең бастысы, Kaggle қауымдастырылған сіздің компанияңыздың немесе деректеріңізге қызығушылық таңытқанын күтудің қажеті жоқ. Бұғынғы күні сізде бар деректермен жұмыс істеуді бастауға болады: өз деректеріңізде аналитиканы іске қосу, өз деректеріңізде модельдерді оқыту, қайталануларды, аномалияларды және паттерндерді анықтау. Бұрын жылдар бойы эксперименттер мен қымбат консалтинг қажет болса, бұғынғы күні бастамашылық, LLM, деректерге ашық көзқарас және оқуға дайындық жеткілікті.

- ☞ Пассажирлардың train.csv оқу деректер жиынтығына негізделген, жолаушылардың аман қалуын болжау үшін машиналық оқыту алгоритмін құру үшін LLM-ді осы тапсырманы шешүге шақырайық:

Титаник жолаушыларының оқу деректер жиынтығына негізделген, аман қалуды болжау үшін машиналық оқыту моделін құрыңыз ↴

 LLM-нің жауабы:



```

1 import pandas as pd
2 import seaborn as sns
3 import matplotlib.pyplot as plt
4 from sklearn.model_selection import train_test_split
5 from sklearn.ensemble import RandomForestClassifier
6 from sklearn.metrics import accuracy_score

1 # Load the dataset
2 train = pd.read_csv('train.csv')

1 # Convert categorical data
2 train['Sex'] = train['Sex'].map({'male': 0, 'female': 1})
3 train = train.dropna(subset=['Age']) # Remove missing age values
4
5 # Select relevant features
6 features = ['Pclass', 'Sex', 'Age', 'SibSp', 'Parch', 'Fare']
7 X = train[features]
8 y = train['Survived']

10 # Split the dataset into training and testing sets
11 X_train, X_test, y_train, y_test = train_test_split(X, y,
12 test_size=0.2, random_state=42)
13
14 # Train the model
15 model = RandomForestClassifier(n_estimators=100, random_state=42)
16 model.fit(X_train, y_train)
17
18 # Make predictions
19 y_pred = model.predict(X_test)
20
21 # Evaluate accuracy
22 accuracy = accuracy_score(y_test, y_pred)
23 print(f'Model Accuracy: {accuracy:.2f}')

```

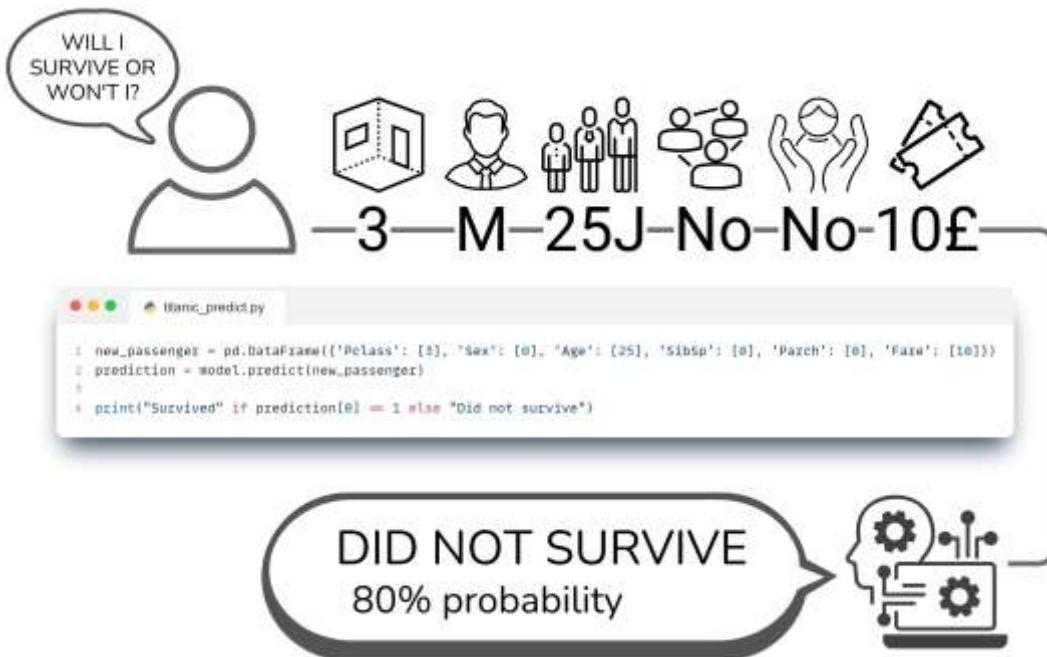
Рис. 9.210 LLM Титаникте аман қалғандарды болжау үшін Random Forest машиналық оқыту алгоритмын қолданды.

LLM-ден алынған код (Рис. 9.210) "Титаник" жолаушылары туралы деректерді жүктейді, оларды тазалайды, категориялық айнымалыларды (мысалы, жынысты сандық форматқа) түрлендіреді және RandomForestClassifier алгоритмы арқылы модельді оқытады, жолаушының аман қалғанын немесе аман қалмағанын болжайды (тәнімал алгоритмдер туралы келесі тарауларда сөйле-семіз).-

Кодтың көмегімен деректер оқыту кезінде оқу және тестілеу жиынтықтарына бөлінеді (Kaggle сайтында оқыту үшін дайын test.csv (Рис. 9.27) және train.csv (Рис. 9.26) файлдары бар), содан кейін модель оқу деректерінде оқытылады және тестілеу деректерінде тексеріледі, осылайша болжам модельдерінің қаншалықты жақсы екенін түсінуге болады. Оқыту аяқталғаннан кейін test.csv файлындағы (амандығы немесе аман қалмағаны туралы нақты деректер) тестілеу деректері модельге беріледі, және ол кімнің аман қалғанын, ал кімнің аман қалмағанын болжайды. Біздің жағдайда, алынған машиналық оқыту модельнің дәлдігі шамамен 80%, бұл оның заңдылықтарды жақсы түсінетінін көрсетеді.-

Машиналық оқытуды балаға тікбұрышты блокты дөңгелек тесікке салуға тырысатын процесс ретінде салыстыруға болады. Алғашқы кезеңдерде алгоритм көптеген тәсілдерді сыйнап көреді, қателіктер мен сәйкесіздіктермен кездеседі. Бұл процесс тиімді емес болып көрініү мүмкін, бірақ ол маңызды оқытуды қамтамасыз етеді: әрбір қателікті талдай отырып, модель өз болжамдарын жақсартады және барған сайын дәл шешімдер қабылдайды.

Енді бұл модельді (Сурет 9.210) жаңа жолаушылардың тіршілік қабілетін болжау үшін пайдалануға болады. Мысалы, егер жолаушы туралы ақпаратты model.predict функциясы арқылы енгізсек: «ер адам», «3-ші класс», «25 жас», «бортта туыстары жоқ», модель 1912 жылы «Титаник» кемесінде болған жолаушының апат кезінде 80% ықтималдықпен тірі қалмайтынын болжайды (Сурет 9.211).-



Сурет 9.211 Жоғарыда құрылған модель енді 80% ықтималдықпен Титаниктің кез келген жаңа жолаушысының тірі қалатынын немесе қалмайтынын болжай алады.

«Титаник» жолаушыларының тіршілік қабілетін болжау моделі әлдеқайда кең концепцияны көрсетеді: күн сайын құрылыш саласындағы мыңдаған мамандар осында «екішты» шешімдер қабылдайды – өмір мен өлім, жоба, смета, құрал, пайда немесе шығын, қауіпсіздік немесе тәуекел. «Тита-

ник» мысалындағыдай, мұнда нәтиже әртүрлі факторларға (жыныс, жас, класс) байланысты, құрылыш саласында шешімнің әр аспектісіне көптеген өз факторлары мен айнымалылары (кесте бағандары) әсер етеді: материалдардың құны, жұмысшылардың біліктілігі, мерзімдер, аяу-райы, логистика, техникалық тәуекелдер, пікірлер және жұздеген мың басқа параметрлер.

Құрылыш саласында машиналық оқыту басқа салалардағыдай принциптер бойынша қолданылады: модельдер тарихи деректерге – жобалар, келісімшарттар, сметалар – негізделіп оқытылады, әртүрлі гипотезаларды тексеру және ең тиімді шешімдерді іздеу үшін. Бұл процесс балаларды сынақ және қателік әдісімен оқытуға үқсас: әр циклмен модельдер бейімделіп, дәлдігі артады.

Жиналған деректер құрылыш үшін жаңа мүмкіндіктер ашады. Еңбек-intensive қолмен есептеулердің орнына, болашақ жобалардың негізгі сипаттамаларын жоғары дәлдікпен болжай алатын модельдерді оқытуға болады. Осылайша, предиктивті аналитика құрылыш саласын жоспарлауға ғана емес, сонымен қатар оқиғалардың дамуын сенімді түрде болжауға мүмкіндік беретін кеңістікке айналдырады.

Тарихи деректер негізінде болжамдар мен предсказаниялар.

Компанияның жобалары туралы жиналған деректер болашақ, әлі жүзеге аспаған объектілердің құны мен уақыттық сипаттамаларын болжауға қабілетті модельдерді құруға мүмкіндік береді – қолмен есептеулер мен салыстыруларды қажет етпей. Бұл процестерді бағалауды едәүір жылдамдатып, жеңілдетуге мүмкіндік береді, субъективті болжамдарға емес, негізделген математикалық болжамдарға сүйене отырып.

Бұрын, кітаптың төртінші бөлімінде, біз жобалардың сметалық құнын есептеудің дәстүрлі әдістерін, оның ішінде ресурс әдісін, параметрлік және сараптамалық тәсілдерді егжей-тегжейлі қарастырық. Бұл әдістер әлі де өзекті, бірақ қазіргі практикада олар статистикалық талдау және машиналық оқыту құралдарымен байытылып, бағалаудың дәлдігі мен қайталаныштығын едәүір арттыруға мүмкіндік береді.

Құн мен уақыттық атрибуттарды қолмен және жартылай автоматты есептеу процестері болашақта тарихи деректерді талдай алатын, жасырын зандағылықтарды таба алатын және негізделген шешімдерді ұсына алатын ML-модельдердің пікірлері мен болжамдарымен толықтырылады. Жаңа деректер мен сценарийлер автоматты түрде бар ақпараттан генерацияланады – ашық көздерден жылдар бойы жиналған деректер негізінде мәтіндер, суреттер және кодтар жасайтын тілдік модельдер (LLM) сияқты.

Бұгінгі таңда адам болашақ оқиғаларды бағалау кезінде тәжірибеге, интуицияға және ішкі статистикаға сүйенгендей, алдағы жылдары құрылыш жобаларының болашағы жинақталған білім мен машиналық оқыту математикалық модельдерінің үйлесімімен анықталатын болады.

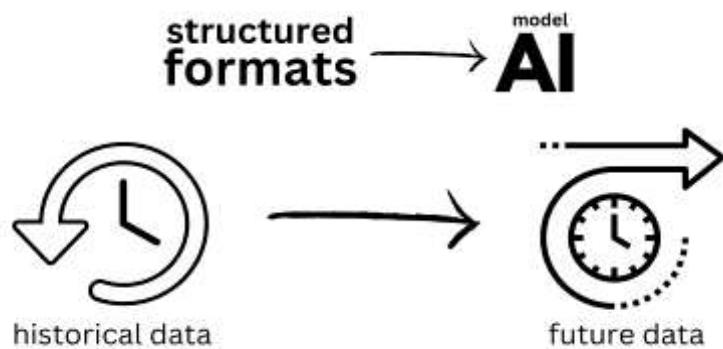


Рис. 9.212 Компанияның сапалы және құрылымдалған тарихи деректері - машиналық оқыту модельдері мен болжамдарын құру үшін негіз болатын материал.

Қарапайым мысалды қарастырайық: үйдің бағасын оның алаңы, участкенің көлемі, бөлмелердің саны және географиялық орналасуы негізінде болжау. Бір тәсіл - осы параметрлерді талдайтын және болжамды бағасын есептейтін классикалық модельді құру (Рис. 9.213). Мұндай тәсіл нақты және алдын ала белгілі формуланы талап етеді, бұл шынайы практикада іс жүзінде мүмкін емес.

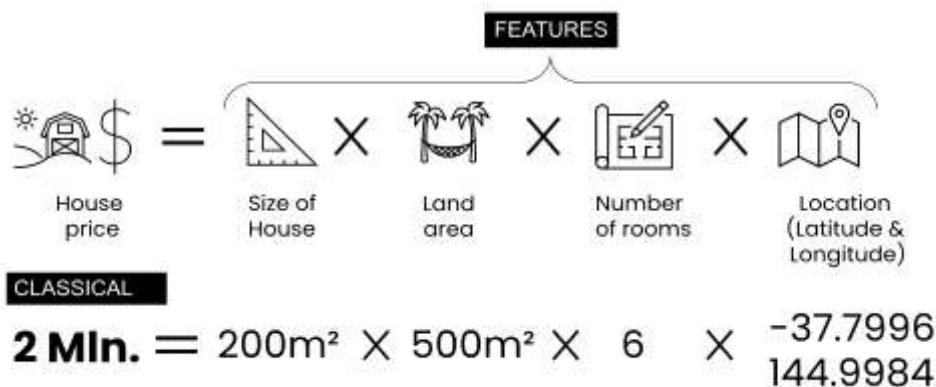


Рис. 9.213 Үйдің құнын бағалау үшін табу қажет фиксиленген формуласы бар классикалық алгоритмді қолдануға болады.

Машиналық оқыту қолмен формулаларды іздеуден бас тартуға және оларды өздігінен тәуелділіктерді анықтайтын оқытылатын алгоритмдермен алмастыруға мүмкіндік береді, бұл кез келген алдын ала берілген теңдеулерден бірнеше есе жоғары дәлдікпен жүзеге асырылады. Альтернативті түрде, мәселені алдын ала түсіну мен тарихи деректер негізінде модельді генерациялайтын машиналық оқыту алгоритмін құрайық, бұл деректер толық болмауы мүмкін (Рис. 9.214).-

Баға қалыптастыру мәселесінде машиналық оқыту нақты бағалау механизмін білмей-ақ әртүрлі математикалық модельдерді құруға мүмкіндік береді. Модель өткен жобалар туралы деректер негізінде «үйренеді», ғимараттардың параметрлері, олардың құны мен орындауда мерзімдері арасындағы нақты занылықтарға бейімделеді.

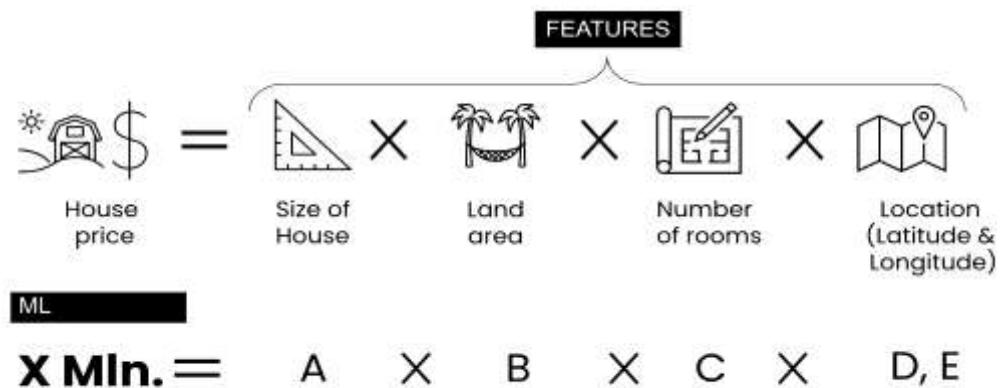


Рис. 9.214 Классикалық формулалар бойынша бағалаудан айырмашылығы, машиналық оқыту алгоритмі тарихи деректерге негізделіп оқытылады.

Машиналық оқытудың контекстінде бақылау (supervised machine learning) бойынша оқыту жиынтығындағы әрбір жоба кіріс атрибуттарын (мысалы, үқсас ғимараттардың құны мен құрылым уақыты туралы деректер) және күтілетін шығу мәндерін (мысалы, құны немесе уақыты) қамтиды. Мұндай деректер жиынтығы машиналық оқыту моделін құру және баптау үшін пайдаланылады (Рис. 9.215). Деректер жиынтығы негұрлым көп және оның сапасы жоғары болса, модель соғұрлым дәл болады және болжамдардың нәтижелері соғұрлым дәл болады.-

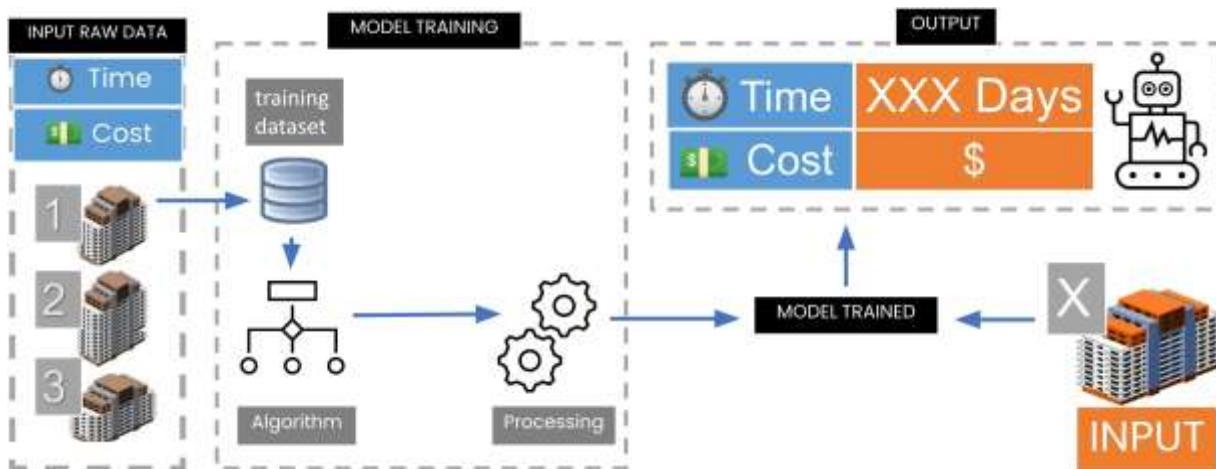


Рис. 9.215 Өткен жобалар туралы деректерге негізделген ML моделі жаңа жобаның құнын және орындалу графигін белгілі бір ықтималдықпен анықтайды.

Модельді құрып, оқытқаннан кейін, жаңа жобаның құрылышын бағалау үшін модельге жаңа жобаның атрибуттарын беру жеткілікті, және модель белгілі бір ықтималдықпен бұрын зерттелген заңдылықтарға негізделген есептеу нәтижелерін ұсынады.

Машиналық оқытудың негізгі ұғымдары

Машиналық оқыту - бұл сиқыр емес, тек математика, деректер және заңдылықтарды іздеу. Ол шынайы интеллектке ие емес, тек деректер негізінде шаблондарды тануға және адамның тұрақты қатысусынсыз шешімдер қабылдауға үйретілген бағдарлама болып табылады.

Машиналық оқытудың құрылымын сипаттау үшін бірнеше негізгі ұғымдарды пайдаланады (Рис. 9.216):-

- Этикеттер (Labels) – бұл модельдің болжауы тиіс мақсатты айнымалылар немесе атрибуттар (Титаник деректер жиынтығындағы «Тірі қалу» параметрі). Мысал: құрылыш құны (мысалы, доллармен), құрылыш жұмыстарының ұзақтығы (мысалы, аймен).
- Ерекшеліктер (Features) – бұл модель үшін кіріс деректері ретінде қызмет ететін тәуелсіз айнымалылар немесе атрибуттар. Болжау моделінде олар меткаларды болжау үшін пайдаланылады. Мысалдар: учаскенің ауданы (шаршы метрмен), ғимараттың қабат саны, ғимараттың жалпы ауданы (шаршы метрмен), географиялық орналасуы (ендік және бойлық), құрылыш кезінде қолданылған материалдардың түрі. Ерекшеліктердің саны деректердің өлшемділігін де анықтайды.
- Модель (Model) – бұл болжау немесе аппроксимациялау қажет мақсатты функцияны жақыннататын әртүрлі гипотезалардың жиынтығы. Мысал: құрылыш құны мен мерзімдерін болжау үшін регрессиондық талдау әдістерін қолданатын машиналық оқыту модель.
- Оқыту алгоритмі (Learning Algorithm) – бұл модельдегі ең жақсы гипотезаны іздеу процесі, ол мақсатты функцияға дәл сәйкес келеді, оқыту деректерінің жиынтығын пайдалана отырып. Мысал: құрылыш құны мен мерзімдері туралы деректерді талдайтын сызықтық регрессия, KNN немесе кездейсоқ орман алгоритмі, байланыстар мен занұдилықтарды анықтау үшін.
- Оқыту (Training) – оқыту процесінде алгоритм оқыту деректерін талдайды, кіріс атрибуттары мен мақсатты меткалар арасындағы байланыстарды сәйкестендіреді. Бұл процестің нәтижесі – болжау жасауға дайын машиналық оқыту модельі. Мысал: алгоритм құрылыш тарихы туралы деректерді (құны, уақыты, объектінің ерекшеліктері) талдайтын процесс, болжау модельін құру үшін.

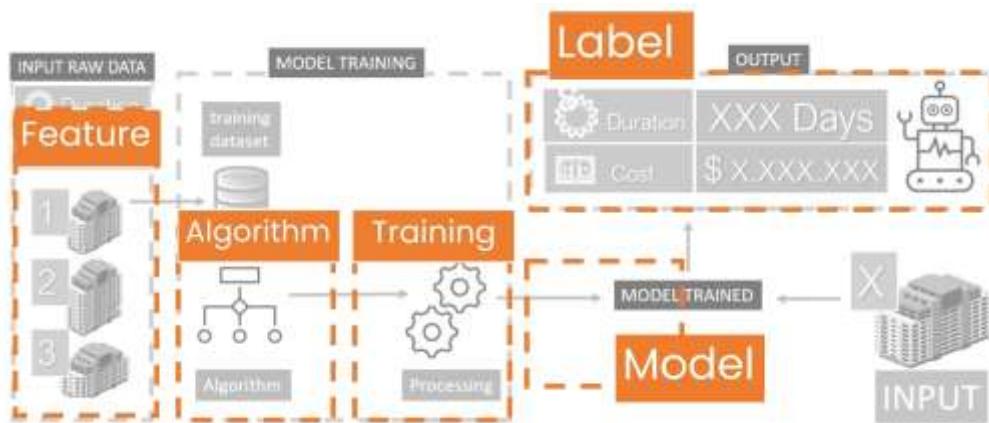


Рис. 9.216 ML меткелер мен атрибуттарды пайдаланып, деректерде алгоритмдер арқылы оқытылатын модельдер құрады, нәтижелерді болжау үшін.

Машиналық оқыту оқшауланған түрде емес, статистика, деректер базалары, интеллектуалды деректерді талдау, үлгілерді тану, үлкен деректер аналитикасы және жасанды интеллект сияқты аналитикалық пәндердің кең әкожүйесінің бір бөлігі болып табылады. Рис. 9.217 осы салалардың қалай қыылсысып, бір-бірін толықтырып, заманауи шешім қабылдау және автоматтандыру жүйелері үшін кешенді негіз құрайтынын көрсетеді.-

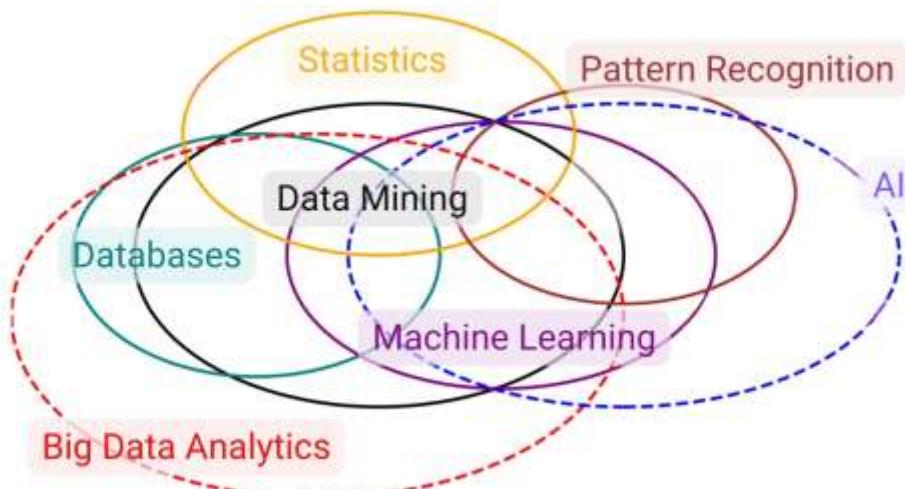


Рис. 9.217 Деректерді талдаудың әртүрлі салалары арасындағы өзара байланыс: статистика, машиналық оқыту, жасанды интеллект, үлкен деректер, үлгілерді тану және интеллектуалды деректерді талдау.

Машиналық оқытудың негізгі маңытасы – компьютерлерді адамның араласуынсыз немесе көмегінсіз автоматты түрде білім алуға және өз әрекеттерін сәйкесінше түзетуге мүмкіндік беру.

Осылайша, болашақта адамның рөлі тек машинаға когнитивтік мүмкіндіктерді беру болып табылады – ол шарттарды, салмақтарды және параметрлерді белгілейді, ал машинады оқыту моделі барлық қалғанын орындаиды.

Келесі тарауда алгоритмдерді қолданудың нақты мысалдары қарастырылады. Нақты кестелер мен қарапайым модельдерде қадам-қадаммен болжау қалай құрылатыны көрсетіледі.



ТАРАУ 9.3.

МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ ҚҰРЫЛЫС ҚҰНЫ ЖӘНЕ МЕРЗІМДЕРДІ БОЛЖАУ.

Машиналық оқытуды жобаның құнын және мерзімдерін анықтау үшін пайдалану мысалы.

Құрылыш мерзімдері мен құнын бағалау – құрылыш компаниясының қызметіндегі негізгі процесстердің бірі. Дәстүр бойынша, мұндай бағалаулар сарапшылар тарапынан тәжірибе, анықтамалықтар мен нормативтік базалар негізінде жүзеге асырылады. Алайда, цифрлық трансформация және деректердің қолжетімділігінің артуы жағдайында, машиналық оқыту (ML) модельдерін пайдалану мүмкіндігі пайда болады, бұл бағалаулардың дәлдігін арттыру мен автоматтандыруға мүмкіндік береді.

Машиналық оқытуды құрылыш құны мен мерзімдерін есептеу процесіне енгізу тек жоспарлаудың тиімділігін арттырып қана қоймай, сонымен қатар, тәуекелдерді басқарудан бастап логистика мен сатып алушы оңтайланудыруға дейінгі басқа бизнес-процестерге интеллектуалдық модельдерді интеграциялаудың бастапқы нүктесіне айналады.

Нысандың құрылыш мерзімін және оның жалпы құнын тез анықтау қабілеті маңызды. Бұл жобаның мерзімдері мен құны туралы сұрақтар, құрылыш саласының пайда болған сәтінен бастап, тапсырыс берушілер мен құрылыш компанияларының санасында орталық орын алады.

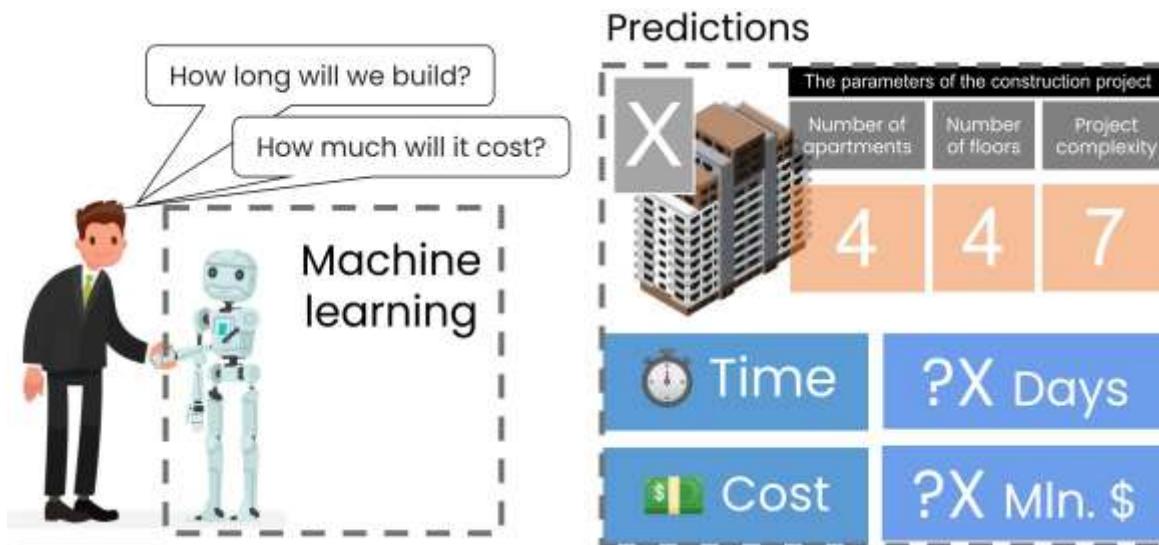


Рис. 9.31 Құрылыш жобаларында табыстың кілтті факторлары – құрылыш мерзімдері мен құнын бағалаудың жылдамдығы мен сапасы.

Келесі мысалда өткен жобалардан негізгі деректер алынып, олардың негізінде машиналық оқыту модельі әзірленеді, бұл модельдің көмегімен жаңа параметрлермен жаңа құрылыш жобаларының

құнын және мерзімдерін бағалауға мүмкіндік береді (Рис. 9.31).-

Уш жобаны үш негізгі атрибутпен қарастырайық: пәтерлер саны (мұнда 100 пәтер 10 санына эквивалентті, визуализацияны жеңілдету үшін), қабаттар саны және құрылым күрделілігінің шартты өлшемі 1-ден 10-ға дейінгі шкала бойынша, мұнда 10 - ең жоғары күрделілік көрсеткіші. Машиналық оқытуда мұндай мәндерді, мысалы, 100-ді 10-ға немесе 50-ді 5-ке айналдыру процесі "нормализация" деп аталады.

Машиналық оқытуда нормализация – әртүрлі сандық деректерді өңдеу мен талдауды жеңілдету үшін бір масштабқа келтіру процесі. Бұл процесс, әсіресе, деректердің әртүрлі масштабтар мен өлшем бірліктеріне ие болған жағдайда маңызды.

Бірінші жобада (Рис. 9.32) 50 пәтер (нормализациядан кейін – 5), 7 қабат және күрделілік бағасы 2 болды, бұл салыстырмалы түрде қарапайым құрылым құнын білдіреді. Екінші жобада 80 пәтер, 9 қабат және салыстырмалы түрде күрделі жоба болды. Мұндай жағдайда бірінші және екінші көпқабатты үйдің құрылым мерзімі 270 және 330 күнди құрады, ал жобаның жалпы құны сәйкесінше 4,5 және 5,8 миллион долларды құрады.

Construction project	The parameters of the construction project			The key parameters of the project	
	Number of apartment	Number of floors	Project complexity	Time	Cost
1	5	7	2	270	\$ 4.502.000
2	8	9	6	330	\$ 5.750.000
3	3	5	3	230	\$ 3.262.000
X	4	4	7	?X	\$?X. XXX.XXX

Рис. 9.32 Өткен жобалар жиынтығы, олар болашақ жоба X-тің уақытын және құнын бағалау үшін пайдаланылатын болады.

Машиналық оқыту моделін құру кезінде мұндай деректер үшін негізгі міндет - болжая үшін критикалық атрибуттарды (немесе белгілерді) анықтау, бұл жағдайда - құрылымының құнын көп болуы мүмкін - бірнеше ондағаннан жүздеген факторларға дейін. Оларға: құрылымының материалдарының түрі, климаттық аймақ, мердігерлердің біліктілік деңгейі, инженерлік желілердің болуы, негіздің түрі, жұмыстардың басталу маусымы, прорабтардың пікірлері және т.б. жатады.-

Прогностикалық машиналық оқыту моделін жасау үшін біз оны құру үшін алгоритмді таңдауымыз керек. Машиналық оқудағы алгоритм - бұл компьютерге деректер негізінде болжам жасауға (параметрлерді дұрыс тәртіпте араластыру) немесе шешім қабылдауға үйрететін математикалық рецепт.

Өткен құрылыш жобалары туралы деректерді талдау және болашақ жобалардың уақытын және құнын болжау үшін (Рис. 9.32) танымал машиналық оқыту алгоритмдерінің бірін қолдануға болады:-

- Линейлік регрессия (Linear regression): бұл алгоритм атрибуттар арасындағы тікелей байланысты табуға тырысады, мысалы, қабаттар саны мен құрылыш құны арасындағы. Алгоритмның мақсаты - осы байланысты ең жақсы сипаттайтын сызықтық теңдеуді табу, бұл болжау жасауға мүмкіндік береді.
- К-жақын көршілер алгоритмі (K-nearest neighbors (k-NN)): бұл алгоритм жаңа жобаны көлемі немесе күрделілігі бойынша үқсас өткен жобалармен салыстырады. k-NN деректерді k (сан) оқыту мысалдарының қайсысы оларға ең жақын екенін негізінде классификациялады. Регрессия контекстінде нәтиже k жақын көршілердің орташа мәні немесе медианасы болып табылады.
- Шешім ағаштары (Decision Trees): бұл деректерді әртүрлі шарттар негізінде подмножестерге бөлөтін болжамдық модель, ағаш тәрізді құрылымды пайдалана отырып. Әрбір ағаш түйіні - бұл деректерді одан әрі бөлуге әкелетін шарт немесе сұрақ, ал әрбір жапырақ - соңғы болжам немесе нәтиже. Алгоритм деректерді әртүрлі сипаттамалар негізінде, мысалы, алдымен қабаттар саны бойынша, содан кейін күрделілік бойынша және т.б. болжам жасау үшін кішірек топтарға бөледі.

Жаңа жобаның құнын бағалау үшін машиналық оқыту алгоритмдерін қарастырайық, мысалы, екі танымал алгоритм: линейлік регрессия және К-жақын көршілер алгоритмі.

Жобаның құны мен уақытын болжау үшін сызықтық регрессияны пайдалану.

Линейлік регрессия - деректерді талдаудың негізгі алгоритмі, ол бір немесе бірнеше басқа айнымалылармен сызықтық байланыс негізінде айнымалының мәнін болжауға мүмкіндік береді. Бұл модель тәуелді айнымалы мен бір немесе бірнеше тәуелсіз айнымалылар арасында тікелей сызықтық байланыс бар деп болжайды, ал алгоритмның мақсаты - осы байланысты табу.

Линейлік регрессияның қарапайымдылығы мен түсініктілігі оны әртүрлі салаларда танымал құралға айналдырды. Бір айнымалымен жұмыс істегендеге, линейлік регрессия деректер нүктелері арқылы өтетін ең жақсы сәйкес келетін сызықты табуға негізделеді.

Сызықтық регрессия кіріс айнымалысы X мен шығыс айнымалысы Y арасындағы байланысты аппроксимациялайтын ең жақсы тік сызықты (қызыл сыйық) табады. Бұл сыйық анықталған сызықтық тәуелділік негізінде жаңа X мәндері үшін Y мәндерін болжауға мүмкіндік береді (Рис. 9.33).-

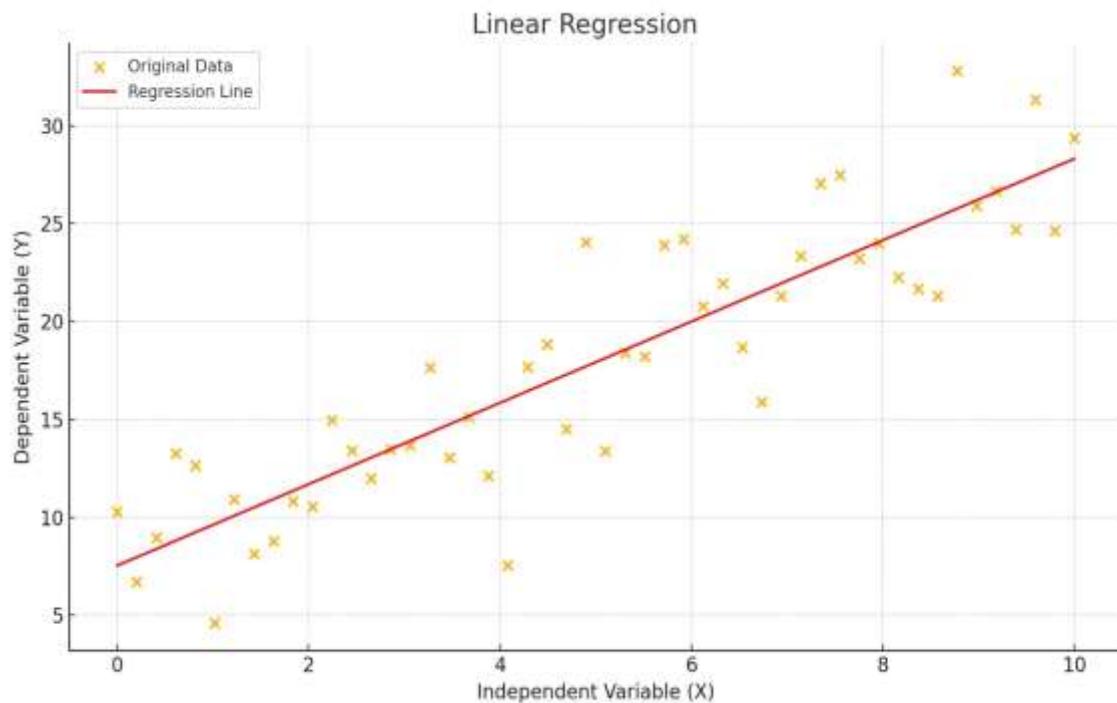


Рис. 9.33 Сызықтық регрессияның жұмыс принципі, ол жаттығу мәндері арқылы өтетін ең жақсы тік сызықты табады.

Бұл сызық тәуелді айнымалы Y -нің болжамды мәнін алу үшін тәуелсіз айнымалының (X) мәнін енгізу арқылы берілген теңдеумен ұсынылады. Бұл процесс X -тің белгілі мәндері негізінде Y -ні тиімді болжауға мүмкіндік береді, олардың арасындағы сызықтық тәуелділікті пайдалана отырып. Сан-Францискодан құрылыш рұқсаттары туралы деректерді бағалау мысалында статистикалық орташа сызықты табу процесін көруге болады (Рис. 9.17), мұнда инфляция сызықтық регрессия арқылы әртүрлі обьектілердің түрлері үшін есептелді.

Келіңіздер, жобаның деректер кестесін (Рис. 9.32 өткен тараудан) LLM-ге жүктеп, бізге қарапайым машиналық оқыту моделін құруды сұрайық.

- ❷ LLM чатында мәтіндік сұраныс жіберініз (ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN):

Жаңа жоба X -тің құны мен іске асыру уақытының болжамын жасау үшін қарапайым машиналық оқыту моделін құруды көрсету қажет (Рис. 9.32 қосымша су-реттінде) ↵ -

LLM-нің жауабы:

The screenshot shows a software interface with the following components:

- Title Bar:** Create in LLM Chat
- Icon Bar:** Includes icons for a cat, a flower, and a whale.
- Code Editor:** A window titled "Linear regression.py" containing Python code for linear regression. The code imports LinearRegression from sklearn.linear_model and uses numpy to handle data. It defines features X and targets y_time, y_cost, creates two models, fits them, and then predicts for a new project's features.
- Execution Area:** Shows the result of running the code: (array([238.44444444]), array([3042337.77777778])).
- Tool Buttons:** Includes icons for a terminal, PC, and file operations.
- Run Button:** Run in IDE

Рис. 9.34 LLM машиналық оқыту моделін құру үшін сызықтық регрессияны таңдады, ол жобаның құны мен уақытын болжау үшін.

LLM қосымша суреттен кестені автоматты түрде танып, деректерді визуалды форматтан кесте мас-сивіне түрлендірді (Рис. 9.34 – 6-шы жол). Бұл массив машиналық оқыту моделін құру үшін негіз ретінде пайдаланылды, онда сызықтық регрессия қолданылды (Рис. 9.34 – 17-22-ші жол).--

"Өте аз" деректер жиынында оқытылған базалық сызықтық регрессия моделінің көмегімен жаңа гипотетикалық құрылышы жобасы, Project X деп аталатын жоба үшін болжамдар жасалды. Біздің тапсырмада бұл жоба 40 пәтер, 4 қабат және 7 күрделілік деңгейімен сипатталады (Рис. 9.32).

Жаңа Project X үшін шектеулі және аз деректер жиынына негізделген сзықтық регрессия мөделимен жасалған болжамдарға сәйкес (Рис. 9.34 - 24-29-шы жол): -

- Құрылыс ұзақтығы шамамен 238 күнді (238,4444444) құрайды.
- Жалпы шығындар шамамен \$3 042 338 (3042337,777) болады.

Жобаның құны туралы гипотезаны тереңірек зерттеу үшін әртүрлі алгоритмдер мен машиналық оқыту әдістерімен эксперимент жасау пайдалы. Сондықтан K-Nearest Neighbours (k-NN) алгоритмының көмегімен жаңа жоба X үшін құн мен уақыттың сол мәндерін болжайық.

Жоба құны мен уақыттыңиң болжамдары К-ең жақын көршілер алгоритмы (k-NN) арқылы.

Қосымша болжам ретінде жаңа жобаның құнын және ұзақтығын бағалау үшін k-Nearest Neighbours (k-NN) алгоритмін қолданамыз. K-Nearest Neighbors (k-NN) алгоритмі – бұл бақылау арқылы машиналық оқыту әдісі, классификация мен регрессия үшін қолданылатын. Сонымен қатар, k-NN алгоритмі векторлық дерекқорларды іздеу контекстінде бұрын қарастырылған (Сур. 8.22), мұнда ол ең жақын векторларды (мысалы, мәтіндер, суреттер немесе техникалық сипаттамалар) табу үшін пайдаланылады. Бұл тәсілде әр жоба көп өлшемді кеңістіктегі нұктеден ұсынылады, мұнда әр өлшем жобаның белгілі бір атрибутына сәйкес келеді.-

Біздің жағдайда, әр жобаның үш атрибутын ескере отырып, оларды үш өлшемді кеңістіктегі нұктелер ретінде көрсетеміз (Сур. 9.35). Осылайша, біздің алдағы жоба X осы кеңістіктегі ($x=4, y=4, z=7$) координаттарымен локализацияланады. Нақты жағдайларда нұктелердің саны мен кеңістіктің өлшемі бірнеше есе көп болуы мүмкін екенін атап өткен жөн.

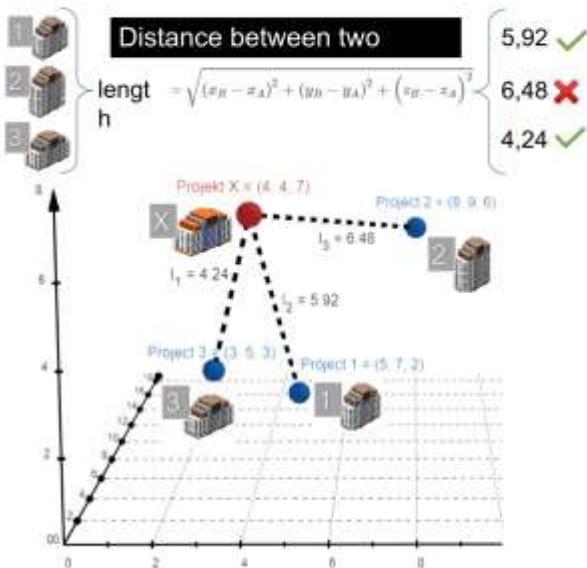
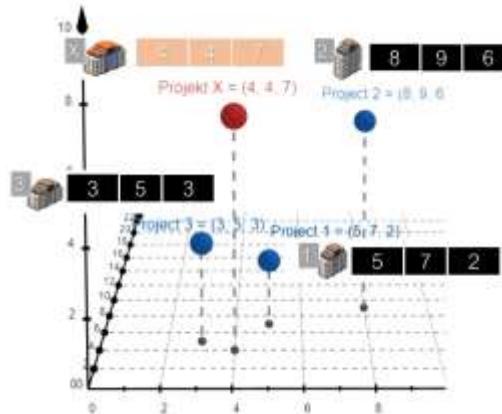
K-NN алгоритмі (k-nearest neighbors) қалаған жоба X пен оқу дерекқорындағы жобалар арасындағы қашықтықты өлшеу арқылы жұмыс істейді. Бұл қашықтықтарды салыстыра отырып, алгоритм жаңа жоба X нұктесіне ең жақын орналасқан жобаларды анықтайды.

Мысалы, егер екінші жоба ($x=8, y=9, z=6$) біздің бастапқы деректер жиынныңда X-тан айтарлықтай алыс орналасқан болса (Сур. 9.35), оны әрі қарайғы талдаудан шығарып таставуға болады. Нәтижесінде, есептеулер үшін тек екі ($k=2$) жақын жобаны пайдалану мүмкін, олардың негізінде орташа мән анықталады.

Мұндай әдіс, көршілерді іздеу арқылы, жобалар арасындағы ұқсастықты бағалауға мүмкіндік береді, бұл өз кезегінде бұрын жүзеге асырылған үқсас жобалар негізінде жаңа жобаның мүмкін болатын құны мен іске асыру мерзімдері туралы қорытындылар жасауға көмектеседі.

k-nearest neighbors algorithm

The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other.



Сур. 9.35 K-NN алгоритмінде жобалар көп өлшемді кеңістікте нұктелер ретінде ұсынылады, және үқсастықты бағалау мен болжам жасау үшін ең жақын жобалар қашықтықтар негізінде таңдалады.

K-NN жұмысы бірнеше негізгі кезеңдерді қамтиды:

- Деректерді дайындау: алдымен оқу және тестілеу деректер жиындары жүктеледі. Оқу деректері алгоритмді "жаттықтыру" үшін, ал тестілеу деректері оның тиімділігін тексеру үшін пайдаланылады.
- К параметрін таңдау: K саны таңдалады, ол алгоритмде қанша жақын көршіні (деректер нұктесін) ескеру керектігін көрсетеді. "K" мәні өте маңызды, себебі ол нәтижеге әсер етеді.
- Тестілеу деректері үшін классификация және регрессия процесі:
- Қашықтықтарды есептеу: тестілеу деректерінің әр элементі оқу деректерінің әр элементіне дейінгі қашықтықты есептейді (Сур. 9.35). Бұл үшін әртүрлі қашықтық өлшеу әдістері, мысалы, евклид қашықтығы (ең көп тараған әдіс), манхэттен қашықтығы немесе Хэмминг қашықтығы қолданылуы мүмкін.
- Сортау және K жақын көршіні таңдау: қашықтықтар есептелгеннен кейін олар сұрыпталып, тестілеу нұктесіне K жақын нұктелер таңдалады.
- Тесттік нұктенің класын немесе мәнін анықтау: егер бұл классификация тапсырмасы болса, тесттік нұктенің класы K таңдалған көршілер арасында ең жиі кездесетін клас негізінде анықталады. Егер бұл регрессия тапсырмасы болса, K көршілердің мәндерінің орташа мәні (немесе басқа орталық тенденция өлшемі) есептеледі.
- Процесті аяқтау: барлық тесттік деректер классификацияланған немесе оларға болжамдар жасалғаннан кейін процесс аяқталады.

K-Nearest Neighbors (k-NN) алгоритмі көптеген практикалық қолданбаларда тиімді және машиналық оқыту мамандарының құралдар арсеналында негізгі құралдардың бірі болып табылады. Бұл алгоритм деректер арасындағы байланыстарды оңай түсіндіруге болатын тапсырмаларда қарапайымдылығы мен тиімділігі арқасында танымал.

Біздің мысалымызда K-близжайших соседей алгоритмін қолданғаннан кейін X жобасына ең жақын екі жобаны (біздің шағын таңдамадан) анықтады. Осы жобалардың негізінде алгоритм олардың бағасы мен құрылыштықтының орташа мәнін анықтайды. Талдаудан кейін алгоритм, жақын көршілерді орташа есеппен, X жобасының құны шамамен 3 800 000 доллар және құрылыштық уақыты 250 күн болады деген қорытындыға келеді.

k-nearest neighbors algorithm



Рис. 9.36 K-Nearest Neighbors алгоритмі X жобасының құнын және кестесін анықтайды, таңдамадан екі ең жақын жобаны талдай отырып.

K-Nearest Neighbors (k-NN) алгоритмі әсіресе классификация және регрессия тапсырмаларында танымал, мысалы, ұсыныс жүйелерінде, мұнда ол нақты пайдалануышының қызығушылықтарына үқсас тауарлар немесе контент ұсыну үшін қолданылады. Сонымен қатар, k-NN медициналық диагностикада науқастың симптомдарына негізделген ауру түрлерін классификациялау, бейнелерді тану және қаржы секторында клиенттердің кредит қабілеттілігін бағалау үшін кеңінен қолданылады.

Деректердің шектеулі көлемі болған жағдайда да, машиналық оқыту модельдері пайдалы болжамдар бере алады және құрылыштық жобаларын басқаруда аналитикалық құрамды едәуір күшайте алады. Тарихи деректерді кеңейту және тазалау кезінде құрылым түрін, орналасуын, құрылыштық басталу маусымын және басқа факторларды ескере отырып, күрделі модельдерге көшу мүмкіндігі бар.

Біздің қарапайым тапсырмамызда үш өлшемді кеңістікте визуализация үшін үш атрибут қолданылды, бірақ нақты жобалар орташа алғанда жүздеген немесе мындаған атрибуттарды қамтиды, бұл кеңістіктің өлшемін және жобаларды векторлар түрінде көрсету күрделілігін едәуір арттырады. -

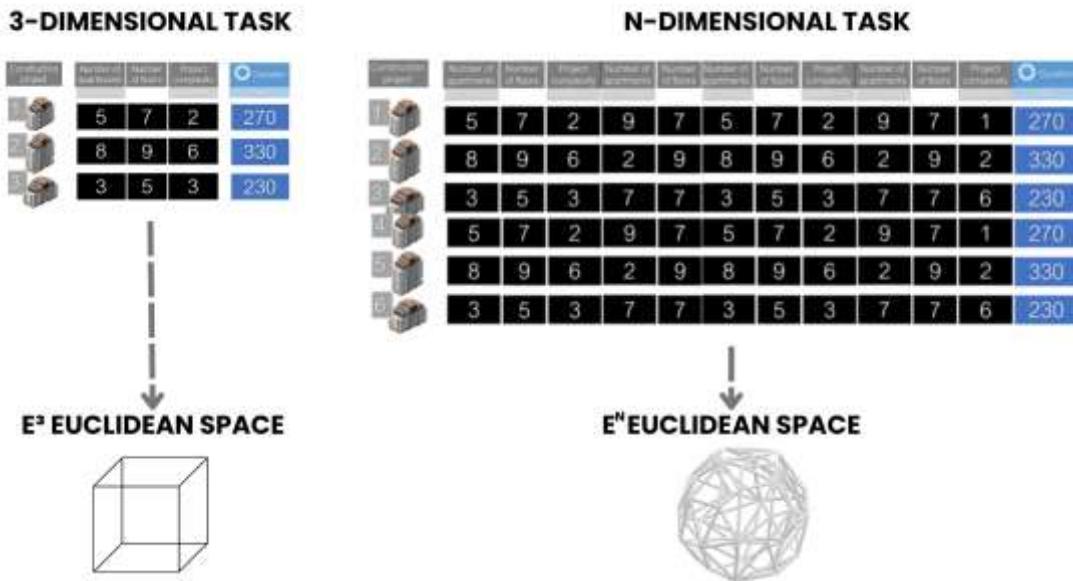


Рис. 9.37 Қарапайым мысалда 3D-визуализация үшін үш атрибут қолданылды, ал нақты жобаларда көп атрибуттар бар.

Бірдей деректер жиынына әртүрлі алгоритмдерді қолдану X жобасында 40 пәтер, 4 қабат және күрделілік деңгейі 7 болғанда әртүрлі болжамдық мәндерді берді. Линейлік регрессия алгоритмі аяқтау уақытын 238 күн және құнын 3 042 338 доллар деп болжады, ал k-NN алгоритмі 250 күн және 3 882 000 доллар деп болжады.-

Болжамдардың дәлдігі машиналық оқыту модельдері арқылы алынған, бастапқы деректердің көлемі мен сапасына тікелей байланысты. Жобалардың саны көп болған сайын, олардың сипаттамалары (белгілері) мен нәтижелері (белгілері) неғұрлым толық және дәл ұсынылған сайын, сенімді болжамдарды алу ықтималдығы жоғарылайды, ал қателік мәндері минималды болады.

Бұл процесте деректерді алдын ала өндөу әдістері маңызды рөл атқарады, соның ішінде:

- Нормализация, белгілерді біркелкі масштабқа келтіруге мүмкіндік беретін;
- Шығындарды анықтау және жою, модельдің бұрмалануын болдырмау;
- Категориялық белгілерді кодтау, мәтіндік деректермен жұмыс істеуге мүмкіндік береді;
- Жоғары тұрақтылықты қамтамасыз ететін жетіспейтін мәндерді толтыру.

Сонымен қатар, модельдің жалпы қабілетін және жаңа деректер жиынтықтарына тұрақтылығын бағалау үшін кросс-валидация әдістері қолданылады, бұл артық үйренуді анықтауға және болжамның сенімділігін арттыруға мүмкіндік береді.

Хаос – бұл шешуді қажет ететін тәртіп. Жозе Сарамаго, «Двойник»

Тіпті сіздің тапсырмаларыңыздың хаосын формалды түрде сипаттау мүмкін емес деп ойласаңыз да, біліңіз – әлемдегі кез келген оқиға, әсіресе құрылыш процестері математикалық заңдылықтарға бағынады, оларды есептеу үшін қатаң формулалар емес, статистика мен тарихи деректер арқылы қолдау көрсету қажет болуы мүмкін.

Дәстүрлі сметалық есептеулер, калькуляциялық бөлімдермен орындалатын, сондай-ақ машиналық оқыту модельдері міндетті түрде белгісіздік пен ықтимал қателік көздерімен бетпе-бет келеді. Дегенмен, сапалы деректердің жеткілікті көлемі болған жағдайда, машиналық оқыту модельдері сараптамалық бағалаулармен салыстырғанда теңдесіз, кейде тіпті жоғары дәлдікпен болжамдар көрсете алады.

Машиналық оқыту, ең алдымен, талдау үшін сенімді қосымша құралға айналатын болады, ол: есептеулердің нақтылауға, баламалы сценарийлерді ұсынуға және жобаның параметрлері арасындағы жасырын байланыстарды анықтауға мүмкіндік береді.

Мұндай модельдер әмбебаптыққа талап қоймайды, бірақ жақын арада жобалық шешімдер қабылдау процесінде маңызды орын алуы мүмкін. Машиналық оқыту технологиялары инженерлердің, сметашылардың және аналитиктердің қатысуын жоймайды, керісінше, тарихи деректерге негізделген қосымша көзқарас ұсынып, олардың мүмкіндіктерін көңейтеді.

Дұрыс интеграцияланған жағдайда құрылыш компанияларының бизнес-процестеріне машиналық оқыту басқарушылық шешімдерді қолдау жүйесінде маңызды элементке айналуы мүмкін – адамды алмастыру емес, оның кәсіби интуициясы мен инженерлік логикасын көңейту ретінде.

Келесі қадамдар: сақтау, талдау және болжау.

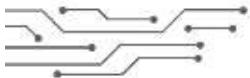
Заманауи деректермен жұмыс істеу тәсілдері құрылыш саласындағы шешім қабылдау принципперін өзгерте бастады. Интуитивті бағалаулардан объективті деректерді талдауға көшу тек дәлдікті арттырып қана қоймай, процестерді оңтайландыру үшін жаңа мүмкіндіктер ашады. Осы бөлімнің қорытындысында, қарастырылған әдістердің күнделікті міндеттерінде қолдануға көмектесетін негізгі практикалық қадамдарды атап өткен жөн:

- Тұрақты деректерді сақтау инфрақұрылымын құру
 - Бөлек құжаттар мен жобалық деректерді біртұтас кестелік модельге біріктіруге тырысыңыз, негізгі ақпаратты бір деректер фрейміне жинақтап, әрі қарай талдау үшін пайдаланыңыз.
 - Деректерді сақтау үшін тиімді форматтарды пайдаланыңыз – мысалы, CSV немесе XLSX орнына Apache Parquet сияқты бағандық форматтар – әсіресе, болашақта машиналық оқыту модельдерін оқыту үшін пайдаланылуы мүмкін деректер жиынтықтары үшін.
 - Жобаның барлық кезеңдерінде өзгерістерді бақылауға мүмкіндік беретін деректердің нұсқаларын басқару жүйесін құрыңыз.
- Талдау және автоматтандыру құралдарын енгізу.
 - Жобалардың тарихи деректерін – құжаттама, модельдер, сметалар – талдай бастаңыз,

занұдыштықтарды, трендтерді және аномалияларды анықтау үшін.

- Деректерді автоматты түрде жүктеу және дайындау үшін ETL (Extract, Transform, Load) процестерін менгеріңіз.
- Түрлі тегін Python визуализация кітапханаларын пайдалана отырып, негізгі метрикаларды визуализациялауды үйреніңіз.
- Репрезентативті және қайта өндіруге болатын аналитикалық қорытындылар алу үшін статистикалық әдістер мен кездейсоқ іріктеулерді қолдануды бастаңыз.
- Деректермен жұмыс істеу деңгейін арттыру.
 - Титаник деректер жынынтысы сияқты қарапайым және түсінікті мысалдарда бірнеше негізгі машиналық оқыту алгоритмдерін зерттеңіз.
 - Қазіргі процестерді талдап, қatal себеп-салдар логикасынан статистикалық болжамдау және бағалау әдістеріне көшу мүмкіндіктерін анықтаңыз.
 - Деректерді қосымша өнім емес, стратегиялық актив ретінде қарастыра бастаңыз: шешім қабылдау процестерін деректер модельдеріне негізделген етіп құрыңыз, нақты бағдарламалық шешімдер айналасында емес.

Деректердің құндылығын түсінген құрылым компаниялары жаңа даму кезеңіне өтуде, мұнда бәсекелестік артықшылық ресурстардың көлемімен емес, аналитика негізінде шешім қабылдау жылдамдығымен анықталады.



БАСПА НҰСҚАСЫМЕН МАКСИМАЛДЫ ҚОЛАЙЛЫЛЫҚ

Сіздер Data-Driven Construction-ның тегін цифрлық нұсқасын ұстап отырсыздар. Материалдарға жылдам қол жеткізу және ыңғайлыштың жұмыс үшін баспа басылымына назар аударуды ұсынамыз:



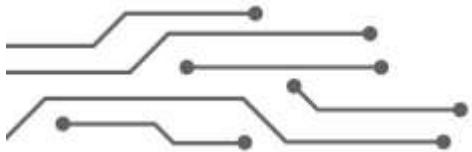
- Эрдайым қол астында: баспа форматындағы кітап сенімді жұмыс құралы болып табылады, қажетті визуализациялар мен схемаларды кез келген жұмыс жағдайында жылдам табуға және пайдалануға мүмкіндік береді.
- Суреттердің жоғары сапасы: баспа басылымындағы барлық бейнелер мен графиктер максималды сапада ұсынылған.
- Ақпаратқа жылдам қол жеткізу: ыңғайлыштың навигация, ескертпелер жасау, бетбелгілер қою және кітаппен кез келген жерде жұмыс істеу мүмкіндігі.



тиімді жұмыс істеуге арналған құрал алдырып: күнделікті тапсырмаларда визуалды материалдарды жедел пайдалану, қажетті схемаларды жылдам табу және ескертпелер жасау мүмкіндігі. Сонымен қатар, сатып алуыңыз ашық білімнің таралуын қолдайды.

Кітаптың баспа нұсқасын мына жерден тапсырыс беруге болады:
datadrivencconstruction.io/books

Кітаптың толық баспа нұсқасын сатып алу арқылы сіз ақпаратпен ыңғайлыштың және



Х БӨЛІМ Цифрлық деректер дәуіріндегі құрылым саласы. Мүмкіндіктер мен сын-тегеуіндер.

Оныншы бөлім цифрлық трансформация дәуіріндегі құрылым саласының болашағына кешенді көзқарас ұсынады. Мұнда себеп-салдарлы талдаудан үлкен деректер корреляцияларымен жұмыс істеуге көшу талданады. Көркем өнердің эволюциясы мен құрылым саласындағы деректермен жұмыс істедің дамуы арасындағы параллельдер жүргізіледі, сала детальды бақылаудан процестерді кешенді түсінуге қалай көшіп жатқанын көрсетеді. Деректердің ашықтығы мен есептеулердің автоматтандырылуы дәстүрлі бизнес модельдерін түбекейлі өзгертуі мүмкін "уберизация" концепциясы қарастырылады, делдалдардың қажеттілігін жоя отырып, спекуляция мүмкіндіктерін азайтады. Универсал класификация мәселелері, құрылым саласындағы жаңа жағдайларға бейімделу үшін уақыт беретін мәселелер егжей-тегжейлі талқыланады. Бөлім цифрлық трансформация стратегиясын қалыптастыру бойынша нақты ұсыныстармен аяқталады, бұл осалдықтарды талдауды және өзгермелі салада бәсекеге қабілеттілікті сақтау үшін қызмет спектрін кеңейтуді қамтиды.

ТАРАУ 10.1.

ТҮРМЫС СТРАТЕГИЯЛАРЫ: БӘСЕКЕЛЕСТІК АРТЫҚШЫЛЫҚТАРДЫ ҚАЛЫПТАСТЫРУ

Корреляциялар есептеулердің орнына: құрылымыс аналитикасының болашағы.

Ақпараттың жедел цифрландыруы (Рис. 1.15) қазіргі құрылымыс саласында түбегейлі трансформацияны тудырып отыр, мұнда деректер тек құрал емес, жобаларды басқару мен бизнесті жүргізудің дәстүрлі тәсілдерін түбегейлі өзгерте алатын стратегиялық активке айналуда.

Мыңжылдықтар бойы құрылымыс қызыметі детерминирленген әдістерге сүйенди – дәл есептеулер, бөлшектеу және параметрлердің қатаң бақылауы. Біздің әрамыздың алғашқы ғасырларында римдік инженерлер акведуктар мен көпірлерді салу үшін математикалық принциптерді қолданды. Орта ғасырларда архитекторлар готикалық соборлардың идеал пропорцияларына ұмтылды, ал XX ғасырдың индустриализация дәүірінде стандартталған нормативтер мен регламенттер жүйелері қалыптасты, олар жаппай құрылымыс негізіне айналды.

Бүгінгі таңда даму векторы қатаң себеп-салдарлық байланыстарды іздеуден ықтималдықты талдауға, корреляцияларды және жасырын занылыштарды іздеуге ауысып отыр. Сала жаңа кезеңге қадам басуда – деректер негізгі ресурсқа айналуда, ал олардың негізінде аналитика интуитивті және жергілікті оңтайланырылған тәсілдерді ығыстырып жатыр.



Рис. 10.11 Құрылымыс деректерінің жасырын әлеуеті: компаниядағы бар есептеулер - менеджментке талдау үшін қолжетімді айсбергтің тек шыңы.

Компанияның ақпараттық жүйесі айсбергке ұқсас (Рис. 10.11): компания менеджментіне деректердің тек кішкене бөлігі көрінеді, ал негізгі құндылық тереңдікте жасырылған. Деректерді тек олардың қазіргі қолданылуы бойынша ғана емес, сонымен қатар олар болашақта ашатын мүмкіндіктері бойынша бағалау маңызды. Деректерден жасырын заңдылықтарды шығарып, жаңа білімдер жасауға үйренетін компаниялар ғана тұрақты бәсекелестік артықшылықты қалыптастыра алады.

Жасырын заңдылықтарды іздеу және деректерді түсіну – бұл тек сандармен жұмыс емес, абстракттылы ойлауды және шашыраңқы элементтердің тұтас көрінісін көру қабілетін талап ететін шығармашылық процесс. Осы тұрғыдан деректермен жұмыс істеудің дамуын бейнелеу өнерінің эволюциясымен салыстыруға болады (Рис. 10.12).

Құрылыш саласының дамуы бейнелеу өнерінің прогресіне таңғаларлықтай ұқсас. Екі жағдайда да адамзат примитивті әдістерден визуализация мен талдаудың күрделі технологияларына дейінгі жолдан өтті. Алғашқы адамдар құнделікті міндеттерді шешу үшін петроглифтер мен примитивті құралдарды қолданды. Орта ғасырлар мен Ренессанс дәүірінде архитектура мен өнердегі күрделілік деңгейі айтартылғатай өсті. Орта ғасырлардың басында құрылыш құралдары қарапайым балталардан техникалық білімнің өсуін білдіретін кең құралдар жиынтығына дейін эволюцияланды.

Реализм дәүірі бейнелеу өнеріндегі алғашқы революция болды: суретшілер ең кішкентай бөлшектерді қайта өндіруді үйреніп, бейненің максималды шынайылығына қол жеткізді. Құрылыш саласындағы осы кезеңнің аналогы дәл инженерлік әдістемелер, егжей-тегжейлі сызбалар және ғасырлар бойы жобалау практикасының негізіне айналған қатаң регламенттелген есептеулер болды.

Кейін импрессионизм өнердің шынайы шындықты қабылдауын өзгерти: форма мен мазмұнды дәл беру орнына суретшілер көңіл-күйді, жарықты және динамиканы фиксируғе көшті, жалпы әсерді көрсетуге тырысты, абсолютті дәлдіктен гөрі. Осылайша, құрылыш аналитикасындағы машиналық оқыту қатал логикалық модельдерден паттерндерді және ықтимал заңдылықтарды тануға ауысады, бұл классикалық талдауда қолжетімсіз деректердегі жасырын тәуелділіктерді "көріп" түсінуге мүмкіндік береді. Бұл тәсіл Bauhaus-тың минимализм мен функционализм идеяларымен үндеседі, мұнда мағына (функция) формаға қарағанда маңызды. Bauhaus артықшылықтарды жоюға, айқындық, утилитарлық және жаппай өндіріс үшін ою-өрнектен бас тартуға ұмтылды. Заттар түсінікті және пайдалы болуы тиіс, артықшылықтарсыз – эстетика конструкцияның логикасы мен мақсатыннан туындаиды.

19 ғасырдың соңында фотографияның пайда болуы өнерге шынайы шындықты бұрын-соңды болмаған дәлдікпен фиксируе үшін жаңа құрал берді және бейнелеу өнеріне деген көзқарасты өзгерти. Сол сияқты, 21 ғасырдағы өнеркәсіптік революция құрылыш саласында роботтандырылған технологияларды, лазерлерді, IoT, RFID және "Connected Construction" сияқты концепцияларды қолдануға әкеледі, мұнда жеке параметрлердің жинақталуы құрылыш алаңының толық шындықты масштабалатын интеллектуалды фиксируеүіне эволюцияланады.

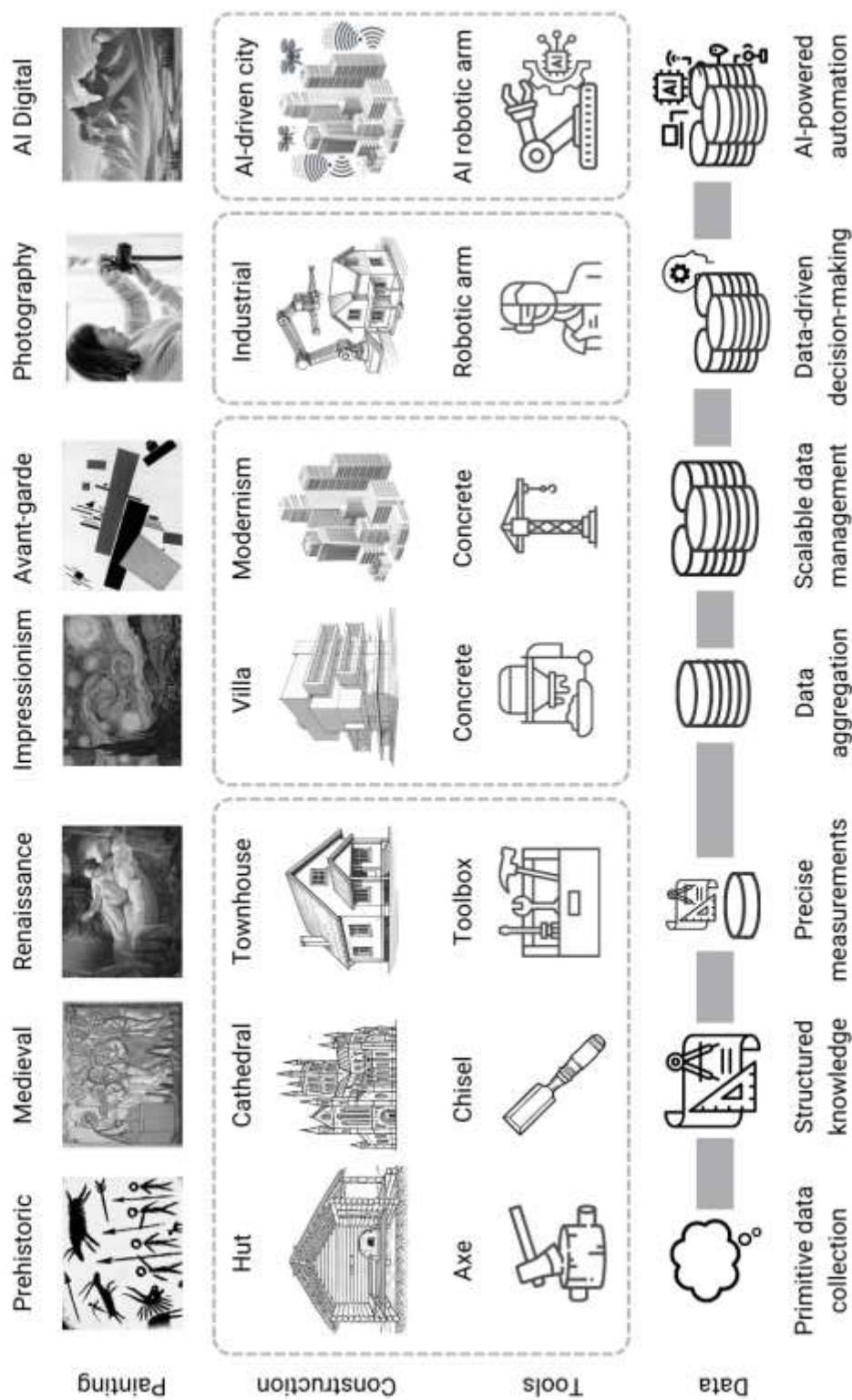


Рис. 10.12 Бейнелеу өнерінің эволюция кезеңдері құрылыш саласындағы деректермен жұмыс істеу тәсілдерінің дамуымен сәйкес келеді.

Бұғынгі күні, дәл солай, бейнелеу өнері AI және LLM құралдарының келуімен қайта қаралу кезеңін бастаң кешірсе, құрылыш саласы да кезекті кванттық секірісін бастаң кешіруде: жасанды интеллект (ИИ) басқаратын интеллектуалды жүйелер, LLM-ченнеллар минималды адам араласуымен шешімдерді болжауға, оңтайландыруға және генерациялауға мүмкіндік береді.

Жобалау мен басқаруда деректердің рөлі радикалды түрде өзгерді. Бұрын білім ауызша беріліп, эмпирикалық сипатта болды – XIX ғасырдан бұрын шындық қолмен жазылған картиналармен фиксируленгендей, ғұғынгі күні құрылыш "картиналарының" толық цифрлық фиксирулеуі орталық назарда. Машиналық оқыту алгоритмдері арқылы бұл цифрлық картина құрылыш шындығының импрессионистік көрінісіне трансформацияланады – дәл көшірме емес, процестердің жалпы, ықтинал түсінігі.

Біз жобалау, құрылыш және ғимараттарды пайдалану процестері тек толықтырылып қана қоймай, сонымен қатар едәуір дәрежеде жасанды интеллект жүйелерімен басқарылатын дәуірге жақындап келеміз. Заманауи цифрлық өнер киesticіз – мәтіндік промпттар мен генеративті модельдер арқылы жасалатын сияқты, болашақтағы архитектуралық және инженерлік шешімдер пайдаланушының қойған негізгі сұраулары мен параметрлері негізінде қалыптасады.

XXI ғасырда деректерге қол жеткізу, олардың интерпретациясы және аналитиканың сапасы жобаның табысының ажырамас шарттарына айналады. Сонымен қатар, деректердің құндылығы олардың көлемімен емес, мамандардың оларды талдау, текстеру және әрекетке айналдыру қабілетімен анықталады.

Деректерге негізделген тәсіл құрылыш саласында: жаңа деңгейдегі инфрақұрылым

Адамзат тарихында әрбір осындағы технологиялық секіріс экономика мен қоғамда түбегейлі өзгерістер әкелді. Бұгін біз XIX ғасырдағы өнеркәсіптік революциямен салыстыруға болатын жаңа трансформация толқынын байқаймыз. Дегенмен, жүз жыл бұрын өзгерістердің негізгі қозғаушы күші механикалық күштер мен энергетикалық технологиялар болса, қазіргі уақытта – бұл деректер мен жасанды интеллект.

Машиналық оқыту, LLM және ИИ агенттері қосыншалардың мәнін өзгертуде, дәстүрлі бағдарламалық стек (кітаптың екінші бөлімінде талқыланған) қажетсіз болып қалуда. Деректермен жұмыс істеудің барлық логикасы ИИ-агенттерінде шоғырланады, ал қатты кодталған бизнес-ережелерде емес.–

Деректер дәуірінде дәстүрлі қосыншалар туралы түсініктер түбегейлі трансформациялануда. Біз көлемді корпоративтік модульдік жүйелердің ашық, жеңіл, мамандандырылған шешімдерге орын беретін модельге қарай жылжудамыз.

Болашақта тек деректердің негізгі құрылымы қалады, ал онымен барлық өзара әрекеттесу дерекқорымен тікелей жұмыс істейтін агенттер арқылы жүзеге асырылады. Мен шын жүректен сенемін, бүкіл қосымша стек жойылады, ейткені жасанды интеллект негізгі дерекқормен тікелей әрекеттескенде оған қажеттілік жоқ. Менің мансабымның барлық кезеңінде SaaS саласында жұмыс істедім – компаниялар құрдым, оларда жұмыс істедім, және шындықты айтсам, қазір жаңа SaaS-бизнесті бастауға, бәлкім, бармас едім. Және, мүмкін, мен қазіргі уақытта SaaS-компанияларға инвестиция салmas едім. Жағдай тым белгісіз. Бұл болашақта бағдарламалық компаниялар болмайды де-генді білдірмейді, тек олар мүлдем басқа түрде болады. Болашақ жүйелері дерекқорлардан тұрады, ал бизнес-логика ИИ агенттеріне шығарылған болады. Бұл агенттер бірнеше дереккөздермен бір уақытта жұмыс істейді, бір дерекқормен шектелмейді. Барлық логика жасанды интеллект деңгейіне ауысады. – Мэттью Берман, Forward Future CEO

Жаңа парадигманың негізгі ерекшелігі - технологиялық жүктемені минимизациялау. Монументалды, құрделі және жабық бағдарламалық кешендердің орнына біз деректер ағынында "тіршілік ететін" икемді, ашық және жылдам бапталатын модульдерді аламыз. Болашақ процестерді басқару архитектурасы микро қосымшаларды – компактты, мақсатты құралдарды пайдалануды көздейді, олар массивті және жабық ERP, PMIS, CDE, CAFM жүйелерінен принципті түрде ерекшеленеді. Жаңа агенттер максималды бейімделгіш, интеграцияланған және нақты бизнес міндеттеріне бағытталған болады.-

Барлық бизнес-логика осы [ЖИ] агенттеріне көшеді, және бұл агенттер бірнеше репозиторийлерде CRUD [Жасау, Оқу, Жаңарту және Жою] операцияларын орындаиды, яғни олар қандай бәкенд қолданылып жатқанын ажыратпайды. Олар бірнеше дерекқорларды жаңартады, ал барлық логика так называемом ЖИ-деңгейінде болады. ЖИ-деңгейі барлық логиканың орналасқан жері болған кезде, адамдар бәкендерді ауыстыра бастайды. Біз Dynamics бәкенд нарығында агенттердің қолданылуында едәуір жоғары женіс пайызын байқап отырмыз, және біз осы бағытта белсенді түрде алға жылжитын боламыз, барлығын біріктіруге тырысамыз. Бұл клиенттерді қызмет көрсету саласында немесе басқа салаларда, мысалы, тек CRM емес, сонымен қатар қаржы және операциялар бойынша шешімдерімізде. Себебі адамдар ЖИ-бағдарланған бизнес-қосымшаларды көбірек қалайды, онда логикалық деңгей ЖИ мен ЖИ агенттерімен басқарылуы мүмкін. [...]. Мен үшін ең қызықты нәрселердің бірі – Python-мен Excel, GitHub-тың Copilot-ымен салыстырғанда. Яғни, біз не жасадық: енді сізде Excel бар, оны ашип, Copilot-ты іске қосып, онымен ойнауды бастауыңыз керек. Бұл тек бар сандарды түсіну емес – ол өзі жоспар құрады. GitHub Copilot Workspace жоспар құрып, содан кейін оны орындастырыңыз сияқты, бұл Excel-ді жолдар мен бағандарды визуализациялау құралы ретінде пайдаланатын деректер аналитигінің жұмысына ұксас. Осылайша, Copilot Excel-ді барлық мүмкіндіктерімен құрал ретінде пайдаланады, себебі ол деректердің генерациялай алады және Python интерпретаторына ие. – Сатья Наделла, CEO, Microsoft, интервью каналу BG2 декабрь 2024 г. [28]

Біз көріп отырған оффис қосымшалары логикасының трансформациясы — модульдік, жабық жүйелерден тікелей ашық деректермен жұмыс істейтін ЖИ-агенттерге көшу — әлдеқайда ауқымды процестің тек бір бөлігі. Бұл интерфейстерді немесе бағдарламалық архитектуралы туралы ғана емес: өзгерістер еңбек ұйымдастыру, шешім қабылдау және бизнес басқару негіздерін қамтиды. Құрылыш саласында бұл деректерге негізделген логиканың қалыптасуына әкеледі, онда деректер процестердің орталық элементі болады — жобалаудан бастап ресурстарды басқару және құрылыш барысын бақылауға дейін.

Келесі үрпақтың цифрлық оффисі: AI жұмыс кеңістігін қалай өзгертеді.

Жұз жыл бұрын адамзат осындай технологиялық революцияны бастап өткерді. Бу машиналарынан электр қозғалқыштарына көшу төрт онжылдықтан астам уақытты алды, бірақ ақырында бұл өнімділіктің бұрын-соңды болмаған өсуінің катализаторы болды — ең алдымен энергия қуаттарын орталықсыздандыру және жаңа шешімдердің икемділігі арқасында. Бұл өзгеріс тек тарихтың ағымын өзгертпей, халықтың негізгі массасын ауылдардан қалаларға көшіруге себеп болды, сонымен қатар қазіргі экономиканың негізін қалады. Технология тарихы — физикалық еңбекпен автоматтандыру мен интеллектуалды жүйелерге дейінгі жол. Трактор оншақты жер өндеушіні алмастырғандай, қазіргі цифрлық технологиялар дәстүрлі оффис басқару әдістерін (Рис. 10.13) ығыстырады. XX ғасырдың басында жер шарының көп бөлігі жерді қолмен өндейді, 1930-шы жылдары машиналар мен тракторлардың көмегімен еңбек механизациясы басталғанға дейін.

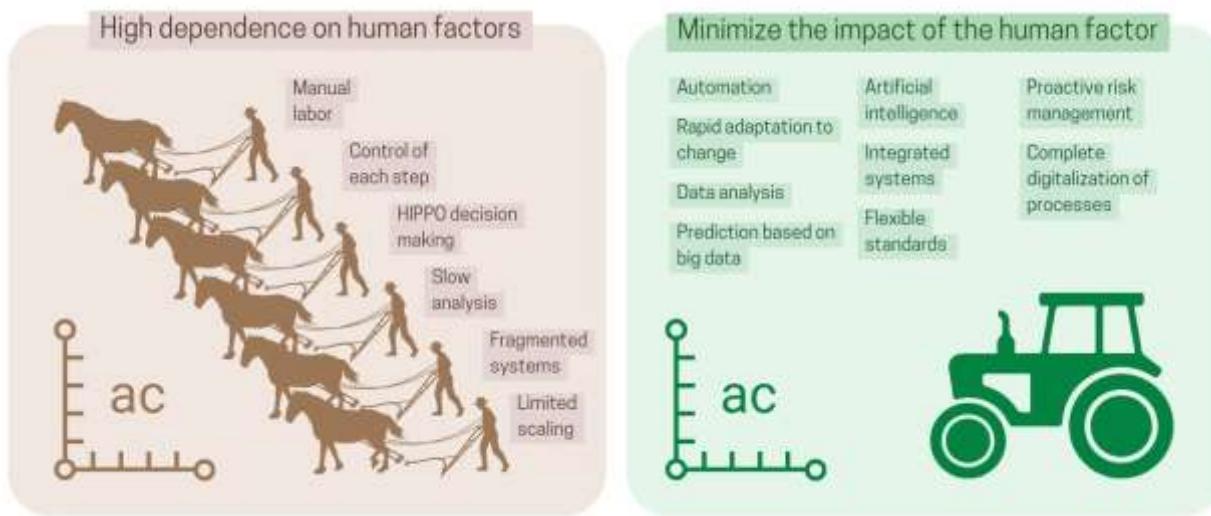


Рис. 10.13 Трактор 20 ғасырдың басында оншақты адамды алмастырғандай, машиналық оқыту 21 ғасырда дәстүрлі бизнес және жобаларды басқару әдістерін алмастырады.

Адамзат жұз жыл бұрын примитивті құралдармен жеке жер участкерін өндеуден, техника қолдану арқылы ауқымды ауыл шаруашылығына көшкендей, бүгін біз ақпараттың шашыраңқы «силостары» өндеуден, ETL-пайплайндары мен жасанды интеллект алгоритмдері арқылы деректер мас-сивтерімен жұмыс істеуге көшудеміз.

Біз цифрлық деңгейде дәстүрлі, қолмен басқарудан деректерге негізделген модельдерге көшу

кезеңінде тұрмыз.

Толыққанды деректерге негізделген архитектураға жету жолы уақыт, инвестициялар және үйимдастырушылық күш-жігерді талап етеді. Бірақ бұл жол тек біртіндеп жақсартуға ғана емес, құрылыш процестерінің тиімділігін, ашықтығын және басқарылуын арттыруға бағытталған сапалы серпіліс ашады. Мұның бәрі цифрлық құралдарды жүйелі енгізу және ескірген бизнес-практикалардан бастарту жағдайында жүзеге асады.

Міндеттерді параметрлеу, ETL, LLM, IoT компоненттері, RFID, токенизация, үлкен деректер және машиналық оқыту дәстүрлі құрылыш саласын деректерге негізделген құрылышқа айналдырады, мұнда жобаның және құрылыш бизнесінің әрбір бөлшегі деректер арқылы бақыланып, оңтайландырылатын болады.

Бұрын ақпаратты талдау үшін мындаған адам-сағат қажет болатын. Енді бұл міндеттер алгоритмдер мен LLM-дер арқылы жүзеге асырылады, олар шашыраңқы деректер массивтерін промпттар арқылы стратегиялық көздерге айналдырады. Технологиялық әлемде ауыл шаруашылығында болған жағдай қайталануда: мотыгадан автоматтандырылған агрокомплексқа көшу. Сол сияқты, құрылыш саласындағы офис жұмысы да Excel файлдарынан және қолмен жинақтаудан интеллектуалды жүйеге өтуде, мұнда деректер жиналып, тазаланып, құрылымдалып, инсайттарға айналады.

Бұғаңғы күні компаниялар сапалы деректер жинау және ақпаратты құрылымдау арқылы ақпараттық алқаптарды «егу» және тазалау мен нормализациялау құралдарымен «тыңайту» жұмыстарын бастау керек, содан кейін «урожай жинау» — болжамды аналитика мен автоматтандырылған шешімдер түрінде. Егер қазіргі заманғы фермер машинамен жүздеген жер жыртушыларды алмастыра алса, онда интеллектуалды алгоритмдер де қызметкерлердің рутиналық жұмыстарын алып, оларды ақпараттық ағындарды стратегиялық басқарушыларға айналдыра алады.

Алайда, нағыз деректерге негізделген үйим құру — бұл жылдам процесс емес. Бұл жаңа орман участкесін өсіру сияқты ұзақ мерзімді стратегиялық бағыт, мұнда әрбір «ағаш» — бұл жеке процесс, құзырет немесе құрал, ол өсіп, дамуға уақытты талап етеді. Және нағыз орман жағдайында, табыс тек отырғызу материалдарының (технологиялардың) сапасына ғана емес, сонымен қатар топырақтың (корпоративтік мәдениет), климаттың (бизнес ортасы) және күтімнің (жүйелі тәсіл) сапасына да байланысты. -

Компаниялар енді тек жабық «қораптан шыққан» шешімдерге сенім арта алмайды. Технологиялық дамудың алдыңғы кезеңдерінен айырмашылығы, қазіргі өтпелі кезең — деректерге ашық қолжетімділікке, жасанды интеллекттің пайдалануға және Open Source-тың таралуына бағытталған, бұл ірі вендорлардың орныққан бизнес модельдеріне және негізгі табыс көздеріне тікелей қауіп төндіреді.

Гарвард бизнес мектебінің зерттеуі [40] көрсеткендегі, бұрын талқыланған төртінші және бесінші технологиялық революциялар туралы тарауда, ең көп қолданылатын Open Source шешімдерін нөлден жасау шығындары барлық компаниялар үшін шамамен 4,15 миллиард долларды құрайды. Дегенмен, егер әрбір компания өздерінің баламаларын әзірлеуді бастаса, бұрыннан бар Open Source құралдарына қол жеткізбей, соңғы онжылдықтарда болғандай, бизнес шығындарының жалпы көлемі 8,8 триллион долларға жетуі мүмкін — бұл бағдарламалық қамтамасыз ету нарығын бағалаудағы рационал емес сұраныстың бағасы.

Технологиялық прогресс міндетті түрде орнықсан бизнес модельдерін қайта қарауға әкеледі. Бұрын компаниялар күрделі, мөлдір емес процестер мен жабық деректер арқылы табыс таба алса, ИИ мен аналитиканың дамуы бұл тәсілді барған сайын өміршең емес етеді.

Нәтижесінде, деректер мен құралдарға қол жетімділіктің демократиялануы дәстүрлі бағдарламалық қамтамасыз ету нарығын едәуір қысқартуы мүмкін. Дегенмен, осыған параллель жаңа нарық – цифрлық сараптама, кастомизация, интеграция және шешімдерді жобалау нарығы өседі. Мұнда құндылық лицензияларды сатудан емес, икемді, ашық және бейімделетін цифрлық процестерді құру қабілетінен қалыптасады. Электрификация мен тракторлардың пайда болуы жаңа өнеркәсіп салаларын тудырғандай, үлкен деректер, ИИ және LLM қолдану бизнес үшін мұлдем жаңа мүмкіндіктер ашады, бұл тек технологиялық инвестицияларды ғана емес, сонымен қатар ойлау, процестер мен үйымдық құрылымдарды терең трансформациялауды талап етеді. Бұл өзгерістерді түсініп, бүгіннен бастап әрекет ететін компаниялар мен мамандар ертеңгі күннің көшбасшылары болады.

Ашық деректер басты активке айналатын әлемде ақпараттың қолжетімділігі ойын ережелерін өзгертерді. Инвесторлар, тапсырыс берушілер мен реттеушілер барған сайын мөлдірлікі талап етеді, ал машиналық оқыту алгоритмдері сметалар, мерзімдер мен шығындардағы сәйкесіздіктерді автоматты түрде анықтай алады. Бұл бізді құрылыш саласындағы "уберизацияға" әкелетін жаңа цифрлық трансформация кезеңіне жағдай жасайды.

Ашық деректер мен уберизация – бұл қазіргі құрылыш бизнесіне қауіп тәндіреді.

Құрылыш ақпаратты басқару процесіне айналуда. Деректер неғұрлым дәл, сапалы және толық болса, жобалау, есептеулер, сметалық калькуляциялар, ғимараттарды салу және пайдалану соғұрлым тиімді болады. Болашақта негізгі ресурс кран, бетон және арматураның болуы емес, ақпаратты жинау, талдау және пайдалану қабілеті болады.

Құрылыш компанияларының клиенттері – құрылышқа қаржыландыру жасайтын инвесторлар мен тапсырыс берушілер болашақта ашық деректер мен тарихи деректер аналитикасының құндылығын міндетті түрде пайдаланады. Бұл сметалық есептеулерді автоматтандыруға, құрылыш компанияларын калькуляция мәселелеріне тартпай-ақ, жобалардың мерзімдері мен құнын есептеуге мүмкіндік береді, бұл шығындарды бақылауға және артық шығындарды тез анықтауға мүмкіндік береді.

Құрылыш алаңын елестетіңіз, мұнда лазерлік сканерлер, квадрокоптерлер және фотограмметрия жүйелері нақты уақыт режимінде пайдаланылатын бетон көлемдері туралы дәл деректер жиһайды. Бұл ақпарат автоматты түрде метадеректермен бірге қарапайым жазық MESH-модельдеріне түрлендіріледі, күрделі CAD (BIM) жүйелерін айналып өтіп, күрделі геометриялық ядроларға, ERP немесе PMIS-ке тәуелділіктерсіз. Құрылыш алаңынан жиналған деректер орталықтандырылған түрде біртұтас құрылымдалған қоймаларға ауыстырылады, олар тапсырыс берушіге тәуелсіз талдау үшін қолжетімді, мұнда әртүрлі құрылыш дүкендерінен нақты бағалар жүктеледі

және, мысалы, кредиттік қаржыландыру ставкасынан бастап, динамикалық түрде өзгеретін факторларға, құрылымдың материалдарының биржалық бағаларына, логистикалық тарифтерге және еңбек күшінің бағасындағы статистикалық маусымдық ауытқуларға дейінгі әртурлі параметрлер жүктеледі. Мұндай жағдайда жобалық және нақты материал көлемдері арасындағы кез келген айырмашылықтар бірден анықталады, бұл жобалау кезеңінде де, объектіні тапсыру кезінде де сметалық есептеулермен манипуляциялауды мүмкін емес етеді. Нәтижесінде құрылымдың процесінің ашықтығы бақылаушылар мен менеджерлердің армиясы арқылы емес, объективті цифрлық де-ректер арқылы қамтамасыз етіледі, мұнда адам факторы мен спекуляция мүмкіндігі минималды деңгейге дейін азаяды.

Мұндай деректерді бақылау жұмысы болашақта тапсырыс беруші тарапынан деректер менеджерлері тарапынан жүргізіледі (Рис. 1.24 CQMS менеджері). Бұл әсіресе жобалардың калькуляциялары мен сметалық есептеулеріне қатысты: бұрынғы сметашылар бөлімінің орнына, ертең машиналық оқыту және болжай құралдары құрылымдық компанияларына сәйкес келуі тиіс бағалық шектерді орнататын болады.

Құрылымдық саласының фрагменттелген сипатын ескере отырып, мұнда көптеген жүйелер мен подсистемалар шағын және орта кәсіпорындармен жеткізілетінін ескеру қажет, цифрлық стратегия тапсырыс берушіден басталуы тиіс. Тапсырыс берушілер цифрлық мүмкіндіктерді жеткізу тізбегінде ашу үшін жағдайлар мен механизмдер құруы қажет.

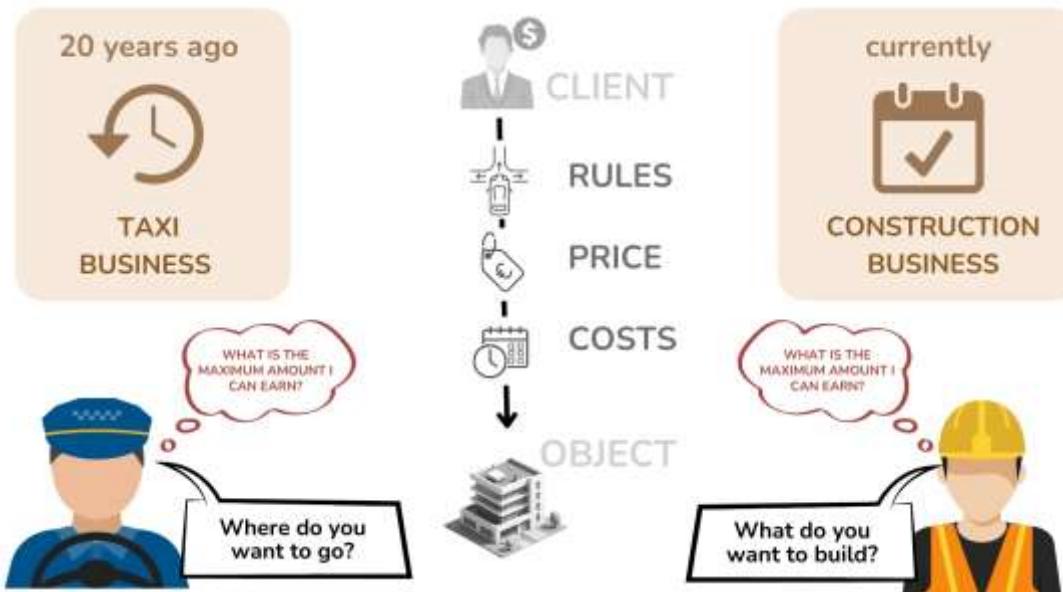
Эндрю Дэвис және Джюлиано Деникол, Accenture «Капиталдық жобалар арқылы үлкен құндылық жасау»

Мұндай деректердің ашықтығы мен мөлдірлігі құрылымдық компанияларына қауіп төндіреді, олар процестердің мөлдір еместігі мен күрделі есептерде табыс табуға үйренген. Мұнда спекуляцияларды және жасырын шығындарды күрделі және жабық форматтар мен модульдік патенттелген деректер беру платформаларының арқасында жасыруға болады. Сондықтан құрылымдық компаниялары, вендорлар тарапынан Open Source шешімдерін ілгерілету жағдайында, өз бизнес-процестеріне ашық деректерді толыққанды енгізуге қызығушылық танытпайды. Егер деректер тапсырыс беруші үшін қолжетімді және оңай өндөлетін болса, олар автоматты түрде тексерілуі мүмкін, бұл көлемдерді асыра көрсету мен сметалармен манипуляциялауды жояды.

World Economic Forum-ның «Құрылымдық болашағын қалыптастыру» (2016) [5] есебіне сәйкес, саладағы негізгі мәселелердің бірі тапсырыс берушінің пассивті рөлі болып қала береді. Дегенмен, тапсырыс берушілер жобалардың нәтижесіне үлкен жауапкершілік алуды тиіс – ерте жоспарлаудан, тұрақты өзара әрекеттесу модельдерін таңдаудан бастап, орындауды бақылауға дейін. Жобалардың иелері тарапынан белсенді қатысу болмаса, құрылымдық саласының жүйелі трансформациясы мүмкін емес.

Бақылауды жоғалту көлемдер мен құндарды есептеуде соңғы 20 жылда басқа экономикалық салаларда трансформацияға әкелді, клиенттерге тікелей, делдалдарсыз, өз мақсаттарына жетуге мүмкіндік берді. Цифрландыру мен деректердің ашықтығы көптеген дәстүрлі бизнес-модельдерді өзгертті, мысалы, Uber пайда болғаннан кейін таксистермен, Airbnb келгеннен кейін қонақ үйлермен, Amazon дамығаннан кейін бөлшек саудагерлермен және дүкендермен, сондай-ақ банктермен

– нео-банктер мен децентрализовандық финтех-екожүйелердің өсүі арқасында, мұнда ақпаратқа тікелей қолжетімділік пен уақыт пен құнды автоматтандыру делдалдардың рөлін едәуір төмендедті.



Строительный бизнес таксистер, қонақ үйлер мен сатушылар 10 жыл бұрын бетпе-бет келген «уберизациям» бетпе-бет келеді.

Деректер мен оларды өңдеу құралдарына қолжетімділіктің демократияландыру процесі 不可避免, және уақыт өте келе жобаның барлық құрамдас бөліктері бойынша ашық деректер клиенттің талаптарына айналып, жаңа стандартқа айналады. Сондықтан ашық форматтарды енгізу және ашық есептеулер мәселелері инвесторлар, тапсырыс берушілер, банктер мен жеке инвестициялық қорлар тарапынан алға жылжытылатын болады – олар, ақырында, салынған объектілердің соңғы пайдаланушылары болып табылады және объектіні ондаған жылдар бойы пайдаланады.

Ірі инвесторлар, тапсырыс берушілер мен банктер құрылымы саласында ашықтықты талап етеді. Accenture-дің «Капиталдық жобалар арқылы үлкен құндылық жасау» (2020) зерттеуіне сәйкес, ашық және сенімді деректер құрылымы саласындағы инвестициялық шешімдер үшін шешуші факторға айналуда. Сарапшылардың айтуынша, сенімді және тиімді жобаларды басқару ашықтықсыз мүмкін емес, әсіресе дағдарыс жағдайында. Сонымен қатар, актив иелері мен мердігерлер деректер алмасуды және бірлескен аналитиканы ынталандыратын келісімшарттарға жиі көшуде, бұл инвесторлар, банктер мен реттешушілер тарапынан жауапкершілік пен ашықтыққа деген өсіп келе жатқан талаптарды көрсетеді.

Инвестор мен тапсырыс берушінің идеядан дайын ғимаратқа дейінгі қозғалысы болашақта автопилотпен саяхатқа ұқсас болады - құрылымы компаниясының делдалысызыз, спекуляциялар мен белгісіздіктерден тәуелсіз болады.

Ашық деректер мен автоматтандыру дәүірі құрылымың бизнесін, банктік іс, сауда, ауыл шаруашылығы және логистика сияқты салаларда болғандай, сөзсіз өзгертеңді. Бұл салаларда делдалдардың рөлі мен дәстүрлі бизнес жүргізу тәсілдері автоматтандыру мен роботтандыруға орын беріп, негізсіз үстемелер мен спекуляцияларға орын қалдырмайды.

Адамның экономикалық қызметінің барлық түрлеріндегі деректер мен процестер құрылымы саласындағы мамандардың бетпе-бет келетіндерінен еш айырмашылығы жоқ. Ұзақ мерзімді перспективада, бүгінгі күні нарықта үстемдік етіп, бағалар мен қызмет сапасының стандарттарын орнатын құрылымы компаниялары тапсырыс беруші мен оның құрылымы жобасы арасындағы негізгі делдал рөлін жоғалтуы мүмкін.

Нерешілген мәселелерді үберизациялау соңғы мүмкіндік ретінде уақытты трансформациялау үшін пайдалану.

Бірақ құрылымы саласындағы шындықтарға оралайық. Қазіргі уақытта кейбір экономикалық секторларда автономды көліктер, орталықтандырылмаған қаржылық жүйелер және жасанды интеллект негізіндегі шешімдер пайда болып жатқанда, құрылымы компанияларының едәуір бөлігі әлі де қағаз жүзіндегі үйімдар болып табылады, онда негізгі шешімдер көбінесе жеке мамандардың интуициясы мен тәжірибесіне негізделеді.

Осы парадигмада заманауи құрылымы компаниясын 20 жыл бұрынғы такси паркімен салыстыруға болады, ол ресурстарды, маршруттарды және жеткізу уақытын бақылап, «сапардың» мерзімі мен құнына жауап береді – жобалық идеядан (логистика және монтаж процесі) объектіні тапсыруға дейін. GPS (құрылымы саласында IoT, RFID) және машиналық оқыту алгоритмдері көлік саласын қалай өзгерткен болса, деректер, алгоритмдер және ИИ-агенттер құрылымы басқаруды интуитивті бағалаудан болжамды, басқарылатын модельге айналдыруы мүмкін. Соңғы 20 жылда қаржы, ауыл шаруашылығы, бөлшек сауда және логистика сияқты көптеген салаларда деректердің мәлдір еместігі арқылы спекуляция жасау мүмкіндігі біртіндеп жойылды. Бағалар, жеткізу немесе қаржылық транзакциялардың құны автоматты түрде және статистикалық негізде есептеледі – бар болғаны бірнеше секунд ішінде цифрлық платформаларда.

Болашаққа көз жүгірте отырып, құрылымы компаниялары деректер мен оларды талдау құралдауна қол жеткізудің демократиялануы жобалардың құны мен мерзімдерін бағалаудағы дәстүрлі тәсілді бұзатынын және көлемдер мен бағалар туралы мәлдір емес деректерге спекуляция жасау мүмкіндігін жоятынын түсінуі тиіс.

Драйвердің араласуынсыз реттелетін жолмен жүру сияқты, болашақтағы құрылымы процестері «уберленген» жүйеге ұқсайтын болады – мерзімдер мен құнды автоматтандырылған бағалау, тапсырмалардың мәлдір маршрутизациясы және адам факторының минималды тәуелділігімен. Бұл «сапардың» идеядан жүзеге асуына дейінгі табиғатын өзгертеңді – оны болжамды, басқарылатын және деректерге негізделген етеді.

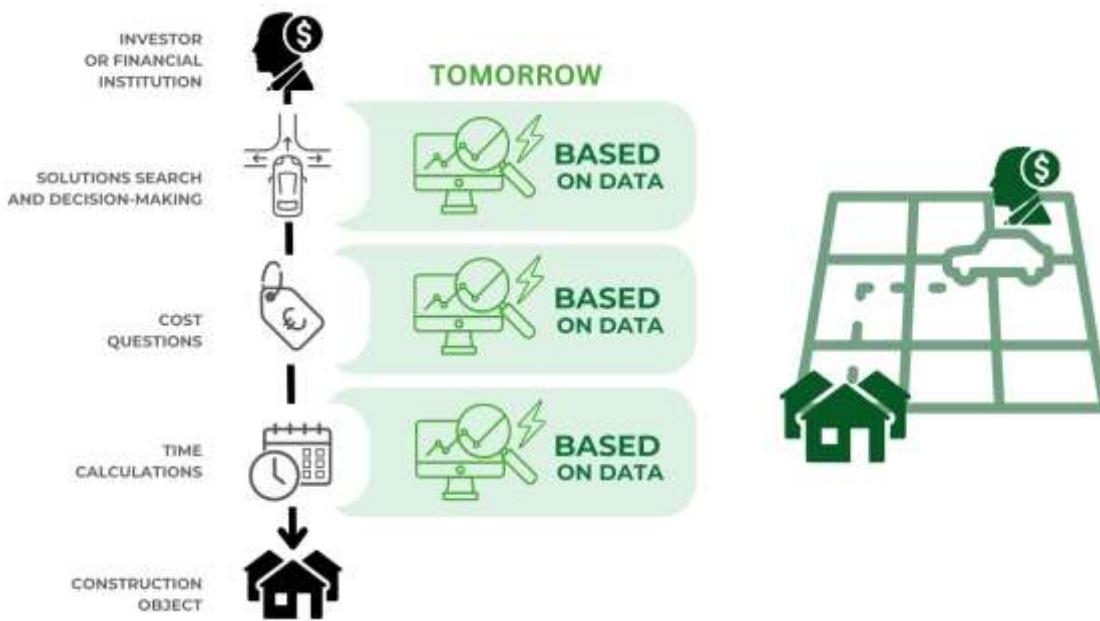


Рис. 10.15 Құрылыс процесіндегі «жолдың» құны мен уақыты машиналық оқыту және статистикалық құралдар арқылы анықталады.

Әлемнің әртүрлі елдерінде CAD (BIM) модельдерін клиенттерге немесе құрылыс жобаларын қаржыландыратын банктерге беру міндеттемелерін енгізу барысында клиент пен тапсырыс беруші деректердің құны мен көлемдерін есептеудің мөлдірлігін қамтамасыз ету мүмкіндігіне ие болады. Бұл әсіресе жеткілікті біліктіліктері мен нарықтық бағаларды жедел талдау құралдары бар ірі тапсырыс берушілер мен инвесторлар үшін өзекті. Көлемді стандартты жобаларды жүзеге асыратын компаниялар – дүкендер, офис ғимараттары, тұрғын үй кешендері – мұндай практикаларды стандартқа айналдыруда.

Модельдердің ақпараттық мазмұны толық және стандартталған сайын, манипуляциялау және спекуляция жасау мүмкіндігі дерлік жойылады. Цифрлық трансформация құрылыс саласындағы ойын ережелерін біртіндеп өзгертуде, және осы өзгерістерге бейімделмеген компаниялар ауыр қындықтарға тап болуы мүмкін.

Бәсекелестіктің күшеюі, технологиялық алшақтық және маржиналдылықтың төмендеуі бизнестің тұрақтылығына әсер етуі мүмкін. Шектеулі ликвидтілік жағдайында саладағы көптеген қатысушылар автоматтандыру, аналитика және деректерді өңдеу технологияларына тиімділікті және процестердің ашықтығын арттыру тәсілі ретінде жүгінеді. Бұл құралдар өзгермелі экономикалық ортада бәсекеге қабілеттілікті сақтау үшін маңызды ресурсқа айналуда.

Сыртқы жағдайлардың шұғыл шаралар қабылдауға мәжбүрлеуін күтудің қажеті жоқ – бұғыннен бастап цифрлық құзыреттілікті нығайтып, заманауи шешімдерді енгізіп, деректермен жұмыс істеуге бағытталған мәдениетті қалыптастыру әлдекәйда тиімдірек.

Құрылыс саласындағы ауқымды цифрлық трансформацияға кедергі келтіретін соғы негізгі технологиялық тосқауылдардың бірі – құрылыс жобалары элементтерін автоматты түрде классификациялау мәселесі.

Сенімді, дәл және масштабталатын классификациясыз толыққанды аналитика, процесстерді автоматтандыру және объектілердің өмірлік циклын басқару үшін ИИ мен предиктивті модельдерді қолдану негізін құру мүмкін емес. Объектілердің классификациясы әлі күнге дейін тәжірибелі мамандар – прорабтар, жобалаушылар, сметашылар – тарапынан қолмен интерпретациялауға тәуелді болғандықтан, құрылыш саласында мүмкіндіктер терезесі ашық. Бұл уақытты ашықтыққа, құралдар мен деректердің демократиялануына, сондай-ақ автоматты классификация жүйелерінің пайда болуына байланысты міндетті өзгерістерге дайындық үшін пайдалануға болады, олар ойын ережелерін түбөгейлі өзгертеді.

Құрылыш әлемінің элементтерін автоматты классификациялау міндеті құрделілігі жағынан автономды жүргізу жүйелеріндегі объектілерді тану мәселесімен салыстыруға болады, бұл негізгі сынқатерлердің бірі. Автономды көлікті А нұктесінен Б нұктесіне қозғалатын деп елестетейік. Заманауи автоматты жүргізу жүйелері лидараторлар мен камералар арқылы танылатын объектілерді классификациялау мәселесіне тап болады. Көлікке кедергі немесе нысанды «көріп» қана қою жеткіліксіз – ол алдында не тұрғанын: жаяу жүргінші, жол белгісі немесе қоқыс контейнері екенін қате түсінбеуі керек.

Сол сияқты, құрылыш саласының алдында да негізгі проблема тұр. Жобаның элементтері – тере зелер, есіктер немесе бағандар – құжаттамада тіркелуі, CAD-модельдерінде ұсынылуы, құрылыш алаңында суретке түсірілуі немесе лазерлік сканерлеу нәтижесінде нұктелер бұлтында танылуы мүмкін. Алайда, шын мәнінде автоматтандырылған жобаны басқару жүйесін құру үшін тек визуалды немесе жалпы геометриялық тану жеткіліксіз. Эр элементтің типі бойынша дәл және тұрақты классификациясын қамтамасыз ету қажет, ол барлық кейінгі процестерде – сметалар мен спецификациялардан бастап логистика, қойма есебі және ең бастысы – пайдалану кезінде біржақты анықталуы тиіс.

Дәл осы кезеңде – танудан мағыналы классификацияға ету кезінде – негізгі кедергілердің бірі пайда болады. Цифрлық жүйелер техникалық тұрғыдан объектілерді модельдер мен құрылыш алаңында анықтап, идентификациялауға қабілетті болса да, негізгі қыныңдық әртүрлі бағдарламалық орта үшін элемент түрін дұрыс және контекстуалды түрде анықтауда жатыр. Мысалы, есік жобалаушы тарапынан CAD-модельде «есік» категориясына жататын элемент ретінде белгіленуі мүмкін, алайда ERP немесе PMIS жүйесіне берілгенде, ол жобалаушының қателігі немесе жүйелер арасындағы сәйкесіздіктер салдарынан дұрыс типизацияланбауы мүмкін. Сонымен қатар, элемент деректерді экспорттау және импорттау кезінде маңызды атрибуттарының бір бөлігін жоғалтуы немесе мүлдем жүйелік есептен жоғалуы жиі кездеседі. Бұл деректер ағынында үзіліс тудырады және құрылыш процестерінің толық цифрандыру принципін бұзады. Осылайша, «көрінетін» және «түсінікті» семантикалық мағына арасында критикалық алшақтық қалыптасады, бұл деректердің тұтастығын бұзады және құрылыш объектісінің өмірлік циклі бойында процестерді автоматтандыруды айтартылғатай қынданады.

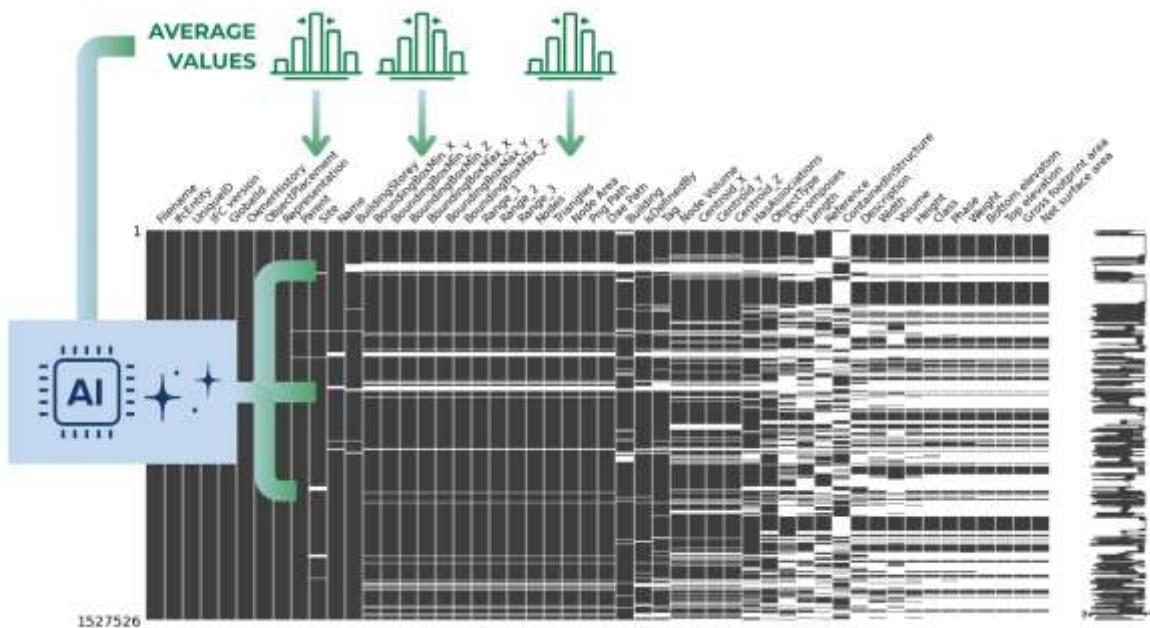
Құрылыш элементтерін универсалды классификациялау міндетін үлкен деректер мен машиналық оқыту технологияларын қолдану арқылы шешу (Рис. 10.16) бүкіл саланың трансформациясының катализаторы болады – және, мүмкін, көптеген құрылыш компаниялары үшін күтпеген жаңалық. Біртұтас, үйретілетін классификация жүйесі масштабталатын аналитика, цифрлық басқару және

ИИ-ді құрылыс үйімдарының күнделікті практикасына енгізу үшін негіз болады.-

NVIDIA және басқа технологиялық көшбасшылар бүгінгі күні басқа экономиканың салаларында автоматты түрде үлкен көлемдегі мәтіндік және визуалдық ақпаратты классификациялау және құрылымдау шешімдерін ұсынуда.

NVIDIA-ның NeMo Curator [161] моделі, мысалы, автоматты классификация мен деректерді алдын ала белгіленген категорияларға бөлуге маманданған, генеративті ИИ модельдерін алдын ала оқыту және нәзік баптау тапсырмалары үшін ақпаратты өңдеу конвейерлерін оңтайландыруды маңызды рөл атқарады. Cosmos платформасы нақты бейнелер мен 3D-сценаларда оқытылады [162], автономды жүйелер мен цифрлық егіздерді құру үшін негіз қалыптастырады, олар NVIDIA әкожүйесінде қазірдің өзінде жасалуда. NVIDIA Omniverse 2025 жылға қарай USD форматымен жұмыс істеудің жетекші құралына айналды – жобалық ақпаратты беру процестерінде IFC форматын алмастыруға қабілетті универсалды сахналарды сипаттау. Isaac Sim – роботтандырылған процестерді симуляциялау [163] – NeMo Curator, Cosmos және Omniverse сияқты шешімдермен бірге деректерді тазалау мен сұзу, оқыту жинақтарын генерациялау, объектілердің қасиеттерін модельдеу және құрылыш алаңында роботтарды оқыту сияқты автоматтандырудың жаңа деңгейін ұсынады. Сонымен қатар, барлық осы құралдар тегін және ашық қолжетімді түрде таратылады, бұл инженерлік және құрылыш практикаларына енгізу үшін кедергілерді айтартылғанды.

Деректерді автоматты түрде жіктеу құрылымдық кестелер деңгейінде – бұл бірінші көзқарас бойынша қындық тудыратын міндет емес. Алдыңғы тарауда көрсеткеніміздей, жинақталған тарихи деректер болған жағдайда, ұқсас параметрлер негізінде сыныптардың жетіспейтін немесе дұрыс емес мәндерін толықтыру мүмкіндігі бар. Егер бірнеше аяқталған жобаларда ұқсас сипаттамалары бар элементтер дұрыс жіктелген болса, жүйе жаңа немесе толық емес элемент үшін сәйкес мәнді жоғары ықтималдықпен ұсына алады. Мұндай логика, орташа мәндер мен контекстті талдауға негізделген, сметалар, спецификациялар немесе CAD модельдерінен келетін кестелік деректерді жаппай өңдеу кезінде әсіресе тиімді болуы мүмкін.-

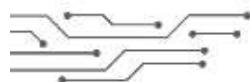


Машиналық оқыту өткен жобалар негізінде толтырылмаған (ақ орындар) кесте параметрлері үшін орташа мәндерді автоматтаты түрде табуға көмектеседі.

Машиналық оқудағы осындағы жылдам прогрестің аясында, 2025 жылы құрылышы элементерін автоматтаты түрде жіктеу мәселесі ұзақ уақыт бойы шешілмейді деп надандықпен ойлау айқын. Иә, қазіргі алгоритмдер толық жетілмеген, әсіресе толық емес немесе гетерогенді деректер жағдайында, бірақ бейімделу мүмкіндіктері тез жабылууда.

Қазіргі уақытта деректерді жинауға, тазартуға және жүйелеуге инвестиция салып жатқан компаниялар, сондай-ақ ETL автоматизацияция құралдарын игеріп жатқан ұйымдар, айқын түрде артықшылыққа ие болады. Басқалары уақытынан кеш қалу қаупіне ұшырайды – дәл сол сияқты, бұрын цифрлық трансформацияның шақыруларына жауап берे алмаған компаниялар көлік және қаржы салаларында.

Деректерді қолмен басқаруға және дәстүрлі бағалау әдістеріне сүйенетіндер 2000-шы жылдардағы такси парктерінің жағдайына тап болуы мүмкін, олар 2020-шы жылдардың басында мобиЛЬДІ қосымшалар мен автоматтандырылған маршрут есептеулер дәүіріне бейімделе алмады.



ТАРАУ 10.2.

ДЕРЕККЕ НЕГІЗДЕЛГЕН ЖАҚЫНДЫ ЖҰЗЕГЕ АСЫРУ БОЙЫНША ПРАКТИКАЛЫҚ НҰСҚАУЛЫҚ

Теориядан практикаға: құрылыс саласындағы цифрлық трансформацияның жол картасы.

Құрылыс саласы біртіндеп жаңа даму кезеңіне өтуде, мұнда дәстүрлі процестер жиі цифрлық платформалармен және ашық өзара әрекеттесу модельдерімен толықтырылады – кейде тіпті ауыстырылады. Бұл компанияларға тек шақырулар ғана емес, сонымен қатар елеулі мүмкіндіктер ашады. Бұғынгі күні ұзақ мерзімді цифрлық стратегияны қалыптастырып жатқан үйымдар нарықтағы өз орнын сақтап қана қоймай, заманауи тәсілдер мен сенімді, технологиялық түрғыдан негізделген шешімдерді ұсынып, оны кеңейте алады.

Сонымен қатар, концепциялар мен технологияларды білу – тек бастапқы нүктө. Басшылар мен мамандар алдында практикалық сұрақ туындаиды: енгізуі қайдан бастау керек және теориялық идеяларды нақты құндылыққа қалай айналдыруға болады. Сонымен қатар, дәстүрлі бағалау әдістері тапсырыс беруші тарапынан кез келген уақытта қайта қаралуы мүмкін болғандықтан, бизнес қандай негізде құрылатыны туралы сұрақ жиі туындаиды.

Жауап, мүмкін, технологияларда емес, деректермен жұмыс істеуді күнделікті практикаға кіріктіруде жаңа кәсіби мәдениетті қалыптастыруда жатыр. Цифрлық технологияларға және инновацияларға жеткіліксіз назар аудару құрылыс саласының соңғы онжылдықтардағы елеулі артта қалуына әкелді.

McKinsey деректеріне сәйкес, құрылыс саласындағы ғылыми-зерттеу және тәжірибелік-конструкторлық жұмыстарға жұмсалатын шығындар табыстың 1%-ынан аз, ал автомобиль және аэроғарыш өнеркәсібінде бұл көрсеткіш 3,5-4,5%-ға жетеді. Сол сияқты, құрылыс саласындағы IT-технологияларға жұмсалатын шығындар жалпы табыстың 1%-ынан төмен деңгейде қалып отыр [107].

Нәтижесінде, автоматтандыру деңгейі ғана емес, еңбек өнімділігі де құрылыс саласында төмендейде, және 2020 жылға қарай құрылыс жұмысшысы жарты ғасыр бұрынғыдан да аз өнім шығарады (Рис. 10.21).

Құрылыс секторындағы өнімділік мәселелері дамыған және дамушы елдердің көпшілігіне тән (құрылыс өнімділігі 29 ОЭСР елінің 16-да төмендеді (Рис. 2.21)), және бұл тек технологиялардың жетіспеушілігін емес, сонымен қатар басқару, оқыту және инновацияларды енгізу тәсілдерінде жүйелі өзгерістер қажеттілігін көрсетеді.

Цифрлық трансформацияның табысы құралдардың саны мен қолжетімділігіне емес, ұйымдардың өз процестерін қайта қарау және өзгерістерге ашық мәдениетті дамыту қабілетіне байланысты. Технологиялардың өздері емес, адамдар мен олардың тиімді қолданылуын қамтамасыз ететін құрылымдалған процестер, үздіксіз оқытуды қолдайтын және жаңа идеяларды қабылдауға ықпал ететін процестер маңызды рөл атқарады.

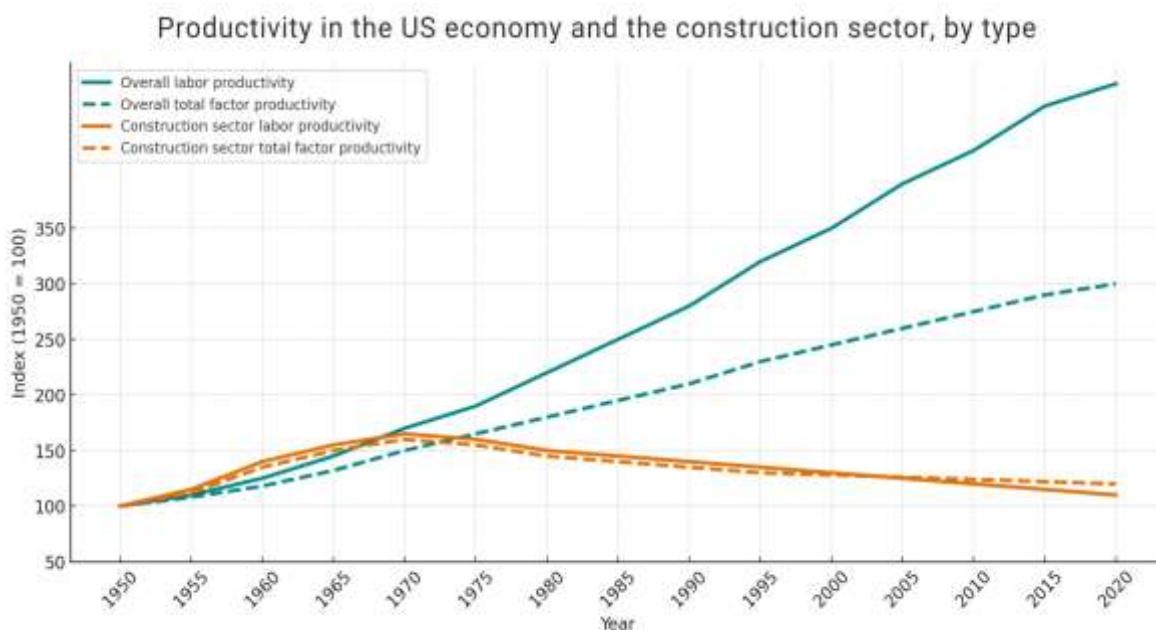


Рис. 10.21 АҚШ экономикасындағы еңбек өнімділігі мен жалпы ресурстық өнімділіктің парадоксы және құрылымдық секторындағы жағдай (1950-2020) (материалдар бойынша [43]).

Кітаптың алғашқы бөлімдерінде бизнес-орта моделі орман экожүйесімен салыстырылды (Рис. 2.12, Рис. 1.24, Рис. 1.32). Денсаулықты орманда кезең-кезеңімен болатын өрттер, олардың жойқын күшіне қарамастан, ұзақ мерзімді жаңаруда маңызды рөл атқарады. Олар ескі өсімдіктерден то-пышрақты тазартады, жинақталған қоректік заттарды қайтарады және жаңа өмір үшін орын жасайды. Кейбір өсімдік түрлері тіпті осылай эволюцияланып, олардың тұқымдары тек өрттің жоғары температурасының әсерінен ашылады – бұл табиғи механизм, өсу үшін тамаша уақытты қамтамасыз етеді.-

Бизнеске де осылай: дағдарыстар "бақыланатын өртеу" рөлін атқара отырып, ескірген жүйелермен байланысы жоқ жаңа тәсілдер мен компаниялардың пайда болуына ықпал етеді. Мұндай кезеңдер тиімді емес практикалардан бас тартуға мәжбүр етеді, инновациялар үшін ресурстарды босатады. Орман өртінен кейінгі өсімдіктер-пионерлер сияқты, дағдарыстан кейін бизнес жаңа, икемді процестерді қалыптастырады, олар жетілген ақпараттық орта үшін негіз болады.

"Сигналдық өрттерді" дұрыс интерпретациялап, олардың жойқын энергиясын конструктивті өзгерістерге айналдыруға компаниялар жаңа тиімділік деңгейіне шығады – деректерді өндөудің

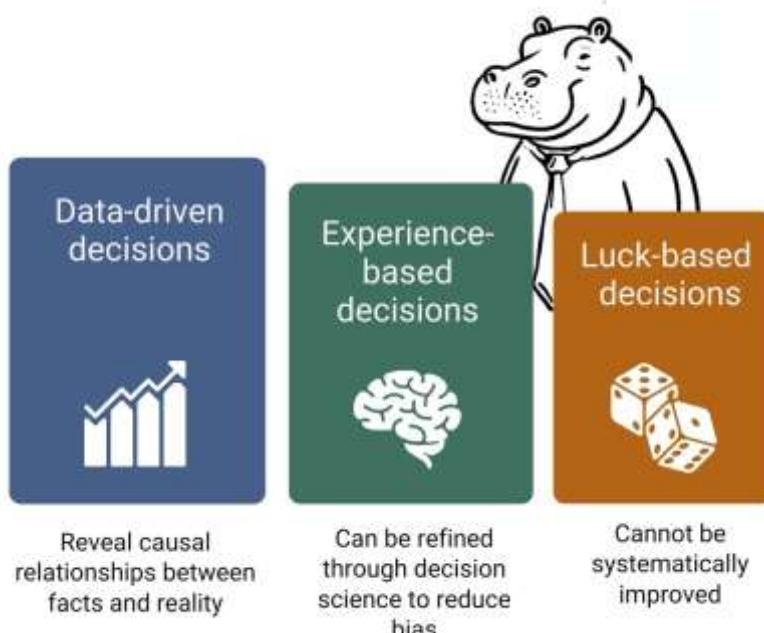
негұрлым ашық, бейімделгіш процестерімен, үйымның жаңару және өсу қабілетін күшайтетін.

Жасанды интеллект пен машиналық оқытудың бизнес-ортаға әсері күмән тудырмайды. Бұл тек уақытша тенденция емес, стратегиялық қажеттілік. Жасанды интеллектіні елемейтін компаниялар инновация мен икемділікті белсенді түрде ынталандыратын нарықта бәсекелестігін жоғалту қаупіне ұшырайды.

Болашақ жасанды интеллектіні тек құрал ретінде емес, қызметтің әр аспектісін қайта ойластыру мүмкіндігі ретінде көретіндерге тиесілі – процестерді оңтайландырудан бастап басқарушылық шешімдер қабылдауға дейін.

Цифрлық негізді қалау: 1-5 қадам цифрлық жетілуге.

Бұл бөлімде біз цифрлық трансформацияның жол картасын қарастырамыз және деректерге негізделген тәсілді енгізу үшін қажетті негізгі қадамдарды анықтаймыз, бұл корпоративтік мәдениетті және компанияның ақпараттық экожүйесін трансформациялауға көмектеседі.



Сурет 10.22 Бақыланатын жаңарту және стратегияны таңдау: жағдай, тәжірибе немесе деректер.

McKinsey зерттеуіне сәйкес «Цифрлық стратегиялардың неге сәтсіздікке ұшырайтыны» (2018), компаниялардың цифрлық трансформация мақсаттарына жетпеуінің кемінде бес себебі бар.

- Түсініксіз анықтамалар: басшылар мен менеджерлер "цифрлық технологиялар" дегеніміз не еkenін әртүрлі түсінеді, бұл түсініспеушілік пен әрекеттердің үйлесімсіздігіне әкеледі.
- Цифрлық технологиялардың экономикасын дұрыс түсінбей: көптеген компаниялар

цифрландырудың бизнес модельдері мен салалардың динамикасына енгізетін өзгерістердің ауқымын төмен бағалайды.

- Экожүйелерді елемеу: компаниялар жеке технологиялық шешімдерге (деректер силостары) назар аударады, кеңірек цифрлық экожүйелермен интеграция қажеттілігін ескермейді.-
- Бәсекелестер тарапынан цифрландыруды төмен бағалау: басшылар бәсекелестердің де цифрлық технологияларды белсенді енгізіп жатқанын ескермейді, бұл бәсекелестік артықшылықтарды жоғалтуға әкелуі мүмкін.
- Цифрландырудың екіжақтылығын елемеу: бас директорлар цифрлық трансформацияға жауапкершілікті басқа басшыларға жүктейді, бұл бақылауды бюрократияландырады және өзгерістер процесін баяулатады.

Бұл мәселелерді шешу үшін ұйымның барлық деңгейлерінде цифрлық стратегияларды нақты түсіну және келісу қажет. Цифрлық стратегияны құрудан бұрын, бастапқы нұктені түсіну маңызды. Көптеген ұйымдар жаңа құралдар мен платформаларды енгізуге тырысады, бірақ қазіргі жағдайдаңың толық көрінісін білмейді.

1-қадам. Қазіргі жүйелер мен деректерді аудит жүргізіңіз.

Процестерді өзгертуге дейін, не бар екенін түсіну маңызды. Аудит жүргізу деректерді басқарудағы әлсіз тұстарды анықтауға және қандай ресурстарды пайдалануға болатынын түсінуге мүмкіндік береді. Мұндай аудит – бизнес-процестердің «рентгені». Ол сізге қауіп аймақтарын анықтауға және жобаңыз немесе бизнесіңіз үшін қандай деректердің маңызды, ал қайсысының екінші дәрежелі екенін анықтауға көмектеседі.

Негізгі әрекеттер:

- IT-ортаны картага түсіріңіз (Draw.io, Lucidchart, Miro, Visio немесе Canva). Сіздің процестеріңізде қолданылатын жүйелерді (ERP, CAD, CAFM, CPM, SCM және басқалар) тізімдеңіз, оларды «Заманауи құрылыш басқару технологиялары мен жүйелері» бөлімінде талқыладыңыз.-
- Әр жүйе үшін деректердің сапасы мәселелерін бағалаңыз, дубликаттардың, мүмкін болатын жіберілген мәндердің және әр жүйедегі форматтардағы сәйкесіздіктердің жиілігін тексеріңіз.
- "Ауыртпалық нұктелерін" анықтаңыз – процестердің үзілуі немесе жиі қолмен араласуды талап ететін жерлер – импорт, экспорт және қосымша тексеру процестері.

Егер команда есептерге сенетін болса, деректердің дұрыстығын басынан бастап бақылау қажет.

Сапалы жүргізілген деректер аудиті қандай деректердің:

- Жетілдіруді қажет ететінін (автоматты тазалау немесе қосымша трансформациялау процестерін баптау қажет) көрсетеді.
- «Қоқыс» болып табылатын деректер, олар тек жүйелерді ластап, процестерде одан әрі

пайдаланбай тастауға болады.

- Мұндай аудитті өз бетіншіше жүргізуге болады. Бірақ кейде сыртқы консультантты тарту пайдалы – әсіреле басқа экономикалық салалардан: жаңа көзқарас пен құрылыш "ерекшеліктерінен" тәуелсіздік статус-квоны объективті бағалауға және белгілі бір шешімдер мен технологияларға қатысты типтік алдын ала болжамдардан аулақ болуға көмектеседі.

2-қадам. Деректерді унификациялау үшін негізгі стандарттарды анықтаңыз.

Аудит жүргізілгеннен кейін деректермен жұмыс істеудің жалпы ережелерін жасау қажет. «Стандарттар: кездейсоқ файлдардан ойластырылған деректер моделіне» тарауында талқыланғандай, бұл ақпараттық ағындардың бөлшектенуін жоюға көмектеседі.

Біртұтас стандартсыз әр команда «өз жолымен» жұмыс істей береді, және сіз «зоопар» интеграцияларын сақтайсыз, мұнда деректер әрбір түрлендіру кезінде жоғалады.

Негізгі әрекеттер:

- Жүйелер арасында ақпарат алмасу үшін деректер стандарттарын таңдаңыз:
- Кестелік деректер үшін бұл құрылымдалған форматтар, мысалы, CSV, XLSX немесе тиімдірек форматтар, мысалы, Parquet болуы мүмкін.
- Әлсіз құрылымдалған деректер мен құжаттарды алмасу үшін: JSON немесе XML.
- Деректер модельдерімен жұмыс істеуді үйреніңіз:
- Концептуалды деректер модельнің деңгейінде тапсырмаларды параметреуден бастаңыз – бұл «Деректерді модельдеу: концептуалды, логикалық және физикалық модель» тарауында сипатталған.
- Бизнес-процестердің логикасына тереңдеген сайын, логикалық және физикалық модельдерде параметрлерді пайдалана отырып, талаптарды формализациялауға көшіңіз.
- Процестер шеңберінде негізгі объектілерді, олардың атрибуттарын және өзара байланыстарын анықтаңыз, сондай-ақ осы байланыстарды визуализациялаңыз – объектілер арасында да, параметрлер арасында да.
- Деректерді валидациялау және стандарттау үшін тұрақты өрнектерді (RegEx) пайдаланыңыз, бұл «Структурированные требования и регулярные выражения RegEx» тарауында талқыланған. RegEx – физикалық деректер модельдері деңгейінде талаптарды жасау кезінде өте маңызды, бірақ күрделі емес тақырып.-
- Деректер деңгейінде стандарттар мен процестерді визуализациялаусыз келісілген және масштабталатын цифрлық орта қамтамасыз ету мүмкін емес. Есіңізде болсын: «жаман деректер қымбатқа түседі». Жобаның немесе ұйымның күрделенуімен қателік бағасы артады. Форматтарды унификациялау, атау ережелерін, құрылым мен валидацияны анықтау – бұл болашақ шешімдердің тұрақтылығы мен масштабталуына инвестициялар.

3-қадам. DataOps енгізіңіз және процестерді автоматтандырыңыз.

Компанияның нақты құрылымдалған архитектурасы болмаса, деректердің бөлшектенгендігіне, оқшауланған ақпараттық жүйелерде қамтылған деректерге тап болу міндettі. Деректер интеграцияланбай, әртүрлі жерлерде дублирленеді және қолдауға елеулі шығындарды талап етеді.

Деректерді су ретінде, ал деректер архитектурасын - деректерді сақтау көзінен пайдалану орнына тасымалдайтын күрделі құбыр жүйесі ретінде елестетініз. Деректер архитектурасы ақпараттың қалай жиналатынын, сақталатынын, түрлендірілетінін, талданатынын және соңғы пайдаланушыларға немесе қосымшаларға жеткізілетінін анықтайды.

DataOps (Деректер Операциялары) - деректерді жинау, тазалау, тексеру және пайдалану процесстерін автоматтандырылған ағынға біріктіретін методология, бұл туралы біз кітаптың сегізінші бөлімінде ежей-тегжейлі талқыладық.

Негізгі әрекеттер:

- Процестерді автоматтандыру үшін ETL-конвейерлерін жасаңыз және баптаңыз:
- Extract: PDF құжаттарынан (Сурет 4.12, Сурет 4.15, Сурет 4.17), Excel кестелерінен, CAD-модельдерінен (Сурет 7.24), ERP жүйелерінен және сіз жұмыс істейтін басқа көздерден деректерді автоматты түрде жинауды үйімдастырыңыз.---
- Transform: деректерді бірегей құрылымдық форматқа автоматты түрде түрлендіру процестерін баптаңыз және жабық қосымшалардан тыс есептеулерді автоматтандырыңыз (Сурет 7.28).-
- Load: деректерді қорытынды кестелерге, құжаттарға немесе орталықтандырылған қоймаларына автоматты түрде жүктеуді жасауды тәжірибеден өткізіңіз (Сурет 7.29, Сурет 7.213, Сурет 7.216).--
- Есептеу және QTO (Quantity Take-Off) процестерін автоматтандырыңыз, бұл туралы "QTO Quantity Take-Off: жобалық деректерді атрибуттар бойынша топтастыру" бөлімінде талқыладық.
- CAD-модельдерінен көлемдерді автоматты түрде алу үшін API, плагиндер немесе кері инженерия құралдарын пайдаланыңыз (Сурет 5.25). -
- Әртүрлі сыныптар үшін элементтерді атрибуттар бойынша топтастыру ережелерін кесте түрінде жасаңыз (Сурет 5.212).
- Модульдік жабық жүйелерден тыс көлемдер мен құнды есептеулерді автоматтандыруды тәжірибеден өткізіңіз (Сурет 5.215).
- Деректерді өндіру үшін Python және Pandas-ты пайдалана бастаңыз, бұл туралы "Python Pandas: деректермен жұмыс істеудің алмастырылмайтын құралы" бөлімінде қарастырдық.
- XLSX файлдарымен жұмыс істеу және кестелік деректерді автоматтандыру үшін DataFrame-ді қолданыңыз (Сурет 3.46).-
- Python-ның әртүрлі кітапханалары арқылы ақпаратты агрегациялау және түрлендіруді автоматтандырыңыз.

- Дайын код блоктарын және толық Pipeline-ды жазуды жеңілдету үшін LLM-ді пайдаланыңыз (Сурет 7.218).-
- Python-да қателерді табатын немесе аномалияларды көретін Pipeline құруды тәжірибеден өткізіңіз және жаупапты тұлғаға (мысалы, жобаның менеджеріне) хабарлама жіберіңіз (Сурет 7.42).-

DataOps принциптеріне негізделген автоматтандыру қолмен және фрагменттелген деректермен жұмыс істеуден тұрақты және қайталанатын процестерге көшуге мүмкіндік береді. Бұл тек күнделікті бірдей түрлендірulerмен айналысатын қызметкерлердің жұқтемесін азайтпайды, сонымен қатар ақпараттық жүйенің сенімділігін, масштабталуын және ашықтығын айтартықтай арттырады.

4-қадам. Ашық деректерді басқару экожүйесін жасаңыз.

Жабық модульдік жүйелердің дамуына және жаңа құралдармен интеграциялануына қарамастан, компаниялар құрделі жүйелердің өсіу олардың пайдалығын арттырудан асып кететін маңызды мәселені шешуге тап болады. Барлық бизнес-процестерді қамтитын бірегей меншікті платформаны құрудың бастапқы идеясы, кез келген өзгерістердің едәуір ресурстар мен уақытты талап ететін шамадан тыс орталықтандыруға әкелді.

Біз "Корпоративтік мицелий: деректер бизнестік процестерді қалай байланыстырады" тарауында қарастыранымыздай, деректермен тиімді жұмыс істеу ашық және біртұтас экожүйені құруды талап етеді, ол барлық ақпарат көздерін біріктіреді.

Экожүйенің негізгі элементтері:

- Деректерді сақтау үшін қолайлы шешімді таңдаңыз:
- Кестелер мен есептеулер үшін деректер базаларын пайдаланыңыз – мысалы, PostgreSQL немесе MySQL (Рис. 3.17) -
- Құжаттар мен есептер үшін бұлттық сақтау жүйелері (Google Drive, OneDrive) немесе JSON форматты қолдайтын жүйелер жарамды болуы мүмкін.
- Data Warehouse, Data Lakes және үлкен көлемдегі ақпаратты орталықтандырылған сақтау мен талдау үшін басқа құралдардың мүмкіндіктерімен танысыңыз (Рис. 8.18) -
- Патентtelgen деректерге қол жеткізу шешімдерін енгізіңіз:
- Егер сіз патентtelgen жүйелерді пайдалансаңыз, API немесе SDK арқылы оларға қол жеткізуді орнатыңыз, сыртқы өңдеу үшін деректерді алу үшін (Рис. 4.12)-
- CAD форматтары үшін кері инжиниリング құралдарының әлеуетімен танысыңыз (Рис. 4.113)-
- Деректерді қосымшалардан немесе серверлерден периодты түрде жинап, оларды ашық құрылымдық форматтарға түрлендіріп, сақтау үшін ETL-Pipeline орнатыңыз (Рис. 7.23) -
- Команда ішінде патентtelgen бағдарламалық қамтамасыз етуді пайдаланбай деректерге қол жеткізуді қамтамасыз ету мәселелерін талқылаңыз.
- Есінізде болсын: деректер интерфейстерден маңызды. Үзақ мерзімді құндылық ақпараттың құрылымы мен қолжетімділігінде, нақты пайдаланушы интерфейсінің

құралдарында емес.

- Деректер бойынша үздік тәжірибе орталығын (CoE) құруды немесе деректермен жұмыс істеу бойынша сараптаманы қамтамасыз етудің басқа жолдарын қарастырыңыз, біз "Деректерді модельдеу бойынша үздік тәжірибе орталығы (CoE)" тарауында талқыладық (Рис. 4.39) -

Деректерді басқару әкожүйесі жобаның барлық қатысуышылары келісілген, өзекті және тексерілген ақпаратпен жұмыс істейтін біртұтас ақпараттық кеңістік құрады. Бұл масштабталатын, икемді және сенімді цифрлық процестердің негізі.

Деректердің әлеуетін ашу: Цифрлық жетілуге 5-10 қадам.

Техникалық интеграциядан басқа, цифрлық шешімдерді сәтті енгізудің маңызды факторы – олардың соңғы пайдаланушылар тарапынан қабылдануы. Клиенттерді немесе пайдаланушыларды тиімділікті бағалау мәселелеріне тарту – бұл пайдаланушы тәжірибесін жақсарту және компаниядағы өзгерістерді басқару міндепті. Егер шешім жұмыс процесіне сәйкес келмесе немесе пайдаланушылардың немесе клиенттердің нақты мәселелерін шешпесе, ол пайдаланылмайды, және ешқандай қосымша шаралар мен ынталандырулар бұл мәселені түзетпейді.

Трансформация – бұл пайдаланушылардың жаңа процестермен өзара әрекеттесуіне қатысты деректерді талдауға негізделген итеративті процесс, жиі тестілеу циклдарымен, тұрақты кері байланыспен және жетілдірulerмен.

5-қадам. Деректермен жұмыс істеу мәдениетін қалыптастырыңыз, персоналды оқытыңыз және кері байланысты жинаңыз.

Ең озық жүйе де қызметкерлердің қатысуының жұмыс істемейді. Деректер күнделікті пайдаланылатын орта құру қажет, ал команда олардың құндылығын түсінуі тиіс.

2024 жылғы Ұлыбритания үкіметінің "Data Analytics and AI in Government Project Delivery" жарияланған есебінде [83] деректер аналитикасы мен ИИ-ді сәтті енгізу үшін деректерді өңдеу және интерпретациялау бойынша қажетті біліктіліктерге ие мамандарды даярлау критикалық маңызды екендігі атап өтілген.

Деректерді талдау саласындағы білімнің жетіспеушілігі цифрлық трансформацияны шектейтін негізгі мәселелердің бірі болып табылады. Басшылар дәстүрлі процедураларға, тоқсандық циклдарға, басымдықты бастамаларға және жобаларды жылжытудағы дәстүрлі жолдарға үрленіп қалған. Өзгерістер үшін ерекше көшбасшы қажет – ықпал ету үшін жеткілікті жоғары дәрежеде, бірақ ұзақ мерзімді трансформация жобасын жүргізуге уақыты мен мотивациясы болмауы тиіс.

Негізгі әрекеттер:

- Жоғары жалақы алатын қызметкердің (HiPPO) пікірі негізінде субъективті шешімдерден фактілер мен деректерге негізделген шешім қабылдау мәдениетіне көшу қажеттілігін түсіну.-
- Жүйелі оқытуды ұйымдастырыңыз:
- Структурланған деректерді пайдалану бойынша тренингтер өткізіңіз және құрылыш

саласында бұғінгі күні танымал өнімдер мен концепцияларға бейтараптық танытатын басқа экономикалық салалардың сарапшыларын шақырыңыз.

- Деректерді талдау тәсілдері мен құралдарын әріптермен талқылаңыз, сондай-ақ Python, pandas және LLM сияқты құралдармен практикалық жұмыс жасауды өз бетіңізде меңгеріңіз.-
- Деректерді құрылымдау және модельдер жасау тақырыбында оқу материалдарының кітапханасын жасаңыз (қысқа бейнелермен жақсы).--
- Заманауи оқыту технологияларын пайдаланыңыз:
- Код пен деректермен жұмыс істеу кезінде LLM тілдік модельдерін қолданыңыз, соның ішінде кодты генерациялау, рефакторингтеу және талдау, сондай-ақ кестелік ақпаратты өңдеу және интерпретациялау.
- LLM арқылы сгенерленген кодтың оффлайн-дамыту ортасында (IDE) толыққанды Pipeline шешіміне қалай бейімделіп, интеграцияланатынын зерттеңіз.-

Басшы «ескі әдіспен» шешім қабылдауды жалғастырса, ешқандай тренинг адамдарды аналитикаға шынайы қарауға сендіре алмайды.

Деректермен жұмыс істеу мәдениетін қалыптастыру тұрақты кері байланыстысыз мүмкін емес. Кері байланыс процестер, құралдар және стратегиялардағы кемшіліктерді анықтауға мүмкіндік береді, оларды ішкі есептер немесе формалды KPI метрикалары арқылы анықтау мүмкін емес. Сіздің шешімдеріңізді пайдаланушылардың комплиментарлы пікірлері практикалық құндылық әкелмейді. Нақты бақылаулар мен фактілерге негізделген сындарлы кері байланыс құнды. Алайда, мұндай ақпаратты алу күш-жігерді талап етеді: қатысушылар – ішкі және сыртқы – пікірлерімен бөлісе алатын процестерді құру қажет (мүмкін, анонимді түрде) бұрмалаусыз және өз пікірлерінің өз жұмысына әсер етуінен қорықпай. Олар бұрмалаусыз және теріс салдардан қорықпай пікір білдіруі маңызды.

Кез келген оқыту, ақырында, өзін-өзі оқыту болып табылады.

– Милтон Фридман, американдық экономист және статистик.

Аналитикалық құралдарды енгізу олардың тиімділігін практикада тұрақты түрде верификациялаумен (ROI, KPI) қатар жүруі тиіс, бұл тек қызметкерлер, клиенттер және серіктестер тарапынан құрылымдық кері байланыс арқылы жүзеге асырылуы мүмкін. Бұл компанияларға тек қателіктерді қайталаудан аулақ болуға ғана емес, сонымен қатар ортадағы өзгерістерге тез бейімделуге мүмкіндік береді. Кері байланыс жинау және талдау механизмі үйымның жетіліуінің бір белгісі болып табылады, ол эпизодтық цифрлық бастамалардан тұрақты үздіксіз жетілдіру моделіне өтуде.

6-қадам. Пилоттық жобалардан масштабтауға.

Достаточно ірі шайқастарды таңдаңыз, олар маңызды болу үшін, және жеткілікті ұсақ шайқастарды, оларды жеңу үшін. Джонатан Козол

Цифрлық трансформацияны "бірден және барлық жерде" іске қосу өте қауіпті. Нәтижелі тәсіл – пилоттық жобалардан бастау және сәтті тәжірибелерді біртіндеп кеңейту.

Негізгі әрекеттер:

- Сынақ үшін қолайлыш жобаны таңдаңыз:
- Нақты бизнес-мәселені немесе өлшенетін нәтижелері бар процесті анықтаңыз (KPI, ROI) (Сурет 7.15).-
- ETL автоматизация процесін таңдаңыз, мысалы, Python және Pandas көмегімен деректерді автоматты тексеру немесе жұмыс көлемін есептеу (QTO) (Сурет 5.210).-
- Нақты табыс метрикаларын орнатыңыз (мысалы, тексеру спецификацияларын немесе деректерді тексеру бойынша есептерді дайындау уақытын бір аптадан бір күнге дейін қысқарту).
- Применяйте итеративные подходы:
- Деректерді түрлендірудің қарапайым процестерінен бастаңыз және әртүрлі форматтағы деректерді сіздің процестеріңзеге қажетті форматтарға ағынды түрлендіруді жасаңыз (Сурет 4.12, Сурет 4.15).-
- Постепенно тапсырмалардың күрделілігін арттырып, процестердің автоматизациясын кеңейтіңіз, құжатталған код блоктарына негізделген толыққанды Pipeline-ды әзірлеу ортасында (IDE) қалыптастырыңыз (Сур. 4.17, Сур. 7.218).-
- Құжаттаңыз және жазып алыңыз (ең жақсысы қысқа бейнемазмұн арқылы) табысты шешімдерді, оларды әріптестеріңізben немесе кәсіби қауымдастықтармен бөлісіңіз.
- Шаблондарды және осындай шешімдерді тираждауға арналған құжаттаманы әзірлеңіз, осылайша сіздің әріптестеріңіз (немесе кәсіби қауымдастықтың мүшелері, әлеуметтік желілердегі пайдаланушыларды қоса алғанда) оларды тиімді пайдалана алатын болады.

Кезең-кезеңімен енгізу жоғары сапалы өзгерістерді сақтауға мүмкіндік береді және параллельді енгізулердің хаосына түспеуге көмектеседі. "Кішігірімнен үлкенге" стратегиясы тәуекелдерді минимизациялап, ұсақ қателіктерден сабак алуға мүмкіндік береді, оларды критикалық мәселелерге айналдырмайды.

Жобалық тәсілден, қызметкерлердің тек жартылай қатысуымен, тұрақты командаларды (мысалы, саралтама орталықтары – CoE) қалыптастыруға көшу өнімнің алғашқы нұсқасы шыққаннан кейін де тұрақты дамуын қамтамасыз етеді. Мұндай командалар тек бар шешімдерді қолдап қана қоймай, сонымен қатар оларды жетілдіруді жалғастырады.

Бұл ұзақ келіссөздерге тәуелділікті азайтады: команда мүшелері өз жауапкершілік аймағында шешім қабылдау құқығына ие болады. Нәтижесінде менеджерлер микроменеджмент

қажеттілігінен босатылады, ал командалар нақты құндылық жасауға назар аудара алады.

Жаңа шешімдерді әзірлеу - бұл спринт емес, марафон. Оған ұзақ мерзімді, жүйелі жұмысқа бастапқыда бағытталғандар ғана жетістікке жетеді.

Технологиялардың тұрақты дамуын түсіну маңызды. Технологиялық шешімдердің ұзақ мерзімді дамуына инвестиция салу - табысты жұмыстың негізі.

7-қадам. Ашық деректер форматтары мен шешімдерін пайдаланыңыз.

Модульдік платформаларға (ERP, PMIS, CAFM, CDE және т.б.) арналған тарауларда талқыланғандай, вендор шешімдерінен тәуелсіздікті қамтамасыз ететін және процесс қатысушылары үшін ақпараттың қолжетімділігін арттыратын ашық және әмбебап деректер форматтарына назар аудару маңызды.

Негізгі өрекеттер:

- Жабық форматтардан ашық форматтарға өтіңіз:
- Проприетарлық форматтардың орнына ашық форматтарды пайдаланыңыз немесе жабық форматтарды ашықтарға автоматты түрде экспорттау немесе конверсиялау мүмкіндігін табыңыз.
- Көптеген заманауи жүйелер арасында ақпарат алмасу стандарттары болып табылатын Parquet, CSV, JSON, XLSX форматтарымен жұмыс істеу құралдарын енгізіңіз.-
- Егер 3D-геометриямен жұмыс сіздің процестеріңізде маңызды рөл атқарса, USD, glTF, DAE немесе OBJ сияқты ашық форматтарды пайдалануды қарастырыңыз.
- Ақпаратты тиімді талдау және іздеу үшін векторлық деректер базаларын пайдаланыңыз:
- 3D-геометриямен жұмыс істеуді жеңілдету үшін Bounding Box және басқа әдістерді пайдаланыңыз.-
- Деректерді векторизациялау - мәтіндерді, объектілерді немесе құжаттарды сандық көріністерге айналдыруды енгізу мүмкіндігін қарастырыңыз.
- Үлкен деректерді талдау құралдарын қолданыңыз:
- Жиналған тарихи деректерді (мысалы, PDF, XLSX, CAD) талдауға қолайлы форматтарда (Apache Parquet, CSV, ORC) сақтау ұйымдастырыңыз. -
- Негізгі статистикалық әдістерді қолдануды бастаңыз және репрезентативті таңдамалармен жұмыс істеңіз - немесе, ең болмағанда, статистиканың негізгі принциптерімен танысыңыз.
- Деректер мен деректер арасындағы байланыстарды визуализациялау құралдарын енгізіп, зерттеңіз, талдау нәтижелерін көрнекі түрде көрсету үшін. Сапалы визуализациясыз деректерді және оларға негізделген процестерді толық түсіну мүмкін емес.

Ашық деректер форматтарына көшу және ақпаратты талдау, сақтау және визуализациялау құралдарын енгізу тұрақты және тәуелсіз цифрлық басқарудың негізін қалайды. Бұл тек вендорларға тәуелділікті азайтпайды, сонымен қатар процесс қатысушылары үшін деректерге тең

қолжетімділікті қамтамасыз етеді.

8-қадам. Болжам жасау үшін машиналық оқытуды енгізуді бастаңыз.

Көптеген компанияларда үлкен көлемдегі деректер жинақталған – бір түрі «ақпараттық гейзерлер», әлі күнге дейін пайдаланылмаған. Бұл деректер жүздеген және мындаған жобалар аясында жиналған, бірақ көбінесе тек бір рет қолданылып немесе мүлдем келесі процестерге тартылмаған. Жабық форматтар мен жүйелерде сақталған құжаттар мен модельдер көбінесе ескірген және пайдасыз жүк ретінде қабылданады. Алайда, шын мәнінде, олар қателіктерді талдау, рутиналық операцияларды автоматтандыру және болашақ жобаларда элементтерді автоматты түрде классификациялау мен тану бойынша инновациялық шешімдерді әзірлеу үшін құнды ресурс болып табылады.

Негізгі міндет – осы деректерді шығарып, практикалық пайдаға айналдыруды үйрену. «Машиналық оқыту және болжамдар» бөлімінде талқыланғандай, машиналық оқыту әдістері құрылышқа қатысты әртүрлі процестерде бағалаулар мен болжамдардың дәлдігін едәуір арттыра алады. Жиналған деректерді толыққанды пайдалану тиімділікті арттыруға, тәуекелдерді азайтуға және тұрақты цифрлық процестерді құруға жол ашады.

Негізгі әрекеттер:

- Жеңіл алгоритмдерден бастаңыз:
- Линейлік регрессияны қолданып көрініз – LLM-ден алынған нұсқауларды пайдалана отырып, деректер жиынтықтарындағы қайталараптың көрсеткіштерді болжау үшін, мұнда көптеген факторларға тәуелділіктер жоқ немесе минималды (Сурет 9.34).
- Процестеріңіздің қандай кезеңдерінде k-ближайших соседей (k-NN) алгоритмін қолдануға болатындығын қарастырыңыз – мысалы, классификация, объектілердің үқастығын бағалау немесе тарихи аналогтар негізінде болжау үшін (Сурет 9.35).
- Модельдерді оқыту үшін деректерді жинап, құрылымдаңыз:
- Жобалар туралы тарихи деректерді бір жерде және бір форматта жинаңыз (Сурет 9.110).
- Автоматты ETL арқылы оқыту жиынтықтарының сапасы мен репрезентативтілігіне жұмыс жасаңыз (Сурет 9.28).-
- Деректерді оқу және тестілеу жиынтықтарына бөлу әдісін үйреніңіз, біз Титаник деректер жиынтығымен мысал ретінде жасағанымыздай (Сурет 9.26, Сурет 9.27).-
- Машиналық оқыту әдістерін кең ауқымды міндеттерді шешу үшін қолдану мүмкіндіктерін қарастырыңыз – жобалардың жүзеге асырылу мерзімдерін болжаудан бастап, логистиканы оңтайландыру, ресурстарды басқару және әлеуетті проблемаларды ерте анықтауға дейін.

Машиналық оқыту – архивтік деректерді болжамдау, оңтайландыру және негізделген шешімдер қабылдау үшін құнды активке айналдыруға мүмкіндік беретін құрал. Кішігірім деректер жиынтықтарынан (Сурет 9.25) және қарапайым модельдерден бастаңыз, біртіндеп күрделілікти арттырыңыз.

9-қадам. IoT және заманауи деректер жинау технологияларын интеграциялаңыз.

Құрылыс әлемі тез цифрануда: әрбір құрылыш алаңындағы фото, әрбір Teams хабарламасы – бұл шындықты параметрлеу және токенизациялау процесінің бір белгі. GPS логистиканы қалай өзгертсе, IoT, RFID және автоматты деректер жинау технологиялары құрылыш саласын да өзгертуде. «IoT Интернет заттары және смарт-контракттар» белімінде қарастырылғандай, датчиктер мен автоматтандырылған мониторингпен жабдықталған цифрлық құрылыш алаңы – саланың болашағы.

Негізгі әрекеттер:

- IoT құрылғыларын, RFID белгілерін енгізің және оларға қатысты процестерді нақтылаңыз:
- Датчиктерді орнатудың қай аймақтарында немесе жобаның кезеңдерінде ең жоғары қайтарымдылық (ROI) алуға болатынын бағалаңыз – мысалы, температура, діріл, ылғалдылық немесе техника қозғалысын бақылау үшін.
- Логистикалық тізбектің барлық кезеңдерінде материалдарды, құралдарды және жабдықтарды бақылау үшін RFID қолдану мәселесін қарастырыңыз.
- Жиналған деректерді автоматтандырылған өндіріс және нақты уақыттағы талдау үшін Apache NiFi сияқты біртұтас ақпараттық жүйеге интеграциялаудың жолдарын ойластырыңыз (Рис. 7.45).-
- Нақты уақыт режимінде мониторинг жүйесін құрыңыз:
- Процесс немесе жобаның негізгі көрсеткіштерін бақылау үшін Streamlit, Flask немесе Power BI сияқты визуализация құралдарын пайдаланып, дашбордтар жасаңыз.
- Жоспардан немесе нормадан критикалық ауытқулар туралы сигнал беретін автоматты хабарламаларды орнатыңыз (Рис. 7.42).
- Жиналған деректер мен анықталған заңдылықтарға негізделген жабдықтың предиктивті қызмет көрсету әлеуетін бағалаңыз (Рис. 9.36).
- Эр түрлі дереккөздерден деректерді біріктіріңіз:
- Деректер моделін физикалық деңгейде визуализациялаудан бастаңыз – CAD жүйелерінен, IoT құрылғыларынан және ERP платформаларынан келетін ақпарат ағындарының құрылымын және негізгі параметрлерін көрсетіңіз (Рис. 4.31).
- Деректерді талдау және басқарушылық шешімдерді қолдау үшін арналған біртұтас платформаның алғашқы сипаттамасын жасаудан бастаңыз. Негізгі функцияларды, дереккөздерді, пайдаланушыларды және болжамды қолдану сценарийлерін бекітіңіз (Рис. 4.37).

Нағыз процестерді цифрлық әлеммен байланыстыра бастаған сайын, деректер арқылы оларды тиімді, ашық және нақты уақыт режимінде басқаруға мүмкіндік аласыз.

10-қадам. Саладағы өзгерістерге дайындалыңыз.

Құрылыш компаниялары сыртқы ортадан: экономикалық дағдарыстардан, технологиялық секірістерден, нормативтік өзгерістерден үнемі қысым көріп отырады. Орманның жаңбыры, қар, құргақшылық және күйдіргіш күн сәулесіне төтеп беруі сияқты, компаниялар үздіксіз бейімделу

жағдайында өмір сүреді. Және ағаштар аяз бен құрғақшылықта төзімділікті терең тамыр жүйесі арқылы алатын болса, тек автоматтандырылған процестерге, өзгерістерді болжау қабілетіне және стратегияларды икемді түрде бейімдеуге ие үйымдар ғана өміршеңдік пен бәсекеге қабілеттілікті сақтай алады.

«Тіршілік стратегиялары: бәсекелестік артықшылықтарды қалыптастыру» бөлімінде атап өтілгендей, құрылыс саласы радикалды трансформация кезеңіне енуде. Тапсырыс беруші мен орындаушы арасындағы өзара әрекеттесу уберизация моделі бойынша дамып келеді, мұнда ашықтық, болжамдылық және цифрлық құралдар дәстүрлі тәсілдерді ығыстырады. Осы жаңа шындықта ең ірі емес, ең икемді және технологиялық жағынан жетілгендер жеңіске жетеді.

Негізгі әрекеттер:

- Ашық деректер контекстінде бизнестің осал тұстарын талдаңыз:
- Уберизация аясында деректерге қол жеткізуіндік демократиялануы сіздің бәсекелестік артықшылықтарыңыз бен бизнесіңізге қалай зиян келтіруі мүмкін екенін бағалаңыз (Рис. 10.15).
- Непрозрач және оқшауланған процестерден ашық шешімдерге, жүйелердің үйлесімділігіне және деректердің ашықтығына негізделген бизнес-модельдерге көшү стратегиясын ойластырыңыз (Сурет 2.25)-
- Ұзақ мерзімді цифрлық стратегияны әзірлеңіз:
- Сіз инновациялардың көшбасшысы болуға үмтүласыз ба, әлде ресурстарыңызды үнемдейтін "артта қалушы" сценарийді таңдауды жөн көресіз бе?
- Кезеңдерді жоспарлаңыз: қысқа мерзімді (процестерді автоматтандыру, деректерді құрылымдау), орта мерзімді (LLM және ETL енгізу), ұзақ мерзімді (цифрлық әкожүйелер, орталықтандырылған қоймалар)
- Қызметтер портфелін кеңейту туралы ойланыңыз:
- Энергия тиімділігіне, ESG-ге, деректерді өңдеу қызметтеріне бағытталған жаңа қызметтерді ұсыну мүмкіндігін қарастырыңыз. Жаңа бизнес модельдері туралы келесі тарауда сөйлесеміз
- Сізді жобалаудан бастап пайдалану кезеңіне дейінгі объектінің өмірлік циклімен бірге жүретін сенімді технологиялық серіктес ретінде позициялауға үмтүлыңыз. Сізге деген сенім жүйелі тәсілге, процестердің ашықтығына және тұрақты технологиялық шешімдердің қамтамасыз ету қабілетіне негізделуі тиіс

Трансформация жағдайында өзгерістерге тек жауап беретіндер емес, алға шығатындар жеңеді. Икемділік, ашықтық және цифрлық жетілу – ертеңгі құрылысқа тұрақтылықтың негізі.

Жол картасы трансформация: хаостан деректерге негізделген компанияға.

Келесі жоспар бастапқы нұсқаулық ретінде қызмет ете алады – деректерге негізделген цифрлық трансформация стратегиясын қалыптастырудың бастапқы нұктесі:

- Аудит және стандарттар: ағымдағы жағдайды талдаңыз, деректерді унификациялаңыз
- Деректерді құрылымдау және классификациялау: құрылымдалмаған және нашар құрылымдалған деректерді автоматтандыруды жүзеге асырыңыз
- Топтастыруларды, есептеулерді және калькуляцияларды автоматтандыру: автоматтандыру үшін ашық құралдар мен кітапханаларды пайдаланыңыз
- Экожүйе және СОЕ: компанияда біртұтас деректер экожүйесін қалыптастыратын ішкі команданы құрыңыз
- Мәдениет және оқыту: HiPPO шешімдерінен деректерге негізделген шешімдерге өтіңіз
- Пилоттар, кері байланыс және масштабтау: итеративті әрекет етіңіз: жаңа әдістерді шектеулі ауқымда тестіленіз, негізделген кері байланыс жинаңыз және шешімдерді біртіндеп масштабтаңыз
- Ашық форматтар: бағдарламалық қамтамасыз етуші вендорлардан тәуелсіздік үшін әмбебап және ашық форматтарды пайдаланыңыз
- Машиналық оқыту: болжамдау және оңтайландыру үшін ML алгоритмдерін процестерге енгізіңіз
- IoT және цифрлық құрылыш алаңы: деректерді жинаудың заманауи технологияларын процестерге интеграциялаңыз
- Стратегиялық бейімделу: саладағы болашақ өзгерістерге дайын болыңыз

Ең бастысы, "деректер өздігінен компанияны өзгертпейді: оны деректермен жұмыс істей алатын адамдар өзгертеуді" екенін ұмытпаңыз. Мәдениетке, ашық процестерге және үздіксіз жетілдіруге баса назар аударыңыз.

Жүйелік тәсіл бөлшектелген цифрлық бастамалардан толыққанды деректерге негізделген басқару моделіне өтуге мүмкіндік береді, мұнда шешімдер интуиция немесе болжамдар негізінде емес, деректер, фактілер және математикалық есептелген ықтималдықтар негізінде қабылданады. Құрылышы саласындағы цифрлық трансформация – бұл тек технологияларды енгізу емес, жобалар туралы ақпараттың әртүрлі жүйелер арасында үздіксіз және итеративті түрде берілуін қамтамасыз ететін бизнес-экожүйені қалыптастыру. Машиналық оқыту алгоритмдері автоматты, үздіксіз талдау, болжау және процестерді оңтайландыруды қамтамасыз етеді. Мұндай ортада спекуляциялар мен жасырын деректер өзектілігін жоғалтады – тек тексерілген модельдер, ашық есептеулер және болжамды нәтижелер қалады.

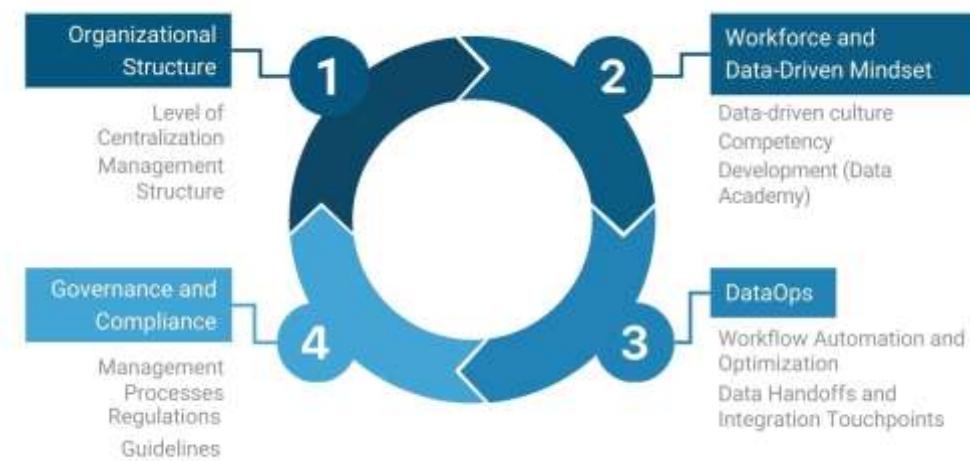


Рис. 10.23 Компания деңгейінде деректерді тиімді басқарудың негізгі элементтері.

Кітаптың әрбір бөлігі құрылымдық жобаларындағы деректерді өндеду және талдау кезеңіне сәйкес келеді (Рис. 2.25). Егер сіз бұрын қарастырылған тақырыптардың біріне оралып, деректерді пайдалану ағынының тұтас түсінігі түрфысынан қарағыңыз келсе – Рис. 10.24-те көрсетілген бөлім атауларына назар аудара аласыз.-

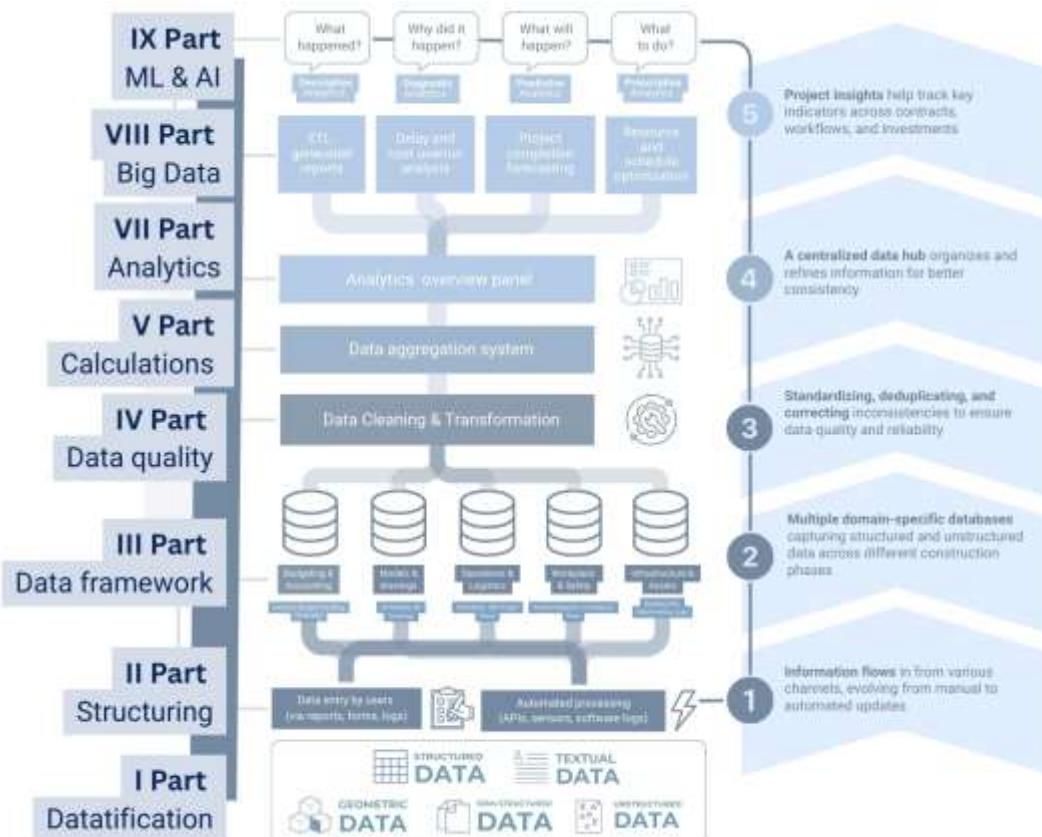


Рис. 10.24 Деректерді өндеду конвейері контекстінде кітаптың бөлімдері (Рис. 2.25): ақпаратты цифрландырудан аналитика мен жасанды интеллектке дейін.-

Сіздің үйіміңдың көлеміне, технологиялық жетілу деңгейіне немесе бюджетіне қарамастан, сіз бүгінгі күні деректерге негізделген тәсілге көшуге кірісе аласыз. Тіпті дұрыс бағытта жасалған шағын қадамдар уақыт өте келе нәтижелерге әкеледі.

Деректерге негізделген трансформация – бұл бір реттік жоба емес, жаңа құралдарды енгізууді, процестерді қайта қарасты және деректер негізінде шешім қабылдау мәдениетін дамытуды қамтитын үздіксіз, итеративті жетілдіру процесі.

Индустрия 5.0-да құрылыш: қалай табыс табуға болады, егер жасыру мүмкін болмаса.

Ұзақ уақыт бойы құрылыш компаниялары процестердің ашық еместігінен пайда тапты. Негізгі бизнес-модель спекуляцияға айналды – жабық ERP-, PMIS-жүйелерінде материалдардың, жұмыстардың көлемдерінің және пайыздық үстемелердің бағасын көтеру. Тапсырыс берушілер мен олардың сенімді тұлғаларының бастапқы жобалық деректерге шектеулі қолжетімділігі тексеру мүмкіндігін іс жүзінде мүмкін емес ететін схемаларға негіз болды.

Алайда, бұл модель тез арада өзектілігін жоғалтуда. Деректерге қолжетімділіктің демократияланымен, LLM-дердің пайда болуымен, ашық деректердің келуімен және ETL автоматизация құралдарының арқасында сала жаңа жұмыс стандарттарына өтуде.

Нәтижесінде, ашық еместік бәсекелестік артықшылық болудан қалады – жақын арада ол одан құтылу қыын жүккө айналады. Ашықтық опциядан нарықта қалу үшін міндетті шартқа айналады.

Клиенттер – банктер, инвесторлар, жеке тапсырыс берушілер, private equity, мемлекеттік тапсырыс берушілер – жаңа цифрлық шындықта кіммен жұмыс істейді? Жауап айқын: тек нәтиже ғана емес, сонымен қатар оған жету жолындағы әр қадамның негізdemесін ұсына алатындармен. Ашық деректер көлемінің өсуі жағдайында, серіктестер мен тапсырыс берушілер ашықтықты, дәлдікті және нәтижелердің болжамдылығын қамтамасыз ететін компанияларды таңдайды.

Осы фонды ескере отырып, спекуляция емес, деректерді басқару және сенім негізінде жаңа бизнес-модельдер қалыптасуда.

- Процестерді сату квадрат метрлерді сатудан гөрі маңызды: негізгі активтер бетонға жеңілдік алу келісімдері емес, сенім мен тиімділік болады. Негізгі құндылық сенімді және тексерілген деректерге негізделген нәтижелердің болжамдылығы болады. Заманауи компаниялар құрылыш объектісін емес, сататын болады:
 - нақты мерзімдер мен ашық жұмыс графиктері;
 - есептеулермен расталған негізделген сметалар;
 - жобаның барлық кезеңдерінде толық цифрлық бақылау мен қадағалауды қамтамасыз ету мүмкіндігі.
- Инжиниринг және аналитика қызмет ретінде: “Деректер қызметі” моделі (интернет арқылы пайдаланушыларға дайын деректерді жеткізу тәсілі), мұнда әр жоба цифрлық деректер тізбегінің бір бөлігіне айналады, ал бизнестің құндылығы осы тізбекті басқару қабілетінде болады. Компаниялар автоматтандыру мен аналитика негізінде шешімдер ұсынатын интеллектуалды платформаларға айналады:
 - автоматтандырылған және ашық сметалар мен жоспарларды құру;
 - машиналық оқыту алгоритмдері негізінде тәуекелдер мен мерзімдерді бағалау;
 - экологиялық көрсеткіштерді (ESG, CO₂, энергия тиімділігі) есептеу;
 - тексерілетін ашық дереккөздерден есептілікті қалыптастыру.

- Инженерлік тәжірибелі өнімге айналдыру: компанияның жинақтаған тәжірибесі ішкі пайдалануда бірнеше рет қолданылуы мүмкін және жеке өнім ретінде таралуы – цифрлық қызметтер арқылы қосымша табыс көзі қалыптастырады. Жаңа жағдайларда компаниялар тек жобалар ғана емес, цифрлық активтер де жасайды:
 - компоненттер мен сметалар шаблондарының кітапханалары;
 - автоматтандырылған верификация модульдері;
 - деректермен жұмыс істеу үшін open-source плагиндер мен скрипттер (консалтинг сату).
- Жаңа типтегі компания: Деректерге негізделген интегратор: бұл нарықтағы қатысушы, ол нақты бағдарламалық қамтамасыз етушілерге немесе модульдік жүйелерге тәуелді емес және бір ғана бағдарламалық қамтамасыз етудің интерфейсінде “қамалмаған”. Ол деректермен еркін жұмыс істейді – және осы негізде бәсекеге қабілеттілігін құрады. Болашақтың құрылыс компаниясы – бұл тек мердігер емес, ақпарат интеграторы, тапсырыс беруші үшін келесі функцияларды орындауға қабілетті:
 - әртүрлі дереккөздерден деректерді біріктіру және аналитика жүргізу;
 - процестердің ашықтығы мен сенімділігін қамтамасыз ету;
 - бизнес процестерін оңтайландыру бойынша кеңес беру;
 - ашық деректер экожүйесінде, LLM, ETL және Pipelines жұмыс істейтін құралдарды әзірлеу.

Индустрия 5.0 (Рис. 2.112) “қолмен орташа коэффициенттер дәүірінің” аяқталуын білдіреді және бас директорлардың сметалық және бухгалтерлік бөліммен кешкі кездесулерінің аяқталуын білдіреді. Бұрын жасырынып келген барлық нәрсе – есептеулер, сметалар, көлемдер – ашық, тексерілетін және тіпті сарапшы емес адамға да түсінікті болады. Алғашқы болып қайта бағдарланатындар ұтысқа ие болады. Басқалары – құрылыс секторындағы жаңа цифрлық экономикадан тысқалады.

ҚОРЫТЫНДЫ

Құрылыс саласы түбөгейлі өзгерістер дәуіріне қадам басуда. Глинялы тақталардағы алғашқы жазбалардан бастап, жобалық серверлер мен құрылыс аландарынан келетін массивті цифрлық деректерге дейін – құрылыс саласындағы ақпаратпен жұмыс істеу тарихы әрқашан өз уақытындағы технологиялардың жетілу деңгейін көрсетіп отырды. Бүтінгі күні автоматизация, ашық форматтар мен интеллектуалды талдау жүйелерінің келуімен, сала біртіндеп эволюциямен емес, жылдам цифрлық трансформациямен бетпе-бет келуде.

Басқа экономикалық секторлар сияқты, құрылыс саласына да құралдарды ғана емес, жұмыс принциптерін қайта қарau қажет. Бұрын нарықты диктейтін және тапсырыс беруші мен жоба арасында негізгі делдал қызметтін атқаратын компаниялар өздерінің ерекше жағдайын жоғалтуда. Деректермен жұмыс істеу қабілеті мен сенім бірінші орынға шығады: деректерді жинаудан және құрылымдаудан бастап, аналитика, болжau және шешімдерді автоматтандыруға дейін.

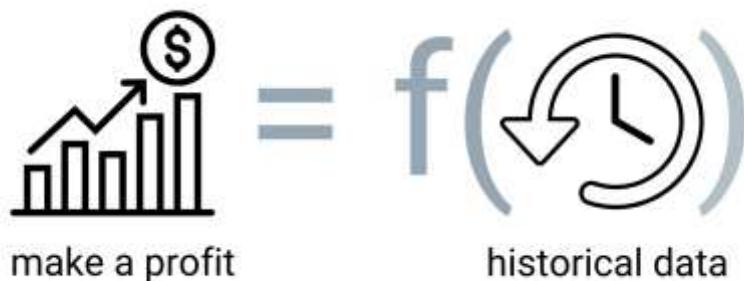


Рис. 10.21 Структурированные исторические данные – эффективный и управляемый бизнес үшін отын.

Бұл кітапта құрылыс саласындағы деректерді басқарудың негізгі принциптері – аудит пен стандарттаудан бастап, процестерді автоматтандыру, визуализация құралдарын пайдалану және интеллектуалды алгоритмдерді енгізуге дейін егжей-тегжейлі қарастырылды. Шектеулі ресурстар болған жағдайда да, жұмыс істейтін деректер архитектурасын құрып, шешімдерді интуицияға емес, тексерілетін фактілерге негізделген қабылдауға болатыны көрсетілді. Деректермен жұмыс тек IT бөлімінің міндеті болмай, компанияның басқару мәдениетінің негізіне айналады, бұл компанияның икемділігіне, бейімделуіне және ұзақ мерзімді тұрақтылығына әсер етеді.

Машиналық оқыту, автоматты өңдеу жүйелері, цифрлық егіздер мен ашық форматтарды қолдану бүгінгі күні адам факторының критикалық маңызды болған жерлерінде жоюға мүмкіндік береді. Құрылыс автономия мен басқаруға қарай жылжуда, мұнда идеядан жобаны жүзеге асыруға дейінгі процесс автопилот режимінде навигациямен салыстыруға болады: субъективті шешімдерге тәуелділіксіз, әр кезеңде қолмен араласудың қажеті жоқ, бірақ толық цифрлық бақылау мен қадағалаумен (Рис. 10.22).-

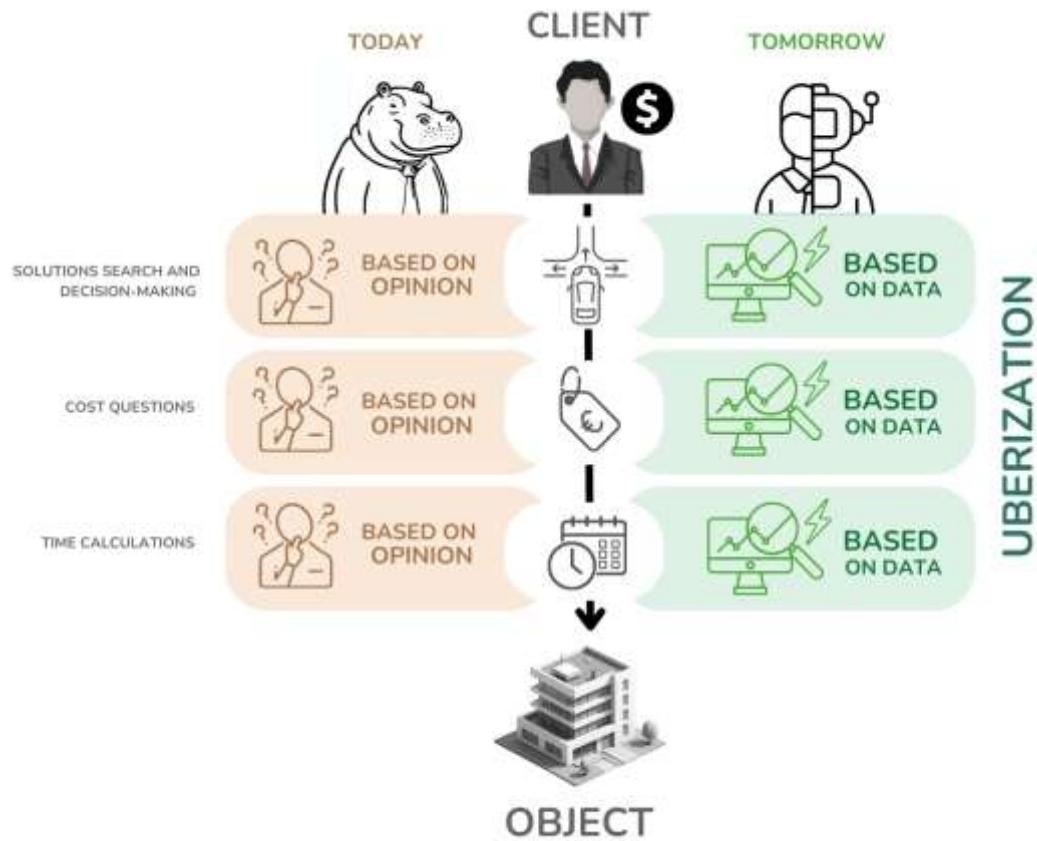


Рис. 10.22 Маңызды мамандардың пікірлеріне (НіPPO) негізделген шешім қабылдаудан деректерді талдауға көшу, ең алдымен, тапсырыс беруші тарарапынан жүзеге асырылатын болады.

Бұл кітапта ұсынылған әдістер, принциптер мен құралдарды зерттей отырып, сіз өз компанияңыңда шешімдерді интуицияға емес, деректерге негізделген қабылдауды бастай аласыз. Сіз LLM-да модульдер тізбегін іске қосып, дайын ETL Pipelines-ті өзініздің даму ортасына (IDE) көшіруге және деректерді автоматты түрде өндеге, қажетті формада ақпарат алуға мүмкіндік аласыз. Алдағы уақытта, үлкен деректер мен машиналық оқытуға арналған тарауларға сүйене отырып, сіз тарихи деректерден жаңа білімдер алуды және процестеріңізді болжау мен оңтайландыру үшін машиналық оқыту алгоритмдерін қолдануды жүзеге асыра аласыз.

Ашық деректер мен процестер жобалардың құны мен мерзімдерін дәл бағалаудың негізі болады, құрылым компанияларын түсініксіз деректер бойынша спекуляция жасау мүмкіндігінен айырады. Бұл сала үшін өз рөлін қайта қараша және жаңа ортаға бейімделу, мұнда ашықтық пен тиімділік табыстың негізгі факторлары болатыны, бір мезгілде шақыру мен мүмкіндік болып табылады.

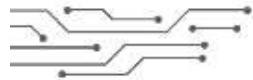
Білімді қабылдап, оны практикада қолдануға дайын болу - цифрлық трансформация дәуіріндегі табыстың кілті.

Бұл туралы алғашқылардың бірі болып түсінетін компаниялар жаңа цифрлық бәсекелестік жағдайында артықшылыққа ие болады. Бірақ, деректердің өздігінен ештең өзгерпейтінін түсіну маңызды. Көптеген адамдардың әдеттегі ойлау тәсілдерін өзгертуі қажет, ал бұл үшін ынталандыру

керек. Сіздің компанияңыз деректер алмасу тәсілін қайта қарастыруы тиіс.

Компанияны өзгертетін - деректермен жұмыс істей алатын, оларды интерпретациялайтын, оңтайландыру үшін пайдаланатын және олардың негізінде жаңа процестер архитектурасын құратын адамдар.

Егер сіз осы жолдарды оқып отырсаңыз, сіз өзгерістерге дайынсыз және сіз уже бір қадам алдасыз. Бұл жолды таңдағаныңыз үшін рақмет. Цифрлық трансформация дәуіріне қош келдіңіз!



Автор туралы

Менің атым Артём Бойко. Менің құрылыш алаңындағы жолым 2007 жылдың басталды - туған қаламдағы сланец шахтасында шахтер болып жұмыс істеген кезімде, Санкт-Петербург тау-кен университетінде «Шахталық және жер асты құрылыш» мамандығы бойынша оқығанымда. Осы кітаптың артқы жағында сіз мені жарылғыш заттармен жұмыс істеп жатқан шахтада көре аласыз, онда біз жүздеген текше метр жанғыш сланецті өндіретін және жарылғыш заттарды қолданатынбыз. Менің мансабым әртүрлі бағыттарда дамыды - шахтада және метрополитен құрылышында жұмыс істегеннен бастап, өнеркәсіптік альпинист, шатыр монтаждаушы және лифт жабдықтарын орнатушыға дейін. Мен әртүрлі ауқымдағы жобаларға қатысу мүмкіндігіне ие болдым: жеке үйлерден бастап, әлемнің әртүрлі аймақтарындағы ірі өнеркәсіптік объектілерге дейін.



Уақыт өте келе, менің жұмысым физикалық құрылыш саласынан ақпаратты басқару және цифрлық процестерге ауысты. 2013 жылдан бері мен Германияның бірнеше аймақтарында шағын, орта және ірі құрылыш компанияларында әртүрлі лауазымдарда жұмыс істедім, жобалаушыдан деректерді басқару менеджеріне дейін. Деректерді басқару бойынша менің тәжірибем ERP, CAD (BIM), MEP, FEM, CMS жүйелеріндегі деректермен жұмыс істеуді қамтиды. Мен процестерді оңтайландыру, автоматтандыру, сондай-ақ құрылыш жұмыстарын жоспарлау, есептеу және орындау кезеңдерінде деректерді талдау, машиналық оқыту, өңдеу жұмыстарымен айналыстым, өнеркәсіптік, тұрғын үй, инфрақұрылымдық және коммуналдық объектілерді құрылышымен айналысатын компанияларда.

2003 жылдан бері мен ашық бағдарламалық қамтамасыз ету және ашық деректермен жұмыс істеп келемін. Осы уақыт ішінде мен веб-жобалардың көптеген түрлерін жүзеге асырдым - сайттар мен интернет-дүкендерден бастап, толыққанды веб-қосымшаларға дейін, open source шешімдері мен ашық CMS пайдалана отырып. Бұл платформалар, қазіргі заманғы құрылыш ERP-мен көп жағынан үқсас, модульдік архитектураға, жоғары бейімделгіштікке және қолжетімділікке ие. Бұл тәжірибе менің кәсіби көзқарасымның негізін қалады - ашық технологияларға және бірлескен дамуға бағдарлану. Ашық кодқа және білімдерді еркін алмасуға деген құрметімді мен құрылыш саласында ілгерілетуге тырысамын. Құрылыш саласындағы деректердің қолжетімділігін арттыру бойынша жұмысым әлеуметтік желілерде деректердің ашықтығы мен құрылыш саласындағы Open Source-ты пайдалану мәселелерін талқылау үшін бірнеше қауымдастықтар құруға, сондай-ақ әртүрлі жабық жүйелер мен платформалардан деректерге қол жеткізу үшін қамтамасыз ететін бірнеше стартаптарды іске қосуға айналды.

Менің кәсіби қоғамдастыққа қосқан үлесім CAD (BIM), ERP, 4D-5D, LLM машиналық оқыту және жасанды интеллект мәселелерін қамтитын конференцияларда баяндамашы ретінде қатысымда, сондай-ақ құрылыш саласына арналған еуропалық басылымдарда жарияланған мақалаларымда көрініс табады. Менің айтартықтай жетістіктерімнің бірі "BIM тарихы" [111] болып табылады, бұл құрылыш саласындағы деректерді басқару үшін маңызды бағдарламалық шешімдердің кешенді картасы. "BIM дамуы және лоббистік ойындар" атты 7 бөлімнен тұратын мақалалар сериям бірнеше тілге аударылып, цифрлық стандарттардың дамуындағы жасырын динамиканы ашуға бағытталған талпыныс ретінде кеңінен танылды.

Мен тау-кен өндірісінен құрылыш деректерін жинау және жүйелеуге дейінгі жолды өттім. Мен әрқашан кәсіби диалогқа, жаңа идеяларға және бірлескен жобаларға ашықтын. Кез келген кері байланысқа ризашылығымды білдіремін және сіздердің хабарламаларыңызды немесе әлеуметтік

желілерде менің жазылушым ретінде көруді қуанышпен қабылдаймын. Бұл кітапты соңына дейін оқығаныңыз үшін көп раҳмет! Менің үмітім, бұл кітап сізге құрылыш саласындағы деректер тақырыбын жақсы түсінуге көмектеседі.

Кері байланыс

Оқырмандардың пікірі жарияланымдардың әрі қарай дамуына және басым тақырыптарды таңдауға маңызды рөл атқарады. Әсіресе, қандай идеялардың пайдалы болғаны, ал қандайларының күмән тудырғаны және қосымша түсініктемелер немесе дереккөздер көрсетуі қажет екендігі туралы ескертулер құнды. Кітап кең ауқымды материалдар мен аналитикалық бағалауларды қамтиды, олардың кейбіреулері даулы немесе субъективті болып көрінү мүмкін. Егер оқу барысында сіз дәлсіздіктер, дұрыс көрсетілмеген дереккөздер, логикалық сәйкесіздіктер немесе баспа қателері тапсаңыз – мен сіздің пікіріңізді, ойларыңызды немесе сын-пікіріңізді boikoartem@gmail.com әлектрондық поштасына немесе LinkedIn арқылы: linkedin.com/in/boikoartem жіберуіңізді ри-зашылықпен қабылдаймын.

Data-Driven Construction кітабы туралы әлеуметтік желілерде кез келген ескертулеріңіз үшін алғысымды білдіремін - оқу тәжірибесімен алмасу ашық деректер мен құралдар туралы ақпаратты таратуға көмектеседі және менің жұмысымды қолдайды.

АУДАРМАҒА ПІКІР

Бұл кітап жасанды интеллект технологияларының көмегімен аударылды. Бұл аударма процесін едәүір жеделдетуге мүмкіндік берді. Алайда, кез келген технологиялық операцияда қателіктер немесе дәлсіздіктер пайда болуы мүмкін. Егер сізде дұрыс емес немесе дұрыс аударылмаған нәрсе байқалса, маған хабарласыңыз. Сіздің ескертулеріңіз аударма сапасын жақсартуға көмектеседі.

DATADRIVENCONSTRUCTION ҚОҒАМДАРЫ

Бұл жерде сіз сұрақтар қоюға және өз мәселелеріңіз бен шешімдеріңізben белісуге еркінсіз:

DataDrivenConstruction.io: <https://datadrivenconstruction.io>

LinkedIn: <https://www.linkedin.com/company/datadrivenconstruction/>

Twitter: <https://twitter.com/datadrivenconst>

Telegram: <https://t.me/datadrivenconstruction>

YouTube: <https://www.youtube.com/@datadrivenconstruction>

БАСҚА ДАҒДЫЛАР ЖӘНЕ КОНЦЕПЦИЯЛАР

DataDrivenConstruction кітабында құрылым саласындағы деректермен жұмыс істеудің негізгі принципперінен басқа, деректермен жұмыс істейтін маманға қажетті қосымша концепциялар, бағдарламалар мен дағдылардың кең ауқымы қарастырылады. Олардың кейбіреулері тек шолумен ұсынылған, бірақ практикада маңызды рөл атқарады.

Қызығушылық танытқан оқырмандар DataDrivenConstruction.io веб-сайтына кіріп, негізгі дағдылар бойынша қосымша материалдарға сілтемелерді таба алады. Бұл материалдар Python және Pandas-ты пайдалану, ETL процестерін құру, CAD құрылымында деректерді өңдеу мысалдары, үлкен деректерді өңдеу жүйелері, сондай-ақ құрылымында деректерін визуализациялау және аналитикаға заманауи тәсілдерді қамтиды.

"DataDrivenConstruction" кітабын дайындау барысында және барлық практикалық мысалдарда ашық кодты көптеген құралдар мен бағдарламалық қамтамасыз етулер пайдаланылды. Автор көлесі шешімдердің әзірлеушілері мен авторларына алғыс білдіреді:

- Python және Pandas – деректермен жұмыс істеу және автоматтандырудың негізі
- Scipy, NumPy, Matplotlib және Scikit-Learn – деректерді талдау және машиналық оқыту үшін кітапханалар
- SQL және Apache Parquet – үлкен көлемдегі құрылымын деректерін сақтау және өңдеу құралдары
- Open Source CAD (BIM) – ашық форматтарда деректермен жұмыс істеуге арналған ашық құралдар
- N8n, Apache Airflow, Apache NiFi – жұмыс процестерін оркестрациялау және автоматтандыру жүйелері
- DeepSeek, LLaMa, Mistral – Open Source LLM

Ашық деректер мен кәсіби қауымдастықтар мен әлеуметтік желілердегі құралдар тақырыбында пікірталастарға қатысқан барлық қатысуышыларға, олардың сыни пікірлері, комментарийлері мен идеялары осы кітаптың мазмұны мен құрылымын жақсартуға көмектескені үшін алғысымды білдіремін.

DataDrivenConstruction.io сайтында жобаның дамуын бақылап отырыңыз, онда кітаптың жаңартулары мен түзетулері ғана емес, сонымен қатар жаңа тараулар, оқу материалдары және сипатталған әдістерді қолдануға арналған практикалық мысалдар жарияланады.

БАСПА НҰСҚАСЫМЕН МАКСИМАЛДЫ ҚОЛАЙЛЫЛЫҚ

Сіздер Data-Driven Construction-ның тегін цифрлық нұсқасын ұстап отырсыздар. Материалдарға жылдам қол жеткізу және ыңғайлы жұмыс үшін баспа басылымына назар аударуды ұсынамыз:



- Эрдайым қол астында: баспа форматындағы кітап сенімді жұмыс құралы болып табылады, қажетті визуализациялар мен схемаларды кез келген жұмыс жағдайында жылдам табуға және пайдалануға мүмкіндік береді.
- Суреттердің жоғары сапасы: баспа басылымындағы барлық бейнелер мен графиктер максималды сапада ұсынылған.
- Ақпаратқа жылдам қол жеткізу: ыңғайлы навигация, ескертпелер жасау, бетбелгілер қою және кітаппен кез келген жерде жұмыс істеу мүмкіндігі.



Кітаптың толық баспа нұсқасын сатып алу арқылы сіз ақпаратпен ыңғайлы және

тиімді жұмыс істеуге арналған құрал аласыз: күнделікті тапсырмаларда визуалды материалдарды жедел пайдалану, қажетті схемаларды жылдам табу және ескертпелер жасау мүмкіндігі. Сонымен қатар, сатып алуыңыз ашық білімнің таралуын қолдайды.

Кітаптың баспа нұсқасын мына жерден тапсырыс беруге болады: datadrivenconstruction.io/books



СТРАТЕГИЯЛЫҚ ПОЗИЦИОНИРОВАНИЯ ҮШІН УНИКАЛЬДЫ МУМКІНДІК

Сіздерге DataDrivenConstruction басылымының тегін нұсқасында жарнамалық материалдарды орналастыруды ұсынамыз. Басылымның ақылы нұсқасы жарияланғаннан кейінгі бірінші жылы Латын Америкасынан бастап Азия-Тынық мұхиты аймағына дейін 50-ден астам елдің мамандарының назарын аударды. Ұнтымақтастықтың жеке шарттарын талқылау және орналастыру мүмкіндіктері туралы толық ақпарат алу үшін, ресми портал datadrivenconstruction.io сайтында көрсетілген байланыс формасын толтырыңыз немесе кітаптың соңында көрсетілген байланыс деректері бойынша жазыңыз.



КІТАП ГЛАВАЛАРЫ DATADRIVENCONSTRUCTION.IO САЙТЫНДА ҚОЛ ЖЕТИМДІ

Сіз Data-Driven Construction кітабының тарауларын сайтта оқи аласыз, онда кітаптың бөлімдері кезең-кезеңімен жарияланады, осылайша қажетті ақпаратты тез таба аласыз және оны жұмысыныңда пайдалана аласыз. Сонымен қатар, сайтта ұқсас тақырыптар бойынша көптеген басқа жарияланымдар, сондай-ақ сіздің дағдыларыңызды дамытуға және құрылыш саласында директордердің қолдануға көмектесетін қосымшалар мен шешімдердің мысалдары бар.



КІТАПТЫҢ СОҢҒЫ НҰСҚАЛАРЫН РЕСМИ САЙТТАН ЖҮКТЕҢІЗ

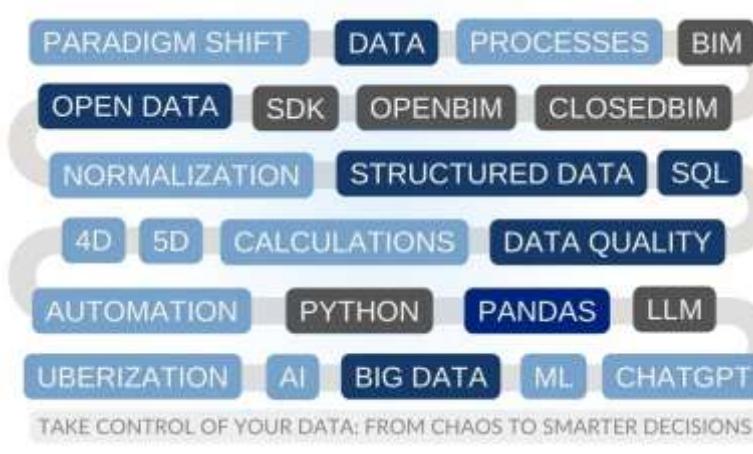
Актуалды және соңғы DataDrivenConstruction кітабының нұсқалары datadrivenconstruction.io сайтында жүктеуге қолжетімді. Егер сіз кітаптың жаңа тарауларымен, практикалық кеңестермен немесе жаңа қосымшалардың шолуларымен жаңартулар алғының келсе, ақпараттық бюллетеньге жазылышыңыз:

- Сіз кітаптың жаңа бөлімдерімен бірінші болып танысасыз
- Құрылыш саласындағы аналитика мен автоматтандыру бойынша практикалық жағдайлар мен кеңестер алуға болады.
- Трендтерді, жарияланымдарды және қосымшалардың мысалдарын бақылау

datadrivenconstruction.io сайтына өтіп, жазылымды рәсімденіз!

DATADRIVENCONSTRUCTION: КОНСАЛТИНГ, СЕМИНАРЛАР ЖӘНЕ ОҚЫТУ

DataDrivenConstruction-ның оқыту бағдарламалары мен консультациялары әлемдегі жетекші құрылым компанияларына тиімділікті арттыруға, шығындарды азайтуға және шешімдердің сапасын жақсартуға көмектесті. DataDrivenConstruction-ның клиенттері арасында миллиард еуро айналымы бар нарықтың ірі ойыншылары, құрылыш, консалтинг және IT-компаниялары бар.



Неліктен бізді таңдау керек?

- Актуалдылық: саладағы негізгі трендтер мен инсайттар турали ақпарат береміз
- Практика: мамандарға күнделікті міндеттерді тиімді шешуге РоС арқылы көмектесеміз.
- Жеке көзқарас: сіздің бизнесіңіздің ерекшеліктерін ескере отырып, оқыту мен консультациялардан максималды пайда алууды қамтамасыз етеміз.

DataDrivenConstruction командасының негізгі жұмыс бағыттары:

- Деректер сапасын басқару: тапсырмаларды параметреуге, талаптарды жинауға, деректерді тексеруге және автоматтандырылған өндеду үшін дайындауға көмектесеміз.
- Деректерді шығару және құрылымдау: ETL процестерін орнатып, деректерді электрондық пошта, PDF, Excel, суреттер және басқа көздерден шығарамыз.
- BIM және CAD аналитикасы: RVT, IFC, DWG және басқа CAD (BIM) форматтарындағы файлдардан ақпарат жинап, құрылымдалған деректерге, аналитикаға, қорытындылар мен шешімдерге айналдырамыз.
- Деректерді аналитика және трансформациялау: шашыраңқы ақпаратты құрылымдалған деректерге, аналитикаға, қорытындылар мен шешімдерге айналдырамыз.
- Деректерді интеграциялау және процестерді автоматтандыру: автоматтандырылған құжаттарды жасаудан бастап, ішкі жүйелермен және сыртқы деректер базаларымен интеграциялауға дейін.

DataDrivenConstruction.io сайтына хабарласыңыз, автоматтандыруды пайдалану сіздің компанияңызға елеулі бизнес нәтижелеріне қалай қол жеткізуге көмектесетінін білу үшін.

ГЛОССАРИЙ

AI (Жасанды Интеллект) – компьютерлік жүйелердің адам интеллектісін талап ететін тапсырмаларды орындау қабілеті, мысалы, бейнелерді тану, оқу және шешім қабылдау.

Apache Airflow – жұмыс процестерін оркестрациялау үшін ашық платформа, DAG (бағытталған ацикльдік графтар) пайдалана отырып, жұмыс процестерін және ETL-ді бағдарламалық түрде жасауға, жоспарлауда және бақылауда мүмкіндік береді.

Apache NiFi – жүйелер арасында деректер ағындарын автоматтандыруға арналған құрал, деректерді маршрутизациялау және түрлендіру бойынша маманданған.

Apache Parquet – деректерді бағандық сақтау үшін тиімді файл форматы, үлкен деректерді талдау жүйелерінде пайдалануға оңтайландырылған. Маңызды қысу мен жылдам өңдеуді қамтамасыз етеді.

API (Қолданбалы Бағдарламалау Интерфейсі) – бір бағдарламаның екіншімен өзара әрекеттесуіне мүмкіндік беретін формализденген интерфейс, деректер мен функционалдылықты стандартталған сұраулар мен жауаптар арқылы алмасуға мүмкіндік береді.

Атрибут – обьектінің ерекшеліктерін сипаттайтын сипаттама немесе қасиет (мысалы, алан, көлем, құн, материал).

Деректер базалары – ақпаратты сақтау, басқару және оған қол жеткізу үшін ұйымдастырылған құрылымдар, деректерді тиімді іздеу және өңдеу үшін пайдаланылады.

BEP (BIM орында жоспары) – ақпараттық модельдеу жоспары, BIM-ді жобада енгізудің мақсаттарын, әдістерін және процестерін анықтайды.

Big Data (Үлкен деректер) – көлемі, әртүрлілігі және жаңартылу жылдамдығы жағынан елеулі ақпарат массивтері, оларды өңдеу және талдау үшін арнайы технологияларды талап етеді.

BI (Бизнес-аналитика) – деректерді шешім қабылдау үшін мағыналы ақпаратқа айналдыруға арналған процестер, технологиялар мен құралдар.

BIM (Ғимараттық ақпараттық моделін жасау) – ғимараттардың цифрлық көріністерін жасау және басқару процесі, ол тек 3D модельдерін ғана емес, сонымен қатар сипаттамалар, материалдар, мерзімдер мен құн туралы ақпаратты қамтиды.

BlackBox/WhiteBox – жүйені түсінудің тәсілдері: бірінші жағдайда ішкі логика жасырын, тек кірістер мен шығыстар көрінеді; екінші жағдайда – өңдеу процесі ашық және талдауға қолжетімді.

Bounding Box – үш өлшемді кеңістікте обьектінің шекараларын минималды және максималды координаттар арқылы сипаттайтын геометриялық құрылым, обьектінің айналасында "қорап" құрады.

BREP (Шекаралық репрезентация) – обьектілерді олардың беттерінің шекаралары арқылы анықтайдын геометриялық репрезентация.

CAD (Компьютерлік жобалау) – архитектура, құрылым, машина жасау және басқа салаларда дәл сыйбалар мен 3D модельдерді жасау, өңдеу және талдау үшін қолданылатын автоматтандырылған жобалау жүйесі.

CAFM (Компьютерлік обьектілерді басқару) – жылжымайтын мүлік пен инфрақұрылымды басқаруға арналған бағдарламалық қамтамасыз ету, кеңістікті жоспарлау, активтерді басқару, техникалық қызмет көрсету және шығындарды мониторингтеуді қамтиды.

CDE (Бірлескен деректер ортасы) – обьектінің өмірлік циклі кезеңдерінде жобалық ақпаратты басқару, сақтау, алмасу және бірлесіп жұмыс істеу үшін орталықтандырылған цифрлық кеңістік.

Озық тәжірибе орталығы (Center of Excellence, CoE) – ұйымдағы белгілі бір білім саласын дамыту, стандарттар мен үздік тәжірибелерді әзірлеу, персоналды оқыту және инновацияларды енгізуі қолдау үшін жаупты мамандандырылған құрылым.

CoClass – құрылыс элементтерінің заманауи үшінші буын классификация жүйесі.

Концептуалдық деректер моделі – деректер базасын жобалаудың бастапқы кезеңдерінде атрибуттардың егжей-тегжейінсіз негізгі объектілер мен олардың өзара байланыстарын жоғары деңгейде көрсету.

CRM (Клиенттермен қарым-қатынасты басқару) – сату және қызмет көрсету процестерін автоматтандыру үшін қолданылатын клиенттермен өзара әрекеттесуді басқару жүйесі.

DAG (Бағытталған ацикличичный граф) – деректерді оркестрациялау жүйелерінде (Airflow, NiFi) тапсырмалардың ретін және тәуелділіктерін анықтау үшін қолданылатын граф.

Dash – деректердің интерактивті веб-визуализацияларын жасау үшін Python фреймворкі.

Дашборд (Dashboard) – нақты уақыт режимінде негізгі тиімділік көрсеткіштері мен метрикаларды визуалды түрде көрсететін ақпараттық панель.

Data-Centric тәсілі – деректерді, бағдарламалар немесе бағдарламалық кодтан гөрі, ұйымның орталық активі ретінде қарастыратын методология.

Data Governance – ұйымдағы деректерді дұрыс және тиімді пайдалануды қамтамасыз ететін практикалар, процестер мен саясаттар кешені, соның ішінде қолжетімділікті, сапаны және қауіпсіздікті бақылау.

Дата көлдері – үлкен көлемдегі өндөлмеген деректерді олардың бастапқы форматында сақтау үшін арналған сақтау орны.

Дата көл үйі – деректер көлдерінің (Data Lake) икемділігі мен масштабталуын деректер қоймаларының (DWH) басқарылуымен және өнімділігімен біріктіретін архитектуралық тәсіл.

Деректерге негізделген құрылыш – объектінің өмірлік циклі кезеңдерінің әрқайсысы – жобалаудан бастап пайдалану кезеңіне дейін – автоматтандырылған, өзара байланысты жүйелермен қолдау көрсетілетін стратегиялық тәсіл. Мұндай тәсіл фактілер негізінде тұрақты оқытуды қамтамасыз етеді, белгісіздікті азайтады және компанияларға салада тұрақты көшбасшылыққа қол жеткізуге мүмкіндік береді.

Деректерге негізделген интегратор – деректерді әртүрлі көздерден біріктіру және оларды басқару шешімдерін қабылдау үшін талдау бойынша маманданған компания.

Деректерге негізделген тәсіл (Дата-орталықты тәсіл) – деректерді стратегиялық актив ретінде қарастыратын және шешімдерді объективті ақпаратты талдау негізінде қабылдайтын методология.

Деректер минимализмі – деректерді ең құнды және маңызды элементтерге дейін қысқартуға бағытталған тәсіл, ақпаратты өңдеу мен талдауды жеңілдетуге мүмкіндік береді.

Деректер батпағы – ақпаратты бақылаусыз жинау және сақтау нәтижесінде пайда болатын ұйымдастырылмаған деректердің шашыраңқы массиві.

DataOps – DevOps, деректер және аналитиканы біріктіретін методология, деректер ағындарын ынтымақтастықты, интеграцияны және автоматтандыруды жақсартуға бағытталған.

Ақпаратты цифрландыру – құрылыш қызметінің барлық аспектілерін талдау, интерпретациялау және автоматтандыру үшін жарамды цифрлық формаға түрлендіру процесі.

DataFrame (Датафрейм) – Pandas кітапханасындағы екі өлшемді кестелік деректер құрылымы, мұнда жолдар жеке жазбаларды немесе объектілерді, ал бағандар олардың сипаттамаларын немесе атрибуттарын білдіреді.

Сипаттамалық аналитика – өткенде не болғанын түсіну үшін тарихи деректерді талдау.

Диагностикалық аналитика – не себепті не болғанын анықтау үшін деректерді талдау.

Гант диаграммасы – жобаларды жоспарлау құралы, тапсырмаларды уақыт шкаласында көлденең жолақтар түрінде көрсетіп, жұмыстардың ретін және ұзақтығын визуалды түрде көрсетуге мүмкіндік береді.

DWH (Деректер қоймасы) – көптеген көздерден ақпаратты жинақтайдын, оны құрылымдайтын және аналитика мен есеп беру үшін қолжетімді ететін орталықтандырылған деректер сақтау жүйесі.

ESG (Экологиялық, Әлеуметтік, Басқару) – компанияның немесе жобаның экологиялық, әлеуметтік және басқарушылық әсерін бағалау үшін критерийлер жиынтығы.

ELT (Extract, Load, Transform) – деректер алдымен көздерден шығарылады және қоймаға жүктелді, содан кейін аналитикалық мақсаттар үшін түрлендіріледі.

ETL (Extract, Transform, Load) – әртүрлі көздерден деректерді шығару, оларды қажетті форматқа түрлендіру және талдау үшін мақсатты қоймаға жүктеу процесі.

ER-диаграмма (Субъект-Қатынас) – деректерді модельдеу кезінде субъектілерді, олардың атрибуттарын және арасындағы байланыстарды көрсететін визуалды схема.

ERP (Кәсіпорын ресурстарын жоспарлау) – құрылым процесінің әртүрлі аспектілерін басқару және оңтайландыру үшін қолданылатын кешенді модульдік жоспарлау жүйесі.

Ерекшеліктер (Характеристики) – машина оқытуында, модель үшін кіріс деректер ретінде пайдаланылатын тәуелсіз айнымалылар немесе атрибуттар.

Деректердің физикалық моделі – деректер базасының құрылымын егжей-тегжейлі көрсету, кестелер, бағандар, деректер түрлері, кілттер мен индекстерді қамтитын, нақты СУБД үшін оңтайландырылған.

FPDF – PDF құжаттарын жасау үшін Python кітапханасы.

Геометриялық ядро – CAD, BIM және басқа инженерлік қосымшаларда геометриялық обьектілерді жасау, өңдеу және талдау үшін негізгі алгоритмдерді қамтамасыз ететін бағдарламалық компонент.

HiPPO (Ең жоғары жалақы алатын адамның пікірі) – үйімдағы ең жоғары жалақы алатын адамның пікірі негізінде шешім қабылдау тәсілі, объективті деректерге емес.

IDE (Интеграцияланған даму ортасы) – код жазу, тестілеу және жөндеу үшін кешенді құрал (мысалы, PyCharm, VS Code, Jupyter Notebook).

IDS (Ақпаратты жеткізу спецификациясы) – жобаның әртүрлі кезеңдерінде деректерге қойылатын талаптарды анықтайдын ақпаратты жеткізу спецификациясы.

IFC (Сала негіздері класы) – BIM деректерін алмасу форматы, әртүрлі бағдарламалық шешімдер арасында үйлесімділікті қамтамасыз етеді.

Industry 5.0 – өнеркәсіпті дамыту концепциясы, цифрандыру, автоматтандыру және жасанды интеллект мүмкіндіктерін адам әлеуеті мен экологиялық тұрақтылықпен біріктіреді.

Деректерді интеграциялау – деректерді әртүрлі көздерден біртұтас, тұтас жүйеге біріктіру процесі, ақпараттың біртұтас көрінісін қамтамасыз ету үшін.

Ақпараттық силостар – басқа жүйелермен ақпарат алмаспайтын оқшауланған деректерді сақтау жүйелері, деректерді тиімді пайдалануға кедегі келтіреді.

IoT (Заттар интернеті) – физикалық обьектілерді интернетке қосу концепциясы, деректерді жинау, өңдеу және беру үшін.

k-NN (k-Жақын көршілер) – обьектілерді оқыту жиындығы жақын көршілерімен ұқсастық негізінде класификациялайтын машина оқыту алгоритмі.

Kaggle – деректерді талдау және машина оқыту бойынша жарыстар өткізу платформасы.

Калькуляция – құрылымдың жұмыстарының немесе процестердің құнын белгілі бір өлшем бірлігіне (мысалы, 1 м² гипсокартон қабырғасы, 1 м³ бетон) есептеу.

KPI (Негізгі өнімділік көрсеткіштері) – компанияның немесе нақты жобаның табыстылығын бағалау үшін қолданылатын сандық түрде өлшенетін метрикалар.

Labels (Меткалар) – машина оқытуында, модель болжауы тиіс мақсатты айнымалылар немесе атрибуттар.

Learning Algorithm (Оқыту алгоритмі) – оқыту деректер жиынын пайдалана отырып, мақсатты функцияға сәйкес модельдегі ең жақсы гипотезаны іздеу процесі.

Linear Regression (Сызықтық регрессия) – тәуелді айнымалы мен бір немесе бірнеше тәуелсіз айнымалылар арасындағы өзара байланысты модельдеу үшін статистикалық әдіс.

LLM (Үлкен тілдік модель) – мәтінді түсіну және генерациялау үшін үлкен деректер массивтері негізінде оқытылған жасанды интеллект, контексті талдай алатын және бағдарламалық код жаза алатын.

LOD (Детализация/Даму деңгейі) – модельдің детализация деңгейі, геометриялық дәлдік пен ақпараттық толтыру дәрежесін анықтайды.

Логикалық деректер моделі – бизнес ақпаратын және ережелерді көрсететін, субъектілер, атрибуттар, кілттер және қатынастардың егжей-тегжейлі сипаттамасы, концептуалды және физикалық модельдер арасындағы аралық кезең.

Машиналық оқыту – деректер негізінде компьютерлік жүйелердің үйренуіне және болжам жасауды мүмкіндік беретін жасанды интеллект әдістерінің класы, айқын бағдарламалаусыз.

Masterformat – құрылымды спецификацияларын бөлімдер мен пәндер бойынша құрылымдауға арналған бірінші буын классификация жүйесі.

МЕР (Механикалық, Электрлік, Сантехникалық) – ғимараттардың инженерлік жүйелері, механикалық, электрлік және сантехникалық компоненттерді қамтиды.

Mesh – шыңдар, қабырғалар және беттерден тұратын 3D объектілердің торлы көрінісі.

Модель – машиналық оқытуда, әртүрлі гипотезалардың жиынтығы, олардың біреуі болжанатын немесе аппроксимацияланатын мақсатты функцияны жақыннатады.

Деректерді модельдеу – ақпараттық жүйелерде жүзеге асыру үшін деректер мен олардың өзара байланыстарын құрылымдық түрде көрсету процесі, концептуалды, логикалық және физикалық деңгейлерді қамтиды.

n8n – төмен кодты (low-code) тәсіл арқылы жұмыс процестерін автоматтандыру және қосымшаларды интеграциялау үшін ашық кодты құрал.

Нормализация – машиналық оқытуда, әртүрлі сандық деректерді өндедеу мен талдауды жеңілдету үшін бір масштабқа келтіру процесі.

Кері инженерия – объектінің құрылымын, функцияларын және жұмысын талдау арқылы оның құрылышын, жұмыс істеуін және өндіріс технологиясын зерттеу процесі. Деректер контекстінде – ақпаратты меншікті форматтардан ашық жүйелерде пайдалану үшін шығару.

OCR (Оптикалық Символдарды Тануға Арналған Технология) – мәтіннің суреттерін (сканерленген құжаттар, фотосуреттер) машинамен оқылатын мәтін форматына түрлендіруге мүмкіндік беретін технология.

OmniClass – құрылымды объектілерінің ақпаратын басқару үшін екінші буын халықаралық классификация стандарты.

Онтология – белгілі бір білім саласын формализациялайтын ұфымдар арасындағы өзара байланыстар жүйесі.

Ашық код – бағдарламалық қамтамасыз етуді әзірлеу және тарату моделі, ашық кодпен қолжетімді, еркін пайдалану, зерттеу және модификациялау үшін.

Ашық BIM – әртүрлі бағдарламалық шешімдер арасында деректерді алмасу үшін ашық стандарттар мен форматтарды пайдалануды көздейтін ашық BIM концепциясы.

Ашық стандарттар – нақты міндеттерді орындау үшін қолжетімді спецификациялар, әртүрлі жүйелердің өзара әрекеттесуіне және деректерді алмасуына мүмкіндік береді.

Pandas – деректерді өңдеу және талдау үшін Python-да ашық кодты кітапхана, кестелік ақпаратпен тиімді жұмыс істеу үшін DataFrame және Series деректер құрылымдарын ұсынады.

Ашық деректер парадигмасы – деректерді өңдеу тәсілі, ақпараттың кез келген адамға еркін қолжетімді, көп рет пайдалануға және таратуға мүмкіндік беретін.

Параметрлік әдіс – жобаның параметрлеріне негізделген құны бағалау үшін статистикалық модельдерді пайдаланатын құрылым жобаларын бағалау әдісі.

PIMS (Жобалық Ақпарат Моделі) – барлық жобалық ақпаратты үйімдастыру, сақтау және алмасу үшін арналған цифрлық жүйе.

Пайплайн – деректерді өңдеу процестерінің тізбегі, алу және түрлендіруден бастап талдау және визуализацияға дейін.

PMIS (Жобалық Ақпаратты Басқару Жүйесі) – жеке құрылым объектісінің деңгейінде тапсырмалардың орындалуын ежей-тегжейлі бақылауға арналған жобаларды басқару жүйесі.

Прогноздық аналитика (Predictive Analytics) – тарихи деректер негізінде болашақ нәтижелерді болжау үшін статистикалық әдістер мен машиналық оқытуды пайдаланатын аналитика бөлімі.

Прескриптивтік аналитика (Prescriptive Analytics) – болашақ нәтижелерді болжаумен қатар, қалаған нәтижелерге қол жеткізу үшін оңтайлы әрекеттерді ұсыннатын аналитика бөлімі.

Проприетарлық форматтар – белгілі бір компания тарапынан бақыланатын жабық деректер форматтары, ақпарат алмасу мүмкіндіктерін шектейді және нақты бағдарламалық қамтамасыз етуге тәуелділікті арттырады.

QTO (Саны алу) – жобалық құжаттамадан элементтердің сандық сипаттамаларын алу процесі, жобаны жүзеге асыру үшін қажетті материалдар көлемін есептеу үшін.

Сапа басқару жүйесі – процестер мен нәтижелердің белгіленген талаптарға сәйкестігін қамтамасыз ететін сапа басқару жүйесі.

RAG (Ақпаратты алу арқылы генерациялау) – тілдік модельдердің генеративтік мүмкіндіктерін корпоративтік деректер базаларынан релевантты ақпаратты алымен біріктіретін әдіс, жауптардың дәлдігі мен өзектілігін арттырады.

RDBMS (Реляциялық Деректерді Басқару Жүйесі) – ақпаратты өзара байланысты кестелер түрінде үйімдастыратын реляциялық деректер базаларын басқару жүйесі.

RegEx (Регулярлы өрнектер) – жолдарды іздеу және өңдеу үшін формализденген тіл, мәтіндік деректерді белгілі бір критерийлерге сәйкестігін тексеру үшін шаблондарды анықтауға мүмкіндік береді.

Регрессия – айнымалылар арасындағы тәуелділікті талдау үшін статистикалық әдіс.

CO₂ есептеулері – құрылымынан шығатын CO₂ газын бағалау әдісі.

Ресурс әдісі – құрылымынан шығатын CO₂ газын бағалау әдісі.

RFID (Радиожиілікпен сәйкестендіру) – радиосигналдар арқылы объектілерді автоматты түрде сәйкестендіру технологиясы, материалдарды, техниканы және персоналды бақылау үшін қолданылады.

ROI (Инвестициядан қайтарым) – пайда мен салынған қаражат арасындағы қатынасты көрсететін көрсеткіш, инвестициялардың тиімділігін бағалау үшін пайдаланылады.

SaaS (Бағдарламалық қамтамасыз етуді қызмет ретінде) – бағдарламалық қамтамасыз етуді қызмет ретінде ұсыну моделі, мұнда қосымшалар провайдермен орналастырылады және интернет арқылы пайдаланушыларға қолжетімді.

SCM (Жеткізу тізбегін басқару) – материалдарды сатып алушан дайын өнімді жеткізуге дейінгі барлық процестерді координациялау және оңтайландыруды қамтитын жеткізу тізбегін басқару.

Деректер силостары – ұйымдағы ақпаратты оқшауланған сақтау орындары, басқа жүйелермен интеграцияланбаған, бұл деректер алмасуды қынданатады және тиімділікті төмендетеді.

SQL (Структурланған сұраныс тілі) – реляциялық деректер базаларымен жұмыс істеу үшін қолданылатын құрылымдық сұраныс тілі.

SQLite – жеке серверді қажет етпейтін, жеңіл, кірістірілетін, кроссплатформалық СУБД, негізгі SQL функцияларын қолдайды, мобильді қосымшалар мен кірістірілген жүйелерде кеңінен қолданылады.

Структурланған деректер – реляциялық деректер базаларында немесе кестелерде анықталған форматта ұйымдастырылған ақпарат.

Әлсіз құрылымдалған деректер – жартылай ұйымдастырылған және икемді құрылымы бар ақпарат, мысалы, JSON немесе XML, мұнда әртүрлі элементтер әртүрлі атрибуттар жиынтығын қамтуы мүмкін.

Субъект (ағылш. entity) – нақты немесе абстрактілі шындық объектісі, оны біржақты бөліп, сипаттап, деректер түрінде ұсынуға болады.

Бақыланатын оқыту – машиналық оқытудың түрі, мұнда алгоритм белгіленген деректерде оқытылады, онда әрбір мысал үшін қажетті нәтиже белгілі.

Таксономия – элементтерді жалпы белгілеріне негізделген санаттарға жүйелі түрде бөлу үшін қолданылатын иерархиялық класификация жүйесі.

Titanic деректер жинағы – машиналық оқыту модельдерін оқыту және тестілеу үшін танымал деректер жинағы.

Оқыту – машиналық оқыту алгоритмі деректерді талдап, заңдылықтарды анықтап, модельді қалыптастыру процесі.

Трансферлік оқыту – машиналық оқытуudyң әдісі, мұнда бір тапсырма үшін оқытылған модель басқа тапсырма үшін бастапқы нүктө ретінде пайдаланылады.

Деректерді түрлендіру – деректердің форматын, құрылымын немесе мазмұнын өзгерту процесі, оларды кейінгі пайдалану үшін.

Деректерге қойылатын талаптар – бизнес-процестерді қолдау үшін қажетті ақпараттың құрылымын, форматтарын, толықтығын және сапасын анықтатын формализденген критерийлер.

Құрылыш саласын «уберизациялау» – дәстүрлі бизнес-модельдерді цифрлық платформалардың әсерінен трансформациялау процесі, бұл тапсырыс берушілер мен орындаушылардың дедалдардың тікелей өзара әрекеттесуін қамтамасыз етеді.

Uniclass – Ұлыбританияда кеңінен қолданылатын екінші және үшінші буын құрылыш элементтерінің класификация жүйесі.

USD (Universal Scene Description) – компьютерлік графика үшін әзірленген деректер форматы, бірақ қарапайым құрылымы мен геометриялық ядролардан тәуелсіздігі арқасында инженерлік жүйелерде қолданылуда.

Деректерді валидациялау – ақпаратты белгіленген критерийлер мен талаптарға сәйкестігін тексеру процесі, деректердің дәлдігін, толықтырын және үйлесімділігін қамтамасыз етеді.

Векторлық база деректері – объектілерді тиімді семантикалық іздеу және салыстыру үшін деректерді көп өлшемді векторлар түрінде сақтайтын мамандандырылған база деректері.

Векторлық ұсыныс (эмбеддинг) – деректерді көп өлшемді сандық векторларға түрлендіру әдісі, бұл машиналық алгоритмдерге ақпаратты тиімді өңдеуге және талдауға мүмкіндік береді.

VectorOps – көп өлшемді векторлық деректерді өңдеуге, сақтауға және талдауға бағытталған методология, әсіресе цифрлық егіздер мен семантикалық іздеу сияқты салаларда өзекті.

Визуализация – ақпаратты тиімді қабылдау және талдау үшін деректердің графикалық көрінісі.

Терминдердің алфавиттік бөлінуі олардың ағылшын тіліндегі атаулары бойынша жүзеге асырылды.

ӘДЕБИЕТТЕР ТІЗІМІ ЖӘНЕ ONLINE-МАТЕРИАЛДАР

- [1] Gartner, IT Key Metrics Data 2017: Index of Published Documents and Metrics, 12 Желтоқсан 2016. [Интернетте]. Қолжетімді: <https://www.gartner.com/en/documents/3530919>. [Қаралған күні: 1 Наурыз 2025].
- [2] KPMG, Familiar challenges – new approaches. 2023 Global Construction Survey, 1 Қаңтар 2023. [Интернетте]. Қолжетімді: <https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2023/06/familiar-challenges-new-solutions-1.pdf>. [Қаралған күні: 5 Наурыз 2025].
- [3] F. Р. Барнард, «Бір сурет мың сөзге тең», 10 Мамыр 1927. [Интернетте]. Қолжетімді: https://en.wikipedia.org/wiki/A_picture_is_worth_a_thousand_words. [Қаралған күні: 15 Наурыз 2025].
- [4] М. Бастиян, Microsoft компаниясының бас директоры Сатья Наделла өзін-өзі жариялаған ЖИ-ді "мағынасыз стандартты бұзу" деп атады, 21 ақпан 2025. [Интернетте]. Қолжетімді: <https://the-decoder.com/microsoft-ceo-satya-nadella-says-self-claiming-agi-is-nonsensical-benchmark-hacking/>. [Қаралған күні: 15 наурыз 2025].
- [5] W. E. Форум, «Құрылым болашағын қалыптастыру – Өзгерістерге үшыраған ландшафт», 1 Қаңтар 2016. [Интернетте]. Қолжетімді: https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Construction.pdf. [Қаралған күні: 2 Наурыз 2025].
- [6] С. Д. Гиллеспи, Глина: Жердің глина дәүіріндегі күрделілігі, 2024. [Интернетте]. Қолжетімді: <https://ufl.pb.unizin.org/imos/chapter/clay/>.
- [7] Папирус б.з.д. III ғасыр. Тілі – грекше, 2024. [Интернетте]. Қолжетімді: <https://www.facebook.com/429710190886668/posts/595698270954525>.
- [8] Мониторинг: қолжетімді құралдарды пайдалану, 1980. [Интернетте]. Қолжетімді: <https://pubmed.ncbi.nlm.nih.gov/10246720/>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [9] PWC, «Деректерге негізделген: Студенттердің тез өзгеріп жатқан бизнес әлемінде табысқа жету үшін қажеттіліктері», 2015 жылғы 15 ақпан. [Интернетте]. Қолжетімді: <https://www.pwc.com/us/en/faculty-resource/assets/PwC-Data-driven-paper-Feb2015.pdf>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [10] Skanska USA, Құрылым нарығының күзгі трендтері, 2023 жылғы 2 қараша. [Интернетте]. Қолжетімді: <https://x.com/SkanskaUSA/status/1720167220817588714>.
- [11] Oxford Essential Quotations (4-ші басылым), Оксфорд Университеті Баспасы, 2016. [Интернетте]. Қолжетімді:

- <https://www.oxfordreference.com/display/10.1093/acref/9780191826719.001.0001/q-oro-ed4-00006236>. [Қаралған күні: 2025 жылғы 1 наурыз].
- [12] Сондергаардтың Деректерді талдау туралы пікірі, [Интернетте]. Қолжетімді: <https://www.causeweb.org/cause/resources/library/r2493>. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [13] Глобалдық ИИ қызығушылығы деректерді басқару нарығын қалай көтеріп жатыр, 28 Мамыр 2024. [Интернетте]. Қолжетімді: <https://iot-analytics.com/how-global-ai-interest-is-boosting-data-management-market/>. [Қаралған күні: 15 Наурыз 2025].
- [14] И. Маккью, ERP тарихы, 2024. [Интернетте]. Қолжетімді: <https://www.netsuite.com/portal/resource/articles/erp/erp-history.shtml>.
- [15] erpscout, ERP бағасы: ERP жүйесінің құны қанша?, [Интернетте]. Қолжетімді: <https://erpscout.de/en/erp-costs/>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [16] softwarepath, «1,384 ERP жобасы ERP таңдау туралы не айтады (2022 ERP есебі)», 2022 жылдың 18 қаңтары. [Интернетте]. Қолжетімді: <https://softwarepath.com/guides/erp-report>. [Қаралған күні: 2025 жылдың 15 наурызы].
- [17] Deloitte, «Цифрлық капиталдық жобалардағы деректерге негізделген басқару», 2016 жылғы 16 желтоқсан. [Интернетте]. Қолжетімді: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/Real%20Estate/us-engineering-construction-data-driven-management-digital-capital-projects.pdf>. [Қаралған күні: 2025 жылғы 1 наурыз].
- [18] Мккинси, «2025 жылғы деректерге негізделген кәсіпорын», 28 Қаңтар 2022. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-data-driven-enterprise-of-2025>. [Қаралған күні: 22 Мамыр 2024].
- [20] Википедия, «Мур заңы», [Интернетте]. Қолжетімді: https://en.wikipedia.org/wiki/Moore%27s_law. [Қаралған күні: 15 Наурыз 2025].
- [22] Акцентур, «Капитал жобаларымен көбірек құн құрғы», 1 Қаңтар 2020. [Интернетте]. Қолжетімді: <https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-143/accenture-industryx-building-value-capital-projects-highres.pdf>. [Қаралған күні: 3 Наурыз 2024].
- [24] Б. Mapp, «Күн сайын қанша дерек жасаймыз? Әлемді таң қалдыратын статистика», 2018. [Интернетте]. Қолжетімді: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read>.
- [26] «Күн сайын қанша дерек өндіріледі?», 2024. [Интернетте]. Қолжетімді: <https://graduate.northeastern.edu/resources/how-much-data-produced-every-day/>.

- [28] Т. Салливан, «ЖИ және жаһандық "деректер ағыны": 2025 жылға қарай адамзаттың қанша ақпаратқа ие болатыны», 2024. [Интернетте]. Қолжетімді: <https://www.datauniverseevent.com/en-us/blog/general/AI-and-the-Global-Datasphere-How-Much-Information-Will-Humanity-Have-By-2025.html>.
- [30] Статиста, «1454 жылдан 1800 жылға дейінгі Батыс Еуропаның әртүрлі аймақтарында шығарылған кітаптардың жалпы саны», [Интернетте]. Қолжетімді: <https://www.statista.com/statistics/1396121/europe-book-production-half-century-region-historical/>. [Қаралған күні: 1 Наурыз 2025].
- [32] McKinsey, «ҚҰРЫЛЫСЫНЫ ҚАЙТА ЖАҢҒЫРТУ: ЖОҒАРЫ ӨНІМДІЛІККЕ ЖОЛ», 1 Ақпан 2017. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/~/media/mckinsey/business%20functions/operations/our%20insights/reinventing%20construction%20through%20a%20productivity%20revolution/mgi-reinventing-construction-a-route-to-higher-productivity-full-report.pdf>.
- [33] Құрылыс Тапсырыс Тобы, Премьер-министрдің орынбасарына, «Құрылысқа қайта қарая», 1 Қазан 2014. [Интернетте]. Қолжетімді: https://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking_construction_report.pdf.
- [34] Forbes, Дерексіз пікірсіз, сіз тек деректермен шектелген адамсыз, 15 наурыз 2016. [Интернетте]. Қолжетімді: <https://www.forbes.com/sites/silberzahnjones/2016/03/15/without-an-opinion-youre-just-another-person-with-data/>. [Қарая күні: 15 наурыз 2025].
- [35] Wikiquote, Чарльз Бэббидж, [Интернетте]. Қолжетімді: https://en.wikiquote.org/wiki/Charles_Babbage. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [36] SAP, «Жаңа зерттеулер көрсеткендей, атқарушы басшылардың жартысынан көбі AI-ға өздерінен гөрі сенеді», 2025 жылғы 12 наурыз. [Интернетте]. Қолжетімді: <https://news.sap.com/2025/03/new-research-executive-trust-ai/>. [Қарая күні: 2025 жылғы 15 наурыз].
- [37] Канадалық Құрылыс Ассоциациясы және KPMG Канада, 2021, «Цифрлық әлемдегі құрылыш», 1 Мамыр 2021. [Интернетте]. Қолжетімді: <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [Қарая күні: 5 Наурыз 2025].
- [38] ЗЦС, «Бесінші өнеркәсіптік революцияны декодтау», [Интернетте]. Қолжетімді: <https://www.pwc.in/decoding-the-fifth-industrial-revolution.html>. [Қарая күні: 2025 жылғы 15 наурыз].
- [39] М. К, Жеке құқықтар және қоғамдық мәселелер: Глобалдық экономика, Питерсон халықаралық экономика институты, 2012.
- [40] F. N. a. Y. Z. Гарвард бизнес мектебі: Мануэль Хоффман, «Ашық кодты бағдарламалық қамтамасыз етудің құндылығы», 24 Қантар 2024. [Интернетте]. Қолжетімді:

<https://www.hbs.edu/faculty/Pages/item.aspx?num=65230>. [Қаралған күні: 15 Наурыз 2025].

- [41] Тәңіз күштері үшін шығындарды талдау орталығы Әуе күштері шығындарды талдау агенттігі, «Бағдарламалық қамтамасыз етуді дамытуға арналған шығындарды бағалау жөніндегі нұсқаулық», 1 қыркүйек 2008. [Интернетте]. Қолжетімді: <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/SW%20Cost%20Est%20Manual%20Vol%20I%20rev%2010.pdf>.
- [42] McKinsey, «Құрылымдың өнімділігін арттыру,» [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/operations/our-insights/improving-construction-productivity>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [43] A. G. a. C. Syverson, АҚШ құрылымдың секторындағы өнімділіктің таңғажайып және қорқынышты жолы, 19 Қаңтар 2023. [Интернетте]. Қолжетімді: <https://bfi.uchicago.edu/insight/research-summary/the-strange-and-awful-path-of-productivity-in-the-us-construction-sector/>. [Қаралған күні: 1 Наурыз 2025].
- [44] McKinsey, «Құрылымдың өнімділігін қамтамасыз ету енді міндетті емес», 2024 жылғы 9 тамыз. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/operations/our-insights/delivering-on-construction-productivity-is-no-longer-optional>. [Қаралған күні: 2025 жылғы 5 наурыз].
- [45] ING Group, «Құрылымдың саласындағы өнімділіктің төмендігі құрылымдың шығындарын арттыруда», 2022 жылғы 12 желтоқсан. [Интернетте]. Қолжетімді: <https://think.ing.com/articles/lagging-productivity-drives-up-building-costs-in-many-eu-countries/>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [46] М. Берман, «Microsoft бас директорының таңқаларлық болжамы: "Агенттер БАРЛЫҚ Бағдарламалық қамтамасыз етуді алмастырады",» 19 Желтоқсан 2024. [Интернетте]. Қолжетімді: <https://www.youtube.com/watch?v=uGOLYz2pgr8>. [Қаралған күні: 15 Наурыз 2025].
- [47] Business Insider, Антропиктің бас директоры «3-бай ішінде жасанды интеллект бағдарламалық қамтамасыз етушілердің жауапкершілігіндегі кодтың 90%-ын жазатын болады» деп мәлімдеді, 15 наурыз 2025. [Интернетте]. Қолжетімді: <https://www.businessinsider.com/anthropic-ceo-ai-90-percent-code-3-to-6-months-2025-3>. [Қаралған күні: 30 наурыз 2025].
- [48] Statista, «2024 жылдың маусымындағы деректер базасын басқару жүйелерінің (DBMS) әлем бойынша танымалдылық салыстырмасы, категория бойынша,» маусым 2024. [Интернетте]. Қолжетімді: <https://www.statista.com/statistics/1131595/worldwide-popularity-database-management-systems-category/>. [Қаралған күні: 15 наурыз 2025].
- [49] DB-Engines, «DB-Engines Рейтинг,» [Интернетте]. Қолжетімді: <https://db-engines.com/en/ranking>. [Қаралған күні: 15 наурыз 2025].
- [50] «Stack Overflow Дамушы Саялнама 2023,» 2024. [Интернетте]. Қолжетімді:

- [https://survey.stackoverflow.co/2023/.](https://survey.stackoverflow.co/2023/)
- [51] «SQL,» 2024. [Интернетте]. Қолжетімді: <https://en.wikipedia.org/wiki/SQL>.
- [52] «Структурированные и неструктурированные данные: айырмашылық неде?,» 2024. [Интернетте]. Қолжетімді: <https://www.ibm.com/blog/structured-vs-unstructured-data/>.
- [53] DataDrivenConstruction, «ҚҰРЫЛЫС ЖОБАЛАРЫ ҮШІН ДЕРЕК ФОРМАТТАРЫН САЛЫСТАРУ PDF,» 23 сәуір 2024. [Интернетте]. Қолжетімді: <https://datadrivenconstruction.io/wp-content/uploads/2024/10/COMPARISON-OF-DATA-FORMATS-FOR-CONSTRUCTION-PROJECTS-1.pdf>.
- [54] «Building Information Modeling Ақпараттық құжат,» 2003. [Интернетте]. Қолжетімді: <https://web.archive.org/web/20030711125527/http://usa.autodesk.com/adsk/servlet/item?id=2255342&siteID=123112>.
- [55] А. Бойко, «Лоббистік соғыстар және BIM-нің дамуы. 5-бөлім: BlackRock – барлық технологиялардың шебері. Корпорациялар ашық кодты қалай басқарады,» 2024. [Интернетте]. Қолжетімді: <https://bigdataconstruction.com/autodesk-oracle-blackrock-open-source/>.
- [56] Д. Ушаков, «Тікелей Модельдеу - Кім және Неге Оған Қажет? Бәсекелес Технологияларға Шолу,» 14 қараша 2011. [Интернетте]. Қолжетімді: https://isicad.net/articles.php?article_num=14805. [Қарau күні: 02 2025].
- [57] Ч. Истман және А. Кетерс, «Истман, Чарльз; Және Кетерс,» қыркүйек 1974. [Интернетте]. Қолжетімді: <https://files.eric.ed.gov/fulltext/ED113833.pdf>. [Қарau күні: 15 наурыз 2025].
- [58] Д. Ушаков, «Тікелей Модельдеу - Кім және Неге Оған Қажет? Бәсекелес Технологияларға Шолу,» 11 қараша 2011. [Интернетте]. Қолжетімді: https://isicad.net/articles.php?article_num=14805. [Қарau күні: 15 наурыз 2025].
- [59] Д. Уайсберг, «CAD тарихы,» 12 желтоқсан 2022. [Интернетте]. Қолжетімді: https://www.shapr3d.com/blog/history-of-cad?utm_campaign=cadhistorynet. [Қарau күні: 15 наурыз 2025].
- [60] ADSK, «Ақпараттық құжат: Құрылыш Ақпараттық Моделдеу,» 2002. [Интернетте]. Қолжетімді: https://web.archive.org/web/20060512180953/http://images.adsk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand. [Қарau күні: 15 наурыз 2025].
- [61] ADSK, «Ақпараттық құжат: Құрылыш Ақпараттық Моделдеу Практикада,» [Интернетте]. Қолжетімді: https://web.archive.org/web/20060512181000/http://images.adsk.com/apac_sapac_main/files/4525077_BIM_in_Practice.pdf. [Қарau күні: 15 наурыз 2025].
- [62] А. Бойко, Лоббистік соғыстар және BIM дамуы. 2-бөлім: ашық BIM VS жабық BIM. Еуропа VS

қалған әлем, 2024. [Интернетте]. Қолжетімді: <https://bigdataconstruction.com/lobbyist-wars-and-the-development-of-bim-part-2-open-bim-vs-closed-bim-revit-vs-archicad-and-europe-vs-the-rest-of-the-world/>.

- [63] А. Бойко, «Құрылыс саласындағы деректер үшін лоббистік соғыстар | Техно-феодализм және BIM тарихы,» 2024. [Интернетте]. Қолжетімді: https://youtu.be/S-TNdUgfHxk?si=evM_v28KQbGOG0k&t=1360.
- [64] ADSK, «Ақ қағаз BIM,» 2002. [Интернетте]. Қолжетімді: https://web.archive.org/web/20060512180953/http:/images.autodesk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand. [Қараша күні: 2025 жылғы 15 наурыз].
- [65] ADSK, «Интеграция Дизайннан-Өндірісқа: Артықшылықтар мен Негіздеме,» [Интернетте]. Қолжетімді: https://web.archive.org/web/20010615093351/http://www3.adsk.com:80/adsk/files/734489_Benefits_of_MAI.pdf. [Қараша күні: 15 Наурыз 2025].
- [66] М. Шеклетт, «Структурланған және құрылымдалмаған деректер: Негізгі айырмашылықтар,» 2024. [Интернетте]. Қолжетімді: <https://www.datamation.com/big-data/structured-vs-unstructured-data/>.
- [67] К. Вулард, Осмысление неструктурированным деректердің өсуі, 2024. [Интернетте]. Қолжетімді: <https://automationhero.ai/blog/making-sense-of-the-rise-of-unstructured-data/>.
- [68] А. С. О. J. L. D. J. a. L. T. G. Майкл П. Галлахер, «Интероперабельділіктің жеткіліксіздігіне байланысты шығындарды талдау», 2004. [Интернетте]. Қолжетімді: <https://nvlpubs.nist.gov/nistpubs/gcr/2004/nist.gcr.04-867.pdf>. [Қараша күні: 02 2025].
- [69] CrowdFlower, «Деректер Ғылымы Есебі 2016», 2016. [Интернетте]. Қолжетімді: https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf. [Қол жеткізу күні: 15 Наурыз 2025].
- 70. Analyticsindiamag, «Деректер ғалымдарының ең көп уақыт жұмсайтын 6 тапсырмасы», 2019 жылғы 15 мамыр. [Интернетте]. Қолжетімді: <https://analyticsindiamag.com/ai-trends/6-tasks-data-scientists-spend-the-most-time-doing/>.
- [71] BizReport, «Есеп: Деректер ғалымдары уақыттың басым бөлігін тазалауға жұмсайды,» 06 Шілде 2015. [Интернетте]. Қолжетімді: <https://web.archive.org/web/20200824174530/http://www.bizreport.com/2015/07/report-data-scientists-spend-bulk-of-time-cleaning-up.html>. [Қараша күні: 5 Наурыз 2025].
- [72] S. Хокинг, «Ғылым AMA сериясы: Стивен Хокингтің AMA жауаптары!», 27 Шілде 2015. [Интернетте]. Қолжетімді: https://www.reddit.com/r/science/comments/3nyn5i/science_ama_series_stephen_hawking_ama_answers/. [Қараша күні: 15 Наурыз 2025].

- [73] Б. Сайферс және К. Доктороу, Конфиденциалдық монополиясыз: Деректерді қорғау және өзара үйлесімділік, 2024. [Интернетте]. Қолжетімді: <https://www.eff.org/wp/interoperability-and-privacy>.
- [74] McKinsey Глобал Институты, Ашық деректер: Сұйық ақпаратпен инновация мен өнімділікті ашу, 1 Қазан 2013. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>. [Қараяу күні: 15 Наурыз 2025].
75. А. Бойко, «Құрылым саласындағы ашық деректер үшін күрес. AUTOLISP, intelliCAD, openDWG, ODA және openCASCADE тарихы», 15 05 2024. [Интернетте]. Қолжетімді: <https://boikoartem.medium.com/the-struggle-for-open-data-in-the-construction-industry-2b97200e6393>. [Қаралған күні: 16 02 2025].
- [76] Wikipedia, «Microsoft және ашық код», [Интернетте]. Қолжетімді: https://en.wikipedia.org/wiki/Microsoft_and_open_source. [Қараяу күні: 15 Наурыз 2025].
- [77] TIME, «Ашық және жабық ИИ модельдері арасындағы алшақтық қысқаруы мүмкін. Бұл неге маңызды?» 2024 жылғы 5 қараша. [Интернетте]. Қолжетімді: <https://time.com/7171962/open-closed-ai-models-epoch/>. [Қараяу күні: 2025 жылғы 15 наурыз].
- [78] The Verge, «Google-дың жаңа кодтарының төрттен бірінен астамы ЖИ арқылы жасалады», 2024 жылғы 29 қазан. [Интернетте]. Қолжетімді: <https://www.theverge.com/2024/10/29/24282757/google-new-code-generated-ai-q3-2024>. [Қараяу күні: 2025 жылғы 15 наурыз].
- [79] McKinsey Digital, «GPU-ларды аналитикалық өндөуді жеделдету үшін пайдалану бизнес жағдайы», 15 Желтоқсан 2020. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/the-business-case-for-using-gpus-to-accelerate-analytics-processing>. [Қаралған күні: 15 Наурыз 2025].
- [80] PwC, PwC Ашық Код Мониторы 2019, 2019. [Интернетте]. Қолжетімді: <https://www.pwc.de/de/digitale-transformation/open-source-monitor-research-report-2019.pdf>. [Қараяу күні: 15 Наурыз 2025].
- [81] Travers Smith, Ашық құпия: Ашық кодты бағдарламалық қамтамасыз ету, 2024. [Интернетте]. Қолжетімді: <https://www.traverssmith.com/knowledge/knowledge-container/the-open-secret-open-source-software/>. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [82] Deloitte, «Корпоративтік трансформациялардағы деректерді беру процесі», 2021. [Интернетте]. Қолжетімді: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/finance/us-the-data-transfer-process-in-corporate-transformations.pdf>. [Қараяу күні: 2025 жылғы 15 наурыз].
- [83] gov.uk, «Мемлекеттік жобаларды жүзеге асыруда деректер аналитикасы және ЖИ» 2024 жылғы 20 наурыз. [Интернетте]. Қолжетімді: <https://www.gov.uk/government/publications/data>

analytics-and-ai-in-government-project-delivery/data-analytics-and-ai-in-government-project-delivery. [Қаралған күні: 2025 жылғы 5 наурыз].

- [84] Барлық нәрселер мүмкіндігінше қарапайым болуы тиіс, бірақ одан да қарапайым емес. 13 Мамыр 2011. [Интернетте]. Қолжетімді: <https://quoteinvestigator.com/2011/05/13/einstein-simple/>. [Қаралған күні: 15 Наурыз 2025].
- 85. Transformer (терең оқыту архитектурасы), [Интернетте]. Қолжетімді: [https://en.wikipedia.org/wiki/Transformer_\(deep_learning_architecture\)](https://en.wikipedia.org/wiki/Transformer_(deep_learning_architecture)). [Қол жеткізу күні: 15 Наурыз 2025].
- [86] Python пакеттерінің жүктеу статистикасы, 2024. [Интернетте]. Қолжетімді: <https://www.pypy.tech/projects/pandas>.
- [87] Interview Bit, «Топ 10 Python кітапханасы», 2023. [Интернетте]. Қолжетімді: <https://www.interviewbit.com/blog/python-libraries/#:~:text=With%20more%20than%20137%C000%20libraries,data%20manipulation%2C%20and%20many%20more>. [Қол жеткізу күні: 30 Наурыз 2025].
- [88] NVIDIA және HP деректер ғылыми мен генеративті ИИ-ді жұмыс станцияларында қуаттандырады, 7 наурыз 2025. [Интернетте]. Қолжетімді: <https://nvidianews.nvidia.com/news/nvidia-hp-supercharge-data-science-generative-ai-workstations>. [Қаралған күні: 15 наурыз 2025].
- [89] Р. Орак, Как миллион жолдан тұратын DataFrame-ді секундтар ішінде өндеду, 2024. [Интернетте]. Қолжетімді: <https://towardsdatascience.com/how-to-process-a-dataframe-with-millions-of-rows-in-seconds>.
- [90] Ç. Uslu, «Kaggle дегеніміз не?», 2024. [Интернетте]. Қолжетімді: <https://www.datacamp.com/blog/what-is-kaggle>.
- [91] «NVIDIA CEO Джленсен Хуанг COMPUTEX 2024-те жасаған баяндамасы», 2 Маусым 2024. [Интернетте]. Қолжетімді: <https://www.youtube.com/live/pKXDVsWZmUU?si=Z3Rj1Las8wiPII2w>. [Қаралған күні: 15 Наурыз 2025].

«Мүшелер: Негізін қалаушылар мен корпоративтік мүшелер», 2024. [Интернетте]. Қолжетімді: <https://www.opendesign.com/member-showcase>.

А. Бойко, «Өзгерістер дәүірі: IFC өткеннің қалдығы немесе неге ADSK және басқа CAD жеткізушілері IFC-ден USD-ге көшуге дайын?», 24 Қараша 2024. [Интернетте]. Қолжетімді: <https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-are-willing-to-3f9a82cccd10a>. [Қаралған күні: 23 Ақпан 2025].

А. Бойко, «BIM-нен кейінгі әлем. Деректер мен процестерге көшу және құрылыш

индустриясына семантика, форматтар мен өзара әрекеттестік қажет пе?», 20 Желтоқсан 2024. [Интернетте]. Қолжетімді: <https://boikoartem.medium.com/the-post-bim-world-7e35b7271119>. [Қаралу күні: 23 Ақпан 2025].

NIH, «NIH Деректер Ғылымы Стратегиялық Жоспары», 2016. [Интернетте]. Қолжетімді: https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf. [Қаралу күні: 23 Ақпан 2025].

Harvard Business Review, «Нашар деректер АҚШ-қа жылына 3 триллион доллар шығын келтіреді», 22 Қыркүйек 2016. [Интернетте]. Қолжетімді: <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year>.

Delpha, «Деректер сапасының әсерлері», 1 Қаңтар 2025. [Интернетте]. Қолжетімді: <https://delpha.io/impacts-of-data-quality/>.

W. B. D. Guide, «Техникалық қызмет көрсету үшін жобалау: Құрылымы жобаларының жобалау кезеңінде операциялар мен техникалық қызмет көрсету мәселелерінің маңызы», [Интернетте]. Қолжетімді: <https://www.wbdg.org/resources/design-for-maintainability>. [Қаралу күні: 15 Наурыз 2025].

Қорғаныс министрлігі, «Әскери жүйелер мен жабдықтар үшін коррозияны алдын алу және бақылау жөніндегі жоспарлау нұсқаулығы», Сәуір 2014. [Интернетте]. Қолжетімді: <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/CPC%20Planning%20Guidebook%204%20Feb%202014.pdf>. [Қаралу күні: 15 Наурыз 2025].

Gartner, «Деректер сапасы: Дәл ақпарат алу үшін үздік тәжірибелер», 1 Қаңтар 2025. [Интернетте]. Қолжетімді: <https://www.gartner.com/en/data-analytics/topics/data-quality>.

«Тырнақтың жетіспеушілігі», [Интернетте]. Қолжетімді: https://en.wikipedia.org/wiki/For_Want_of_a_Nail. [Қаралу күні: 15 Наурыз 2025].

PWC, «ESG туралы хабардарлық», 2024 жылғы 1 шілде. [Интернетте]. Қолжетімді: <https://www.pwc.com/kz/en/assets/esg-awareness/kz-esg-awareness-rus.pdf>. [Қаралу күні: 2025 жылғы 15 наурыз].

- [105] G. Хэммонд, «Дене көміртегі - Көміртегі мен энергияның инвентаризациясы (ICE)», 2024. [Интернетте]. Қолжетімді: <https://greenbuildingencyclopaedia.uk/wp-content/uploads/2014/07/Full-BSRIA-ICE-guide.pdf>.
- [106] CO2_дene carbon-ды есептеу, 2024. [Интернетте]. Қолжетімді: https://github.com/datadrivenconstruction/CO2_calculating-the-embodied-carbon.
- [107] McKinsey, Құрылымы саласының цифрлық болашағын елестету, 24 маусым 2016. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/operations/our-insights/imaging>

- constructions-digital-future. [Қаралу күні: 25 ақпан 2025].
- [108] Bund der Steuerzahler Deutschland e.V., «Қара кітап», 2024 жылғы 10 қазан. [Интернетте]. Қолжетімді: <https://steuerzahler.de/aktuelles/detail/das-schwarzbuch-202425/>. [Қаралу күні: 2025 жылғы 15 наурыз].
- [109] SAS, «Деректер көлі мен деректер қоймасы – айырмашылықты білініз,» [Интернетте]. Қолжетімді: https://www.sas.com/en_is/insights/articles/data-management/data-lake-and-data-warehouse-know-the-difference.html. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [110] ADSK, «Құрылым Ақпарат Моделдеу», 2002. [Интернетте]. Қолжетімді: https://www.laiserin.com/features/bim/autodesk_bim.pdf. [Қаралу күні: 15 Наурыз 2025].
- [111] А. Бойко, «BIM тарихының картасы,» 2024. [Интернетте]. Қолжетімді: <https://bigdataconstruction.com/history-of-bim/>.
- [112] А. С. Борковский, «Ұйымдар мен стандарттар бойынша BIM анықтамалары,» 2023 жылғы 27 желтоқсан. [Интернетте]. Қолжетімді: <https://encyclopedia.pub/entry/53149>. [Қаралу күні: 2025 жылғы 5 наурыз].
- [113] CAD жеткізушісі, OPEN BIM бағдарламасы, 2012. [Интернетте]. Қолжетімді: https://web.archive.org/web/20140611075601/http://www.graphisoft.com/archicad/open_bim/. [Қаралу күні: 30 Наурыз 2025].
- [114] Wikipedia, «Сала Негіздері Класстары», [Интернетте]. Қолжетімді: https://en.wikipedia.org/wiki/Industry_Foundation_Classes. [Қаралу күні: 2025 жылғы 15 наурыз].
- [115] Wikipedia, «IGES,» [Интернетте]. Қолжетімді: <https://en.wikipedia.org/wiki/IGES>. [Қаралу күні: 30 Наурыз 2025].
- [116] А. Бойко, «CAD (BIM) тарихы,» 15 Желтоқсан 2021. [Интернетте]. Қолжетімді: https://miro.com/app/board/o9J_laML2cs=/. [Қаралу күні: 24 Ақпан 2025].
- [117] T. K. K. A. O. F. B. C. E. L. H. H. E. L. P. N. S. H. T. J. v. L. H. G. D. H. T. K. C. L. A. W. J. S. Франческа Ноардо, «IFC бағдарламалық қамтамасыз етуінің сілтеме зерттеуі: GeoBIM бенчмаркті 2019 – I бөлім», 8 Қаңтар 2021. [Интернетте]. Қолжетімді: <https://arxiv.org/pdf/2007.10951.pdf>. [Қаралу күні: 5 Наурыз 2025].
- [118] И. Рогачёв, «BIM туралы сөйлесейік: Максим Нечипоренко | Renga | IFC | Отандық BIM,» 2021 жылғы 13 сәуір. [Интернетте]. Қолжетімді: <https://www.youtube.com/watch?t=3000&v=VO3Y9uzF9M&feature=youtu.be>. [Қаралу күні: 2025 жылғы 5 наурыз].
- [119] D. Ares, «RETS жылжымайтын мүлікте: Неліктен ол тиімділік пен өсу үшін маңызды», 17 Желтоқсан 2024. [Интернетте]. Қолжетімді: <https://www.realpha.com/blog/rets-importance-in-real-estate-explained>. [Қаралу күні: 5 Наурыз 2025].

- [120] Flex токенінің құны, 2024. [Интернетте]. Қолжетімді: <https://www.adsk.com/buying/flex?term=1-YEAR&tab=flex>.
- [121] А. Бойко, «BIM-ді ұмытып, деректерге қолжетімділікті демократияландыру (17. Инвестор – Жоғары оқу орны – Құрылыш индустриясы),» 2024. [Интернетте]. Қолжетімді: <https://www.bim.bayern.de/wp-content/uploads/2023/06/Kolloquium-17-TUM-Bauprozessmanagement-und-Bay-Bauindustrie.pdf>.
- [122] Д. Хилл, Д. Фолдези, С. Феррер, М. Фридман, Э. Лох және Ф. Плашке, «Құрылыш саласының өнімділік жұмбабын шешу,» 2015. [Интернетте]. Қолжетімді: <https://www.bcg.com/publications/2015/engineered-products-project-business-solving-construction-industries-productivity-puzzle>.
- [123] SCOPE – Жобалық деректер ортасы және көпфункционалды құрылыш өнімдерін модельдеу, ғимараттың сыртқы қабығына назар аудара отырып, 1 Қаңтар 2018. [Интернетте]. Қолжетімді: <https://www.ise.fraunhofer.de/de/forschungsprojekte/scope.html>. [Қаралу күні: 2 Наурыз 2025].
- [124] Apple.com, «Pixar, Adobe, Apple және NVIDIA OpenUSD ашық стандарттарын дамыту үшін альянс құрды», 1 Тамыз 2023. [Интернетте]. Қолжетімді: <https://www.apple.com/newsroom/2023/08/pixar-adobe-apple-adsk-and-nvidia-form-alliance-for-openusd/>. [Қаралу күні: 2 Наурыз 2025].
- [125] AECmag, «ADSK-ның ұсақ деректер стратегиясы», 2024 жылғы 25 шілде. [Интернетте]. Қолжетімді: <https://aecmag.com/technology/autodesks-granular-data-strategy/>. [Қаралу күні: 2025 жылғы 15 наурыз].
- [126] А. Бойко, «Өзгерістер дәүірі: IFC өткеннің еншісінде немесе неге ADSK және басқа CAD жеткізушілері IFC-ны USD-ға ауыстыруға дайын?» 24 қараша 2024. [Интернетте]. Қолжетімді: <https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-3f9a82ccd10a>. [Қаралу күні: 2025 жылдың 23 ақпаны].
- [127] А. Бойко, ENG BIM Кластері 2024 | Деректер үшін курес және Құрылыш саласындағы LLM мен ChatGPT қолдану, 7 Тамыз 2024. [Интернетте]. Қолжетімді: ENG BIM Кластері 2024 | Деректер үшін курес және Құрылыш саласындағы LLM мен ChatGPT қолдану. [Қаралу күні: 15 наурыз 2025].
- [128] Джейфри Зелдман ұсынған, 6 Мамыр 2008. [Интернетте]. Қолжетімді: <https://zeldman.com/2008/05/06/content-precedes-design/>. [Қол жеткізу күні: 15 Наурыз 2025].
- [129] А. Бойко, «DWG талдау ChatGPT-пен | Деректерге негізделген құрылыш,» 2024 жылғы 5 наурыз. [Интернетте]. Қолжетімді: <https://www.kaggle.com/code/artemboiko/dwg-analyse-with-chatgpt-datadrivenconstruction>. [Қаралу күні: 2025 жылғы 15 наурыз].
- [130] McKinsey, «Цифрлық және Жасанды Интеллект дәүірінде бәсекелестікте алға шығу жөніндегі McKinsey нұсқаулығы», 2023. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/featured>

- insights/mckinsey-on-books/rewired. [Қаралған күні: 30 Наурыз 2025].
- [131] Forbes, «Деректерді баяндау: Әркімге қажетті деректер ғылымының негізгі дағдысы», 31 наурыз 2016. [Интернетте]. Қолжетімді: <https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/>. [Қаралған күні: 15 наурыз 2025].
- [132] J. Бертина, «Графика және графикалық ақпаратты өндіреу», 2011 жылғы 8 қыркүйек. [Интернетте]. Қолжетімді: https://books.google.de/books/about/Graphics_and_Graphic_Information_Process.html?id=csqXxnm4tcC&redir_esc=y. [Қаралған күні: 2025 жылғы 15 наурыз].
- [133] CauseWeb, «Wells/Wilks статистикалық ойлау туралы», [Интернетте]. Қолжетімді: <https://www.causeweb.org/cause/resources/library/r1266>. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [134] Ministrymagazine, «Ғылым қалай Жаратылысты ашты,» 1986 жылғы қаңтар. [Интернетте]. Қолжетімді: <https://www.ministrymagazine.org/archive/1986/01/how-science-discovered-creation>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [135] BCG, «Деректерге негізделген трансформация: Қазір ауқымды түрде жеделдетіңіз», 23 Мамыр 2017. [Интернетте]. Қолжетімді: <https://www.bcg.com/publications/2017/digital-transformation-transformation-data-driven-transformation>. [Қаралған күні: 15 Мамыр 2024].
- [136] Деректер архитектурасын инновацияны дамыту үшін қалай құруға болады—бүгін және ертең, 3 маусым 2020. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-to-build-a-data-architecture-to-drive-innovation-today-and-tomorrow>. [Қаралған күні: 15 наурыз 2025].
- [137] Оксфорд, «Вудро Вильсон 1856–1924,» [Интернетте]. Қолжетімді: <https://www.oxfordreference.com/display/10.1093/acref/9780191866692.001.0001/q-oro-ed6-00011630>. [Қаралған күні: 2025 жылғы 15 наурыз].
- [138] Конвертерлер, 2024. [Интернетте]. Қолжетімді: <https://datadrivenconstruction.io/index.php/convertors/>.
- [139] PWC, «Сыйлықты өлшеу: Жасанды интеллекттің бизнесінде үшін шынайы құны неде және қалай пайда табуға болады?», 1 Қаңтар 2017. [Интернетте]. Қолжетімді: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>. [Қаралған күні: 18 Ақпан 2025].
- [140] Трубопровод құрылыш саласында, 2024. [Интернетте]. Қолжетімді: <https://datadrivenconstruction.io/index.php/pipeline-in-construction/>.
- [141] Wikipedia, «Apache NiFi,» 2025 жылғы 1 қаңтар. [Интернетте]. Қолжетімді: https://de.wikipedia.org/wiki/Apache_NiFi. [Қол жеткізу күні: 2025 жылғы 5 наурыз].

- [142] n8n, Gmail AI Авто-жауап беруші: Келіп түсетін электрондық хаттарға арналған жобалық жауаптарды жасау, 1 Мамыр 2024. [Интернетте]. Қолжетімді: <https://n8n.io/workflows/2271-gmail-ai-auto-responder-create-draft-replies-to-incoming-emails/>. [Қараяу күні: 15 Наурыз 2025].
- [143] n8n, «Zillow API, Google Sheets және Gmail арқылы жылжымайтын мүлік күнделікті мәмілелерді автоматтандыру», 2025 жылғы 1 наурыз. [Интернетте]. Қолжетімді: <https://n8n.io/workflows/3030-real-estate-daily-deals-automation-with-zillow-api-google-sheets-and-gmail/>. [Қол жеткізу күні: 2025 жылғы 15 наурыз].
- [144] В. Т. O'Neill, «Аналитика, ЖИ және үлкен деректер жобаларының сәтсіздік деңгейі = 85% – уайымдайтын жайт!», 1 Қаңтар 2025. [Интернетте]. Қолжетімді: <https://designingforanalytics.com/resources/failure-rates-for-analytics-bi-iot-and-big-data-projects-85-yikes/>.
- [145] Ж. Нейман, Өкілдік әдісінің екі түрлі аспектісі туралы: Стратификацияланған іріктеу әдісі және Мақсатты іріктеу әдісі, Оксфорд университеті баспасы, 1934.
- [146] Т. Ж. С. а. Ж. С. Джесси Перла, «Милтон Фридманды таң қалдырыған мәселе», Жулиямен сандық экономика, 2025 жылдың 1 қаңтары. [Интернетте]. Қолжетімді: https://julia.quantecon.org/dynamic_programming/wald_friedman.html. [Қараяу күні: 2024 жылдың 1 мамыры].
- [147] Т. Ландсалл-Велфэр, Үлттық көніл-күйді болжай қазіргі уақытта, Маңыздылығы, 2012.
- [148] А. Бойко, «Сан-Франциско. Құрылыш секторы 1980-2019,» 2024. [Интернетте]. Қолжетімді: <https://www.kaggle.com/search?q=San+Francisco.+Building+sector+1980-2019>.
- [149] А. Бойко, «Kaggle: RVT IFC Файлдары 5000 Жоба,» 2024. [Интернетте]. Қолжетімді: <https://www.kaggle.com/datasets/artemboiko/rvtifc-projects>.
- [150] CFMA, «Байланысқан құрылыш арқылы болашаққа дайындалу,» [Интернетте]. Қолжетімді: <https://cfma.org/articles/preparing-for-the-future-with-connected-construction>. [Қаралған күні: 15 Наурыз 2025].
- [151] Cisco, «Cisco сауалнамасы IoT жобаларының үштен екісінің сәтсіздікке үшірағанын көрсетеді,» 22 Мамыр 2017. [Интернетте]. Қолжетімді: <https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2017/m05/cisco-survey-reveals-close-to-three-fourths-of-iot-projects-are-failing.html>.
- [152] «Зауыттың қазба қалдықтарын сақтау үшін қажетті жағдайлар,» 2024. [Интернетте]. Қолжетімді: <https://ucmp.berkeley.edu/IB181/VPL/Pres/PresTitle.html>.
- [153] «BlackRock-тың облигациялар, бірігу және сатып алу, АҚШ-тың рецессиясы, сайлаулар туралы Финкпен толық сұхбат,» 2023. [Интернетте]. Қолжетімді: <https://www.bloomberg.com/news/videos/2023-09-29/blackrock-s-fink-on-m-a-recession-election->

- full-intv-video.
- [154] cio, «12 танымал AI апаттары,» 02 Қазан 2024. [Интернетте]. Қолжетімді: <https://www.cio.com/article/190888/5-famous-analytics-and-ai-disasters.html>. [Қаралған күні: 15 Наурыз 2025].
- [155] G. Kasparov, Deep Thinking, PublicAffairs, 2017.
- [156] Wikipedia, «Kaggle,» 1 Қаңтар 2025. [Интернетте]. Қолжетімді: <https://en.wikipedia.org/wiki/Kaggle>. [Қаралған күні: 15 Наурыз 2025].
- [157] Kaggle, «Титаник - Апаттан машиналық оқыту,» 1 Қаңтар 2025. [Интернетте]. Қолжетімді: <https://www.kaggle.com/competitions/titanic/overview>. [Қаралған күні: 10 Наурыз 2025].
- [158] Ш. Йохри, «ChatGPT-ны жасау: Деректерден диалогқа,» 2024. [Интернетте]. Қолжетімді: <https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/>.
- [159] П. Домингос, «Машиналық оқыту туралы білуге тиіс бірнеше пайдалы нәрселер,» 2024. [Интернетте]. Қолжетімді: <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>.
- [160] J. Сарамаго, «Цитата,» [Интернетте]. Қолжетімді: <https://www.goodreads.com/quotes/215253-chaos-is-merely-order-waiting-to-be-deciphered>. [Қаралған күні: 17 Наурыз 2025].
- [161] NVIDIA, «Жаңа NVIDIA NeMo Curator Classifier модельдерімен оқу деректерінізді жақсартыңыз,» 19 Желтоқсан 2024. [Интернетте]. Қолжетімді: <https://developer.nvidia.com/blog/enhance-your-training-data-with-new-nvidia-nemo-curator-classifier-models/>. [Қаралған күні: 25 Наурыз 2025].
- [162] «NVIDIA Cosmos World Foundation модельдері мен физикалық AI деректер құралдарының маңызды шығарылымын жариялады,» 18 Наурыз 2025. [Интернетте]. Қолжетімді: <https://nvidianews.nvidia.com/news/nvidia-announces-major-release-of-cosmos-world-foundation-models-and-physical-ai-data-tools>. [Қаралған күні: 25 Наурыз 2025].
- [163] NVIDIA, «NVIDIA Isaac Sim,» [Интернетте]. Қолжетімді: <https://developer.nvidia.com/isaac/sim>. [Қаралған күні: 25 Наурыз 2025].
- [164] М. Төрт айда бір рет, «Цифрлық стратегиялардың сәтсіздікке ұшырау себептері», 25 Қаңтар 2018. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-digital-strategies-fail>. [Қаралған күні: 15 Наурыз 2025].
- [165] М. Дж. Перри, Менің сүйікті Милтон Фридман дәйексөздерім, 17 қараша 2006. [Интернетте]. Қолжетімді: <https://www.aei.org/carpe-diem/my-favorite-milton-friedman-quotes/>. [Қаралған күні: 1 наурыз 2025].
- [166] Д. А. Уилер, «Ақпарат, физика, квант: байланыстарды іздеу,» 1990.

- [169] А. Бойко, «Лоббистік соғыстар және BIM-нің дамуы. 5-бөлім: BlackRock – барлық технологиялардың шебері. Корпорациялар ашық бастапқы кодты қалай бақылауда үстайды», 2024. [Интернетте]. Қолжетімді: <https://boikoartem.medium.com/lobbyist-wars-and-the-development-of-bim-d72ad0111a7d>.
- [170] Т. Крайнин және Ж. Беетц, «Бинарлық форматтағы IFC құрылымы модельдері үшін SPARQL сұраныс қозғалтқышы», Advanced Engineering Informatics, 2024.
- [171] 2021 жылды Ұлыбританиядағы құрылымы секторындағы кәсіпорындардың саны, кәсіпорынның көлемі бойынша, 2024. [Интернетте]. Қолжетімді: <https://www.statista.com/statistics/677151/uk-construction-businesses-by-size/>.
- [172] 5000 IFC&RVT жобасы, 2024. [Интернетте]. Қолжетімді: <https://www.kaggle.com/code/artemboiko/5000-projects-ifc-rvt-datadrivenconstruction-io>.
- [173] М. Попова, «Bit-тен: Ақпарат, шындықтың табиғаты және неге біз қатысушы ғаламда өмір сүреміз туралы әйгілі физик Джон Арчибалд Уилер» 2008. [Интернетте]. Қолжетімді: <https://www.themarginalian.org/2016/09/02/it-from-bit-wheeler/>. [Қараша күні: 2025 жылдың ақпаны].
- [174] Лоббистік соғыстар құрылымы саласындағы деректер үшін / Техно-феодализм және BIM-нің жасырын тарихы. [Фильм]. Германия: Артём Бойко, 2023.
- [175] А. Бойко, «CHATGPT REVIT ЖӘНЕ IFC | Жобалардан құжаттар мен деректерді автоматты түрде алу», 16 Қараша 2023. [Интернетте]. Қолжетімді: https://www.youtube.com/watch?v=ASXolti_YPs&t. [Қараша күні: 2 Наурыз 2025].
- [176] М. & Компания, «Цифрлық трансформацияның толық құнын Capturing үшін үш жаңа мандат», 22 Қаңтар 2022. [Интернетте]. Қолжетімді: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/three-new-mandates-for-capturing-a-digital-transformations-full-value>. [Қараша күні: 15 Ақпан 2025].
- [177] KPMG, «Цифрлық әлемдегі құрылымы», 2021 жылғы 1 мамыр. [Интернетте]. Қолжетімді: <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [Қараша күні: 2024 жылғы 5 сәуір].
- [178] LLP, KPMG, Cue Construction 4.0: Маңызды кезең, 17 Наурыз 2023. [Интернетте]. Қолжетімді: <https://kpmg.com/ca/en/home/insights/2023/05/cue-construction-make-or-break-time.html>. [Қараша күні: 15 Ақпан 2025].
- [179] O. Business, Сатья Наделла «AI агенттерінің SaaS модельдерін қалай бұзатынын жариялады», 2025 жылдың 10 қаңтары. [Интернетте]. Қолжетімді: <https://www.outlookbusiness.com/artificial-intelligence/microsoft-ceo-satya-nadella-reveals-how-ai-agents-will-disrupt-saas-models>. [Қаралған күні: 2025 жылдың 15 наурызы].

- [180] Forbes, «Улкен деректерді тазалау: Деректер ғылымындағы ең көп уақытты алатын, ең жағымсыз тапсырма, сауланама нәтижелері», 23 наурыз 2016. [Интернетте]. Қолжетімді: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>. [Қаралған күні: 15 наурыз 2025].
- [181] Сыртқы істер министрлігі, Коммонвелт және даму істері, Ұлыбритания, «Цифрлық даму стратегиясы 2024-2030», 18 наурыз 2024. [Интернетте]. Қолжетімді: <https://www.gov.uk/government/publications/digital-development-strategy-2024-to-2030/digital-development-strategy-2024-to-2030>. [Қаралған күні: 15 наурыз 2025].
- [182] Кеңістік және стратегия құрылышы жобалау индустриясында, 7 қараша 2003. [Интернетте]. Қолжетімді: <https://web.archive.org/web/20030711125527/http://usa.adsk.com/adsk/servlet/item?id=2255342&siteID=123112>. [Қаралған күні: 5 наурыз 2025].
- [183] М. Бочаров, «Ақпараттық модельдеу,» Наурыз 2025. [Интернетте]. Қолжетімді: <https://www.litres.ru/book/mihail-evgenievich-bocharov/informacionnoe-modelirovanie-v-rossii-71780080/chitat-onlayn/?page=5>. [Қаралған күні: 15 Наурыз 2025].
- [184] Интеграция Дизайн-Өндіріс: Артықшылықтар мен Негіздеме, 2000. [Интернетте]. Қолжетімді: https://web.archive.org/web/20010615093351/http://www3.autodesk.com:80/adsk/files/734489_Benefits_of_MAI.pdf. [Қаралған күні: 2025 жылғы 25 наурыз].
- [185] CAD жеткізушісі, Open BIM бағдарламасы маркетингтік науқан, 12 наурыз 2012. [Интернетте]. Қолжетімді: <https://web.archive.org/web/20120827193840/http://www.graphisoft.com/openbim/>. [Қаралған күні: 30 наурыз 2025].

ТАҚЫРЫПТЫҚ ІЗДЕУ

3D, 8, 14, 71, 73, 84, 191, 210, 215, 232, 234, 263, 276, 277, 287, 298, 299, 302, 303, 306, 307, 337, 375, 393, 394, 448, 468, 480

4D, 84, 172, 196, 199, 210, 229, 234, 237, 287

4-өнеркәсіптік революция, 43

5D, 84, 172, 196, 210, 237, 287, 492

6D, 172, 196, 229, 232, 233, 234, 235

7D, 232, 233, 234, 287

8D, 172, 196, 229, 232, 233, 234, 287

A

Жасанды интеллект, 3, 50, 52, 100, 102, 103, 106, 107, 116, 395, 457, 459, 461, 477

AIA, 289

AIM, 3, 289

AMS, 14, 84, 151, 153

Apache Airflow, 311, 361, 362, 363, 364, 366, 367, 369, 371, 399, 494

Apache NiFi, 116, 311, 361, 367, 368, 371, 399, 482, 494

Apache ORC, 62, 330, 378

Apache Parquet, 62, 67, 380, 381, 494

API, 54, 94, 95, 97, 109, 137, 138, 154, 168, 218, 219, 256, 257, 260, 271, 273, 294, 295, 296, 297, 300, 308, 326, 329, 342, 366, 369, 370, 488

B

БДС, 258, 259

Үлкен деректер, 9, 67, 245, 383

БИМ, 2, 3, 4, 6, 3, 14, 17, 24, 56, 58, 60, 63, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 84, 92, 137, 138, 139, 140, 141, 142, 144, 150, 154, 156, 166, 172, 183, 184, 186, 187, 190, 195, 196, 208, 210, 211, 213, 216, 217, 218, 219, 220, 221, 222, 227, 228, 237, 238, 239, 242, 243, 246, 250, 251, 252, 255, 256, 257, 258, 259, 260, 261, 262, 263, 266, 271, 272, 273, 275, 276, 277, 278, 279, 280, 282, 285, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 308, 309, 324, 328, 337, 351, 356, 358, 361, 367, 372, 376, 383, 389, 402, 403, 405, 406, 413, 416, 431, 448, 462, 466, 475, 476, 480, 492, 494, 497

475, 476, 492, 494, 497.

BlackBox, 240, 242, 243

BMS, 8

Бокех, 320, 337

БОМ, 76, 77, 79, 257, 263

Шекаралық жәшік, 234, 373, 392, 393, 394, 414, 415, 480

БРЭП, 142, 234, 263, 264, 276, 283, 284

C

CAD, 6, 14, 18, 24, 56, 57, 58, 63, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 84, 85, 95, 97, 111, 126, 137, 138, 139, 140, 141, 142, 144, 146, 147, 152, 153, 155, 156, 166, 172, 175, 183, 184, 186, 187, 190, 195, 196, 206, 208, 210, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 224, 227, 228, 232, 234, 237, 238, 239, 242, 243, 251, 252, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 308, 309, 324, 328, 337, 344, 351, 356, 358, 361, 367, 372, 376, 383, 389, 402, 403, 405, 406, 413, 416, 431, 448, 462, 466, 475, 476, 480, 492, 494, 497

CAE, 16, 283

CAFM, 14, 24, 62, 84, 151, 153, 172, 233, 278, 326, 356, 387, 458, 473

Кампания, 16, 78

Капиталдық шығындар, 14, 82

CDE, 84, 175, 388, 389, 390

ЧатГПТ, 103, 104, 109, 110, 124, 129, 221, 303, 307, 343, 349, 421

Клод, 103, 104, 109, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 429, 443

CO₂, 72

CO₂, 229, 234, 235, 236, 237, 238, 239, 243

Кобие, 156, 289, 292

CoE, 56, 168, 169, 170, 477

Копилот, 114, 116, 459

CPIXML, 143, 272, 273, 276, 277, 279, 285, 296

CPM, 14, 17, 62, 166, 175, 233, 326, 473

CQMS, 14, 84, 177, 178, 462

CRM, 109, 369, 459

CRUD, 51, 459

CSG, 263

CSV, 61, 62, 88, 89, 120, 123, 128, 129, 130, 131, 135, 168,

268, 272, 280, 329, 333, 344, 354, 356, 373, 377, 378,

379, 380, 401, 407, 414, 474, 480

D

DAE, 276, 277, 278, 280, 281, 284, 285, 414
 DAG, 362, 363, 365, 366
 Dash, 320, 336, 337
 Деректерді басқару, 373, 395, 396, 398, 400, 401
 Деректер көлі, 214, 373, 376, 384, 385, 386, 387, 388, 389, 390, 400
 Деректер көлтабан, 67, 373, 386, 387, 388
 Деректер минимализмі, 373, 395, 396, 400, 401
 Деректер батпағы, 373, 395, 397, 401
 Деректер қоймасы, 382, 383, 400
 Деректер қызметі, 487
 деректерге негізделген, 50, 170, 460, 461, 484, 486
 DataFrame, 67, 117, 121, 122, 123, 125, 129, 130, 131, 133, 134, 135, 137, 219, 220, 224, 237, 308, 328, 329, 330, 332, 333, 344, 345, 347, 348, 354, 365, 377, 381, 409, 414, 415, 475
 DataOps, 170, 373, 398, 399, 400, 401, 475
 DeepSeek, 103, 104, 107, 109, 110, 124, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 421, 429, 443, 494
 DGN, 8, 140, 186, 227, 357
 DWG, 8, 70, 71, 73, 97, 140, 186, 211, 227, 272, 287, 302, 303, 304, 307, 357, 376, 497
 DWH, 67, 373, 376, 382, 383, 384, 386, 387, 388, 389, 390
 DXF, 8, 73, 277

E

ECM, 58, 175
 ECS, 142
 EIR, 289
 eLOD, 289
 ELT, 384, 385
 EPM, 14, 166, 198
 ERP, 2, 11, 12, 14, 17, 18, 24, 25, 58, 62, 109, 153, 166, 172, 175, 196, 198, 210, 232, 239, 240, 241, 242, 243, 244, 245, 246, 247, 249, 272, 277, 278, 279, 282, 326, 351, 356, 361, 369, 387, 388, 389, 390, 458, 462, 473, 475, 487, 492
 ESG, 196, 235, 236, 238
 ETL, 6, 1.1-8, 32, 49, 81, 113, 116, 119, 128, 188, 193, 219, 291, 311, 312, 317, 323, 324, 325, 326, 327, 328, 330, 331, 333, 338, 339, 340, 343, 344, 348, 349, 350, 351, 353, 354, 356, 361, 362, 363, 364, 365, 367, 371, 372, 381, 382, 383, 384, 385, 399, 409, 430, 475, 481, 494
 Excel, 57, 61, 62, 65, 66, 85, 88, 111, 120, 123, 125, 132, 154, 167, 187, 210, 223, 224, 226, 227, 228, 278, 291, 329, 333, 342, 351, 365, 376, 378, 459, 475, 497
 Экстракт, 81, 128, 134, 193, 311, 323, 324, 326, 328, 330, 345, 361, 362, 363, 365, 383, 384, 450, 475

F

Құстар, 62, 123, 330, 378
 ФПДФ, 339, 340, 341, 342, 343

G

GDPR, 109
 Географиялық ақпараттық жүйе, 58
 GLTF, 143, 278
 Google Sheets, 368, 370
 Грок, 103, 104, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 429, 443.

H

HDF5, 62, 67, 123, 329, 330, 378, 379, 380
 ХиППО, 29, 37, 95, 424, 477, 484, 490
 HTML, 123, 340, 365, 370

Мен

ИДС, 289, 290, 291
 ИФК, 8, 73, 138, 142, 186, 227, 261, 262, 263, 264, 265, 266, 267, 268, 272, 273, 276, 277, 278, 279, 280, 284, 286, 292, 296, 302, 329, 357, 414, 415, 417, 497
 ИГЕС, 262, 263, 276
 iLOD, 289
 IoT, 10, 18, 67, 271, 367, 369, 405, 413, 417, 418, 419, 455, 460, 465, 482, 484
 ISO 19650, 388

Ж

JavaScript, 320, 369, 378
 JSON, 88, 89, 90, 92, 123, 128, 142, 269, 272, 280, 329, 330, 333, 378, 474, 480
 Jupyter Notebook, 114, 115, 116, 130, 187, 224, 330, 346, 417, 425

К

Кагл, 115, 121, 130, 187, 224, 303, 307, 330, 346, 408, 415, 417, 425, 426, 430, 431, 433
 k-NN, 392, 393, 442, 445, 446, 447, 448
 KPI, 245, 311, 317, 318, 319, 320, 321, 324, 353, 372, 389, 478, 479

Л

LEED, 235, 236, 238
 ЛлаMa, 103, 104, 120, 124, 129, 133, 166, 219, 300, 329,

343, 349, 381, 421, 429, 443, 494
 LLM, 3, 4, 24, 29, 50, 51, 52, 55, 56, 92, 95, 99, 102, 103,
 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114,
 115, 116, 118, 120, 124, 125, 126, 129, 131, 133, 134,
 135, 166, 187, 219, 220, 221, 222, 224, 225, 226, 231,
 237, 238, 251, 294, 300, 301, 302, 303, 304, 305, 306,
 307, 308, 309, 328, 329, 332, 333, 334, 335, 336, 338,
 339, 340, 343, 344, 345, 346, 347, 348, 349, 354, 370,
 372, 376, 381, 382, 392, 394, 399, 413, 425, 429, 430,
 432, 433, 435, 443, 444, 457, 460, 461, 476, 488, 494.
 Жүктеу, 81, 128, 134, 193, 311, 323, 325, 326, 333, 334,
 338, 339, 342, 343, 361, 362, 363, 365, 367, 383, 384,
 450, 475
 LOD, 287, 289
 LOI, 287
 LOMD, 287
 Тәмөнгі код, 368, 369

M

Matplotlib, 123, 306, 320, 333, 335, 342, 372, 409, 411,
 494
 MCAD, 77, 78, 257, 284
 MEP, 14, 175, 492
 MESH, 234, 283, 284, 285, 296, 462
 Microsoft SQL, 65
 Mistral, 103, 104, 107, 110, 129, 133, 166, 219, 221, 300,
 329, 343, 349, 381, 421, 429, 443, 494
 MRP, 11, 12
 MS Project, 70
 MySQL, 63, 64, 65, 291, 329, 330

H

n8n, 116, 311, 361, 368, 369, 370, 371, 372
 NLP, 69, 394
 No-Code, 368, 369
 NURBS, 142, 282, 283, 284, 285
 NWC, 8, 276

O

OBJ, 143, 273, 276, 277, 278, 280, 281, 284, 285
 OCCT, 273
 OCR, 69, 128, 132, 134
 OmniClass, 154, 155, 156
 ашық BIM, 142, 216, 219, 256, 261, 278, 291
 Ашық код, 45, 55, 97, 98, 107, 108, 273, 275, 494
 OWL, 267, 268, 269

П

Pandas, 56, 67, 103, 117, 118, 119, 120, 121, 122, 123, 125,
 130, 134, 138, 186, 187, 220, 221, 225, 227, 269, 300,
 303, 328, 329, 330, 377, 378, 380, 381, 403, 408, 409,
 410, 411, 412, 414, 475, 479, 494
 Parquet, 67, 123, 329, 330, 373, 378, 379, 380, 381, 382,
 386, 401, 403, 414, 474, 480
 PDF, 69, 70, 71, 85, 111, 126, 127, 128, 129, 130, 131, 132,
 134, 146, 147, 177, 186, 190, 191, 211, 215, 278, 338,
 339, 340, 341, 342, 344, 345, 346, 347, 354, 355, 356,
 357, 365, 376, 475, 480, 497

PDM, 16

PHP, 63, 339

Tы pipeline, 44, 53, 115, 128, 183, 303, 307, 308, 311, 312,
 349, 350, 351, 352, 354, 355, 356, 367, 370, 417, 430,
 476

PLM, 16, 246

PLN, 8, 272, 296

Plotly, 320, 336, 337, 372

PMIS, 3, 24, 32, 62, 152, 196, 198, 211, 239, 240, 245, 246,
 247, 248, 249, 250, 272, 326, 351, 387, 388, 389, 390,
 458, 462, 487

PMS, 84, 151, 233

PostgreSQL, 63, 64, 65, 329, 395

Power BI, 320, 372, 482

жеке капитал, 464, 487

Питон, 56, 63, 103, 105, 112, 114, 115, 117, 118, 119, 129,
 130, 131, 133, 134, 135, 166, 167, 179, 187, 219, 224,
 225, 227, 303, 304, 308, 320, 329, 330, 332, 333, 339,
 340, 346, 348, 356, 367, 369, 372, 378, 409, 417, 426,
 459, 475, 476, 479, 494.

Q

QTO, 72, 196, 214, 215, 216, 217, 218, 219, 221, 223, 225,
 226, 228, 237, 238, 242, 243, 251, 301, 302, 475, 479

QWEN, 103, 104, 124, 129, 133, 166, 219, 221, 300, 329,
 343, 349, 381, 429, 443

R

RAG, 111, 116

RDBMS, 63, 64, 65, 82, 89

РДФ, 267, 268, 269

РегExс, 126, 136, 177, 179, 331, 332, 333, 357, 474

RFID, 8, 18, 58, 84, 418, 419, 455, 460, 465, 482

ROI, 311, 317, 319, 321, 351, 370, 389, 479

RPM, 14, 84, 151, 331, 333, 334

PBT, 8, 73, 77, 140, 186, 227, 272, 296, 300, 302, 329, 357,
 376, 414, 415, 417, 497

S

SaaS, 24, 50, 51, 52, 458
САЯCAT, 273, 277
SDK, 139, 141, 257, 264, 273, 281, 286, 296, 329, 414
Сиборн, 123, 320, 336, 337, 408, 412
SPARQL, 269
SQL, 63, 65, 66, 88, 89, 103, 105, 119, 123, 166, 168, 268, 269, 276, 277, 300, 329, 376, 392, 494
SQLite, 63, 64, 65, 166, 167, 296, 329, 363
КАДАМ, 261, 262, 263, 266, 272, 276, 277, 292
Стримлит, 336, 337
СВФ, 142, 276, 285

T

Трансформация, 128, 193, 311, 323, 325, 326, 330, 331, 333, 334, 340, 341, 342, 344, 347, 361, 362, 363, 365, 367, 383, 384, 450, 475

U

Уникласс, 154, 155, 156
АҚШ доллары, 142, 143, 255, 276, 277, 278, 279, 280, 281, 284, 285, 286, 292, 296

V

Векторлық операциялар, 373, 398, 400, 401
VR, 84, 271, 285

W

WhiteBox, 240, 242, 243

X

XLSX, 8, 61, 62, 123, 128, 129, 231, 268, 276, 277, 280, 296, 299, 308, 351, 373, 377, 378, 379, 380, 401, 474, 480
XML, 61, 88, 89, 92, 128, 143, 269, 272, 277, 280, 291, 299, 329, 378, 414, 474