**Лекція 5. Світові та національні стандарти якості ПЗ**

Сучасне індустріальне суспільство засноване на використанні стандартів, оскільки найбільш ефективним методом скорочення сукупних витрат на розробку, виробництво і експлуатацію виробів, а так само підвищення їх якості є стандартизація та уніфікація. Стандарти в галузі програмного забезпечення (ПЗ) мають величезне значення і охоплюють різні сфери розробки і застосування цього виду продуктів. Вони, зокрема, забезпечують можливість розробникам ПЗ використовувати дані і програми інших розробників, здійснювати експорт / імпорт даних, регламентують взаємодію між різними програмами. Завдяки стандартам ми маємо можливість запускати одні і ті ж програми на комп'ютерах різних виробників. Мови програмування - є теж не що інше, як стандарти. Мережа Інтернет - це по суті теж набір стандартів, використання яких дозволяє комп'ютерам і встановленим на них програмами обмінюватися даними по каналах зв'язку.

Стандартизація сприяє розвитку конкуренції та розширює ринки збуту продукції, дозволяє і виробникам і споживачам зменшувати ступінь залежності від конкретних осіб і організацій.

Стандарти виникають "де-юре" і "де-факто". Стандарт "де-факто" - термін, що позначає продукт будь-якого постачальника, який захопив більшу частку ринку і який інші постачальники прагнуть емулювати, копіювати або використовувати для того, щоб захопити свою частину ринку. Стандарт "де-юре" створюється формально визнаної стандартизує організацією. Він розробляється при дотриманні правил консенсусу в процесі відкритої дискусії, в якій кожен має шанс взяти участь. Жодна група не може діяти незалежно, створюючи стандарти для промисловості.

Сучасна індустрія ПЗ характеризується дуже високим ступенем конкуренції і для успішної роботи на цьому ринку компанія повинна розробляти, впроваджувати та супроводжувати ПЗ швидко, в строк і з необхідною якістю. Практика досягнення виробничими компаніями якості продукції, що випускається довела, що ця якість визначається якістю **процесу** виробництва. Тому багато компаній вкладають гроші в поліпшення якості процесу, пам'ятаючи про те, що подібне вкладення грошей обов'язково окупається. Вивчення прикладів поліпшення процесів розробки ПЗ показує, що в успішних випадках спостерігається істотне поліпшення продуктивності і якості, а вкладення в цю діяльність окупаються в рази.

Стандарти якості ПЗ або окремі розділи в стандартах розробки ПЗ, присвячені вимогам до якості ПЗ, в основному стосуються процесу розробки. При цьому потрібно мати на увазі, що існують і інші напрямки стандартизації в області ПЗ. Існує кілька різних підходів і відповідних стандартів до забезпечення якості ПЗ. Найбільш відомими з них є стандарти **ISO** і **CMMI**.

Однією з перших моделей якості став стандарт ISO (Міжнародної організації зі стандартизації) серії **9000**, перша версія якого була випущена в 1987 році. Серія стандартів ISO 9000 торкається різних аспектів управління якістю і включає кілька стандартів. Стандарти містять керівництва і інструментарій для компаній і організацій, які хочуть, щоб їх продукція і послуги постійно відповідали вимогам замовника, а якість постійно поліпшувалося. Один з основних на сьогоднішній день стандартів цієї серії ISO 9001: 2008 встановлює вимоги до системи менеджменту якості. Організація, яка впровадила цей стандарт, може пройти процедуру незалежного аудиту і отримати міжнародний сертифікат відповідності вимогам стандарту, який підвищує її статус на ринку.

Стандарти ISO 9000 універсальні і застосовуються в самих різних галузях промислового виробництва. За даними ISO цей стандарт запроваджено в понад один мільйон компаніях і організаціях більш ніж в 170 країнах світу (http://www.iso.org/iso/ru/home/standards/management-standards/iso\_9000.htm).

В області розробки програмного забезпечення в США найбільшого поширення набув стандарт Capability Maturity Model (**CMM**), що зазвичай перекладають як "модель зрілості процесу розробки ПЗ". Його історія починається в 1991 році, коли американський інститут SEI (Software Engineering Institute - Інститут системного програмування при університеті Карнегі-Меллон) опублікував першу версію цього стандарту. Початковою метою розробки стандарту було створення методики, що дозволяє великим урядовим організаціям США (SEI був створений і фінансувався коштом Міністерства оборони США) вибирати найкращих постачальників ПЗ. Для цього передбачалося створити вичерпний опис способів оцінки процесів розробки ПЗ та методики їх подальшого удосконалення. В результаті, авторам вдалося домогтися такого ступеня подробиці і деталізації, що стандарт виявився придатним і для звичайних компаній-розробників, які бажають поліпшити існуючі процеси. В даний час дана модель більше відома під абревіатурою CMMI. Вона використовується не тільки для розробки програмних систем, а й багатьма великими компанії, які випускають вироби, куди ПЗ входить як складова частина. Наприклад, авіаційна, аерокосмічна індустрії, де розробка ПЗ відбувається разом з інженерними роботами інших видів. Завдання CMMI - надати таким проектам і компаніям єдину платформу організації процесу розробки.

Головним поняттям стандарту є **зрілість організації**. Незрілої вважається організація, в якій процес розробки ПЗ залежить тільки від конкретних виконавців і менеджерів, і рішення часто просто імпровізують "на ходу". У цьому випадку велика ймовірність перевищення бюджету або завалювання термінів здачі проекту, і тому менеджери змушені займатися тільки дозволом найближчих проблем. В зрілої організації є чітко визначені процедури створення програмних продуктів і управління проектами. Ці процедури в міру необхідності уточнюються і удосконалюються. Оцінки часу і вартості виконання робіт ґрунтуються на накопиченому досвіді і досить точні. Нарешті, в компанії існують стандарти на процеси розробки, тестування та впровадження ПЗ, правила оформлення кінцевого програмного коду, компонент, інтерфейсів і т. Д. Все це становить інфраструктуру і корпоративну культуру, яка підтримує процес розробки програмного забезпечення.

Основу вітчизняної нормативної бази в області документування ПЗ становить комплекс стандартів Єдиної системи програмної документації (ЕСПД). Це система міждержавних стандартів країн СНД (ГОСТ), що діють на території Російської Федерації на основі міждержавної угоди по стандартизації. На сьогоднішній день ця система стандартів залишається чинною.

**Єдина система програмної документації** - це комплекс державних стандартів, що встановлюють взаємопов'язані правила розробки, оформлення та обігу програм і програмної документації. Стандарти ЕСПД в основному охоплюють ту частину документації, яка створюється в процесі розробки ПЗ, і пов'язані, здебільшого, з документуванням функціональних характеристик ПЗ.

Ці стандарти зачіпають досить формальні аспекти розробки ПЗ і не враховують різноманіття сучасних процесів і методологій розробки ПЗ, тому доцільно орієнтуватися на міжнародні стандарти, які використовують у своїй діяльності провідні світові виробники ПЗ, а до . ЕСПД звертатися як до довідкової інформації.

Стандарт **ISO 9002** використовується в ситуаціях, коли відповідність установленим вимогам до якості повинна підтримуватися тільки в процесах виробництва, монтажу й обслуговування продукції. Він подає модель системи якості, що забезпечує якість під час виробництва та випробування продукції, а також її встановлення у споживача.

Стандарт **ISO 9003** можна використовувати тоді, коли достатньо забезпечити відповідність установленим вимогам до якості лише в процесі випробувань готової продукції у споживача. Таким чином, стандарт **ISO 9001** найбільш повний і фактично містить вимоги стандарту ISO 9002, а той, у свою чергу, вимоги стандарту ISO 9003].

Стандарт **ISO 9004** може використовуватися як методична допомога у побудові систем якості, оскільки містить рекомендовану структуру, системи якості, характеристики її основних функціональних елементів, певні вимоги до організаційної структури, складу й змісту даних, які повинні або можуть застосовуватися в системі якості. У ньому також розглядаються економічні аспекти якості, різні види витрат і статті витрат на якість, а також даються вказівки з проведення внутрішніх перевірок якості, що дозволяють керівництву організації оцінити ступінь готовності своїх підрозділів до стабільного постачання продукції, яка відповідає вимогам специфікацій, стандартів і очікуванням споживачів. Але цей стандарт застосовується, насамперед, для вирішення завдань щодо внутрішнього забезпечення якості і не використовується для цілей сертифікації.

Усі згадані вище стандарти були прийняті ISO ще в 1994 р., частково застаріли, однак продовжували діяти до 2004 р. Замість сукупності стандартів ISO 9001, ISO 9002 і ISO 9003 прийнято нову редакцію стандарту ISO 9001 2000 р. (і відповідно нову редакцію стандарту ДСТУ ISO 9001 (2001 р.)), яким і повинні керуватися організації, що не мають раніше сертифікованих систем  
менеджменту якості. Це стосується також стандарту ISO 9004 і гармонізованого з ним ДСТУ ISO 9004.

Потрібно відзначити, що ISO не тільки реструктурувала серію стандартів 9000, але й переглянула загальну частину їх назви. Тепер це стандарти не для систем якості (quality systems), а для систем управління якістю (quality management systems). Це вказує на важливість такого аспекту якості, як досягнення вищих цілей організації — задоволення потреб споживача щодо якості продукції, а не просте забезпечення відповідності продукції встановленим вимогам до якості.

Стандарти **ISO 9001:2000** і **ISO 9004:2000** розроблені як погоджена пара стандартів на системи менеджменту якості, що доповнюють один одного.

ISO 9001:2000 встановлює вимоги до системи управління якістю, які призначені для внутрішнього застосування організаціями з метою сертифікації або укладання договорів (контрактів). Він формулює мінімальний набір умов, який повинна задовольняти система управління якістю, що забезпечує гарантії випуску продукції, що відповідає встановленим вимогам і сподіванням споживачів.

ISO 9004:2000 надає інструкції з упровадження та застосування системи управління якістю для вдосконалення роботи організації в цілому. Методичні вказівки охоплюють створення, функціонування, підтримання в робочому стані та постійне поліпшення системи управління якістю.

Структуру і зміст нових стандартів змінено. Тепер вони відповідають сучасному загальноприйнятому процесо-орієнтованому підходу до створення продукції. Крім процесів, що безпосередньо стосуються випуску продукції, у стандартах чітко виділені процеси вимірювання, перевірки (аналізу, аудита) і вдосконалення як продукції, так і самої системи управління якістю, а також процеси управління ресурсами, включаючи трудові ресурси. Крім того, відзначається посилення ролі вищого керівництва організації в розвитку та поліпшенні системи управління якістю.

Під час побудови систем управління якістю ПЗ організації-розробники повинні керуватися стандартами ДСТУ ISO 9001:2001 і ДСТУ ISO 9004:2001, а також можуть використовувати ДСТУ ISO 9000-3:1998, що містить посібник із застосування стандарту ДСТУ ISO 9001:1995 (відповідає ISO 9001:1994) під час розроблення, постачання й обслуговування ПЗ.

Стандарт ISO 9000-3:1997 [74] має перехресні посилання на ISO/IEC 12207:1995, що дозволяє спільно використовувати ДСТУ ISO 9000-3:1998 і ДСТУ 3918 (гармонізований з ISO/IEC 12207:1995 [1]). Перевірка систем якості регулюється стандартами ДСТУ ISO 10011:1997 (частини 1, 2, 3). У них містяться вимоги до процедур перевірки, кваліфікації аудиторів систем якості, а також управління програмою перевірок.

***Сертифікація програмних продуктів і систем управління якістю в Україні***

Програмно-інформаційна продукція — це невід’ємна частина  
будь-якої інформаційної системи та інформаційної технології.

У 1993–1995 рр. у межах приблизно 6000 проектів державних науково-технічних програм (ДНТП) розроблялися понад 1000 програмних засобів. Однак, на думку фахівців, не було створено жодного програмного продукту (ПП), придатного для постачання користувачам. Наказ Державного комітету з науки і техніки (ДКНТ) № 47 від 17.03.94 р. передбачав певні заходи щодо забезпечення необхідного рівня якості ПЗ, що розроблялися у межах ДНТП. Затверджено тимчасовий порядок організації розроблення, випробування, постачання та супроводження ПП. Передбачалася реєстрація всіх ПП, що успішно пройшли приймальні випробування, а також організація незалежних  
випробувань ПЗ (третьою стороною) за заявками керівників проектів з виділенням для цих цілей додаткових засобів та ресурсів.

Однак в 1994–1995 рр. жодної заявки на такі випробування не надійшло, як і ніяких даних про успішні приймальні випробування будь-якого ПП не надійшло, а спеціально проведена нарада-семінар з активом керівників проектів остаточно розвіяла ілюзії щодо можливості успішного доведення до завершення будь-якого з виконуваних проектів.

Починаючи із середини 1995 р., почалася перманентна організаційно-штатна перебудова управлінських структур ДНТП (тобто ДКНТ як Державного комітету Міністерства науки), а з нею завершилися спроби впорядкування і підвищення ефективності ро біт зі створення програмної продукції в межах ДНТП. Причини та кого стану полягають навіть не в недостатньому фінансуванні, а у відсутності належної організації робіт в індустрії ПЗ, подібної до прийнятої на Заході. Чи можна, наприклад, очікувати конкуренто спроможний багатофункціональний ПП, що має необхідні споживчі властивості, якщо на початку його розроблення не встановлені:

− повні й однозначні вимоги до ПЗ, що відображають його споживчі властивості;

− план забезпечення якості ПЗ, включаючи методи та критерії контролю якості продуктів на всіх етапах розроблення;

− часові й інші ресурси для підготовки і проведення випробувань кінцевого продукту та його компонентів.

Однак у жодному з п’ятдесятьох (зі згаданої 1000), перевірених за завданням ДКНТ проектів, цих документів не виявилося. Результати проектування відповідні. Такими будуть очевидно і всі наступні результати доти, доки організації, що займаються програмуванням, не впровадять сучасну концепцію забезпечення якості ПЗ, відображену в декількох серіях актуальних міжнародних стандартів.

***Сучасна концепція забезпечення якості ПЗ***

Для забезпечення необхідного рівня якості ПЗ у міжнародній практиці знайшли застосування два підходи: продукто-орієнтований і процесо-орієнтований. Обидва підходи потребують системи менеджменту якістю. Така система визначає зобов’язання управління стосовно якості, встановлює його політику і деталізує необхідні дії.

За першого підходу акцент робиться на контролюванні якості шляхом випробування готового ПП. Цей підхід базується на припущенні, що чим більше виявлено й усунуто помилок у ПЗ під час випробувань, тим вища його якість. Недоліки такого підходу:

1) усунення помилок у готовому продукті на етапі випробувань обходиться в десятки разів дорожче, ніж якби ці помилки були відвернені або усунуті вчасно на ранніх етапах ЖЦ ПЗ;

2) відсутні методи й засоби випробування ПЗ, що гарантують повне виявлення у випробуваних ПЗ помилок і недоліків.

За другого підходу акцент робиться на вживання заходів щодо запобігання, оперативного виявлення й усунення помилок у ПЗ шляхом завчасного визначення відповідальності, планів забезпечення, основних процедур із забезпечення якості розроблюваних ПП і проведення відповідних заходів неухильно й послідовно, починаючи із найбільш ранніх етапів ЖЦ. Такий підхід натепер можна вважати загальноприйнятим. Його покладено в основу концепції якості ПЗ Міжнародних організацій зі стандартизації ISO/IEC і він реалізується в численних міжнародних стандартах, проектах стандартів і робочих матеріалах цих організацій

***Стандартизація в інженерії якості П***З

У міжнародному масштабі спеціалізовані системи стандартів формують ISO (Міжнародна організація зі стандартизації) і IEC (Міжнародна електротехнічна комісія).

Для створення системи стандартизації у галузі інформаційних технологій (ІТ) цими організаціями був сформований спільний технічний комітет ISO/IEC JTC1, до складу національних груп якого з 1994 р. входить і Україна. До складу JTC1 входять близько 30 підкомітетів (SC), що охоплюють всі основні компоненти ІТ, зокрема SC7 «Програмна інженерія». Підкомітет SC7 складається із групи планування і робочих груп (WG) за такою тематикою:

WG2: Документація системних ПП;

WG4: Інструменти і CASE-засоби;

WG6: Оцінювання ПП і метрики;

WG7: Управління ЖЦ;

WG8: Підтримання процесів ЖЦ;

WG9: Забезпечення гарантій і цілісність ПЗ;

WG10: Оцінювання і контроль процесів ЖЦ;

WG11: Подання і визначення даних в інженерії ПЗ;

WG12: Функціональні вимірювання ПЗ;

WG19: Мови моделювання, метадані, схема та компоненти відкритого розподіленого оброблення;

WG20: Ядро знань у сфері програмної інженерії;

WG21: Процес управління напрацюванням у галузі ПЗ;

WG23: Управління якістю систем;

WG24: Життєві цикли ПЗ для малих підприємств.

В Україні під егідою Держстандарту створено технічний комітет 22 (ТК22) «Інформаційні технології». У Держстандарті України сформовано загальні принципи розроблення стандартів у галузі інженерії якості та вимоги до них (документ JTC1/SC7 № 1345 від 24.03.95 р.). У ньому зокрема визначено, що документи повинні:

− орієнтуватися на процеси, продукти та ресурси;

− сприяти забезпеченню необхідного рівня якості ПЗ;

− сприяти підвищенню продуктивності праці програмістів;

− бути гнучкими стосовно об’єктивних можливостей;

− сприяти відтворюваності процесів ЖЦ ПЗ;

− захищати суспільну безпеку, здоров’я людей і економіку;

− базуватися на реаліях, а не на неперевіреній теорії;

− ініціювати або мотивувати забезпечення інструментального підтримання;

− бути актуальними.

Розроблено, обговорено і введено в дію стандарт. ISO/IEC 9126–1991 «Оцінювання програмного продукту. Характеристики якості та посібник з їх застосування». Документ специфікує модель якості у вигляді ієрархії, на верхньому рівні якої задано шість групових характеристик (функціональність, надійність, зручність використання, ефективність, супроводження і мобільність), що розкладаються надалі на підхарактеристики. Стандарт, таким чином, започаткував розвиток робіт зі встановлення і стандартизації повної номенклатури показників якості (включно до одиничних) та показників (метрик), що підлягають вимірюванню.

Розроблені і пройшли всі стадії узгодження стандарти:

*ISO/IEC 9126-1 «Програмна інженерія — Якість продукту — Характеристики й метрики якості програмних засобів — Частина 1*.

*Модель якості»; ISO/IEC 9126-2 «Програмна інженерія — Якість продукту — Частина 2. Зовнішні метрики»; ISO/IEC 9126-3 «Програмна інженерія — Якість продукту — Частина 3. Внутрішні метрики»*.

Під метрикою розуміють шкалу оцінювання та метод вимірювання кінцевої одиничної властивості ПЗ. Для характеристики споживчих властивостей ПЗ вводиться поняття «зовнішня метрика», а для характеристики властивостей проміжних продуктів розроблення ПЗ із погляду задоволення ними вимог до кінцевого ПП — «внутрішня метрика». У стандартах цієї серії вводяться й інші нові поняття, орієнтовані на інтерпретацію поняття «якість» залежно від фази ЖЦ ПЗ:

1) якість (Quality): сукупність характеристик деякої сутності, що впливає на її здатність задовольняти сформульовані і передба чувані потреби;

2) цільова якість (Goal Quality): необхідна і достатня якість, що відображає реальні потреби користувача;

3) необхідна якість продукту (Required Product Quality) — набір характеристик, що відображають якість, реально встановлену в специфікації вимог;

4) проектна якість (Design Quality) — якість, відображена в основі проекту або його частинах (архітектурі, коді програм);

5) прогнозована якість продукту (Estimated Product Quality): орієнтовна або попередня оцінка якості кінцевого продукту, що прогнозується на основі аналізу результатів, отриманих на певному етапі розроблення;

6) якість продукту, що поставляється (Delivered Product Quality): якість продукту, що поставляється, і який випробуваний у проімітованному середовищі функціонування;

7) якість у використанні (Quality in use): якість, що сприймається користувачами ПП під час використання в реальному середовищі.

В ISO/IEC 9126-2 розглянуто понад 100 базових зовнішніх метрик ПЗ і подано детальні рекомендації щодо їх використання. Однак, оскільки авторам не вдалося досягти повноти й однозначності запропонованих характеристик, стандарти ISO/IEC 9126-2 і ISO/IEC 9126-3 були введені зі статусом «Технічний звіт» (TR) для подальшої експериментальної апробації.

**ISO/IEC 12119:1994** «Інформаційна технологія — Пакети прикладних програм — Вимоги до якості і тестування». У стандарті встановлюються загальні вимоги до програмних продуктів.

*ISO/IEC 12207:1995 «Інформаційна технологія — Процеси життєвого циклу програмного засобу».* Визначаються процеси, що виникають на кожній фазі ЖЦ ПЗ, та встановлюється максимальний обов’язковий перелік дій і завдань, що виконуються під час реалізації кожного процесу.

*ISO/IEC 14102:1995 «Інформаційна технологія Оцінювання й відбір CASE-засобів»*]. Під CASE (Computer Aided Software Engineering) мають на увазі автоматизоване проектування програм. CASE-засоби — це будь-які, реалізовані на комп’ютері, інструментальні засоби, що використовуються у процесах ЖЦ ПЗ, під час збору та аналізу вимог проектування, кодування, тестування, документування тощо.

Стандарт *ISO/IEC 14102:1995* деталізує різноманітні функціональні характеристики CASE-засобів, що має сприяти вибору користувачем найбільш придатних інструментальних засобів із пропонованої сукупності, або сформулювати вимоги до них Серія стандартів ISO/IEC 14958 — частини 1, 2, 3, 4, 5, 6 — «Оцінювання програмного продукту»:

Частина 1. Загальні положення.

Частина 2. Планування й управління.

Частина 3. Процес для розробників.

Частина 4. Процес для покупців.

Частина 5. Процес для оцінювачів.

Частина 6. Документування модулів оцінювання.

Відповідно до стандарту ISO 8402 під оцінюванням розуміють систематичну перевірку ступеня здатності контрольованого об’єкта виконувати заздалегідь установлені вимоги. Такими об’єктами можуть бути як кінцевий програмний продукт, так і проміжні продукти процесу розроблення ПП або його компонентів.

Серія стандартів *ISO/IEC 15504*: частини 1, 2, 3, 4, 5, 6, 7, 8, 9 — «Оцінювання процесів ЖЦ програмних засобів»:

Частина 1. Концепції та вступна настанова (інформаційна частина).

Частина 2. Еталонна модель процесів та потужності процесу (нормативна).

Частина 3. Виконання оцінювання (нормативна).

Частина 4. Настанови щодо виконання оцінювання (інформаційна).

Частина 5. Модель оцінювання та настанови з показників (інформаційна).

Частина 6. Настанови щодо визначення компетенції оцінювачів

ДСТУ ISO/IEC 25012:2016 **Інженерія систем і програмних засобів.** **Вимоги до якості систем і програмних засобів та її оцінювання**

Схематично цей стандарт представлено на рис.1. Цим стандартом необхідно користуватися, коли потрібно протестувати якість даних і ви не хочете нічого пропустити— читайте стандарт. Якщо вам треба «залізні» докази, що ваші дані якісні — посилайтесь на стандарт.

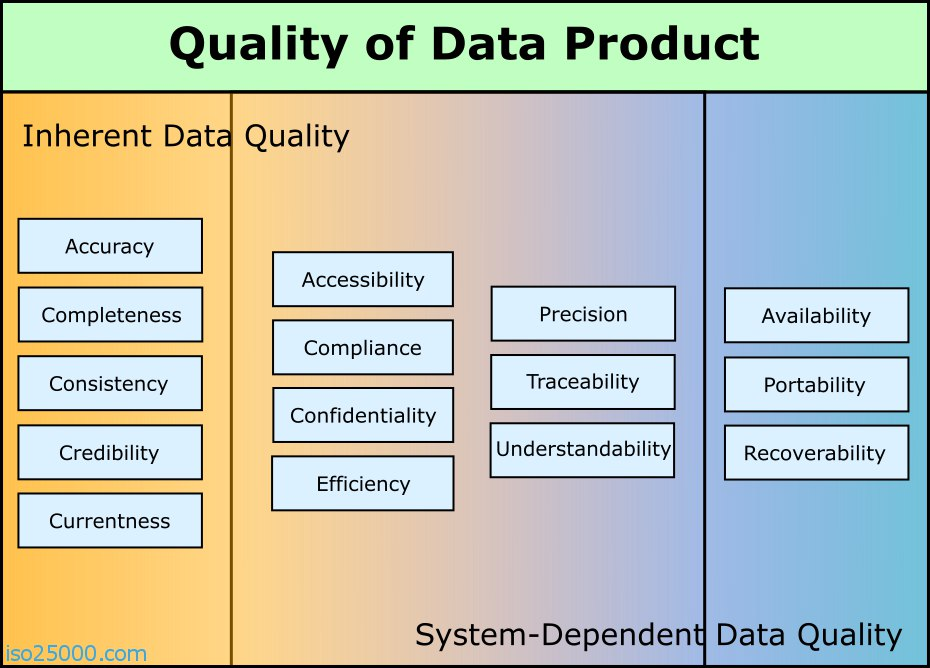


Рисунок 1 - ДСТУ ISO/IEC 25012:2016 Інженерія систем і програмних засобів. Вимоги до якості систем і програмних засобів та її оцінювання

**Визначення**

**Якість даних** — міра відповідності даних визначеним вимогам (майже співзвучно визначенню якості ПЗ). Якість даних складається з 15 характеристик, розділених на 3 категорії:



**Власна якість даних**

**Власна якість даних (Inherent Data Quality)** — ступінь відповідності даних явним і неявним потребам.

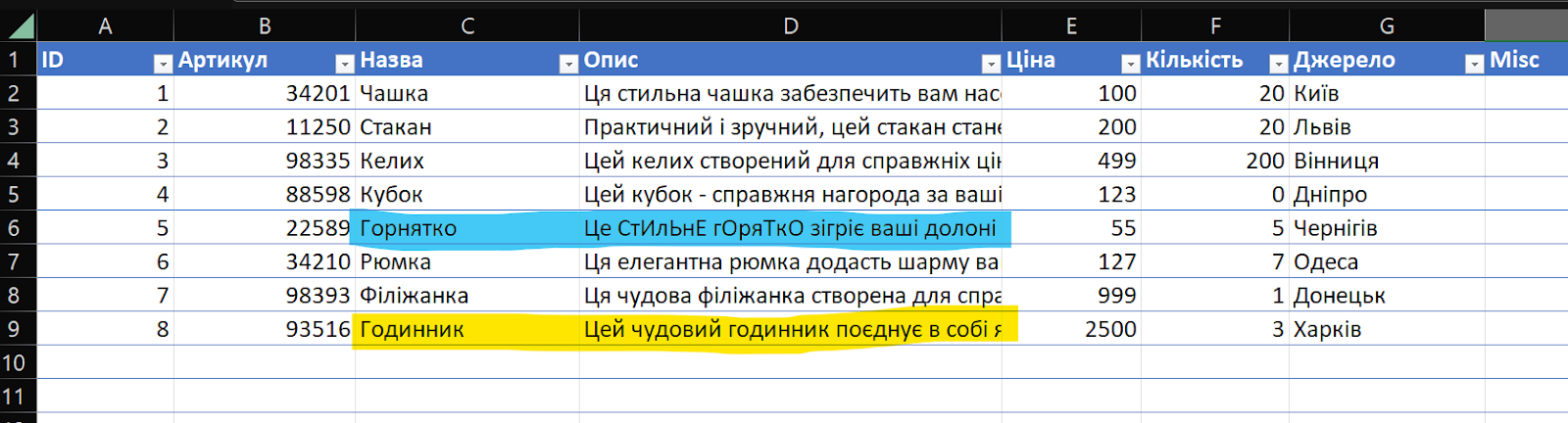
Ця категорія містить 5 характеристик:

**1. Точність (Accuracy)** — міра відповідності атрибутів даних справжньому значенню передбачуваного концепцією в конкретному контексті використання.

Має 2 підатрибути:

✅ **Синтаксична точність (Syntactic Accuracy)** — міра відповідності даних правилам, визначеним у домені, який вважається синтаксично правильним (мають правильну форму).

✅ **Семантична точність (Semantic Accuracy)** — міра відповідності даних правилам, визначеним у домені, який вважається семантично правильним (мають правильний зміст).

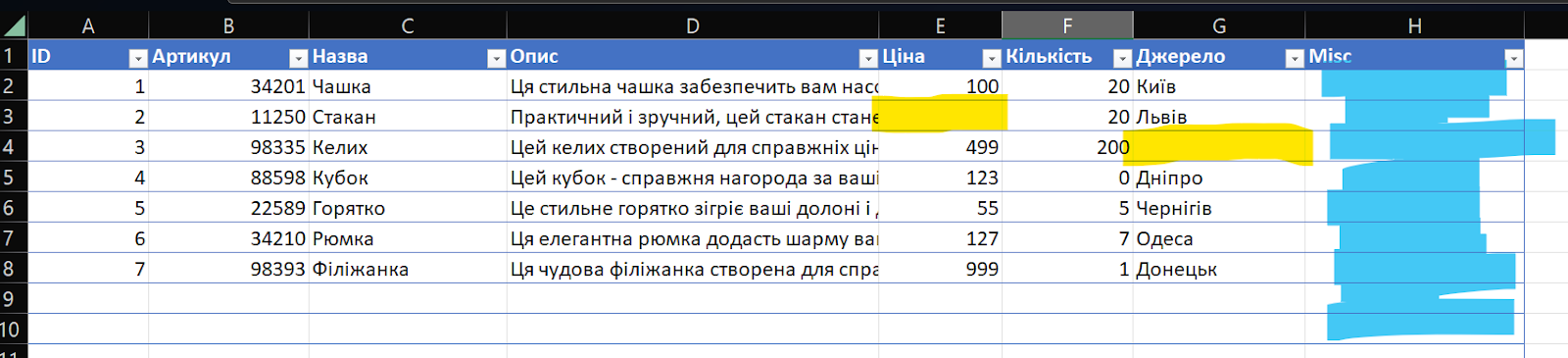


Приклад на зображенні. Є набір даних: товари для інтернет-магазину, який продає різні чашки. Два записи містять дані низької якості:

➡ № 5 Горнятко: проблема з синтаксисом в описі товару.

➡ № 8 Годинник: проблема зі змістом (самою суттю) — зайвий товар в усьому наборі даних.

**2. Повнота (Сompleteness)** — міра відповідності даних, пов’язаних з конкретною сутністю, до еталонної. Дані мають мати значення для всіх очікуваних атрибутів і пов’язаних екземплярів в конкретному контексті використання.



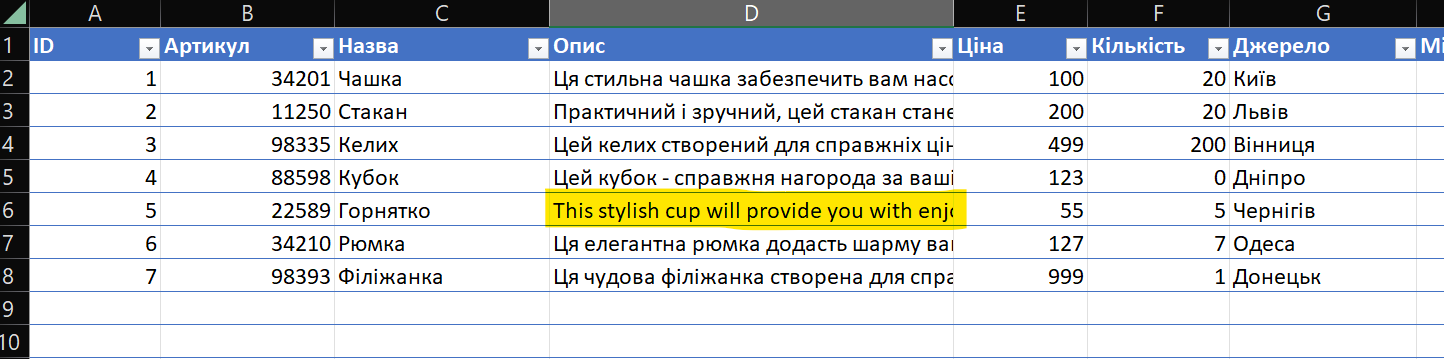
Приклад: товари для інтернет-магазину, який продає різні чашки.

➡ № 2 Стакан: не містить поля ціна, тож дані є не повними.

➡ № 3 Келих: не містить поля джерело, тож дані є не повними.

➡ Для жодного запису не має даних в атрибуті Misc. Або ж всі дані — не повні, або це поле тут зайве і має бути видалене.

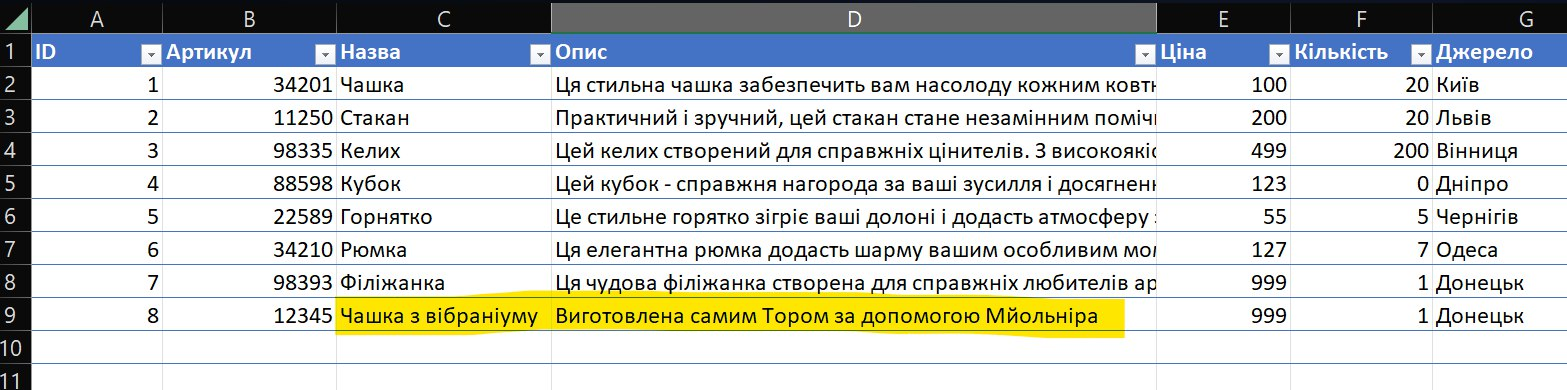
**3. Узгодженість (Сonsistency)** — міра відповідності даних іншим даним в конкретному контексті використання.



Приклад на зображенні. Є набір даних: товари для інтернет-магазину, що продає різні чашки.

➡ Опис товару «горнятко» написано англійською мовою, тож попри те, що дані є, вони синтаксично і семантично правильні, вони не мають бути в цьому наборі.

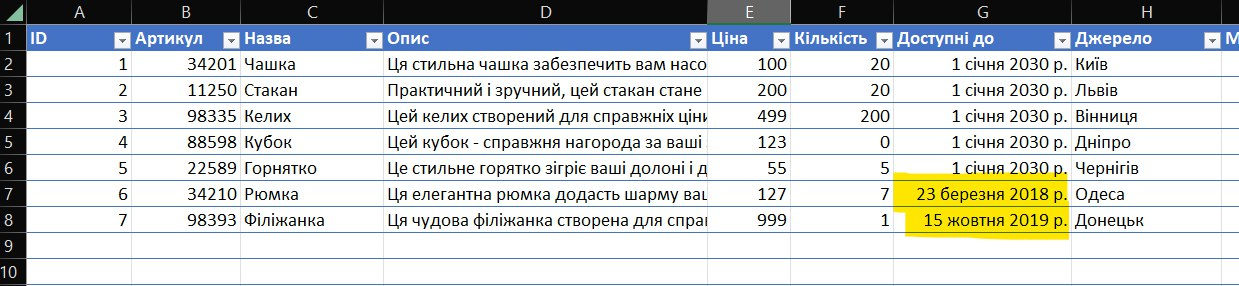
**4. Достовірність (Credibility)** — міра відповідності даних таким, що вважаються правдивими в конкретному контексті використання. Достовірність включає концепцію автентичності (правдивість походження, атрибуції, зобов’язань)



Приклад на зображенні. Є набір даних: товари для інтернет-магазину, що продає різні чашки.

➡Концепція магазину передбачає, що назва і опис товару відповідають дійсності. Саме тому запис № 8 викликає сумніви. Вібраніум — вигаданий метал, а Тор і Мйольнір — персонажі скандинавської мітології і коміксів Марвел.

**5. Актуальність (Currentness)** — міра відповідності даних віку в конкретному контексті використання.



Приклад на зображенні. Є набір даних: товари для інтернет-магазину, що продає різні чашки.

➡ Товари № 6 та № 7 були доступні для продажу у 2018 та 2019 роках відповідно. Тобто зараз втратили свою актуальність. Тут варто зазначити, що визначати актуальність даних потрібно дуже обережно, бо деякі старі дані можуть зберігатись не просто так, а для забезпечення зворотної сумісності чи для історії.

**Якість даних, залежна від системи**

➡ **Якість даних, залежна від системи** (System-Dependent Data Quality) — ступінь досягнення та збереження даних в конкретній комп’ютерній системі.

Ця категорія містить всього 3 атрибути:

**1. Доступність (Availability)** — міра, в якій дані можуть бути отримані авторизованими користувачами чи ПЗ.

Наприклад: кожен користувач, що реєструється в системі, має мати доступ до всіх своїх даних. При цьому система присвоює кожному користувачу прихований ID. Саме цей ID потенційно зменшує якість всіх даних, бо не все доступно користувачу.

**2. Портативність (Portability)** — міра, в якій дані можуть бути перенесені в іншу систему без втрати якості.

Наприклад: типова проблема багатьох інженерів — налаштувати міграцію даних. Гарні системи одразу проєктуються таким чином, щоб мати механізм експорту: в json, csv, xls тощо.

**3. Відновлюваність (Recoverability)** — міра, в якій дані можуть бути збережені для подальшої роботи у випадку помилок.

Наприклад: є бекап — чудово, нема бекапа — проблема.

**Змішана якість даних**

**Змішана якість даних** (Inherent and System-Dependent Data Quality) — ступінь відповідності даних і потребам, і конкретному ПЗ.

Ця категорія містить 7 атрибутів:

**1. Доступність (Accessibility)** — міра, в якій дані можуть бути доступними в конкретному контексті використання; наприклад, людям з обмеженими можливостями.

Наприклад: до кожної картинки на вашому сайті є текстовий опис, який може бути озвучений програмою для людей, що погано бачать.

**2. Відповідність (Compliance)** — міра відповідності атрибутів даних стандартам, домовленостям і правилам.

Наприклад: номером кредитної картки є 16-тизначне число. Це стандарт. Якщо у вас є набори даних з 17-тизначними номерами — це не обов’язково баг, але потенційна проблема.

**3. Конфіденційність (Confidentiality)** — міра, в якій дані мають атрибути, що забезпечують доступ до даних лише авторизованим користувачам.

Наприклад: дані зберігаються в зашифрованому вигляді і доступні лише тим, що має пароль для розшифрування.

**4. Ефективність (Efficiency)** — міра, в якій дані зберігаються у тому виді, в якому їх обробка займає оптимальну кількість часу.

Наприклад: числа можна зберігати у вигляді цифр: 1, 2, 15, 27, а можна у вигляді слів: один, два, п’ятнадцять, двадцять сім. Дані ті ж, але у вигляді слів їх обробляти складніше.

**5. Влучність (Precision)**— тут буде складно не переплутати з Точністю (Accuracy). Міра відповідності можливості чітко їх розрізнити і зрозуміти їхнє призначення.

Наприклад: у вашому інтернет-магазині атрибут товару — «ціна», має назву не «ціна», а абстрактно «гроші». Майже те саме, але погіршує розуміння.

**6. Відстежуваність (Traceability)** — міра відповідності можливості відстежити доступ до даних і зміни даних.

Наприклад: Git — гарний приклад системи, що містить дані та повну історію їх змін, тобто дані в git — якісні з точки зору відстежуваності.

**7. Зрозумілість (Understandability)** — міра, в якій дані можуть бути прочитані та інтерпретовані користувачами, написані відповідними мовами і символами.

Типовий приклад: дані, збережені в кодуванні UTF-8 — якісніші, ніж дані, збережені в національному кодуванні типу CP-1251.