華東師範大學
EAST CHINA NORMAL UNIVERSITY

数据中台

Data Power Platform

第一章 数据应用发展历史

$$(1+x)^n = 1 + \frac{nx}{1!} + \frac{n(n-1)x^2}{2!} + \cdots$$

课程提纲

Content

1 数据库系统

2 数据仓库

3 数据平台

4 数据中台

课程提纲

Content

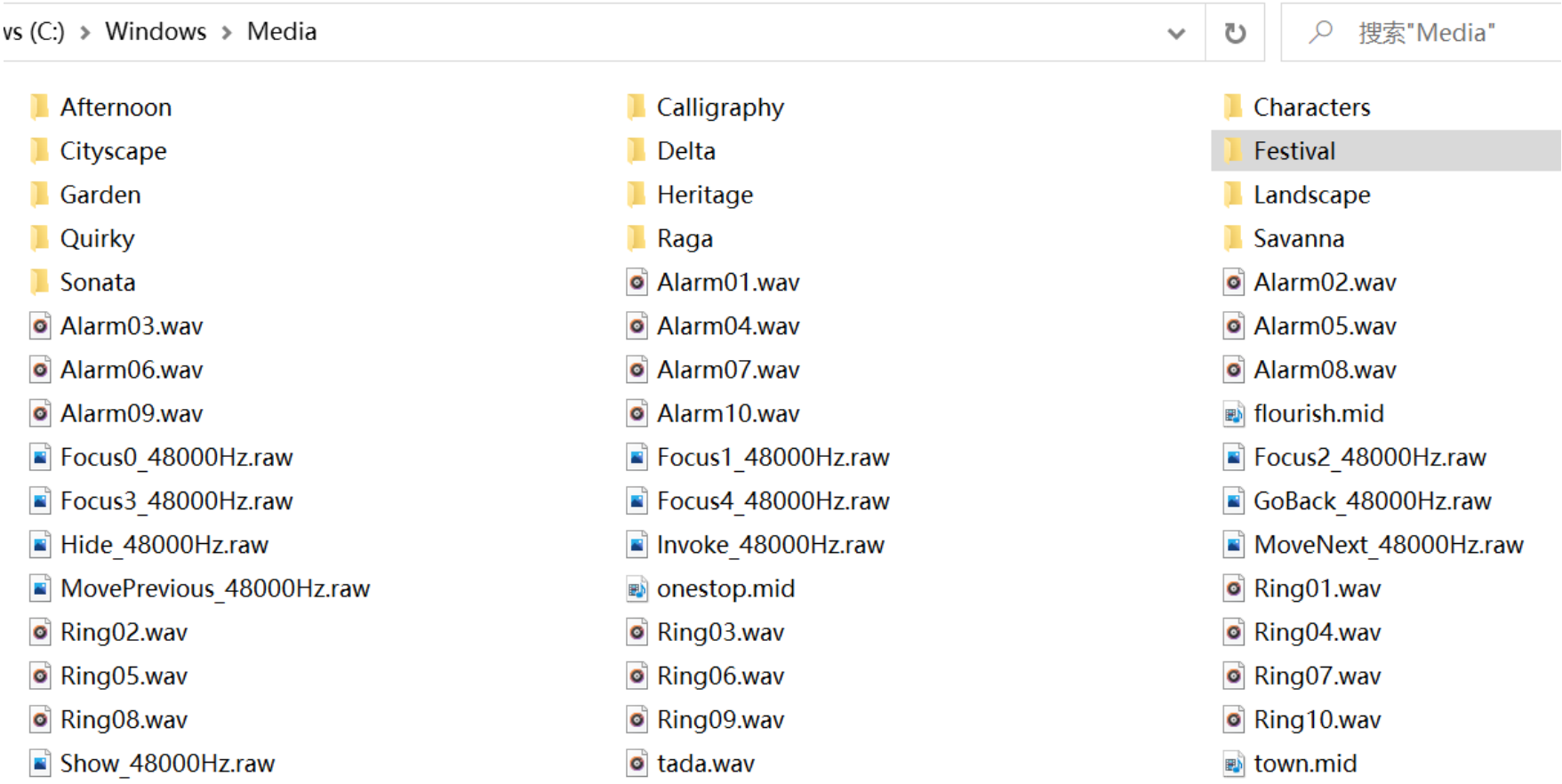
人工管理数据

- 最原始的数据管理：结绳记事
  - “事大，大结其绳；事小，小结其绳，之多少”
- 经济活动中的数据管理
  - 账本记账
  - 会计专业
- 人工数据管理的极致：信息资源编目管理
  - 按照一定的标准和规则，对文献信息资源的外部特征和内容特征进行分析、选择、描述，并记录为款目，继而资源有序组织的过程
  - 目的在于实现**采编流转**
  - 图书、杂志报刊、视音频资料等编目
  - 可以运用在其他领域，如设备编目、植物编目
- 新型存储介质的出现：磁带、卡片、纸片和磁盘改变了人类数据管理的方式

操作系统

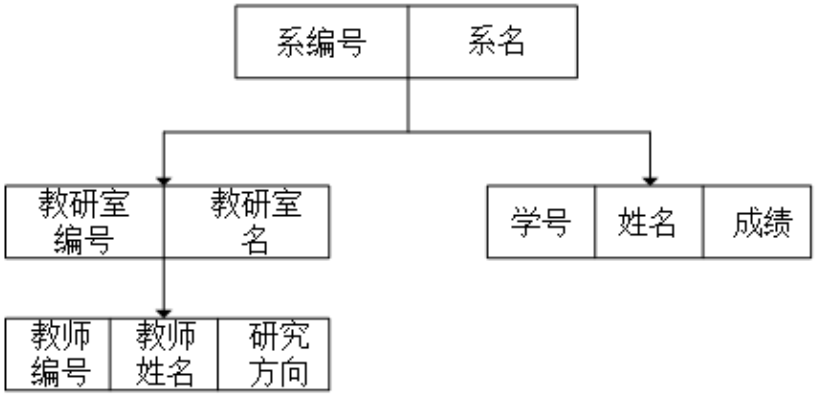
- 操作系统是管理计算机硬件资源，控制其他程序运行并为用户提供交互操作界面的系统软件的集合
  - 管理与配置内存
  - 决定系统资源供需的优先次序
  - 控制输入与输出设备
  - 操作网络
  - 管理文件系统
- 常用的操作系统
  - Android、BSD、iOS、Linux、Mac OS X、Windows、Windows Phone

操作系统的数据管理


vs (C:) › Windows › Media
搜索"Media"
Afternoon
Calligraphy
Characters
Cityscape
Delta
Festival
Garden
Heritage
Landscape
Quirky
Raga
Savanna
Sonata
Alarm01.wav
Alarm02.wav
Alarm03.wav
Alarm04.wav
Alarm05.wav
Alarm06.wav
Alarm07.wav
Alarm08.wav
Alarm09.wav
Alarm10.wav
flourish.mid
Focus0_48000Hz.raw
Focus1_48000Hz.raw
Focus2_48000Hz.raw
Focus3_48000Hz.raw
Focus4_48000Hz.raw
GoBack_48000Hz.raw
Hide_48000Hz.raw
Invoke_48000Hz.raw
MoveNext_48000Hz.raw
MovePrevious_48000Hz.raw
onestop.mid
Ring01.wav
Ring02.wav
Ring03.wav
Ring04.wav
Ring05.wav
Ring06.wav
Ring07.wav
Ring08.wav
Ring09.wav
Ring10.wav
Show_48000Hz.raw
tada.wav
town.mid

数据库发展历史

- 1963年，美国Honeywell公司的IDS (Integrated Data Store) 揭开了DB技术的序幕
- 20世纪70年代，层次和网状DB占据整个商用市场
  - 层次数据模型：用树状<层次>结构来组织数据
    - 按照树的定义，每棵树都有且仅有一个根节点，其余的节点都是非根节点
    - 每个节点表示一个记录类型对应实体的概念
    - 记录类型的各个字段对应实体的属性。

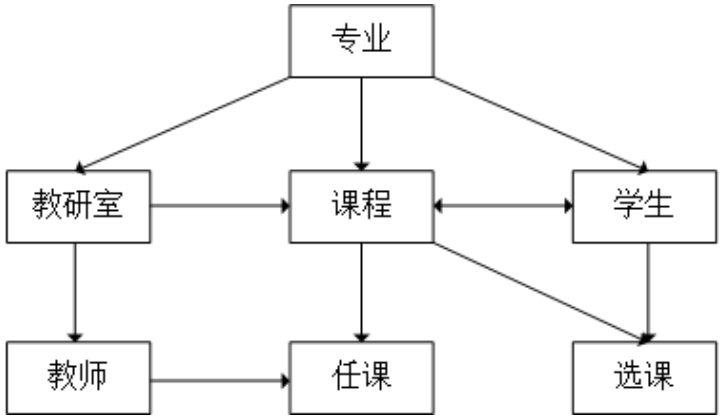


- 学校某个系的组织结构为例

示例: 层次DB

数据库发展历史

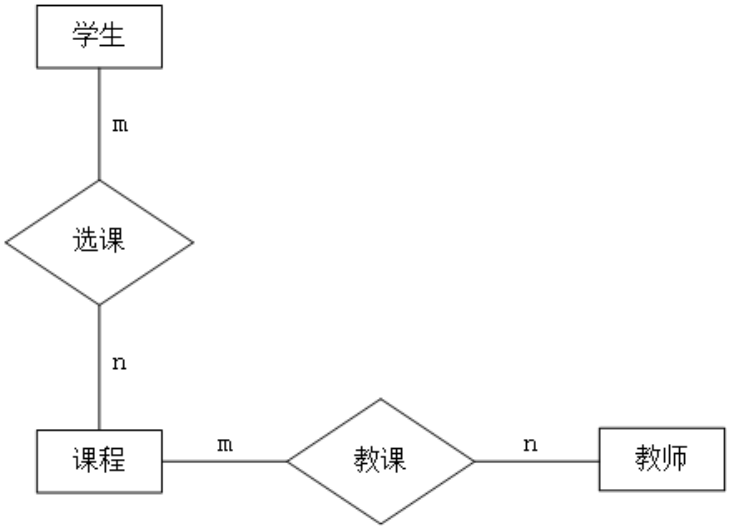
- 1963年，美国Honeywell公司的IDS (Integrated Data Store) 揭开了DB技术的序幕
- 20世纪70年代，层次和网状DB占据整个商用市场
  - 网状数据模型：用有向图表示实体和实体之间的联系的数据
    - 所有的节点允许脱离父节点而存在，允许存在两个或多个没有根节点的节点，同时也允许一个节点存在多个的父节点，成为一种网状的有向图。



示例: 网状DB

数据库发展历史

- 20世纪80年代，关系DB逐步替代网状和层次DB
- 20世纪90年代，关系DB成为主流数据库系统
  - 关系数据模型：使用表格表示实体和实体之间关系
  - 实体以及实体之间的联系都被映射成统一的二维表，在关系模型中，操作的对象和结果都是一张二维表



学生选课系统示意图

学生

stu_id	stu_name	sex	age

课程

cour_id	cour_name	xuefen

教师

tea_id	tea_name	sex	age

选课

stu_id	cour_id	chengji

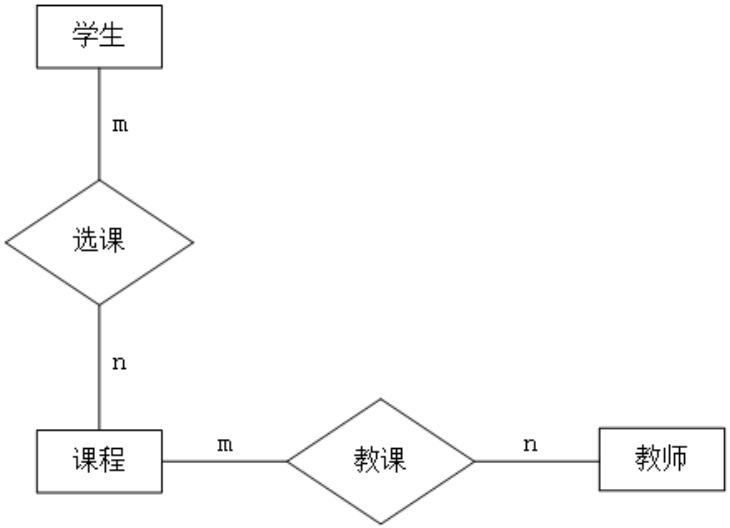
教课

tea_id	cour_id

示例：关系数据模型的表格

数据库发展历史

- 20世纪80年代，关系DB逐步替代网状和层次DB
- 20世纪90年代，关系DB成为主流数据库系统
  - 关系数据模型：使用表格表示实体和实体之间关系
  - 实体以及实体之间的联系都被映射成统一的二维表，在关系模型中，操作的对象和结果都是一张二维表



学生选课系统示意图

学生

stu_id	stu_name	sex	age

课程

cour_id	cour_name	xuefen

教师

tea_id	tea_name	sex	age

选课

stu_id	cour_id	chengji

教课

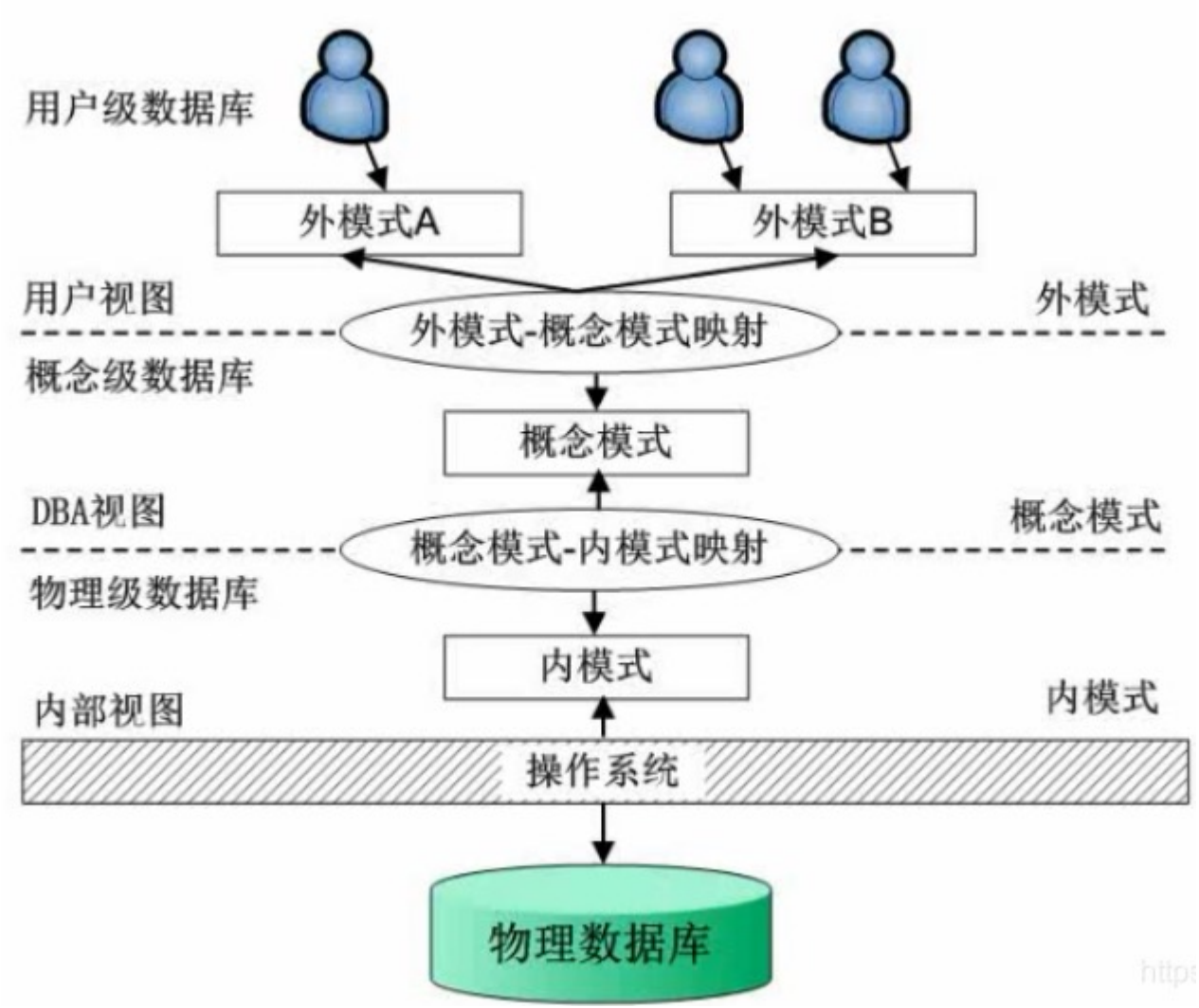
tea_id	cour_id

示例：关系数据模型的表格

数据库发展历史

- 1963年，美国Honeywell公司的IDS (Integrated Data Store) 揭开了DB技术的序幕
- 20世纪70年代，网状和层次DB占据整个商用市场
- 20世纪80年代，关系DB逐步替代网状和层次DB
- 20世纪90年代，关系DB成为主流数据库系统
- 进入21世纪，定制化数据库得到推广和普及
  - **one-size-fits-all**
  - **one-size-fits-a-bunch**

数据库



- 三级模式
- 两层映射
- 两个独立性

高级数据库技术

- 分布式数据库
  - 数据在物理上是分布的，但逻辑上完整的
  - 每个应用可以本地访问DB，也可以异地访问
  - 各节点间通过网络相连
  - 通过复制和镜像技术增加DB系统的可用性
- 对象数据库系统
  - 描述各种类型数据，表达数据间的复杂关系
  - 主要用于非结构化数据的存储，如Ontos、O2
- 图数据库
  - 描述对象间的相互关联，特别应用于数据推理
  - 主要用于对知识图谱的管理，如Neo4J

高级数据库技术

- 分布式数据库
  - 数据在物理上是分布的，但逻辑上完整的
  - 每个应用可以本地访问DB，也可以异地访问
  - 各节点间通过网络相连
  - 通过复制和镜像技术增加DB系统的可用性
- 对象数据库系统
  - 描述各种类型数据，表达数据间的复杂关系
  - 主要用于非结构化数据的存储，如Ontos、O2
- 图数据库
  - 描述对象间的相互关联，特别应用于数据推理
  - 主要用于对知识图谱的管理，如Neo4J

数据库的优势

- 数据库技术具有的优势
  - 采用一定的数据模型实现数据结构化
    - ✓ 概念数据模型：ER图
    - ✓ 逻辑数据模型：关系、层次或网状
  - 程序与数据具有较高的独立性
    - ✓ 逻辑独立性
    - ✓ 物理独立性
  - 控制数据冗余
  - 支持数据共享
  - 数据安全性较高

OS和DB间的区别与联系

文件系统	数据库系统
均为数据组织的管理技术	
数据库系统在文件系统的基础上运行	
文件将数据长期保存在外存上	用数据库统一存储数据
程序和数据有一定的联系	程序和数据分离
用操作系统中的存取方法对数据进行管理	用DBMS统一管理和控制数据
以文件为单位的数据共享	以记录和字段为单位的数据共享

课程提纲

Content

现有数据库的侧重点

- 数据库系统主要侧重于事务处理
  - 订票，如12306，机票，订单
  - 记账，如学生成绩系统，财务系统
  - 转账，如银行转账
- 强调多用户并发、数据一致性和完整性
- 但是
  - 各类信息系统大多属于OLTP系统
  - 多年运行之后，积累了大量的数据
  - 但是这些数据间相互不连通，形成数据烟囱
  - 这些数据不能充分发挥其价值

现实需求

- 持卡人今年的教育情况与以往相比，有怎样的交易特点（存款、取款、转账、消费）是什么？持卡人消费倾向（宾馆、大型商场、超级市场等）是什么？
- 要求
  - 多个子系统中数据的集成
  - 历史数据
  - 汇总和综合的数据
  - 一致的数据视图
- 类似的例子还有很多：今年销售量下降的因素、营收的组成、市场的变化。。。

OLTP系统处理分析型应用存在的问题

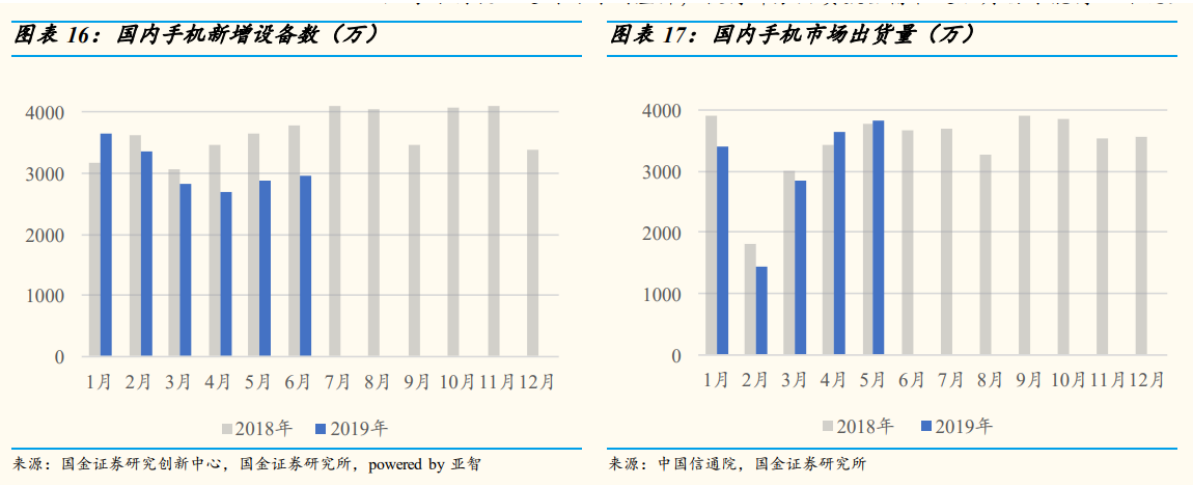
- 数据分散
- 数据动态集成问题
- 历史数据问题
- 数据的综合问题：非细节数据，而是多维度的数据综合
- 数据处理的效率问题（市场竞争）
- 数据质量和可信问题
- 数据共享和隐私保护的难题
- 决策过程的可持续性

数据烟囱

- 以部门为单位的IT建设
- 不同时间不同第三方建设
- 顶层设计缺失

数据烟囱
数据孤岛
数据碎片化

数据质量问题



数据统计口径不一

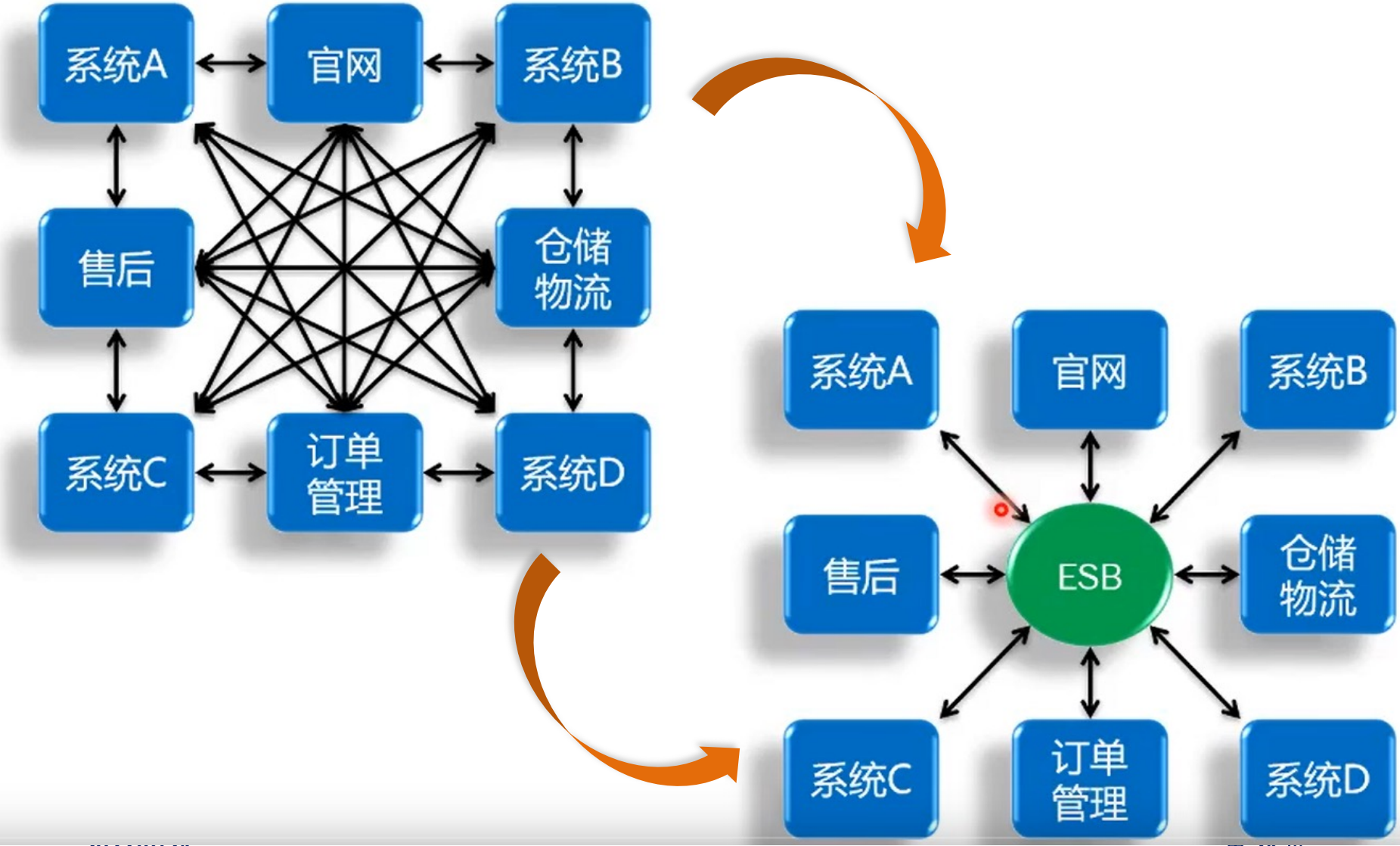
Roll Number	Name	Address	BookTaken
12	Alicia Ruth	12, Temple Street	AC091
14	Jason Darren	123, Sunset Blvd.	AC043
15	Mary Beth	32, Golden Avenue	AC021
12	Alicia Ruth	12, Temple Street	AC043
12	Alicia Ruth	12, Temple Street	AC011
15	Mary Beth	33, Golden Avenue	AC011

数据不一致

	name	toy	born
0	Alfred	fdsfa	NaT
1	Batman	Batmobile	1940-04-25
2	Catwoman	Bullwhip	NaT

数据缺失

数据共享问题


系统A
官网
系统B
售后
仓储
物流
系统C
订单
管理
系统D
系统A
官网
系统B
售后
ESB
仓储
物流
系统C
订单
管理
系统D

数据仓库的定义

- 数据仓库创始人 Bill Inmon:“ A Data Warehouse is a subject oriented, integrated, nonvolatile, and time variant collection of data in support of management's decisions”
- 数据仓库是为支持管理决策建立的面向主题的、综合的、稳定的、随时间变化的数据集合
  - 面向业务（操作）： 财产险、寿险，健康险等
  - 面向主题：客户、保单、保费、理赔等

数据仓库是面向OLAP的系统

- 数据仓库实现了信息的传递
- 数据仓库来源在于决策需求，而OLTP系统不能很好地处理这类需求
- 数据仓库不创造新的数据
  - 使用所有现存的数据
  - 通过清洗、转换
  - 提供企业综合、完整的总结与概括，生成非常有用的决策信息
  - 无需妨碍操作型系统，也能够支持决策需求

OLTP和OLAP的区别 I

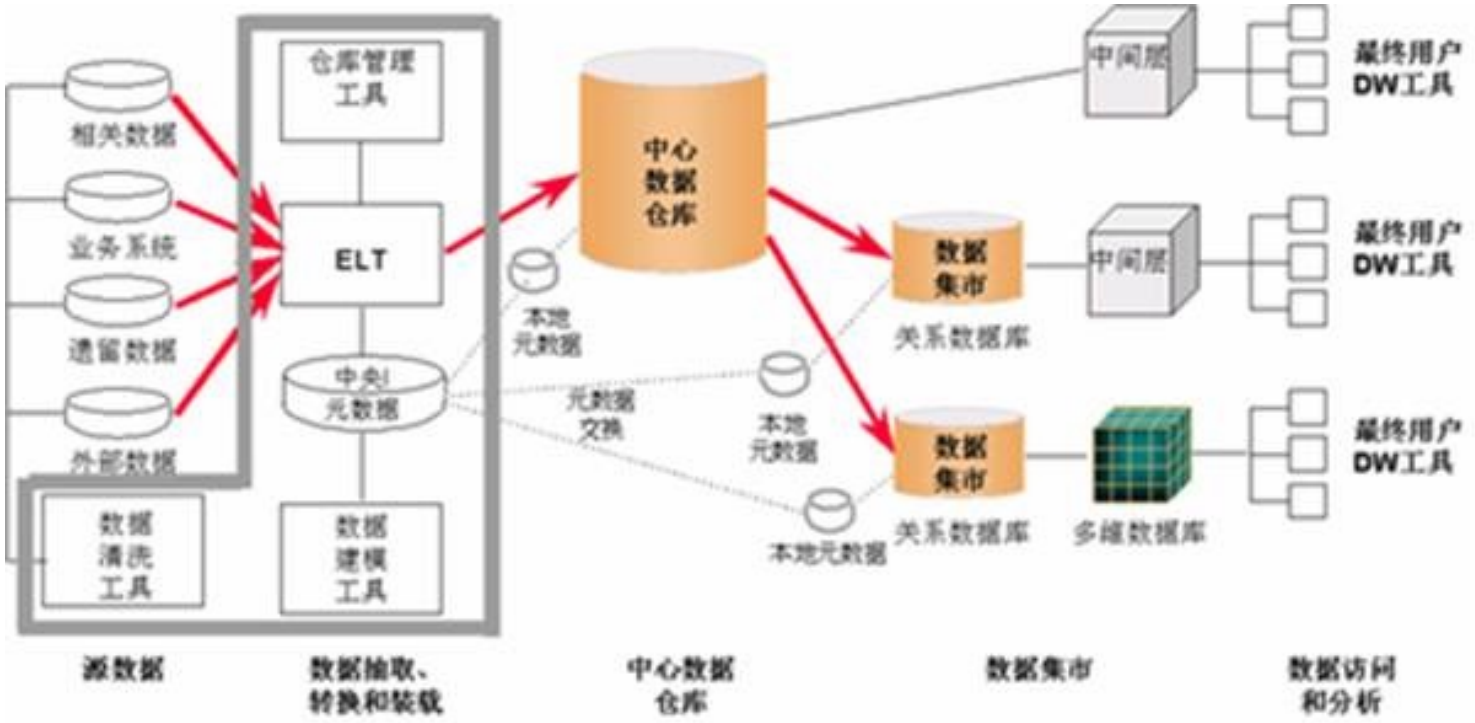
- OLTP
  - 下订单
  - 处理呼叫
  - 装货
  - 开发货单
  - 收起现金
  - 预定座位
  - 用户收藏
- 主要是数据库的写入操作

- OLAP
  - 销售量最好的产品名单
  - 出问题的区域
  - 告诉我为什么（向下钻取）
  - 看看其他数据（横向钻取）
  - 利润预警
- 主要是数据库的查询操作

OLTP和OLAP的区别 II

	OLTP	OLAP
数据内容	当前值，细节数据	历史数据，非细节
数据结构	适应于事务处理	适用于复杂查询
访问频率	高	中、低
访问连接类型	读取、更新、删除	读取
使用方法	可预知、反复性	特别查询、随机性、探索性分析
响应时间	快	一般
用户数量	大量	较少

数据仓库组件



数据仓库结构的元件

课程提纲

Content

数据仓库存在的问题

效率问题

- 需求交付周期长，需要几个小时
- 企业有哪些数据？找到所需要的数据费时费力
- 大量的数据很少被访问，但是却占用了大量的存储资源
- 部分实时数据处理需求得以满足

协助问题

- 数据孤岛问题突出，存在数据质量问题
- 部门内部系统的重复建设
- 没有汇聚部门全量数据
- 业务逻辑混乱

能力问题

- 开发人员多
- 既懂业务又了解技术的人少
- 不直接面对业务，不具备业务创新能力
- 缺乏数据治理，数据产品开发效率低

数据平台阶段的数据处理环节

数据采集	数据存储	数据计算	数据分析
来源单一，以内部结构化数据为主	主要面向结构化数据的OLTP 请求	依赖高性能计算机、单机或并行技术	主要利用统计和机器学习方法
扩展到传感、互联网、交易等多源多类型数据	扩展到面向非结构化数据的汇聚与存储	需要分布式并行计算，scale out的能力	需要更加智能的数据挖掘、AI和机器学习技术

数据平台阶段的三大主流技术

	Hadoop	MPP	数据仓库
数据结构	结构、半结构、非结构	结构化	结构化
数据规模	PB级甚至更高	TB级到PB级	TB级
SQL支持	低	高	高
BI工具支持	低	高	高
计算性能	对非关系型效率高	关系型高	关系型中
是否开源	开源	不开源	不开源
扩展能力	高	中	低
平台开放性	高	低	低
成本	低	中	高
开发维护	高	中	中

数据平台阶段形成数据闭环

我们都需要转接头


USB
AC POWER


ROMOSS
罗马仕
一线三头 共享充电
支持主流苹果机型，安卓和Type-C接口设备同时充电
*实现同时为3台设备正常充电，要求充电头或移动电源输出达2.1A以上

数据治理

- 数据采集规则
  - 按照规则从数据源直接采集，避免重复数据采集
  - 已采集的数据，发掘其剩余价值
  - 未采集的数据，增加采集点并发掘其价值
- 数据应用规则
  - 制定数据标准，保证数据质量
  - 按照规则进行统一数据清洗和数据转换
  - 根据不同需求，进行数据分发与权限控制
  - 制定数据共享机制，保证数据安全共享

课程提纲

Content

数据平台存在的问题

效率问题

- 需求交付周期长，平均一周左右
- 企业有哪些数据？找到所需要的数据费时费力
- 大量的数据很少被访问，但是却占用了大量的存储资源
- 实时数据处理需求难以满足

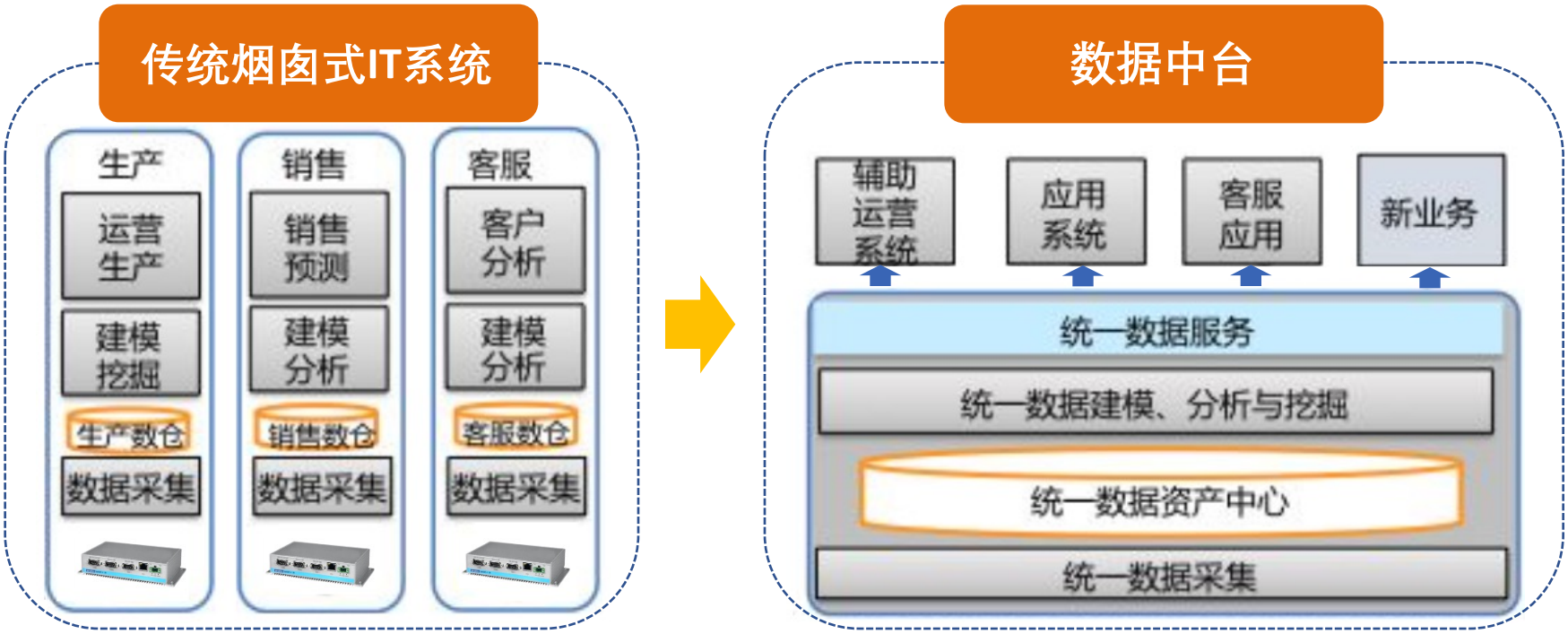
协助问题

- 形成更大的数据孤岛，同样存在数据质量问题
- 部门间系统的重复建设
- 没有汇聚企业全量数据
- 业务逻辑混乱

能力问题

- 开发人员多
- 既懂业务又了解技术的人少
- 业务创新能力弱

数据中台



独：数据孤岛、重复开发
断：数据价值链条断层
缺：缺乏标准、治理
难：理解、获取数据难

融合烟囱式IT系统，建立统一的数据采集和整合大数据处理、计算和数据服务能力，汇聚各领域数据，发挥数据资产价值创新业务能力

数据中台—三个核心认知（1）

- 数据中台需要提升到企业下一代基础设施的高度，进行规模化投入。
  - 数据中台的目标是提供普惠数据服务，在“互联网+”行动计划和“智能+”推动下，数字产业化和产业数字化成为数字经济的两大基础。
  - 数据中台只有被企业提升到下一代基础设施的高度，才能帮助企业从根本上解决数字化转型过程中遇到的瓶颈和痛点，例如数据孤岛林立（其实质是底层计算和存储架构的复杂性和异构造成的）、数据资产化程度低、数据服务提供效率与业务诉求严重不匹配等。
  - 相比于信息化部门把数据中台中的某些功能和特性作为新技术来局部验证和引入，数据中台更需要企业从战略高度进行顶层设计、确定规模化投入政策、设置更合理的组织架构，才能够确保数据中台作为数据应用的基础设施并落地建设，承担起企业数据资产全生命周期的管理。

数据中台—三个核心认知（2）

- 数据中台需要全新的数据价值观和方法论，并在其指引下形成平台级能力。
  - 数据中台所包含的数据技术创新可以在成熟的平台型企业内部孕育，技术的创新和融合应用于很多贴近业务的创新应用场景。
  - 但数据中台不仅仅是技术平台，倘若停留于此，就完全忽略了IT到DT的本质变化是围绕数据资产，企业面临的主要矛是无法解决业务端的灵活性和经营管理稳定性之间的冲突，单纯地增大技术投入和人才投入都无法保障企业经营效能的持续提升。
  - 只有秉持数据价值观和方法论，才可能系统性地解决企业经营发展围绕数据的诸多问题，谁能率先解决面向数字经济特征的全新数据价值观和方法论的问题，并在其指引下打造出平台级能力，谁就能真正意义上帮助企业把数据用起来。

数据中台—三个核心认知（3）

- 数据中台围绕业务、数据、分析会衍生出全新人才素养要求，需要尽快启动人才储备。
  - 人才永远是瓶颈，人才的定义也在动态变化，需要为人才准备成长的土壤。
  - 信息化历程中从简单的搭建网站、单功能系统开发，到复杂系统开发、建设、运营，再到新技术引入等都曾经是人才具体定义的重要关注点。
  - 信息化人才天然趋向两类企业：成熟稳定的平台型企业或有成熟平台潜力的企业。企业只有围绕数据中台明确了人才在企业的定位和职业通道，才可能吸引到或培养出拥有业务、数据、分析等综合素养的新型信息化人才，企业在数据中台人才储备上需要尽快做起来。

数据中台—三个发展阶段（1）

第一阶段：数据中台探索

- 这个阶段会将**数据生命周期各个阶段的技术与现有业务场景或创新业务场景结合**，迅速形成可见、可展示的业务成果。
- 特点是项目短小精悍，容易见效果
- 缺点是由于缺乏数据中台整体规划及让数据用起来的完整流程设计，无法对众多单个数据应用沉淀的数据形成通用数据资产，每个项目都需要从头到尾走一遍，当应用需求爆发式增长时，底层数据支撑的效率会大幅度下降，甚至影响最终的业务效果。

数据中台—三个发展阶段（2）

第二阶段：数据中台整合数据应用提升效率

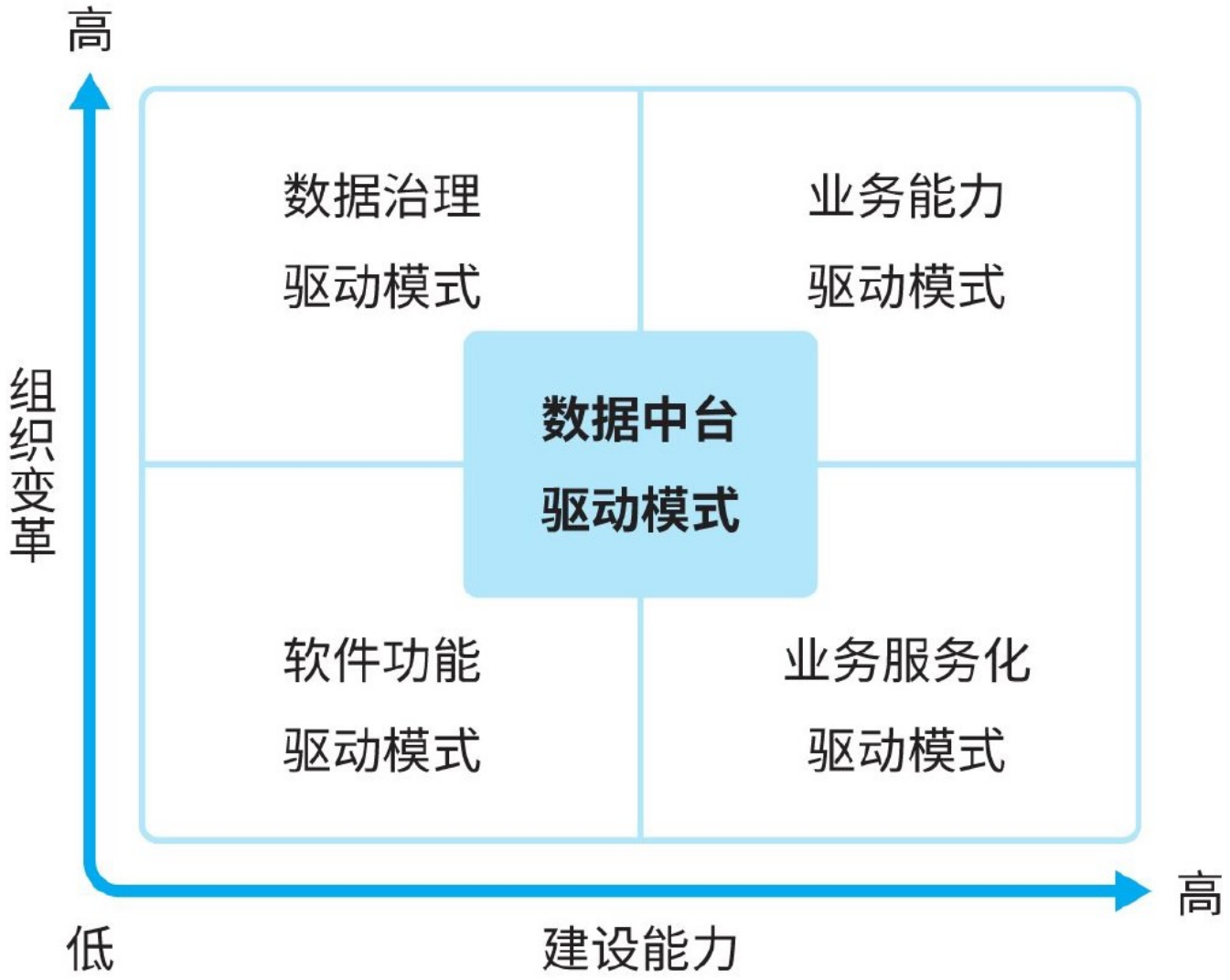
- 这一阶段的特点是构建数据中台的技术、理念、方法论是可复制的，市场上已有成熟的支撑数据中台高效运转的平台级产品。企业通过规划、建设、实施数据中台能够具备三方面的基础能力：
  - ✓ 数据的多样性、多态性、多云连接能力（**汇聚/交换能力**）。交换的能力用来解决企业有哪些数据、数据在哪里等问题。
  - ✓ **数据资产化的能力**是数据中台建设的关键，包括清洗、加工、治理、安全、质量等工具模块及实施方法论。（说明：能直接作用于业务领域，业务能阅读、能理解的数据才叫数据资产。）
  - ✓ **数据服务化的能力**，用数据技术来使用数据的方法。
- 有了这三个能力，就能将上一阶段构建起来的场景级数据应用，甚至是历史建成的系统都整合成企业级数据应用平台，既能满足原有系统对数据的需求，又能快速满足新业务场景对数据的需求，将数据作为资产上架，成为共享的生产要素。

数据中台—三个发展阶段（3）

第三阶段：数据中台重构数据空间和业务空间

- 到了这一阶段，数据中台已经成为企业数据资产的核心能力和基础，通过快速构建数据资产体系，帮助企业真正实现对其全量数据的有效管理。
- 业务和业务流程本身都可以通过适当的颗粒度进行数字化解耦和标准化，企业能够构建更加宏大的产业、行业价值链范围的数据空间和业务空间，以数据编排的方式响应业务需求，业务实现自流程化，数据实现自我管理能力。
- 这里需要引入业务空间和数据空间的基本概念。
  - ✓ **企业业务空间**：企业任何一个业务条线从初始设立到日益精细分化：清晰定义该业务条线内专项业务的“毛细血管”功能体系、建设或升级相应技术支撑系统、生成专项业务数据。当所有业务条线都遵循这个发展规律，纵横交错的业务条线构成了企业实际运营的多维业务空间。企业的业务空间是产生和形成全量数据的根本依据和前提。
  - ✓ **企业数据空间**：在数字化时代，任何一家企业都是市场生态中的一个节点，任何一家企业的数据全集只是整个市场数据生态空间中的一个子集。从企业自身视角来看，依据数据的生成和交互方式，企业全量数据的数据空间大致由三个维度构成：自主生产和消费的数据、外部数据（含单向外部获取数据和单向对外提供数据）、内外部交互数据。

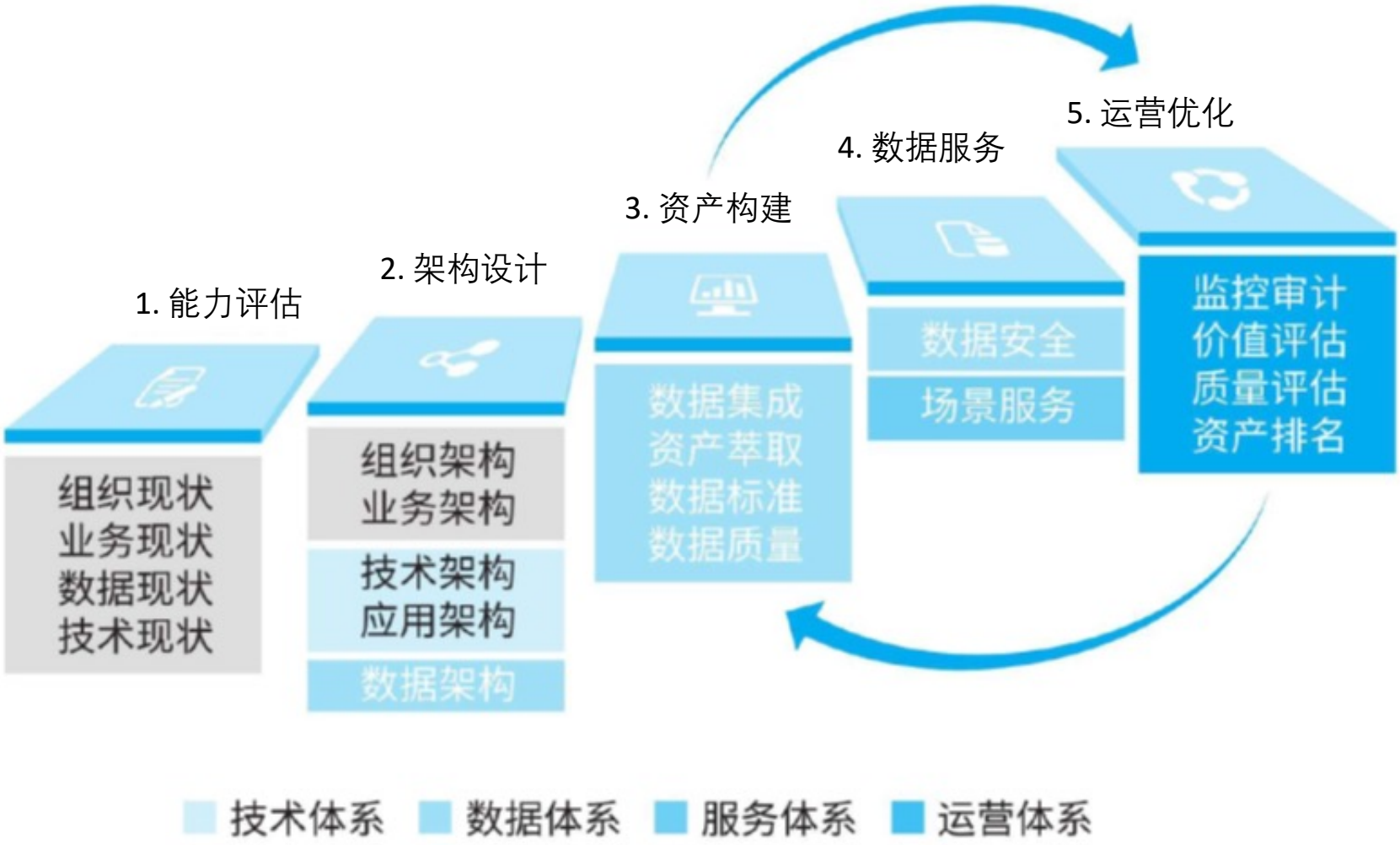
数据中台—建设模式


高
组织变革
数据治理
驱动模式
业务能力
驱动模式
数据中台
驱动模式
软件功能
驱动模式
业务服务化
驱动模式
低
建设能力
高

数据中台—建设模式

- **软件功能驱动模式**：该模式对组织变革和建设能力要求最低，通常以采购和实施成熟产品为主，目标是业务部门直接能用。
- **数据治理驱动模式**：该模式的目标是针对同一数据不同问题或不同数据同一问题进行分类治理，通常是业务上遇到难题，立个专项解决。
- **业务能力驱动模式**：该模式对组织变革和建设能力要求最高，目标基于企业架构（EA）自上而下开展规划建设，覆盖组织从战略到执行全业务过程，从业务设计到IT实现。该建设模式实施难度极高，通常会形成顶层规划设计和一系列实施项目。
- **业务服务化驱动模式**：该模式专注于新技术的引入，通常是面向用户提升体验、面向业务拉通资源调度。

数据中台—关键步骤


1. 能力评估
组织现状
业务现状
数据现状
技术现状
2. 架构设计
组织架构
业务架构
技术架构
应用架构
数据架构
3. 资产构建
数据集成
资产萃取
数据标准
数据质量
4. 数据服务
数据安全
场景服务
5. 运营优化
监控审计
价值评估
质量评估
资产排名
技术体系
数据体系
服务体系
运营体系

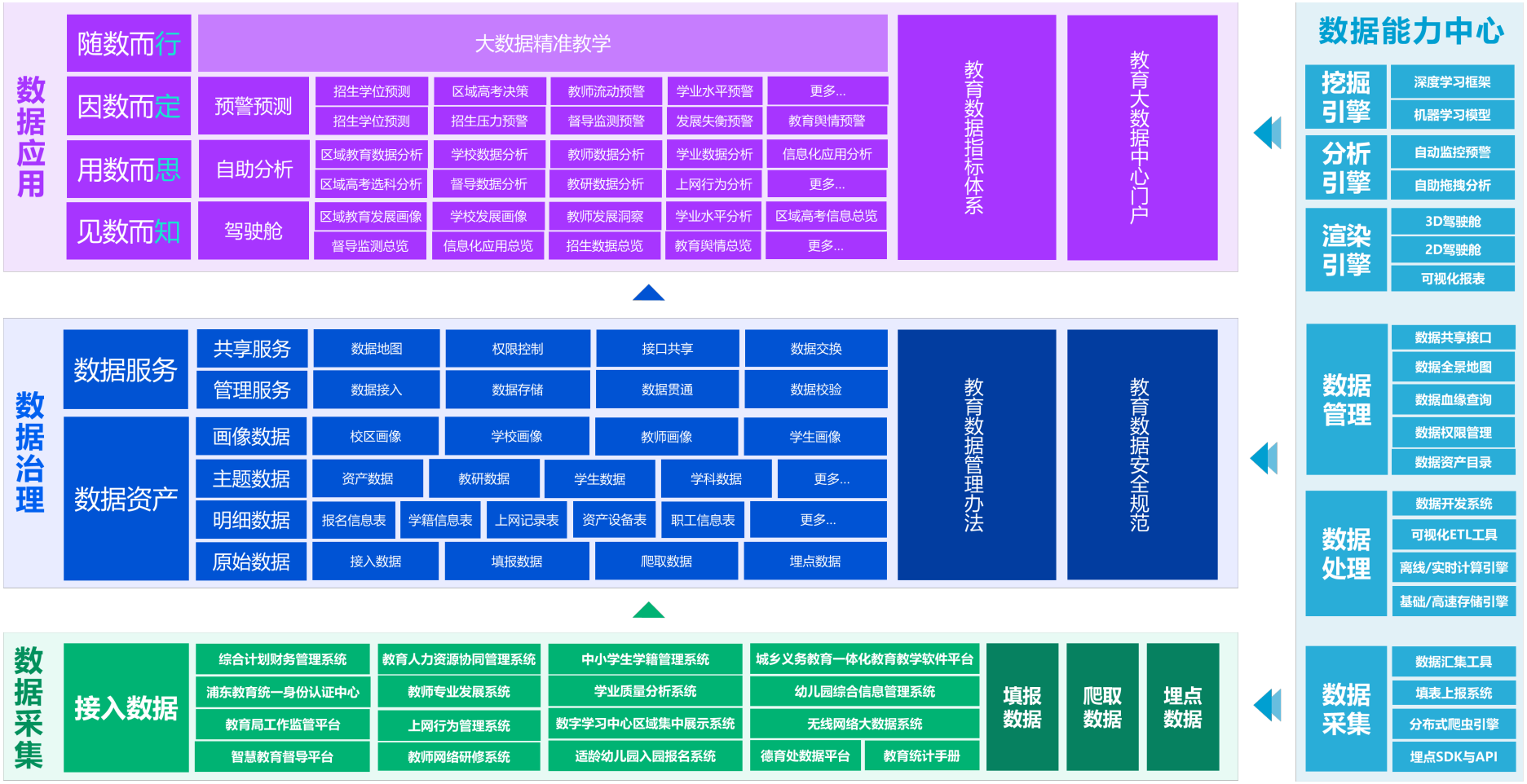
数据中台


前台
进行敏捷开发，不断创新
数据API接口
快速提供服务
数据集中
数据采集
数据接入
数据同步
数据关联
业务创新
数据输出
数据
数据质量
数据管理
中台
提升数据质量
后台
数据开发，稳步向前推进

数据中台—阿里案例


赋能小二-阿里数据
阿里数据PC端
阿里数据APP
数据化运营
社会影响-数据大屏
媒体数据大屏
接待数据大屏
技术大屏
赋能商家-生意参谋
零售电商版
1688版
AE版
CP版
.....
数据资产管理
资产地图
资产分析
资产管理
资产应用
资产运营
面向应用及开放->统一数据服务中间件（OneService）
以业务/自然对象+萃取标签为架构构建（OneID体系）
萃取数据中心
消费者数据体系
企业数据体系
内容数据体系
商品数据体系
位置数据体系
以业务板块+业务过程+分析维度为架构构建（OneData体系）
公共数据中心
电商
文娱
营销
物流
金融
出行
社交
健康
采集/接入
爬取/采购
垂直数据中心
淘宝
天猫
聚划算
AE
妈妈
优土
UC
高德
神马
.....
天气
.....
智能数据研发
数仓规划
模型构建
指标规范
数据同步
数据开发
任务调度
监控告警
计算与存储平台
离线计算
实时计算

数据中台—教育案例


数据应用
随数而行
因数而定
用数而思
见数而知
大数据精准教学
预警预测
自助分析
驾驶舱
招生学位预测
区域高考决策
教师流动预警
学业水平预警
更多...
招生学位预测
招生压力预警
督导监测预警
发展失衡预警
教育舆情预警
区域教育数据分析
学校数据分析
教师数据分析
学业数据分析
信息化应用分析
区域高考选科分析
督导数据分析
教研数据分析
上网行为分析
更多...
区域教育发展画像
学校发展画像
教师发展洞察
学业水平分析
区域高考信息总览
督导监测总览
信息化应用总览
招生数据总览
教育舆情总览
更多...
教育数据指标体系
教育大数据中心门户
数据治理
数据服务
数据资产
共享服务
管理服务
画像数据
主题数据
明细数据
原始数据
数据地图
权限控制
接口共享
数据交换
数据接入
数据存储
数据贯通
数据校验
校区画像
学校画像
教师画像
学生画像
资产数据
教研数据
学生数据
学科数据
更多...
报名信息表
学籍信息表
上网记录表
资产设备表
职工信息表
更多...
接入数据
填报数据
爬取数据
埋点数据
教育数据管理办法
教育数据安全规范
数据采集
接入数据
综合计划财务管理系统
浦东教育统一身份认证中心
教育局工作监管平台
智慧教育督导平台
教育人力资源协同管理系统
教师专业发展系统
上网行为管理系统
教师网络研修系统
中小学生学籍管理系统
学业质量分析系统
数字学习中心区域集中展示系统
适龄幼儿园入园报名系统
城乡义务教育一体化教育教学软件平台
幼儿园综合信息管理系统
无线网络大数据系统
德育处数据平台
教育统计手册
填报数据
爬取数据
埋点数据
数据能力中心
挖掘引擎
分析引擎
渲染引擎
深度学习框架
机器学习模型
自动监控预警
自助拖拽分析
3D驾驶舱
2D驾驶舱
可视化报表
数据管理
数据处理
数据采集
数据共享接口
数据全景地图
数据血缘查询
数据权限管理
数据资产目录
数据开发系统
可视化ETL工具
离线/实时计算引擎
基础/高速存储引擎
数据汇集工具
填表上报系统
分布式爬虫引擎
埋点SDK与API

总结：数据复用的演进过程

数据库	• 克服文件系统不能数据共享和复用的缺陷 • 典型应用：**ERP、CRM、WMS、。。。**	• 数据无法互联互通 • 无法支持企业内部经营决策
数据仓库	• 打通企业内部的信息系统，仅支撑经营决策 • 典型应用：**OLAP**	• 形成更大的信息孤岛 • 不直接面向业务
数据平台	• 数据与业务紧密结合 • 形成局部的数据闭环	• 业务数据无法融合 • 数据应用相互独立
数据中台	• 形成数据、资产、服务、业务全域的闭环 • 在企业内部最大限度地实现数据复用	• 形成数据资产 • 实现数据业务化

总结：迎接数据中台新时代

- 数据中台的需求不是来源于外部，而是来自内部，来自企业对自身未来发展的担忧。数据中台是增援未来，是以发展的观点解决企业面临的问题，面对不确定的未来，企业无法确认今天的数据未来会怎么用，会产生什么样的价值，所以才需要数据中台。
- 现在把数据源源不断地接进来，源源不断地进行资产化、服务化，未来当企业看清楚业务场景，把对数据的需求输入数据中台时，才知道原来数据可以这样使用，才知道怎么去适配。数据中台是对未来场景的能力支撑，是增援未来的能力。