



DTCC

数 / 造 / 未 / 来

第十二届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2021



2021年10月18日 - 20日 | 北京国际会议中心





数 / 造 / 未 / 来
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

滴普数据智能平台建设与应用

滴普科技 谢辉

DTCC
2021



北京国际会议中心

2021/10/18-10/20



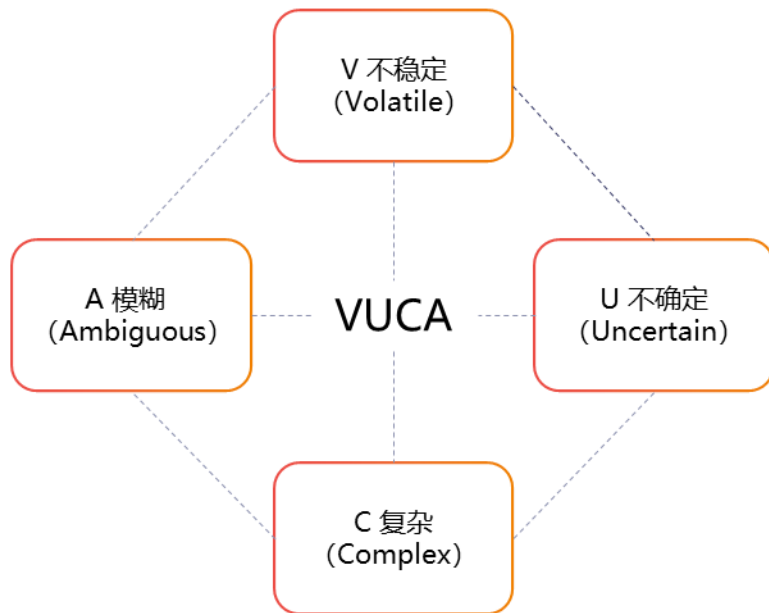
ChinaUnix.net

ITPUB

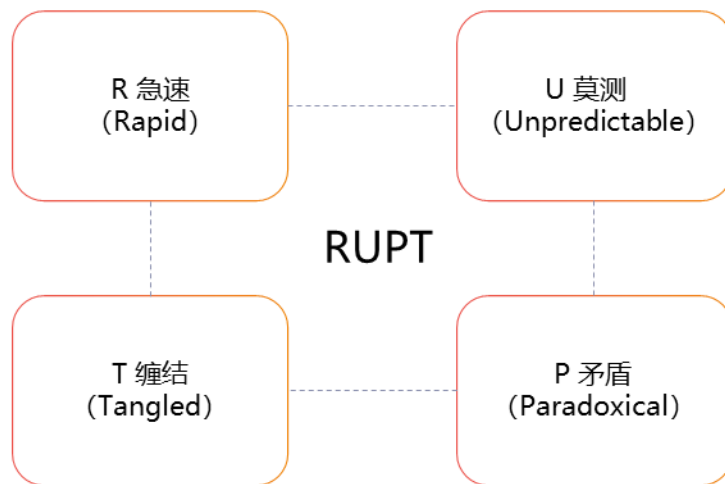
时代的新要求

- 当前的商业环境就是VUCA时代，并且在不久的将来会升级为RUPT时代。在这样一个市场需求高度不确定和不稳定的时代，企业需要打造快速反应市场、指挥前线作战、协调整合资源、管理盈亏的能力，所以需要通过数字化手段构建企业的核心能力

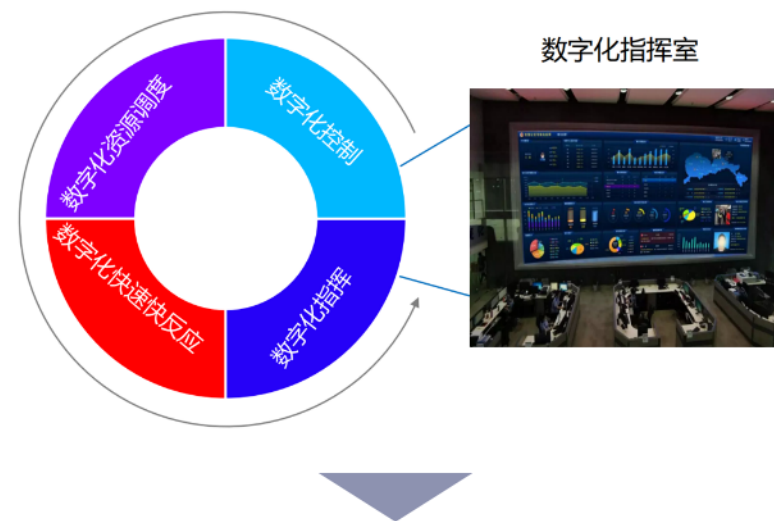
VUCA时代



RUPT时代



应对核心能力



✓ 数据智能平台是支撑敏捷高效的必要基础设施



数 / 造 / 未 / 来



IT168.com

ChinaUnix.net

ITPUB

我们的方案





数 / 造 / 未 / 来
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

数字化咨询

CHAPTER ONE

DTCC
2021



北京国际会议中心

2021/10/18-10/20



ChinaUnix^{net}

ITPUB

P-CBSD(D-ASM)



基于业务场景的数据平台建设方法

Project management-项目管理 (P) 数据智能平台业务咨询项目的项目管理全生命周期。项目管理纳入了预期管理、变更管理、质量管理、商机管理作为基础因素。

Consumer Lifecycle-消费者旅程 (C) 以消费者视角切入，消费者在企业业务的全场景体验调研，以获知-交易-履约-权益作为划分消费者旅程的关键阶段进行拆解梳理。

Run Business-业务运营 (B) 梳理消费者视角的全场景中业务运营在关键环节中具备的运营能力情况。

Empower by System-系统赋能 (S) 对项目范围内的系统现状进行梳理，与业务运营能力进行匹配对齐，梳理系统与业务运营之间的赋能关系。

Data Platform-数据支撑 (D) 对当前业务运营需求的数据支撑现状进行梳理，整理数据在支撑过程中遇到的问题，并与统一大数据平台的数据服务支撑能力作对比，分析二者之间差距。

Technology Architecture of Data Platform-数据平台技术架构 (A) 根据业务、技术、数据输入，识别企业当前数据流向，梳理企业当前数据平台（或传统数仓）技术架构，并结合实际情况规划面向未来的全场景数据智能平台所应包含的大数据技术架构。数据平台技术架构纳入了传统数仓工具、商业数据中间件、开源大数据生态组件、多种编程语言、数据挖掘、AI等技术栈。

Data Structure-数据结构 (S) 涵盖了适用于体现企业当前数据平台并支撑面向未来的全场景数据智能平台能力所应包含的全部数据结构。

Data Model-数据模型 (M) 涵盖了适用于企业面向未来的全场景数据智能平台能力所应构建的数据模型。数据模型纳入了概念模型、逻辑模型、物理模型作为基础因素。

数据智能平台业务咨询

数据治理咨询

数据管理组织

数据管理委员会(DGC)

DGO领导

IT主管

业务数据所有者

业务数据管理者

数据管理办公室(DGO)

数据治理领导

数据治理经理

IT组织人员

数据专家

数据分析师

数据生产者&使用者

数据生产者

数据使用者

数据管理流程

管理数据管理组织

管理政策和流程标准

管理项目启动&变更

管理升级&决议

管理数据定义

管理数据质量&度量

数据管理内容

考核体系

过程考核体系

数据标准

标准定义

标准分类

标准属性

标准编码

标准命名

数据质量

数据监控

数据合规性

数据跟踪&追溯

主数据&元数据

属性定义

采集定义

模板定义

数据安全

数据隐私

数据防护

结果考核体系

数据架构

数据转移

数据存储

数据访问

数据归档

数据封存

数据建模

建模规范

看护规范

概念模型

逻辑模型

物理模型

数据管理规范

数据生命周期

数据创建

数据存储

数据迁移

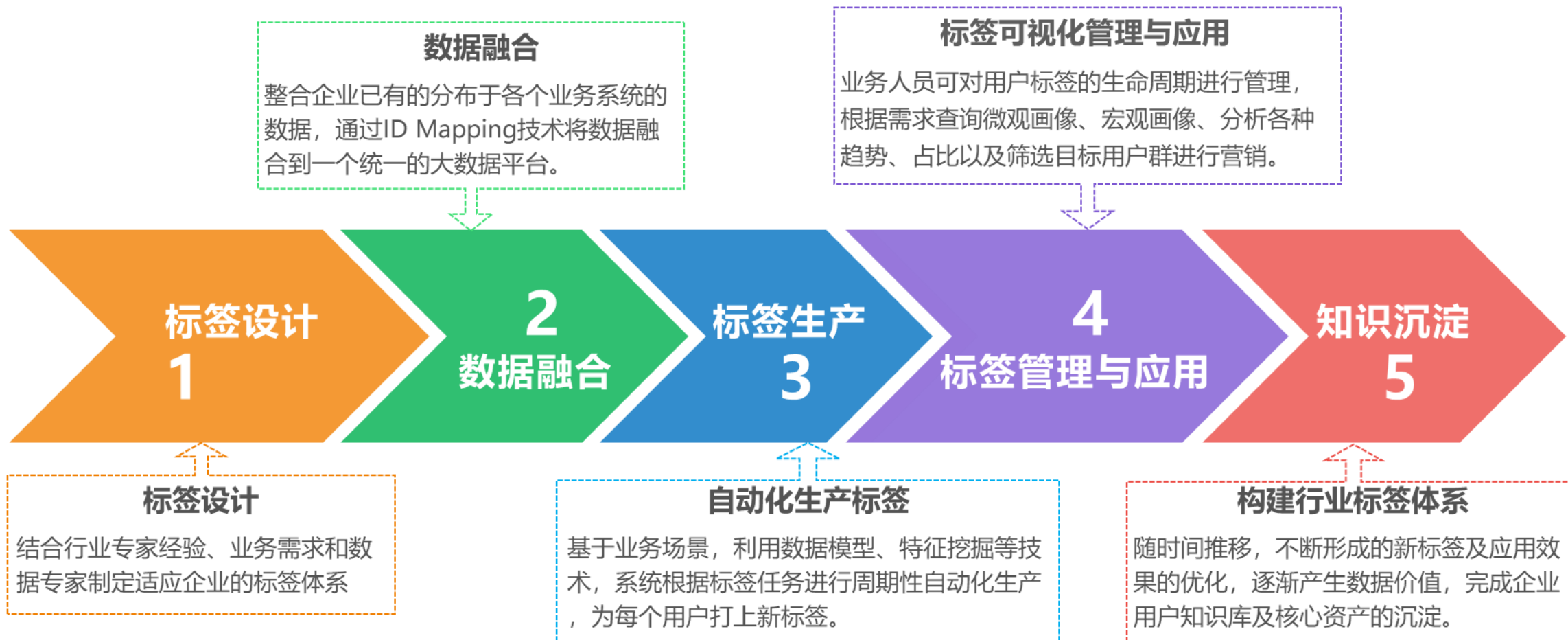
数据使用

数据停用

标签体系设计咨询



DTCC 2021
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021



数，造，未，来





数 / 造 / 未 / 来
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

数据智能平台建设

CHAPTER TWO

DTCC
2021



北京国际会议中心

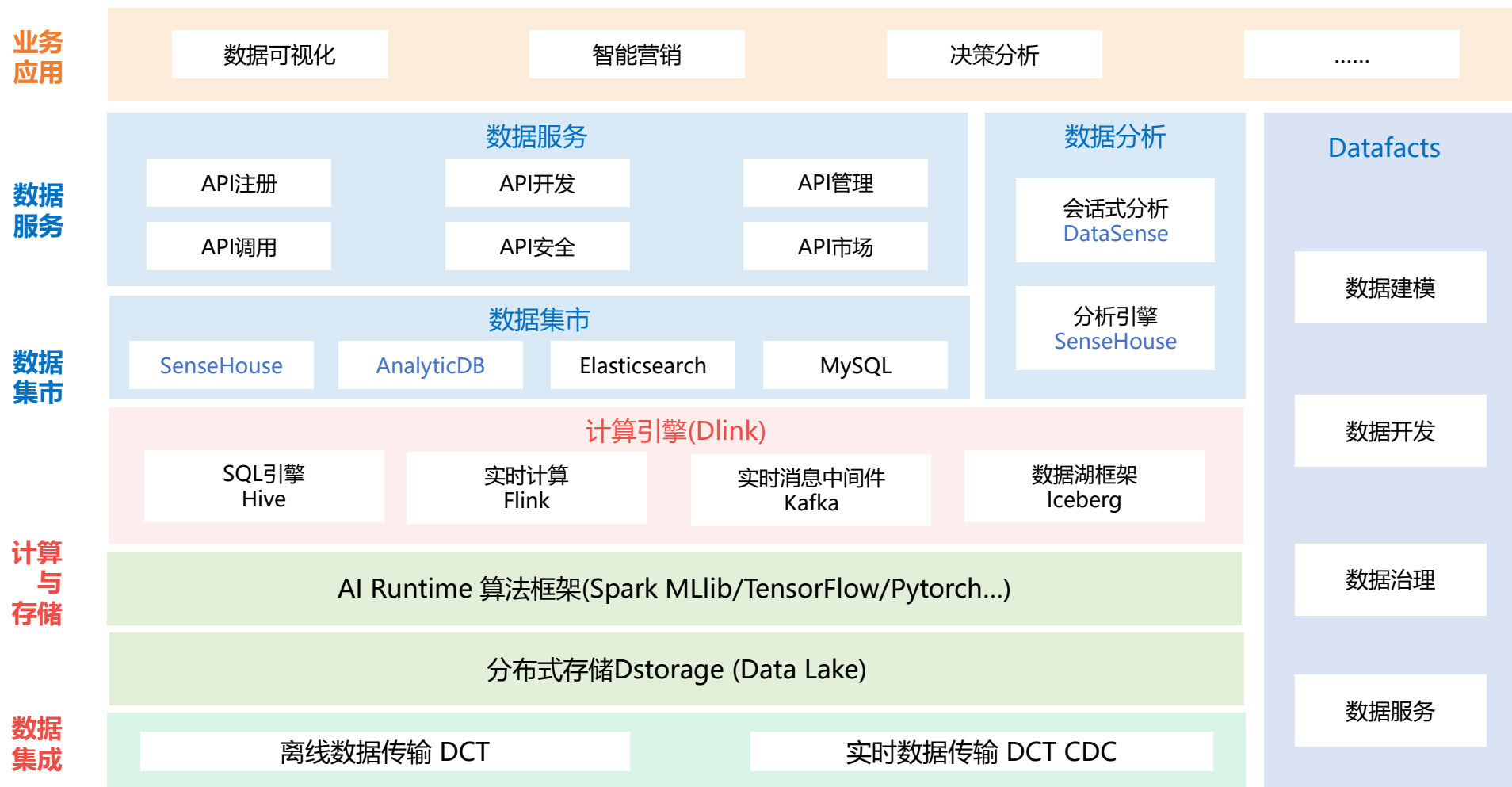
2021/10/18-10/20



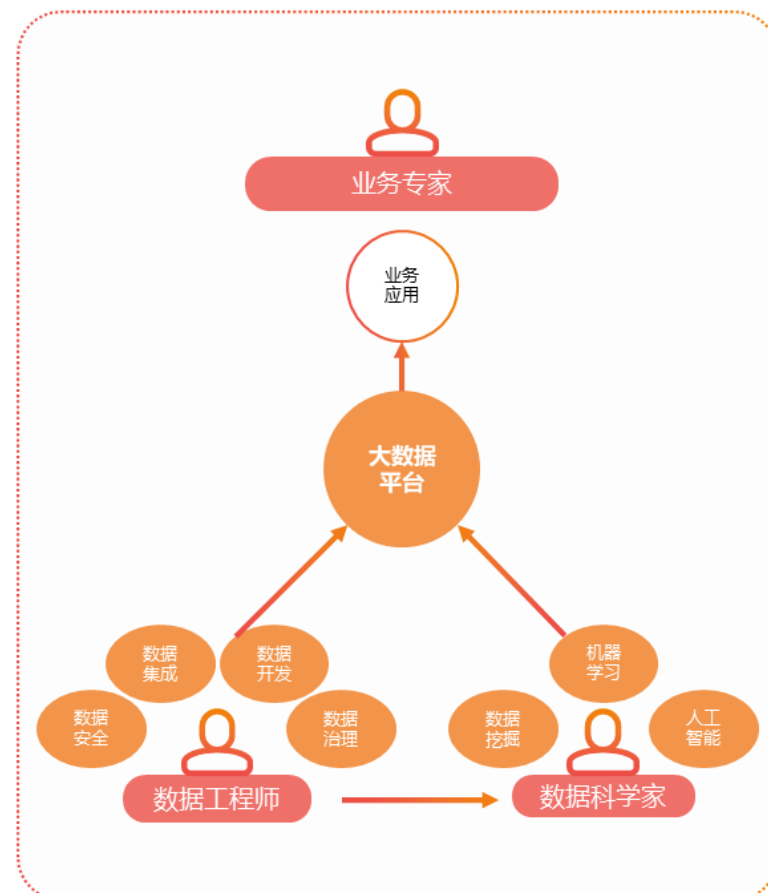
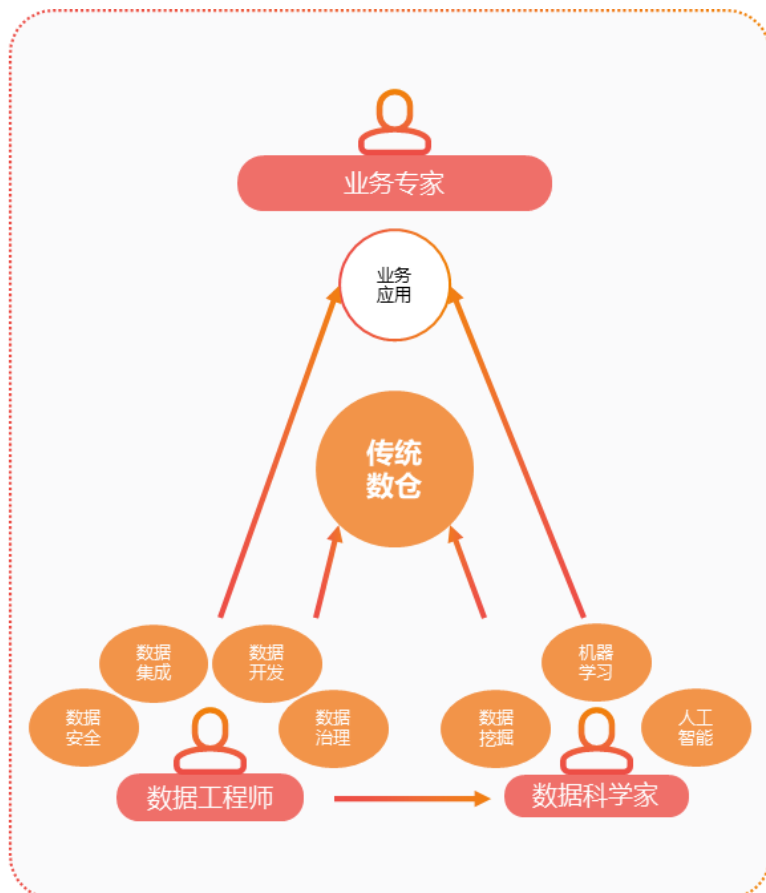
ChinaUnix^{.net}

ITPUB

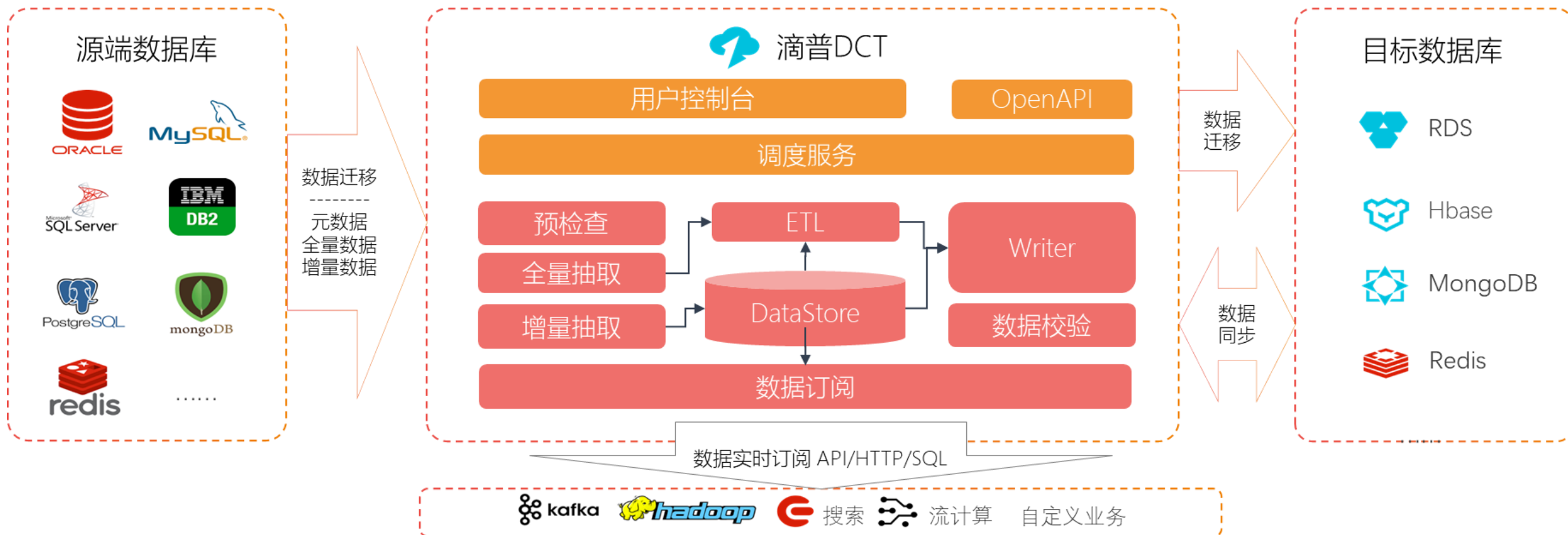
数据智能平台：FastData



数据智能平台搭建后协同的变化



DCT: 数据采集传输



集数据迁移、数据订阅及数据实时同步于一体的数据采集传输服务，解决远距离、毫秒级异构数据传输难题

数据一致性检查、数据压缩、线上运维监控



数，造，未，来





DataFacts: 数据建模、开发和调度监控的智能工厂



产品价值

① 简单易用

提供一站式操作界面和大量辅助型功能, 降低使用门槛

② 任务丰富

支持Jar、shell、python等任务类型, 数据开发高度灵活。

③ 调度可靠

支持千个节点规模的业务流程/主流程可靠调度, 支持动态扩容。

④ 运维便捷

离线、实时运维界面化, 任务异常处理便捷。可接入多种告警方式。



指标一致性难题

一致性指标定义就是定义原子指标、修饰词、时间周期和派生指标的含义、类型、命名、算法，被用于模型设计，是建模的基础；

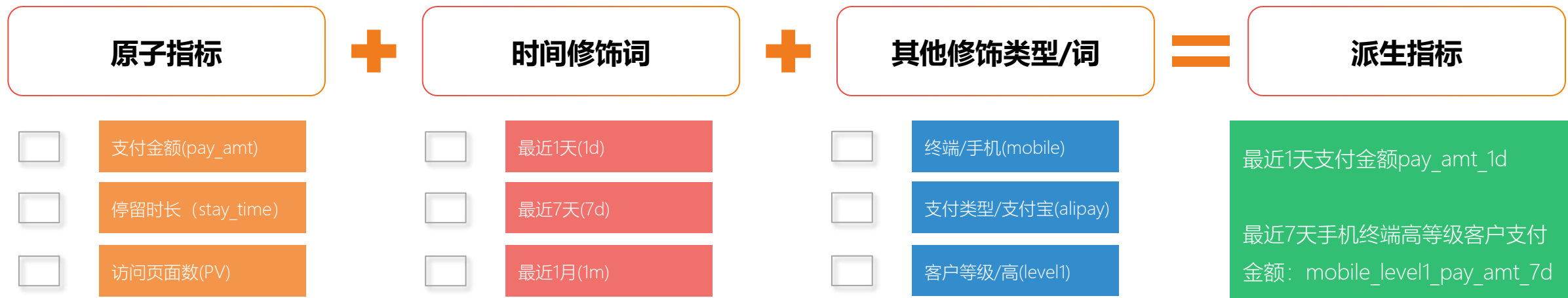
原子指标：是在某业务过程下具有明确的业务含义最小度量单元，原子指标不可再拆分，在设计时以**动作+度量**构成，比如支付金额(pay_amt)、停留时长(stay_time)。

时间修饰词：数据统计的时间范围或者时间点，如最近1天、最近一周、截止当天。

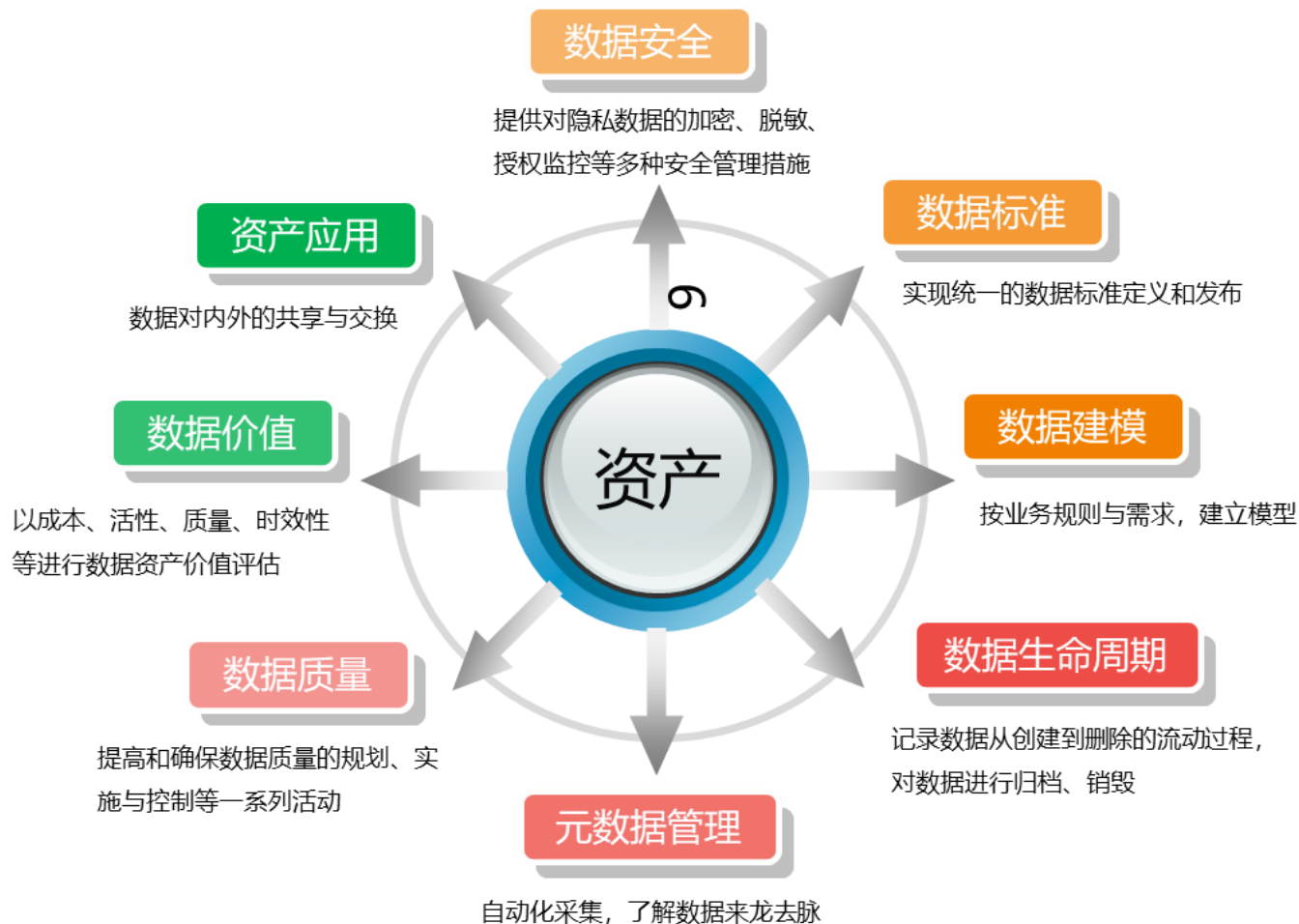
其他修饰词：是对指标业务场景的限定，如客户等级高、终端为PC；修饰词归属到某个抽象的修饰类型，如：客户等级、终端类型。

派生指标：由**原子指标**、**时间周期修饰词**、以及**其他修饰词**组合而成。其中原子指标为单一选择，时间修饰词为单一选择，而其他修饰词可以多选。派生指标继承了原子指标所在的数据域、数据类型、算法，保障指标的一致性。原子指标、时间修饰词、其他修饰词的命名确定后，派生指标的命名也被确认下来，如下面示例。

上述原子指标、时间修饰词、其他修饰词定义规范有**专业团队负责，沉淀到工具**，工具作为模型定义入口，保障一致性指标的落地。



DataFacts: 数据资产管理中心



① 智能数据建模结构化管理

支持数仓规划设计、制定并沉淀企业数据标准、维度建模、数据指标定义，将建模设计产出的维度表、事实表和汇总表物化到计算引擎中并进一步应用

② 流动数据生命周期管理

从数据产生、存储、应用到销毁的整个流动过程，对数据进行定义存储策略、冷热计算、数据归档、销毁等全生命周期监控

③ 全域元数据汇集及数据跟踪

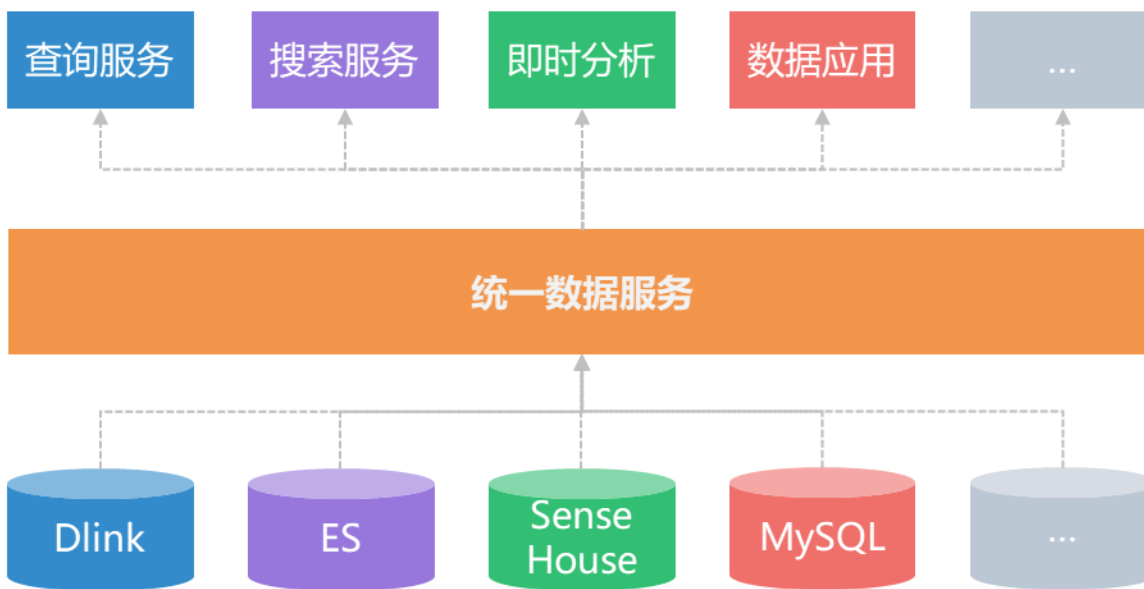
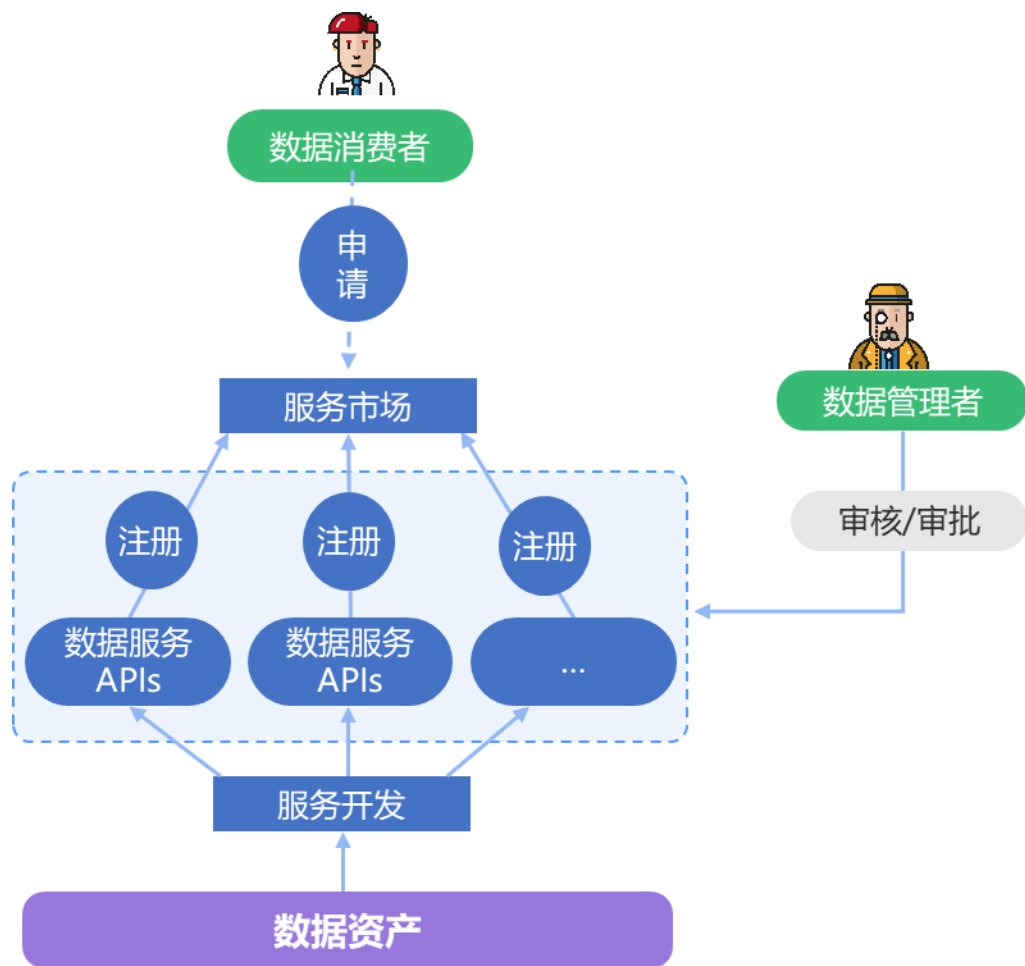
支持多源数据采集，多维度管理，基于元数据，提供血缘分析，影响分析，快速定位数据，跟踪元数据变更

④ 全链路监控数据质量

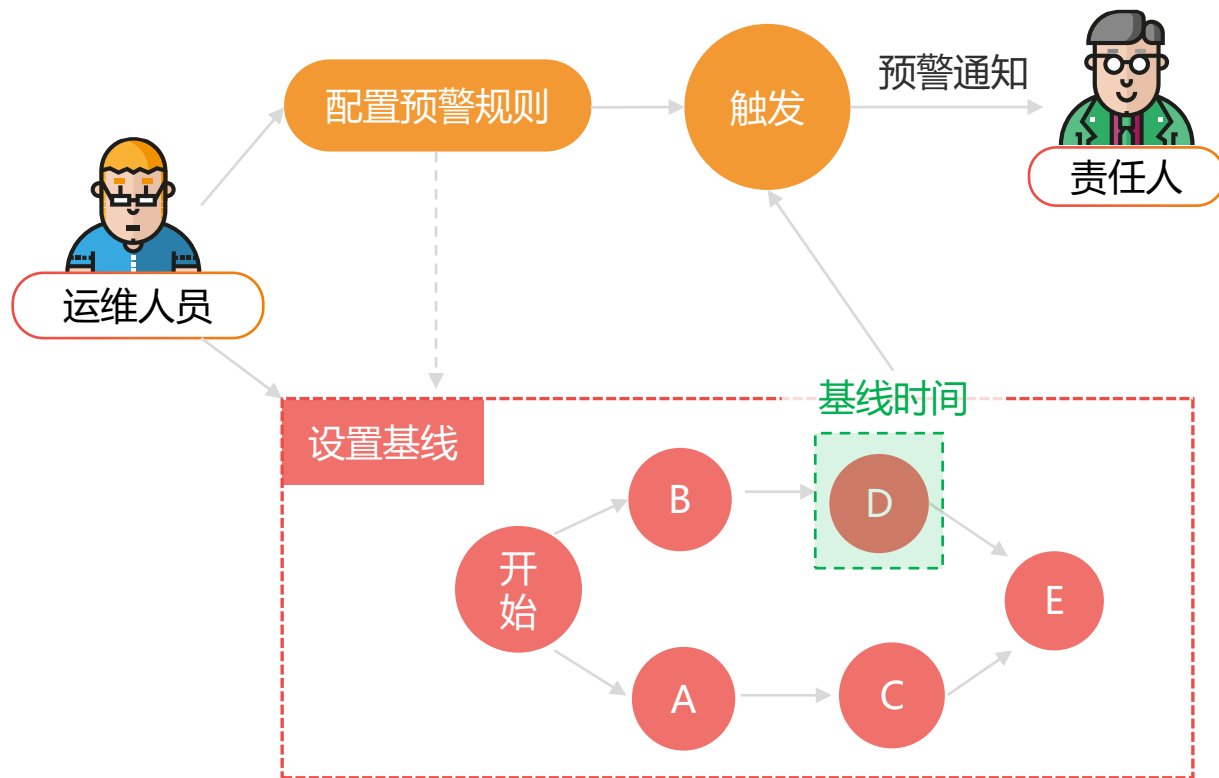
支持以数据标准做质量检核，可检测检查或关联ETL任务做数据稽查，将质量检核、问题告警、指派责任人、质量整改等工作环节进行流程整合，形成完整的数据质量管理闭环



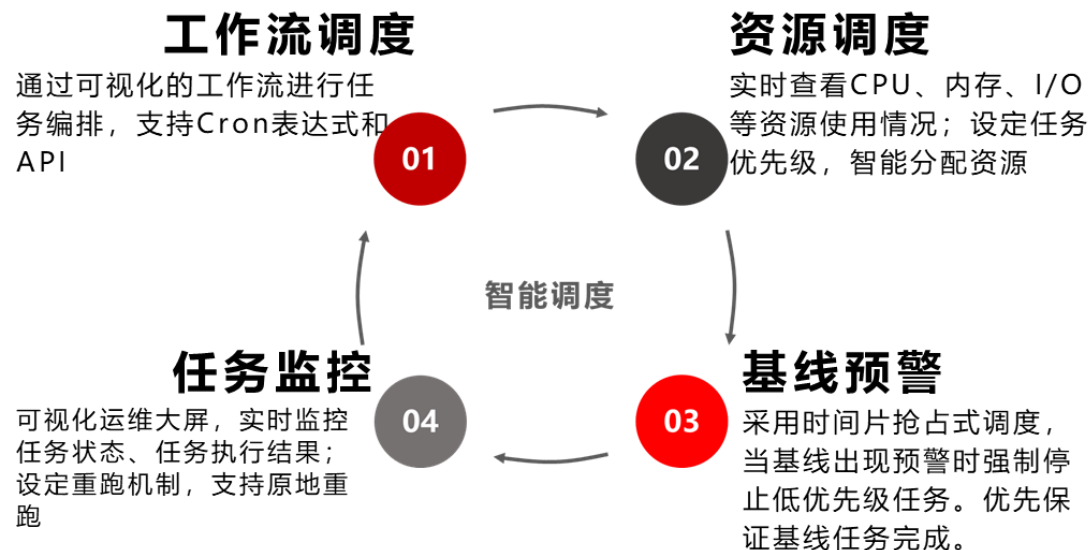
数据智能服务



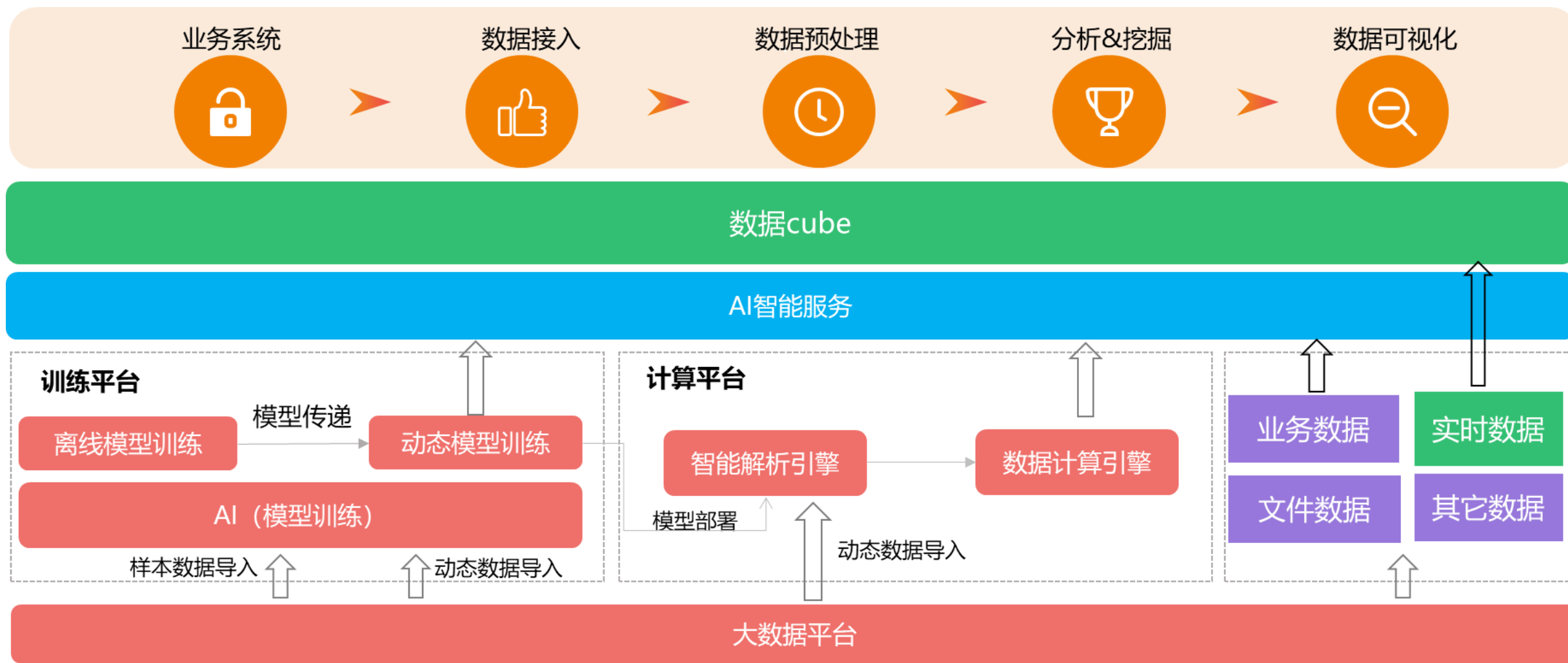
平台智能运维



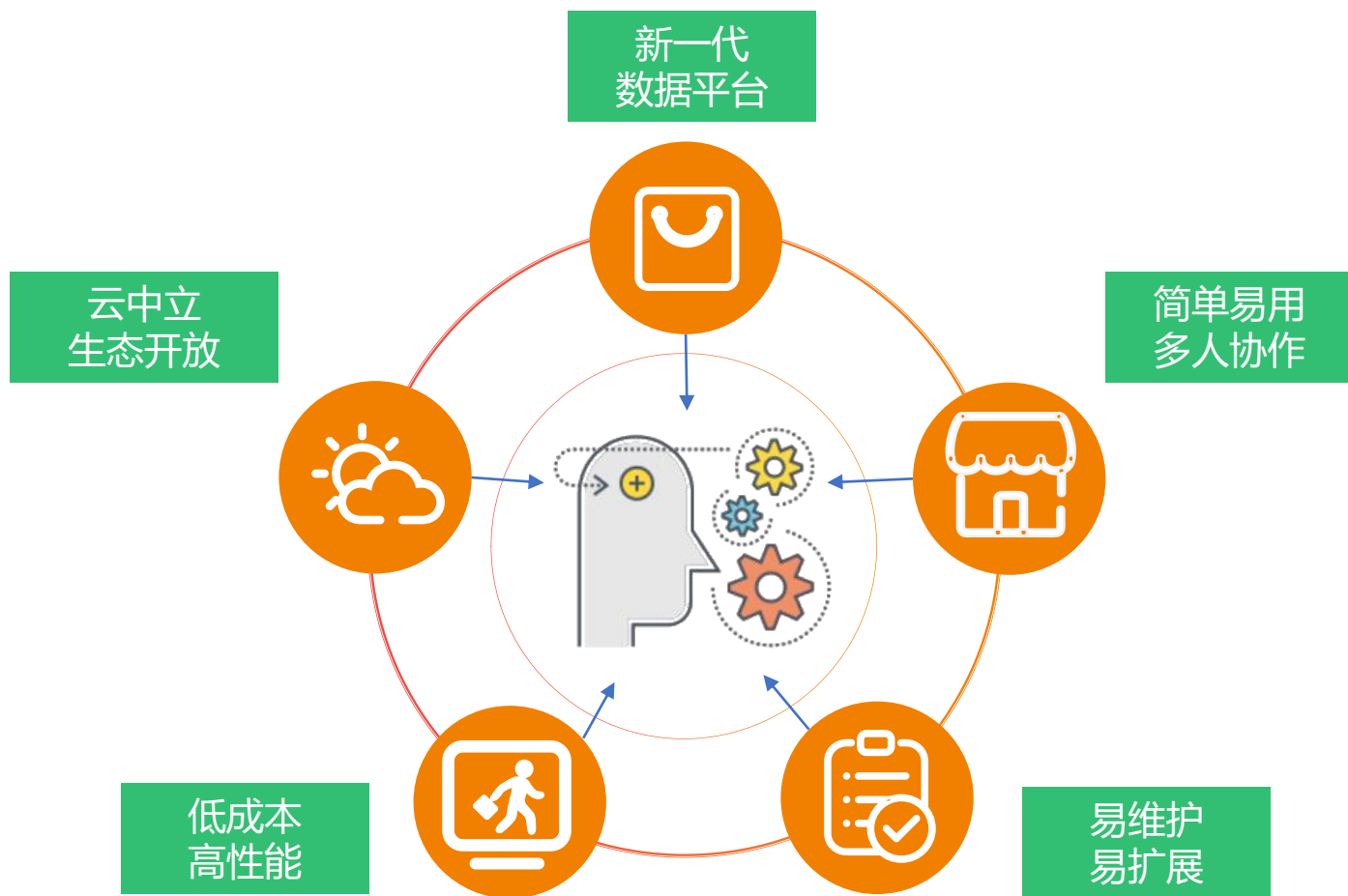
智能调度技术



DataSense: 会话式搜索



产品优势





数 / 造 / 未 / 来
第十二届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

数字化运营共创

CHAPTER THREE

DTCC
2021



北京国际会议中心

2021/10/18-10/20



ChinaUnix^{net}

ITPUB

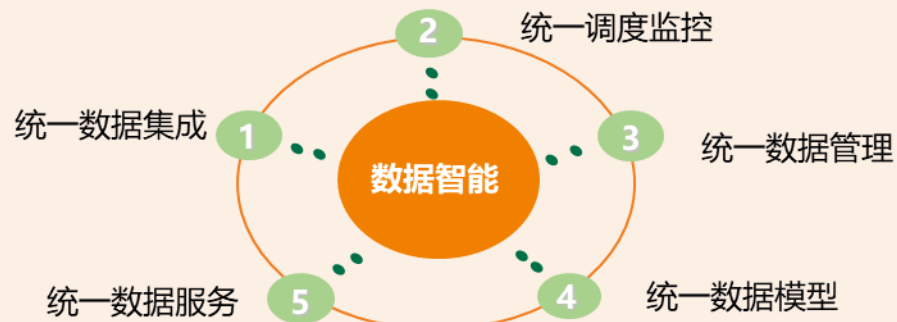
运营共创最佳实践

原则

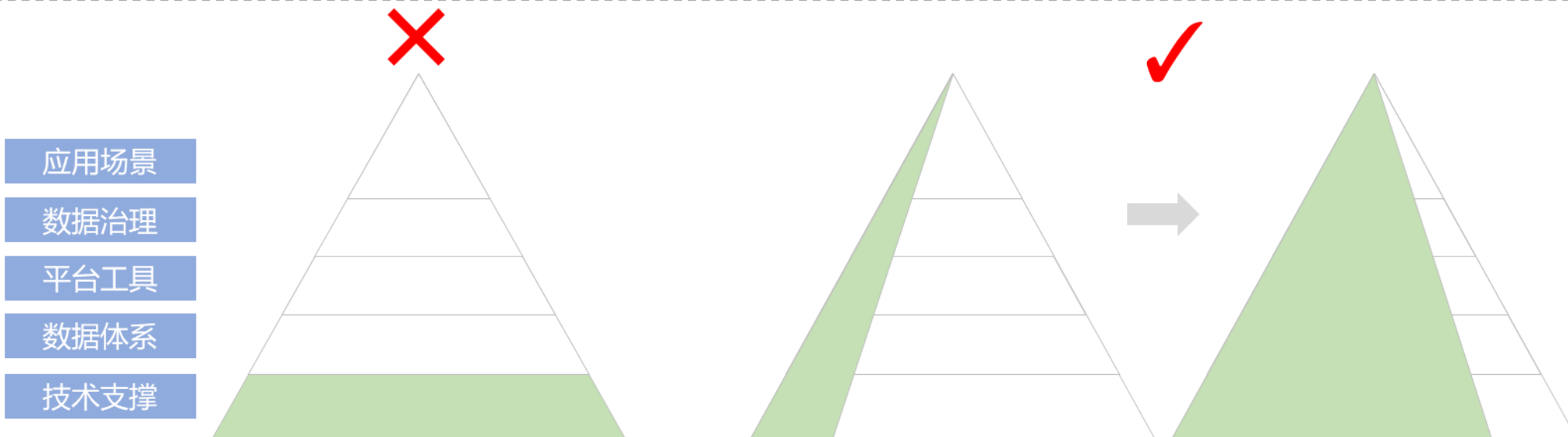
建设原则：

- ✓ 不影响现有业务系统的情况下进行，从**解决实际问题的**角度出发，从始至终围绕明确的目标；
- ✓ 先设计蓝图，落地从局部到全面，逐步优化，**循序渐进**的改进；
- ✓ 产品和技术是基础，建设的核心是**数据资产**和**数据应用**

设计原则：“一个系统、五个统一”



路径



某品牌数据智能平台建设

业务数据字典管理

通过统一的平台提供统一的数据信息和协作模式，管理业务的指标、标签和维度，实现业务与技术团队的协作共创

数据资产治理

通过数据资产管理平台，显示数据标准、数据质量，数据安全的规范管理，实现数据资产的可见、可管理、可运营

数据服务管控

通过统一的数据服务管理平台，实现数据服务流程的在线协作，并提供安全可靠的服务管控



数据采集

1. 实时同步
2. 离线同步
3. 数据清洗
4. 传输监控



数据开发

1. 实时开发
2. 离线开发
3. 工作流调度
4. 运维监控



数据资产

1. 业务数据字典
2. 数据质量
3. 数据模型
4. 元数据管理
5. 数据地图
6. 数据权限



数据服务

1. API统计
2. API管理
3. 应用管理
4. 策略管理



商品智能配补调

计划

采购

出厂

干线
运输

仓储

货品
管理

支线
运输

销售

售后



数，造，未，来



某医药集团数据智能平台建设

客户背景：年销售收入超过400亿元的高度专业化创新型医药上市集团，位居中国制造业500强。

客户诉求 or 痛点

- 1、减少手工报表，提升数据质量；
- 2、提升业务响应时效性；
- 3、核心业务指标梳理、构建；
- 4、规划建设大数据平台。

客户价值

- ✓ **管控集中化**：构建集团级数据仓库，涵盖企业内外部数据，实现公司总部、基地、医贸的数据集中。
- ✓ **分析标准化**：建立标准化的数据分析，统一数据标准，提升数据质量，提高协同运营效率。
- ✓ **管理精细化**：通过数据平台的建设，准确反映财务核算，库存状态，物流效率等企业经营情况，提升决策效率与水平。
- ✓ **响应灵活化**：建立统一数据平台，及时响应与满足业务端灵活、多样化的业务需求，不断适应市场环境政策的变化。

实现场景

报表分析体系

实现分析指标，分析方法的标准化，跨越企业内部复杂的业务与组织层级，实现钻取分析，强化业务整合与协同能力。

数据资产管理及应用

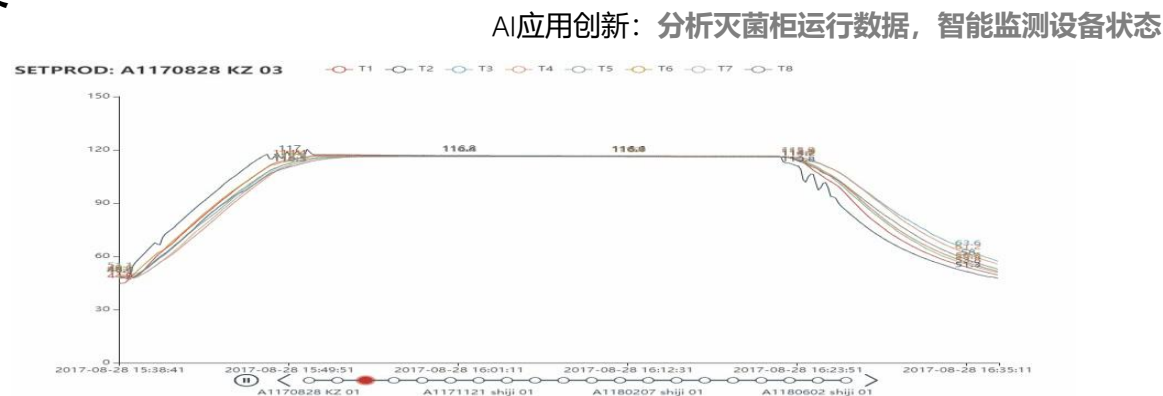
构建统一的大数据平台，实现科伦数据资产的集中管理和运营，有效支持与提升企业管理决策能力。

智能监测预警

图形化灭菌柜监测数据，比理想曲线，预警差异信息；通过算法平台发现异常批次，对设备主动采取维护活动。

实际效果

管理指标体系



口径统一

5 大业务版块
10+ 系统打通

管理提效

229 张管理报表
17 个生产基地

研发降本

开发效率提升 30%
人力成本节省 45%

AI创新

7 亿多条监测数据
2 天找到“黄金曲线”



THANKS