

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





快手数据治理实践 – 基于评估体系实现可持续治理

快手 孙伟



孙伟



快手

- 数据平台部-数据管治组负责人



阿里巴巴

- 数据技术及产品部-数据技术专家



百度

- 商业基础数据团队-数据研发工程师



01

为什么需要数据治理？

02

怎么做好数据治理？

03

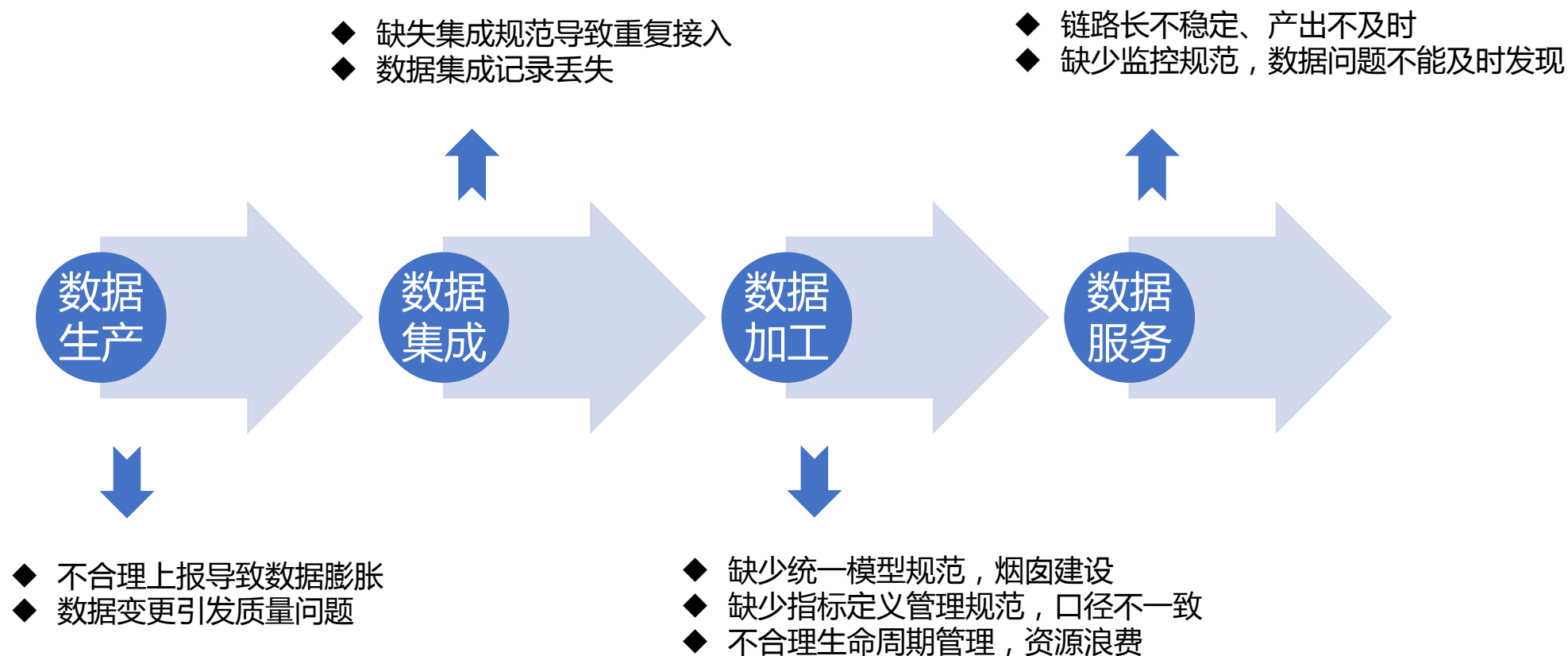
如何建设治理评估体系？

04

总结&展望



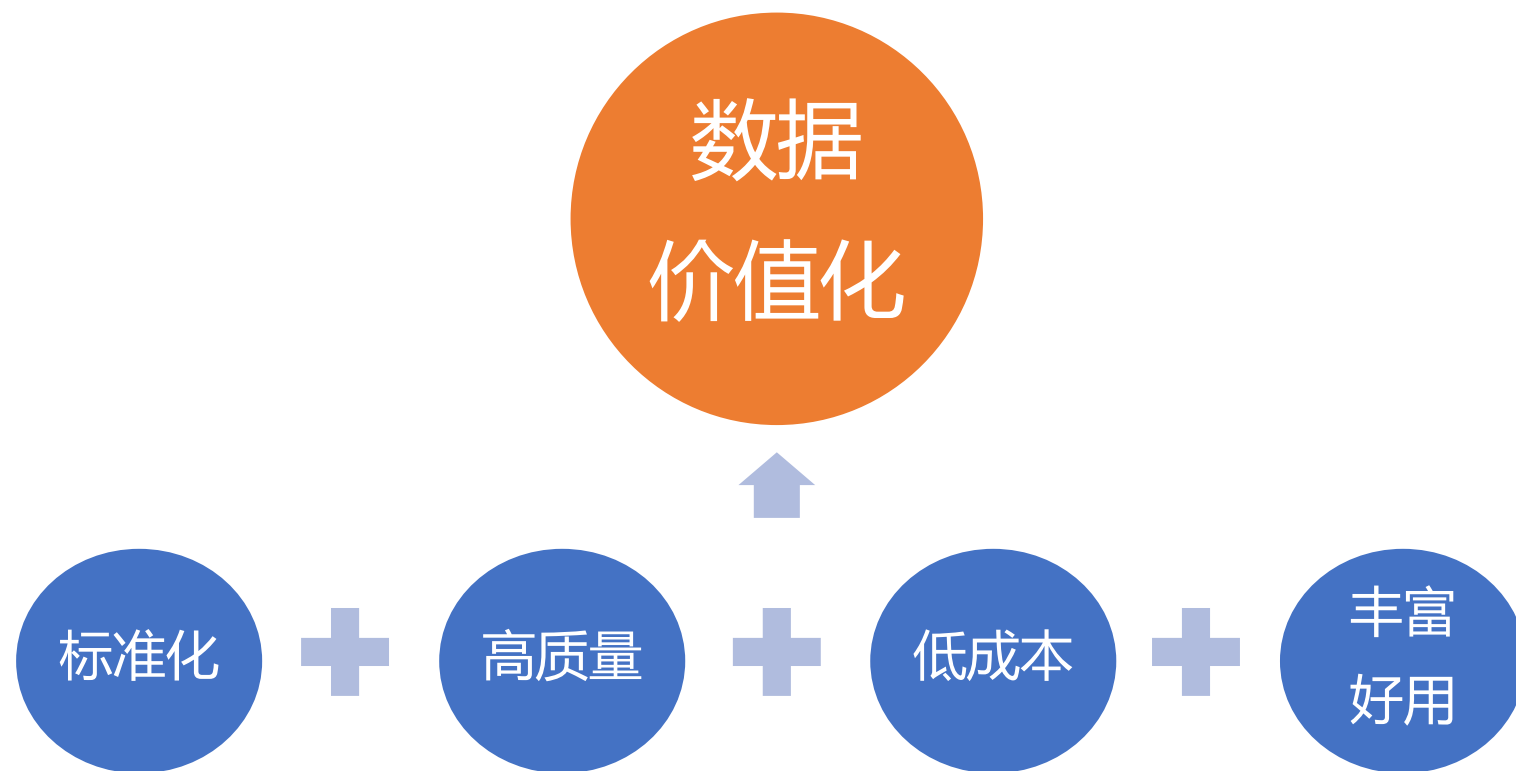
01 为什么需要数据治理？



01 为什么需要数据治理？



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



数，造，未，来



01

为什么需要数据治理？

02

怎么做好数据治理？

03

如何建设治理评估体系？

04

总结&展望



02 怎么做好数据治理 - 快手遇到了哪些挑战

链路长问题多

目标不清晰

运营缺少抓手

治理动力不足



如果无法度量，
则无法治理

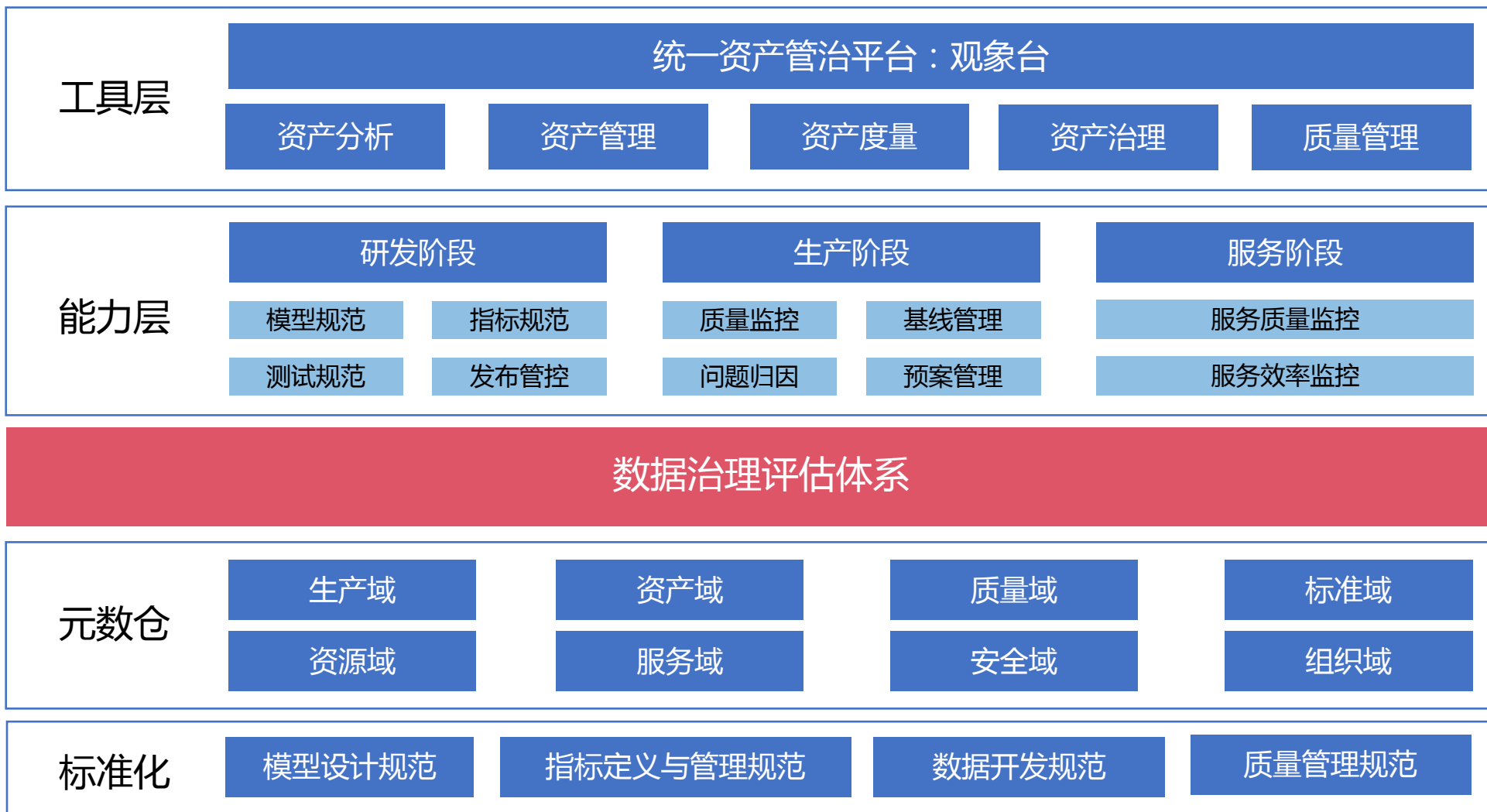
不可持续



02 怎么做好数据治理 - 快手解决方案

数据治理组织 & 流程

数据治理运营机制



数，造，未，来



01

为什么需要数据治理？

02

怎么做好数据治理？

03

如何建设治理评估体系？

04

总结&展望

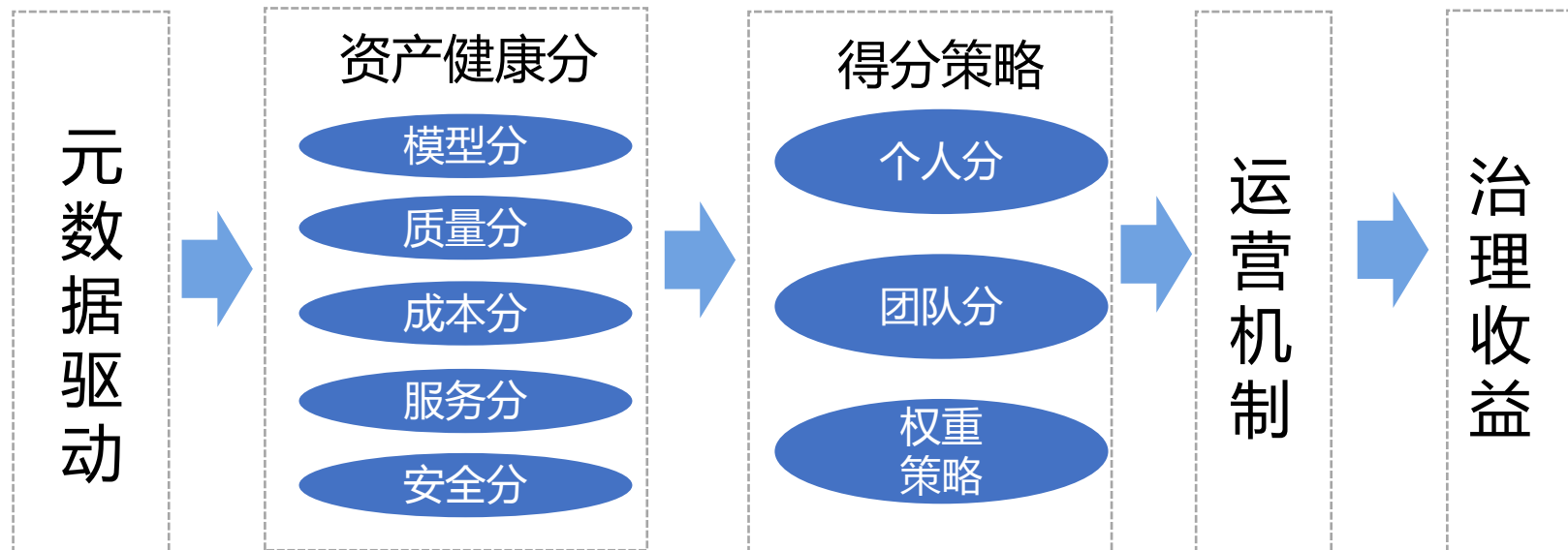


03 如何建设治理评估体系 - 目标



目标

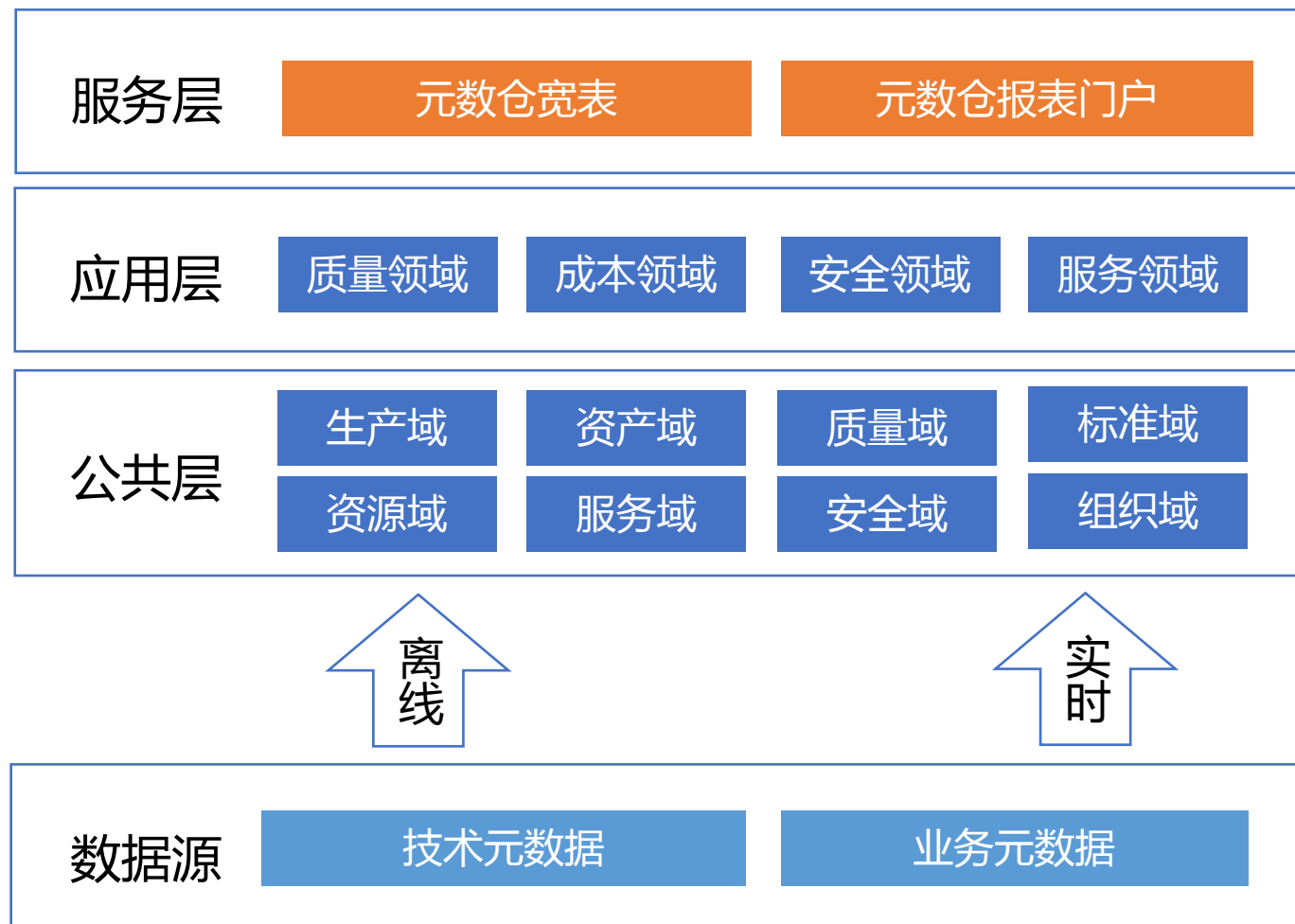
- ◆ 问题标准化
- ◆ 资产可量化
- ◆ 过程策略化
- ◆ 运营有抓手
- ◆ 治理可持续



03 如何建设治理评估体系 - 元数仓

问题

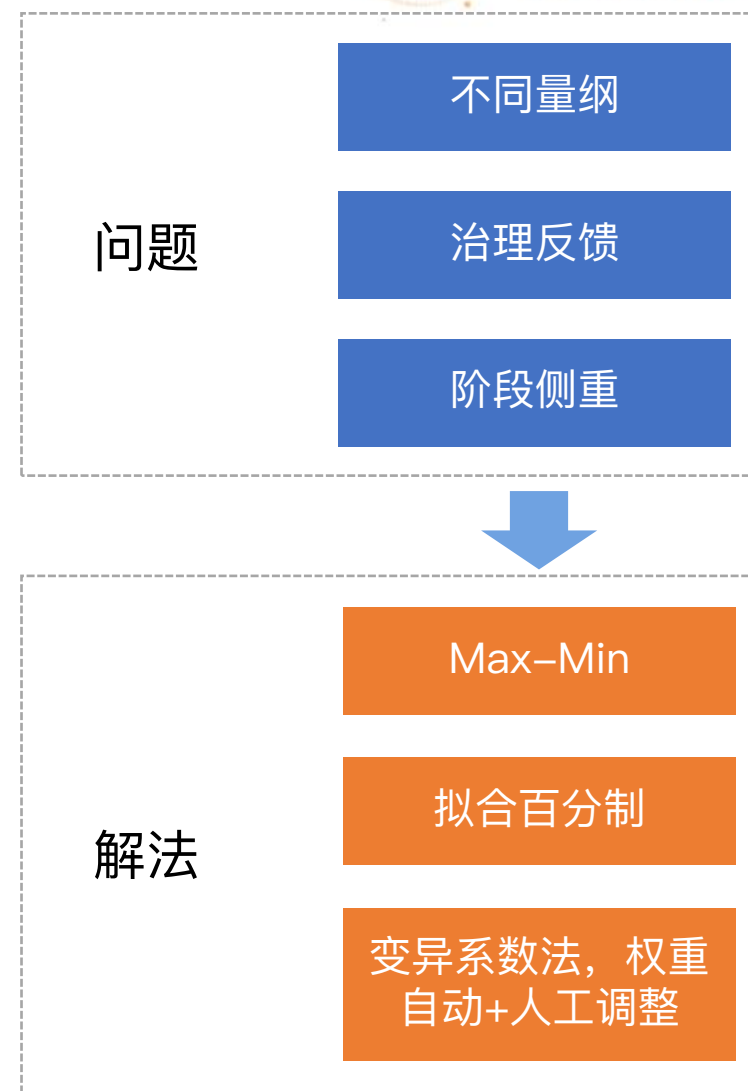
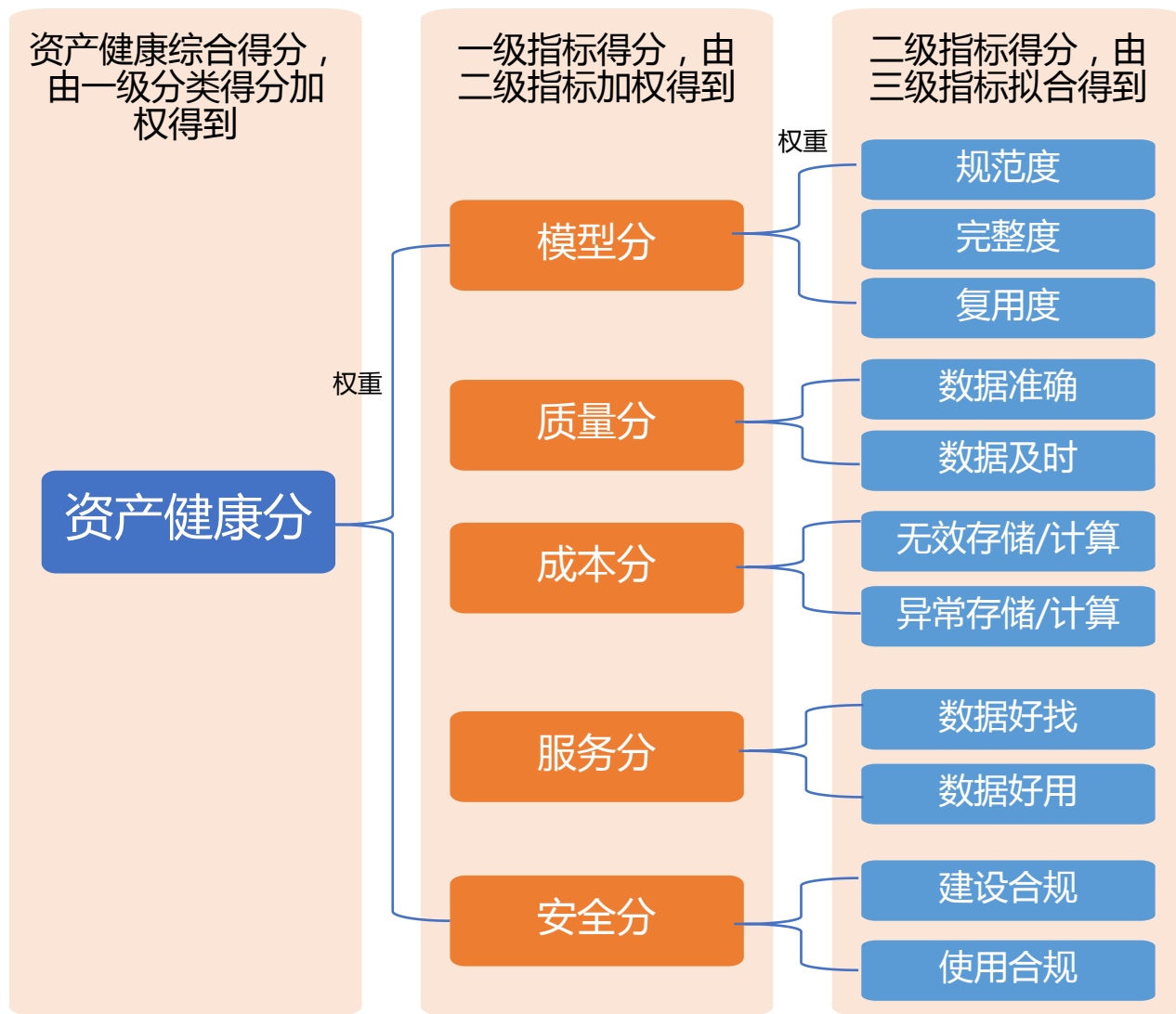
- ◆ 「标准」本身是否标准
- ◆ 系统分散口径不统一
- ◆ 时效性不够
- ◆ 历史数据缺少沉淀



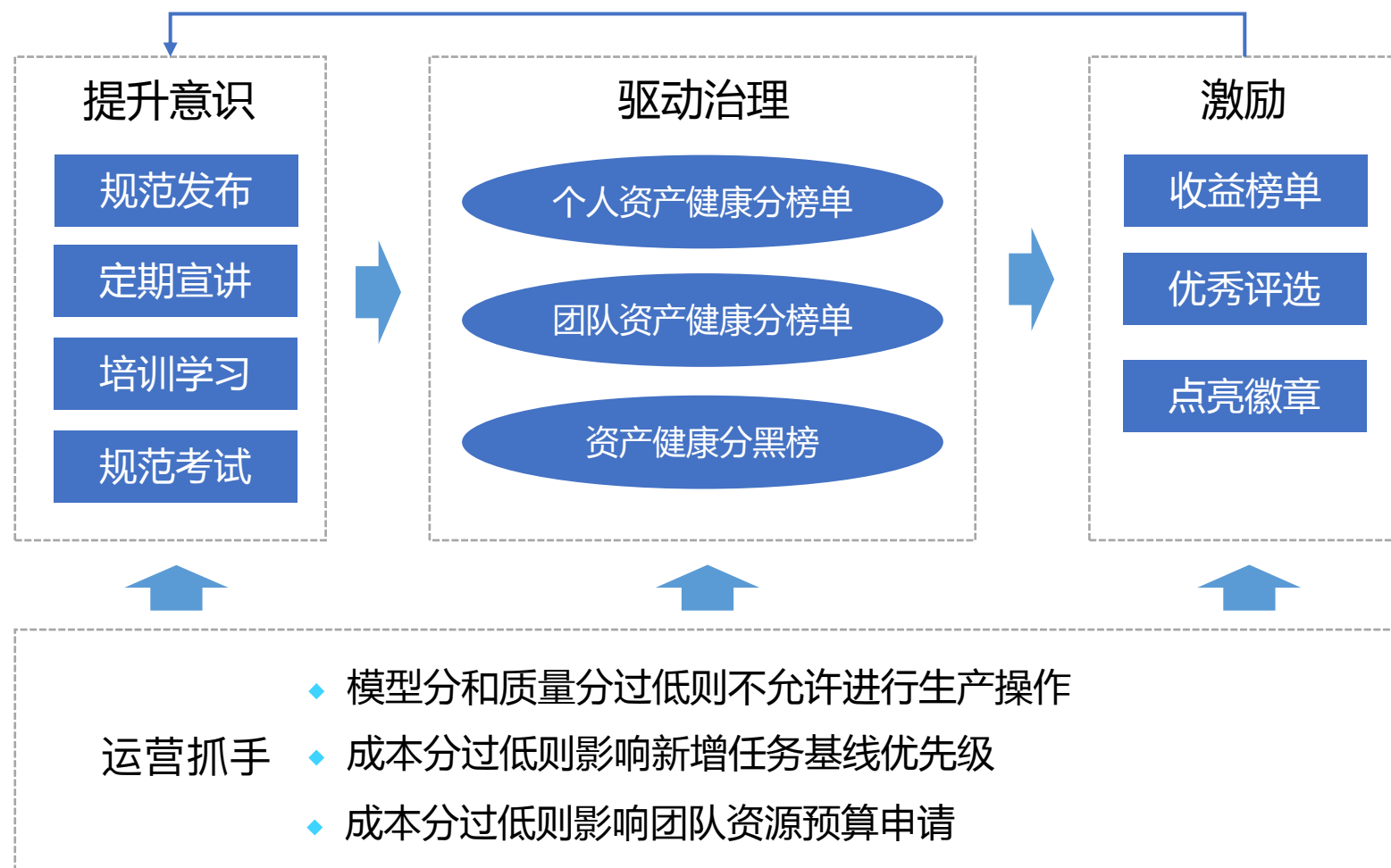
元数仓建设标准规范



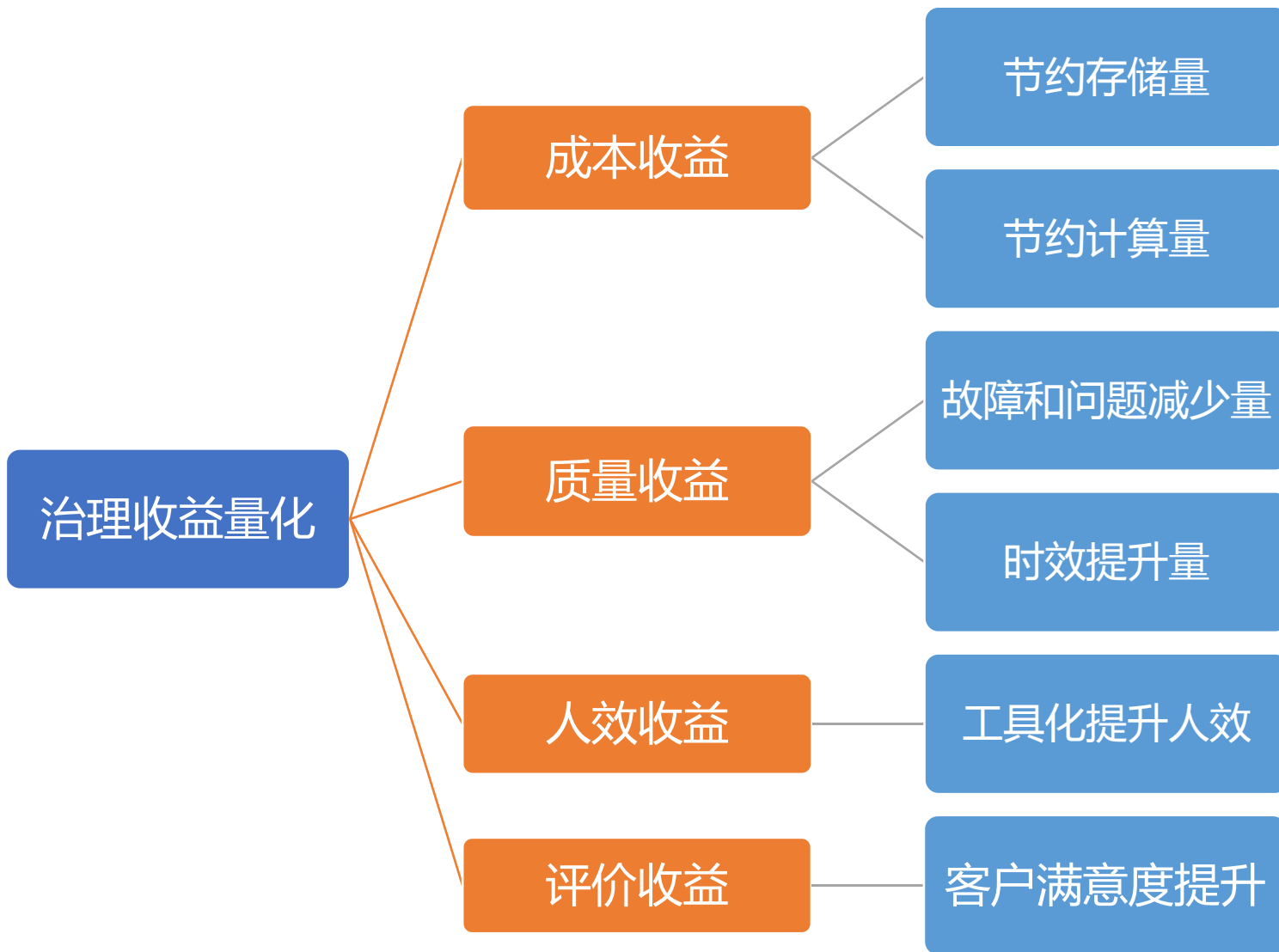
03 如何建设治理评估体系 - 资产健康分



03 如何建设治理评估体系 - 运营机制



03 如何建设治理评估体系 - 治理收益量化



03 如何建设治理评估体系 - 快手降本提质实践

问题

- ◆ 数据源和加工链路质量问题多、发现不及时
- ◆ 模型不合理建设，导致资源浪费严重
- ◆ 治理手段原始，效率低，收益无法回收



整体收益：资产健康分整体从58分提升到了80+分，治理参与度超过95%

- ◆ 存储资源优化超过300PB，计算资源优化超过100TB

**成本
收益**

- ◆ 落地平台化的治理能力，并结合榜单形成自驱式治理，提升专项治理效率

**人效
收益**

**质量
收益**

- ◆ 经过阶段性治理，质量故障和问题数降低超过50%

**评价
收益**

- ◆ 业务对数据部门的满意度调查得到显著提升



03 如何建设治理评估体系 - 快手降本提质实践

标准化流程

Plan

- ◆ 问题定义
- ◆ 问题量化
- ◆ 治理方案

P

D

A

C

DO

- ◆ 元数据驱动
- ◆ 专项治理
- ◆ 自驱治理

Check

- ◆ 量化目标持续监控
- ◆ 榜单持续运营

Act

- ◆ 资产健康分覆盖治理规则
- ◆ 工具化手段提升治理效率

能力建设

完善标准

- ◆ 质量管理流程规范
- ◆ 模型设计规范
- ◆ 指标管理规范
- ◆ 成本管理规范

量化标准

- ◆ 资产质量分
- ◆ 资产成本分

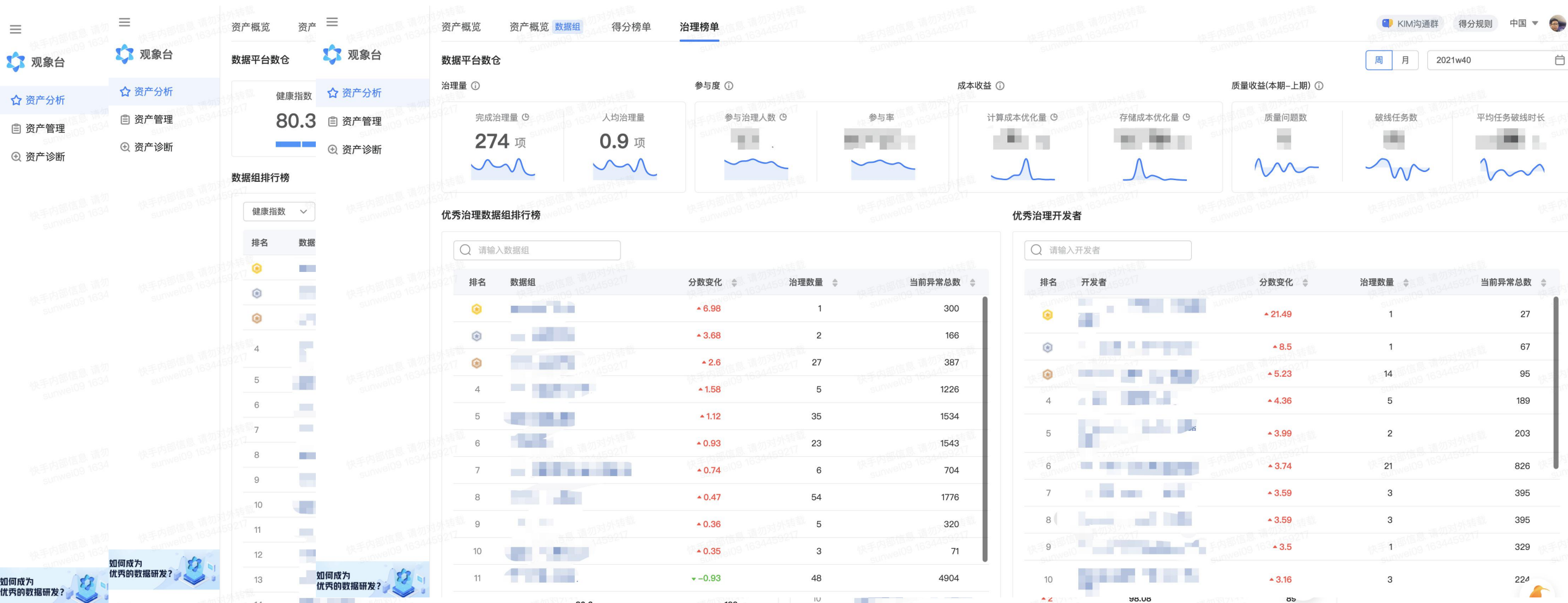
工具化

- ◆ 数据源结构/内容监控工具
- ◆ 数据测试工具
- ◆ 数据发布规范检查工具
- ◆ 数据监控工具升级

运营&流程

- ◆ 质量分榜单、待治理问题榜单
- ◆ 横向质量小组
- ◆ 模型评审小组
- ◆ 数据质量考试

03 如何建设治理评估体系 - 快手降本提质实践



01

为什么需要数据治理？

02

怎么做好数据治理？

03

如何建设治理评估体系？

04

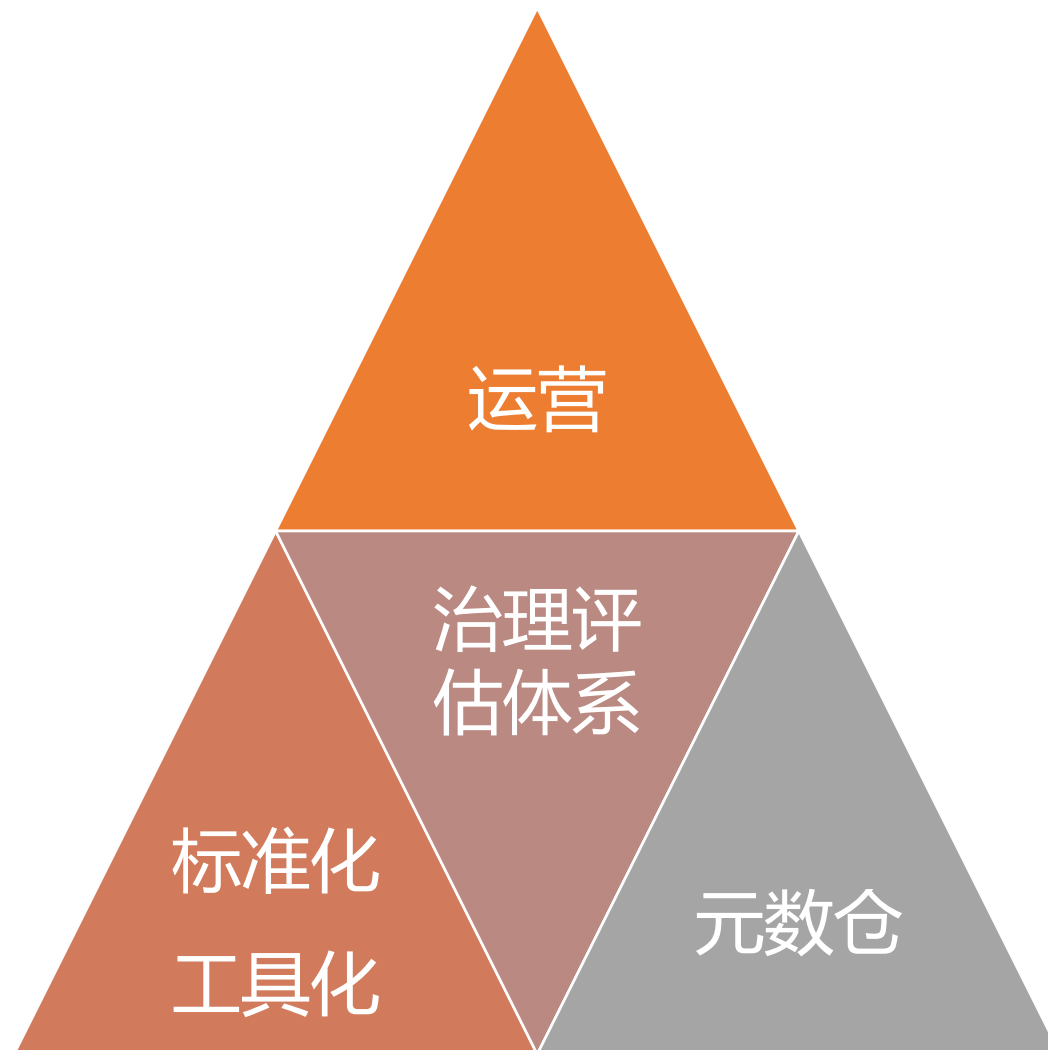
总结&展望



04 总结&展望 - 可持续治理的关键因素总结



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



数 / 造 / 未 / 来



IT168.com

ChinaUnix.net

ITPUB

04 总结&展望 - 未来规划



- ◆ 最优的治理就是不需要治理，结合规范与治理评估体系，将部分治理内容前置到生产环节中，例如数据测试工具、模型管理工具、自动化监控配置工具等

- ◆ 打通系统边界，实现平台化一键治理能力，提高治理效率，例如一键删除/下线表、一键改生命周期等



欢迎交流

快手搜索

生活的各种问题 就上快手搜索

快手搜索 生活的问题 Q

视频搜索上快手

用生活回答每一种生活

- ◆ 海量大数据挑战
- ◆ 丰富业务场景实践
- ◆ 专业大数据团队
- ◆ 行业大牛面对面交流

快手数据平台公众号



快手数仓团队公众号





THANKS