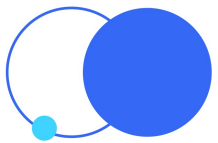


如何打造标准化的数据治理 评估体系

孙伟

大数据研发专家



自我介绍

孙伟 快手大数据研发专家

- 主要负责快手数据管治方向工作
- 主要关注数据架构和数据管治领域



1. 数据治理背景
2. 数据治理评估体系
3. 快手实践效果
4. 未来规划



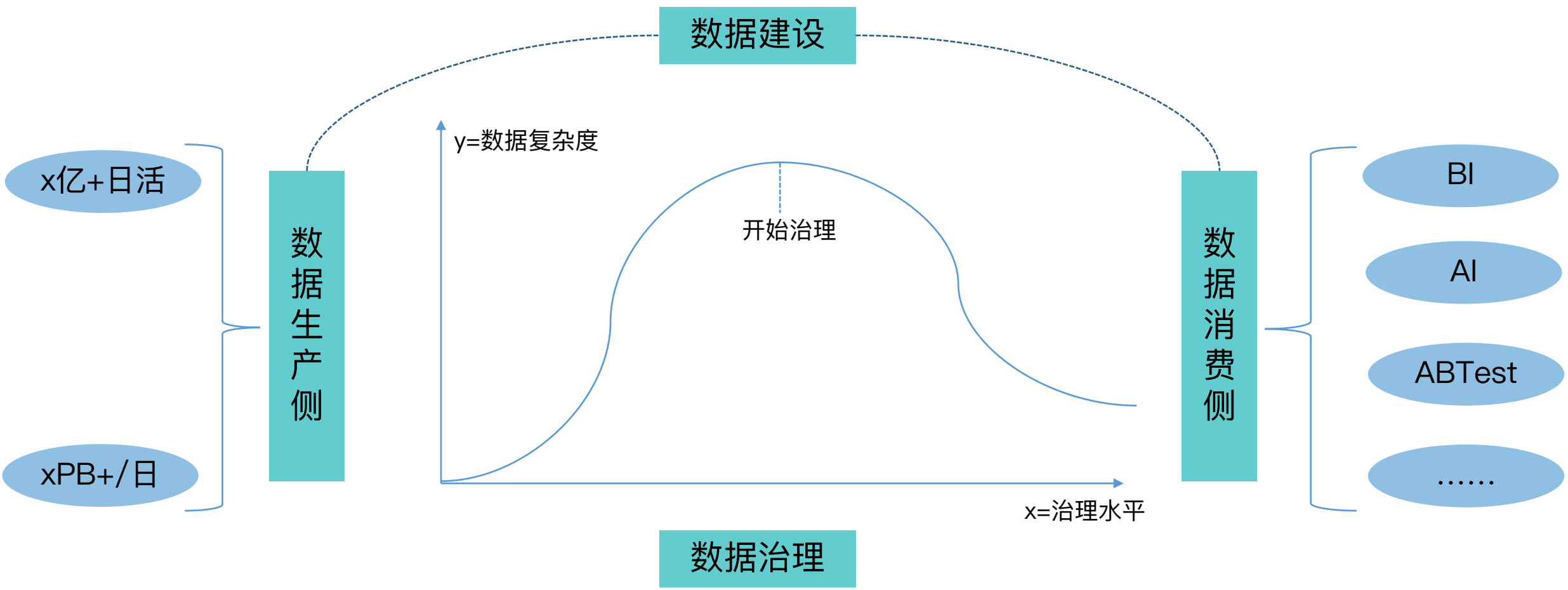
1. 数据治理背景

2. 数据治理评估体系

3. 快手实践效果

4. 未来规划

01 背景：为什么需要数据治理



01 背景：数据治理面临的挑战

链路长，问题多并且分散，无法聚焦

数据从生产、加工到服务整个链路很长，任何一个环节都可能存在需要治理的问题，如何标准化？

怎么衡量资产好坏，治理到什么程度算好

问题如何量化？并且治理了一段时间，结果怎么衡量？

应该重点解决哪些问题，并且不同阶段的治理侧重不同

治理进入深水区了怎么办？如何动态调整治理的方向？

治理动力不足

怎么提高意识并驱动各个团队进行治理？





1. 数据治理背景

2. 数据治理评估体系

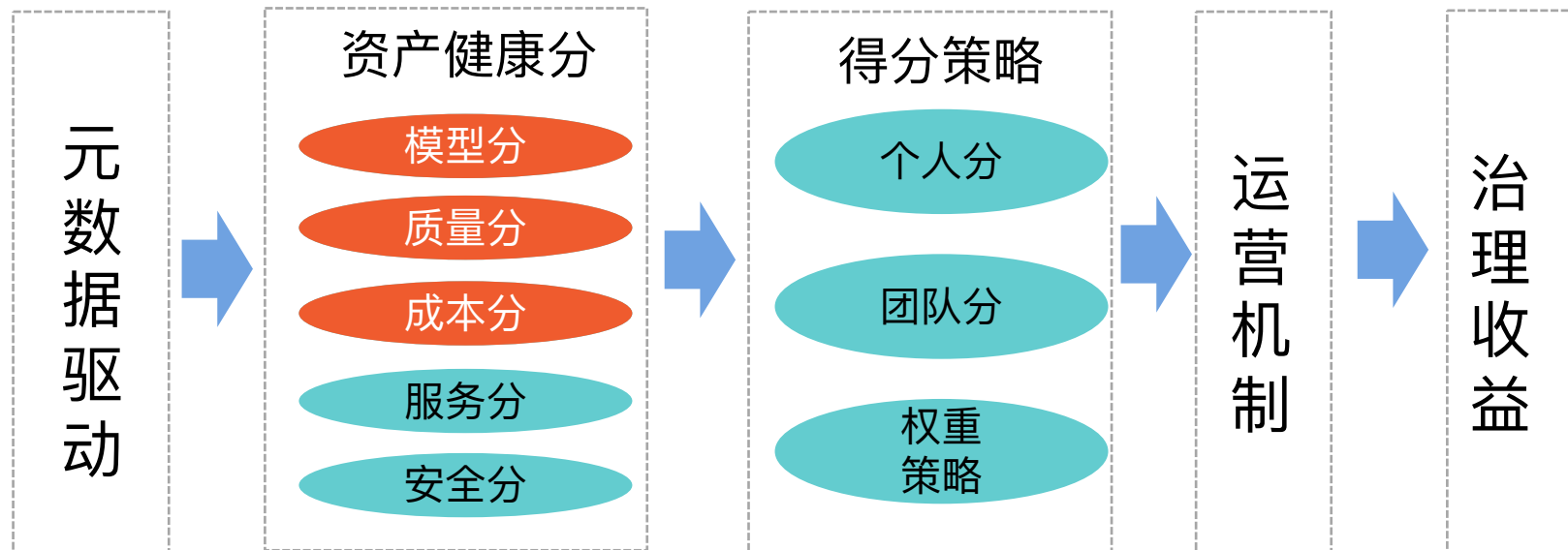
- ◆ 目标与策略
- ◆ 资产健康评估
- ◆ 运营机制
- ◆ 治理收益评估

3. 快手实践效果

4. 未来规划

目标

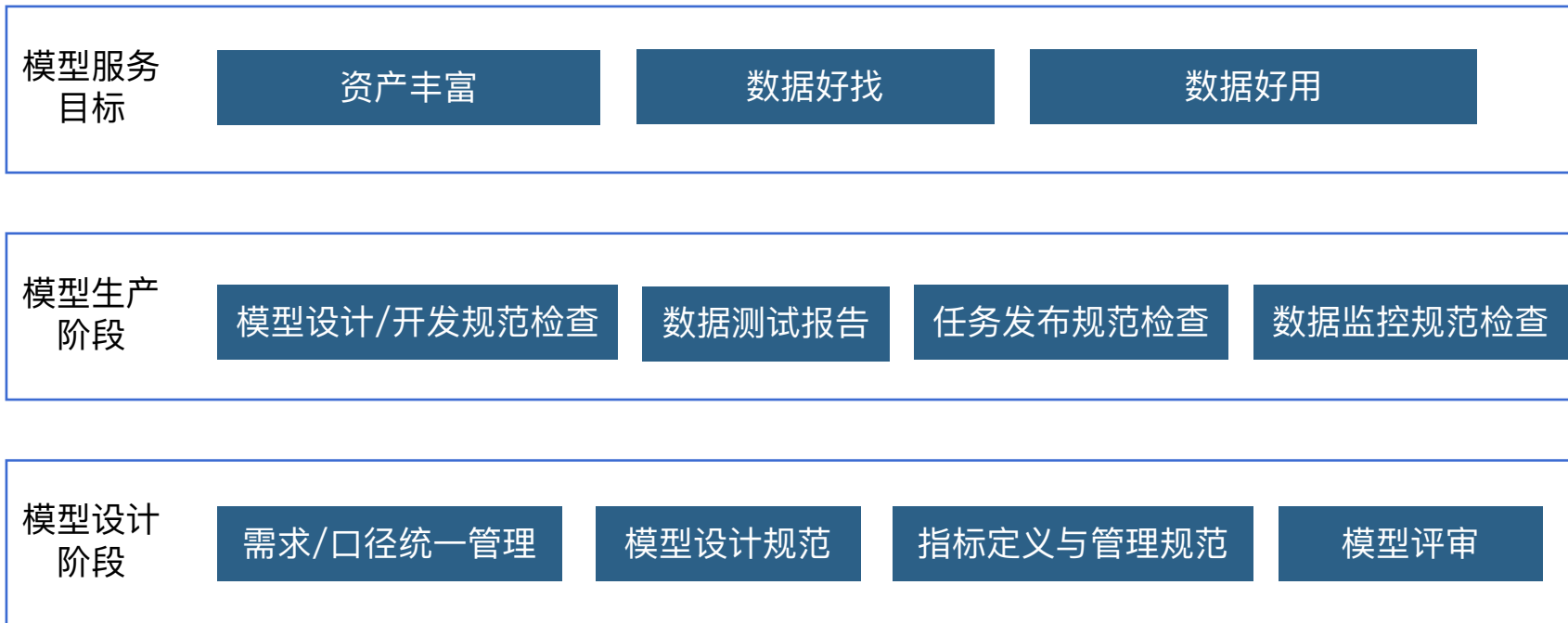
- ◆ 问题标准化
- ◆ 治理可量化
- ◆ 过程策略化
- ◆ 运营有抓手



02 资产健康评估：模型（问题与解决方案）

问题

- ◆ 无数据可用
- ◆ 口径不一致
- ◆ 分析效率慢



数据模型评估体系





02

资产健康评估：模型（评估）

规范性

- 定义规范
- 开发规范
- 测试规范
- 发布规范
- 元数据规范

复用性

- 模型宽度
- 模型深度

完整性

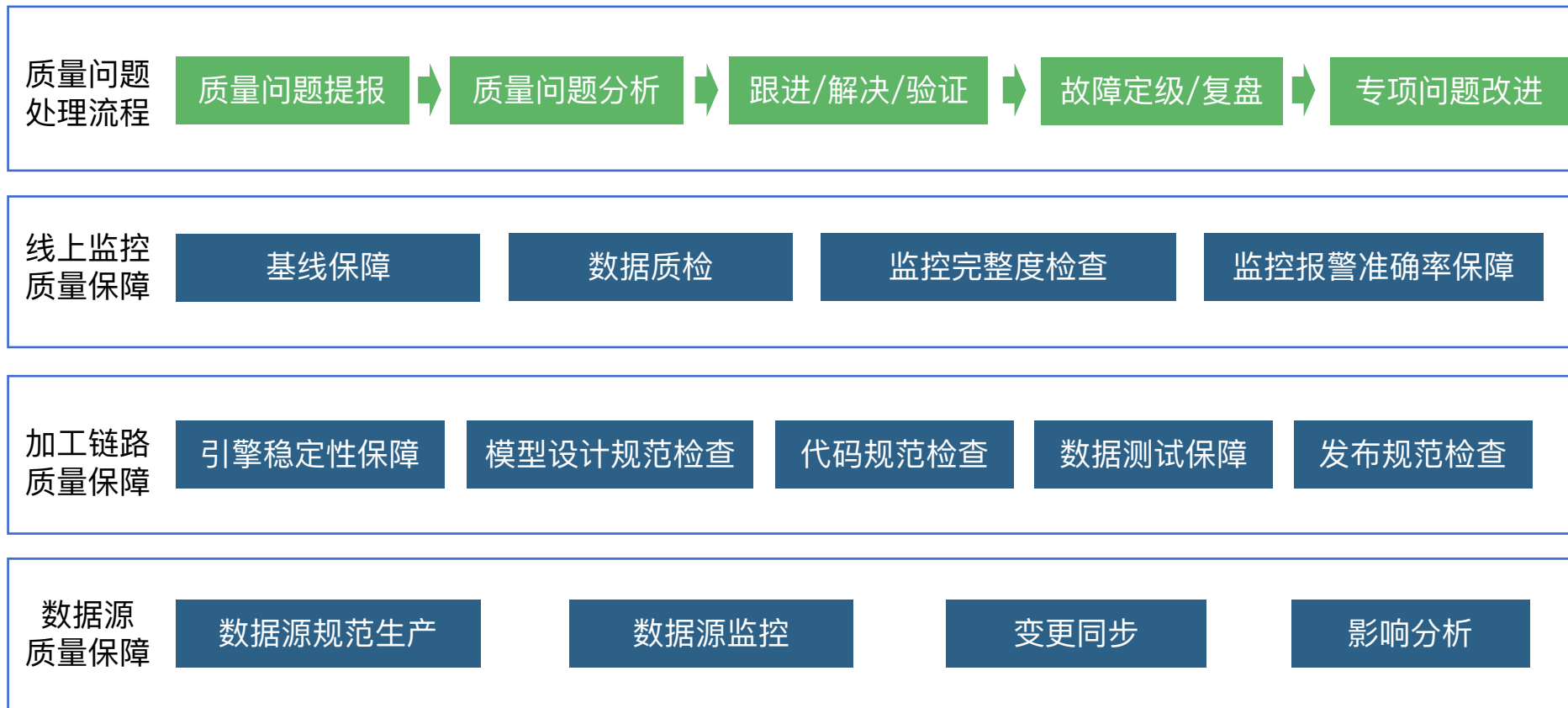
- 跨层引用率
- 明细表查询占比
- 汇总表查询占比



02 资产健康评估：质量（问题与解决方案）

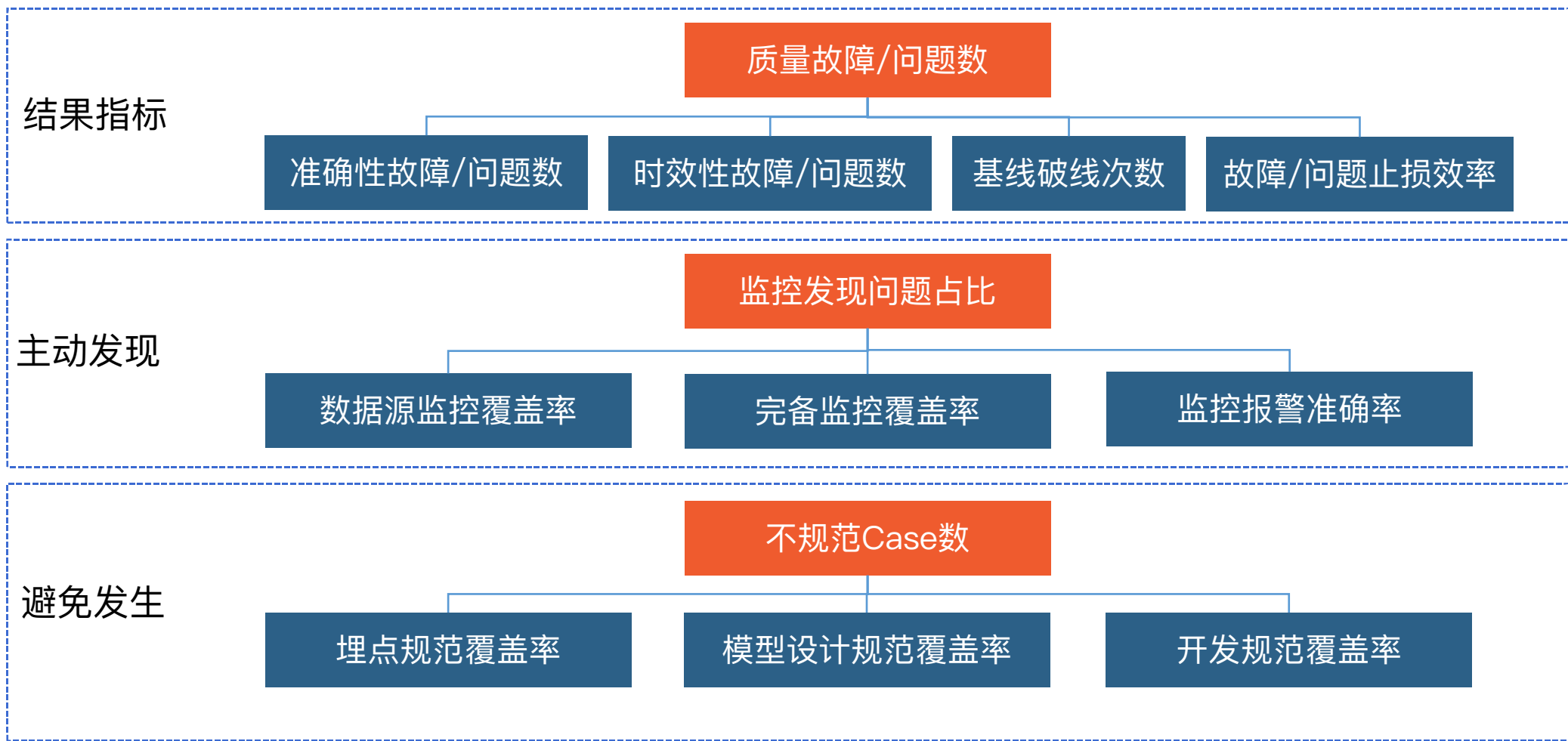
问题

- ◆ 链路长，问题多
- ◆ 意识差，被动发现
- ◆ 跟进止损效率低



数据质量评估体系

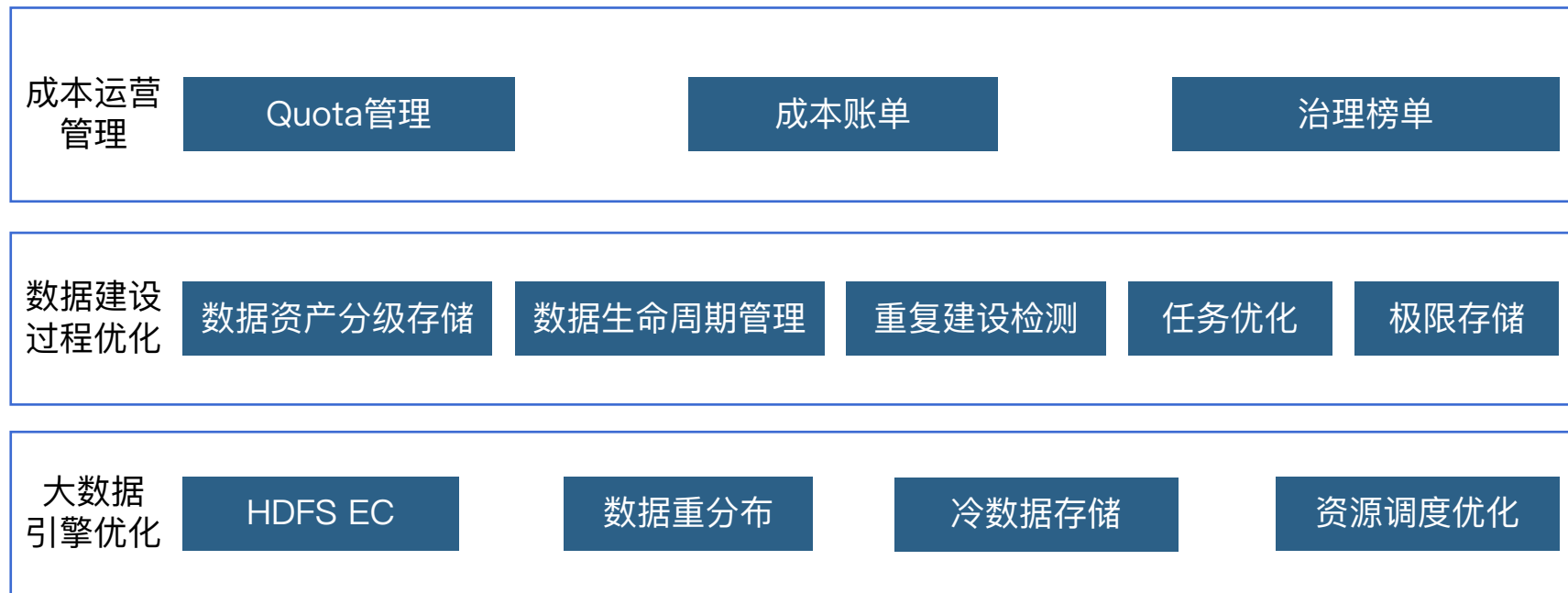
02 资产健康评估：质量（评估）



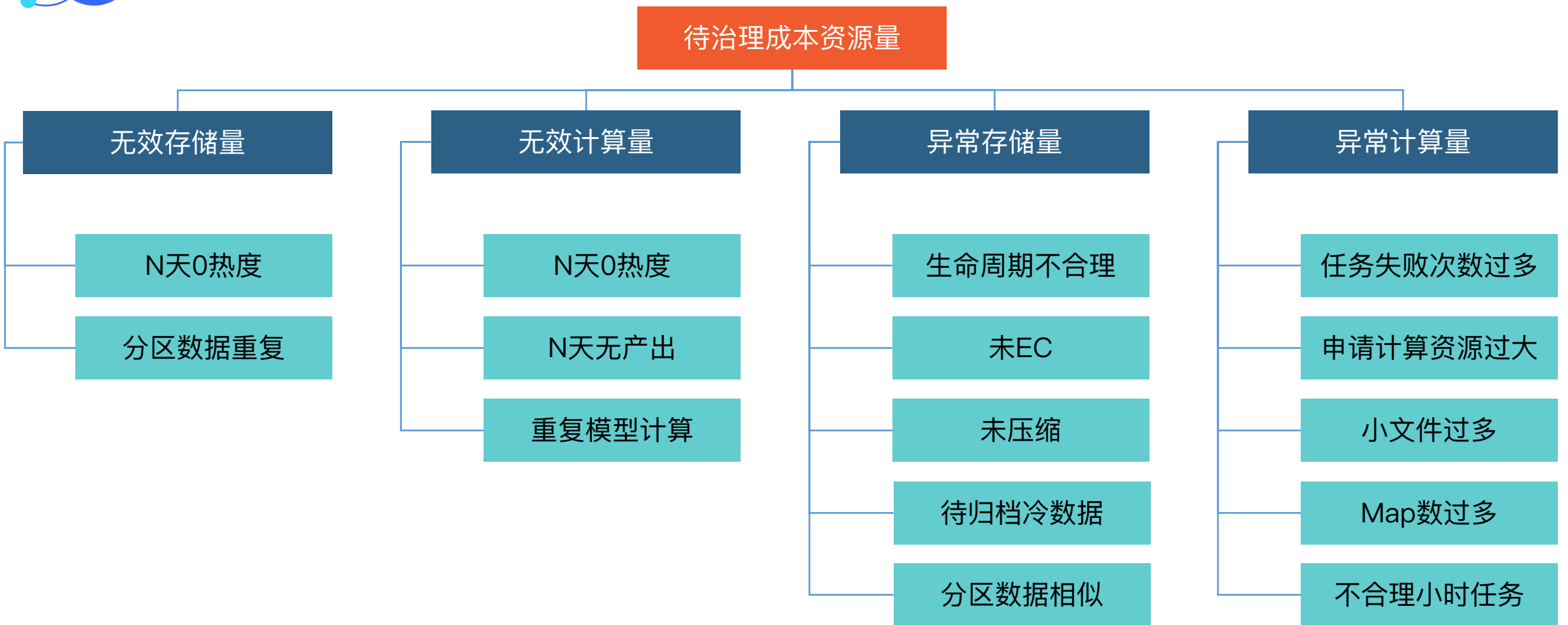
02 资产健康评估：成本（问题与解决方案）

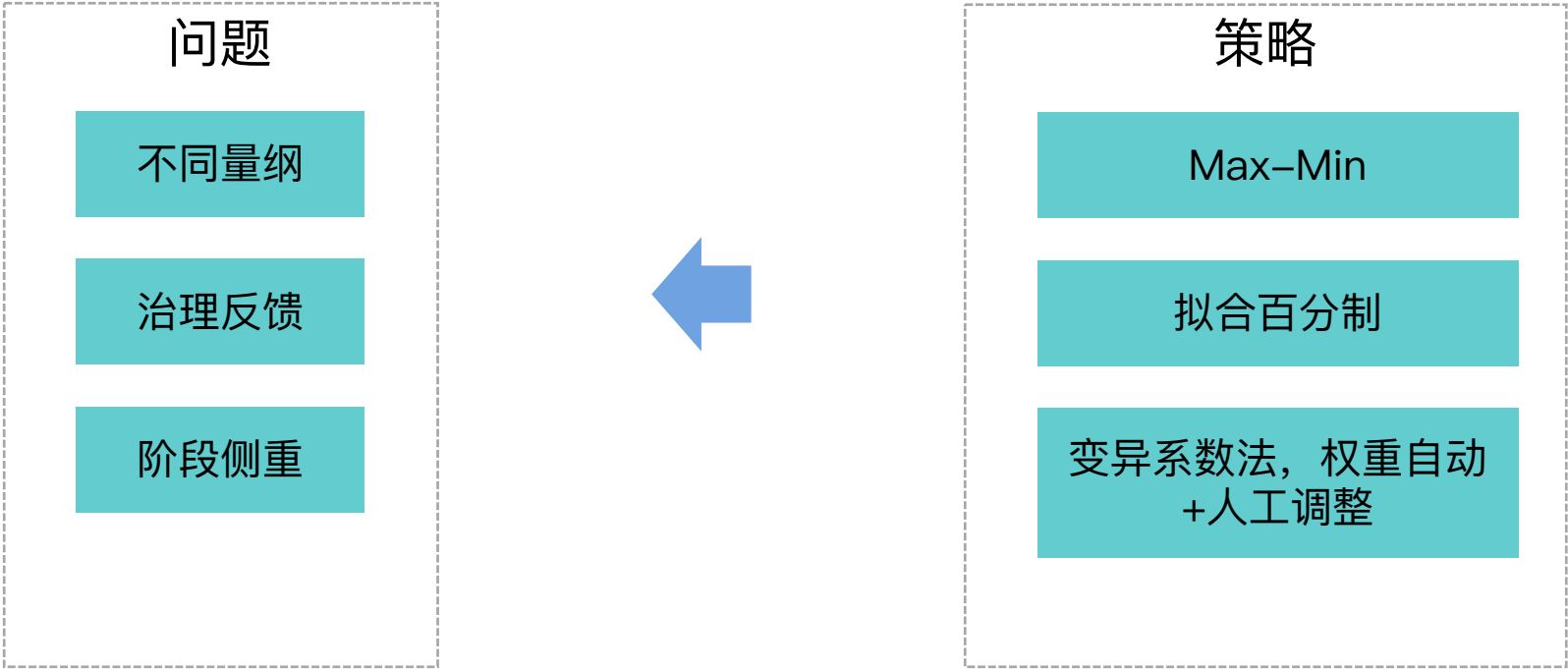
问题

- ◆ 体量大，管理难
- ◆ 缺少成本意识
- ◆ 数据只上不下

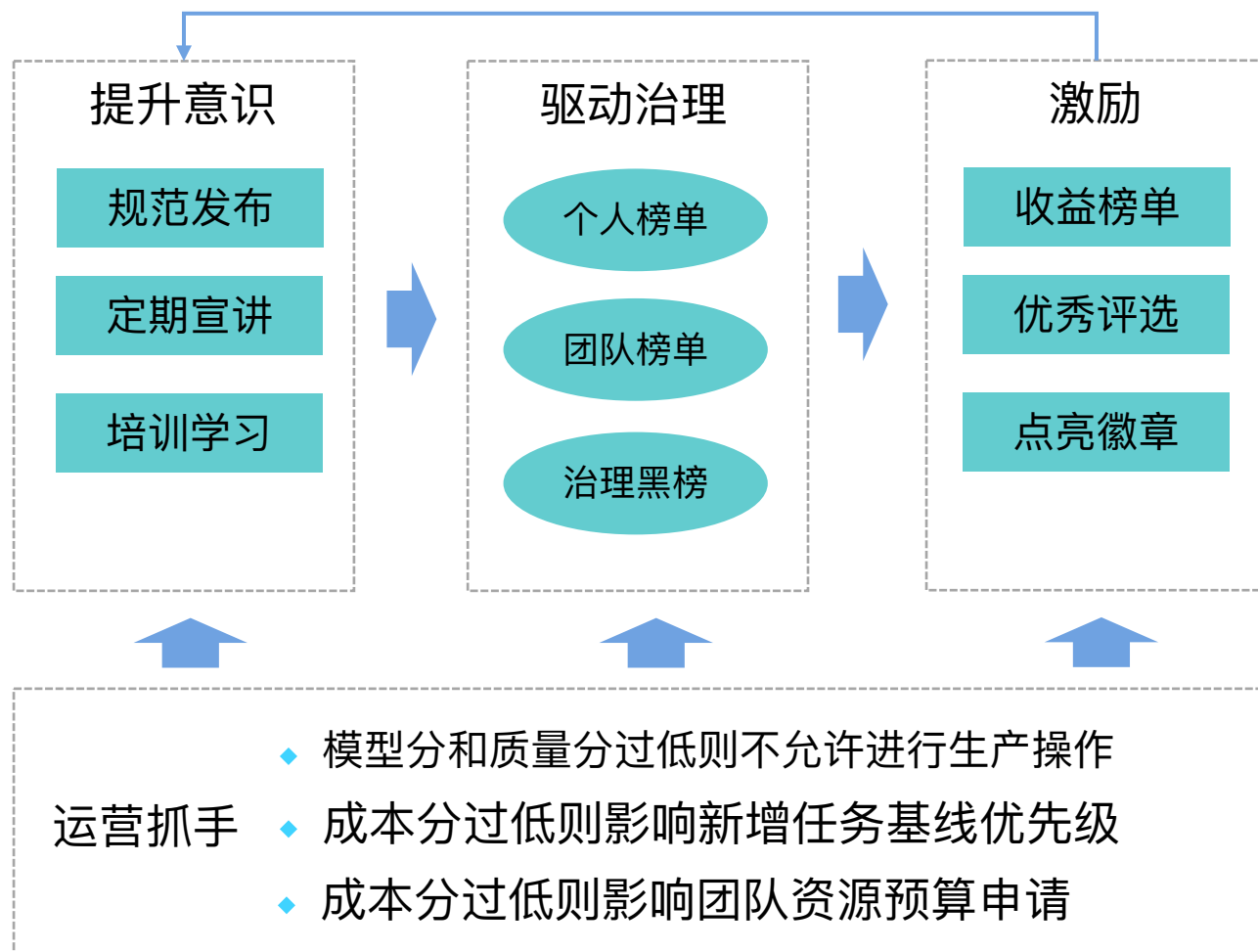


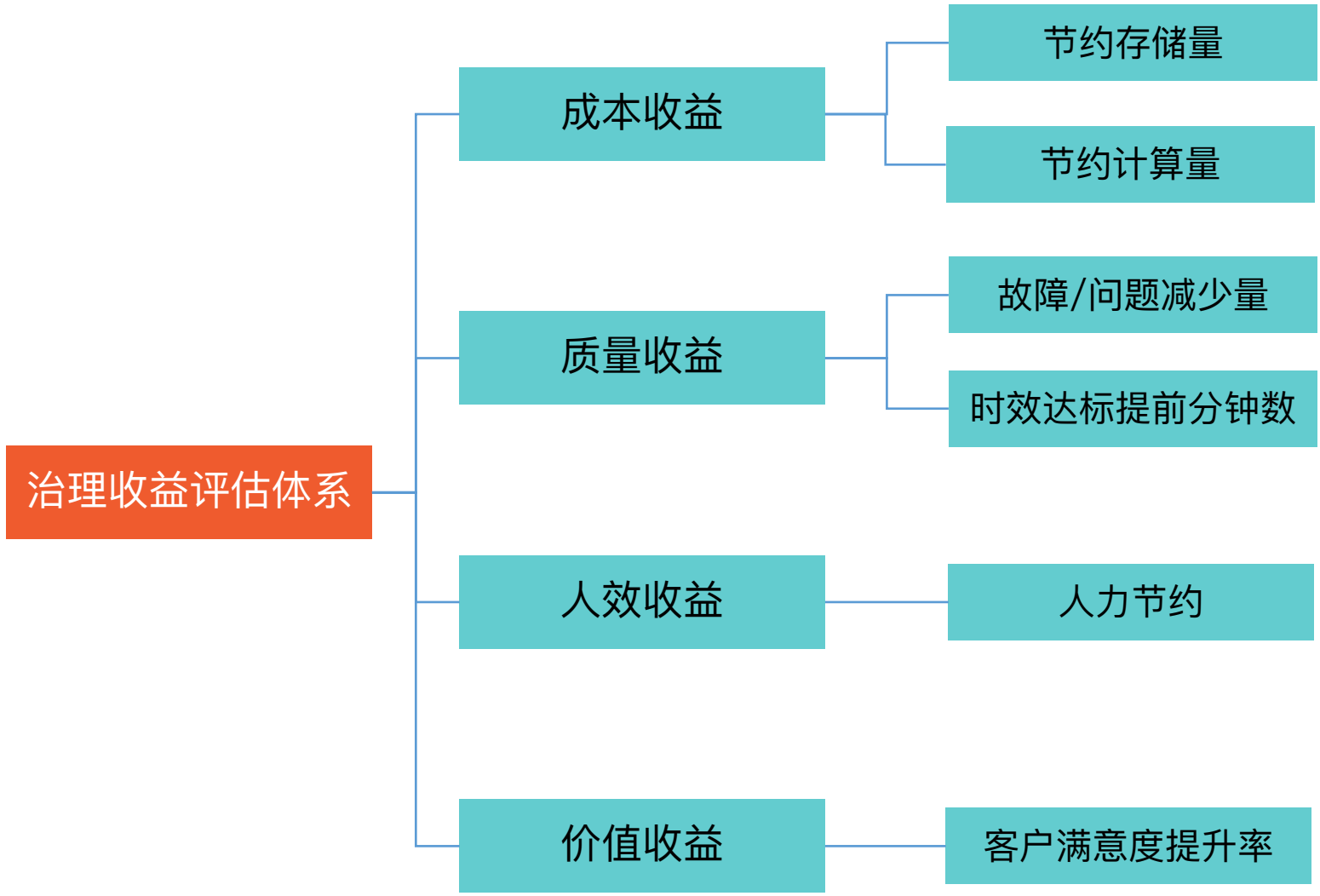
数据成本评估体系





02 结合治理评估体系的运营机制







1. 数据治理背景

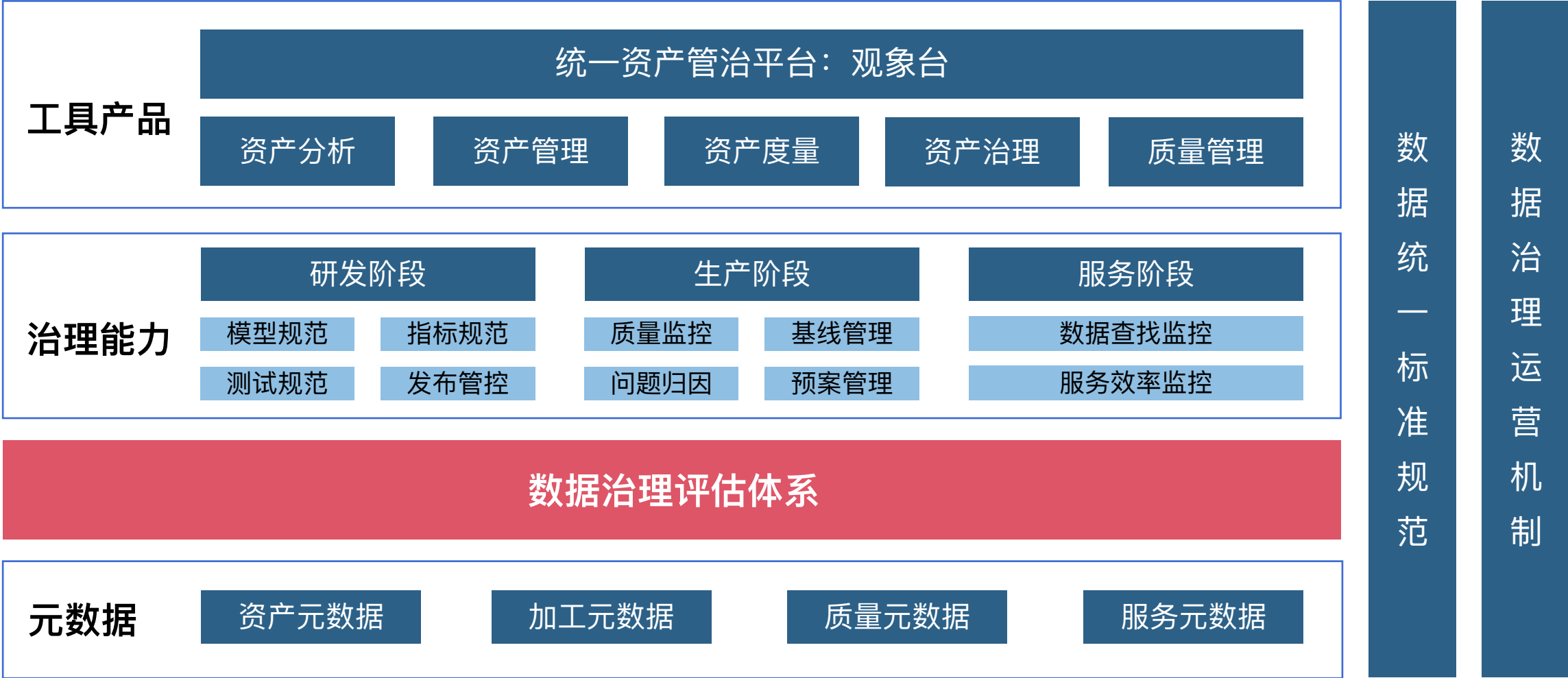
2. 数据治理评估体系

3. 快手实践效果

4. 未来规划

03

快手数据治理架构





1. 数据治理背景
2. 数据治理评估体系
3. 快手实践效果
4. 未来规划

04 未来规划

治理前置

- ◆ 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，例如模型设计管理工具、自动化监控配置等

提效治理动作

- ◆ 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



快手大数据
KUAISHOU DATA

数据内容
建设交流会

THANKS