

DATAFUNCON

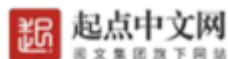
2020大数据 AI的最新技术实践

阅文集团
CHINA LITERATURE

用户画像在阅文实践和探索

阅文集团智能业务中心高级总监 陈炜于

原创文学品牌：



图书出版品牌：



有声阅读品牌：



影视制作与发行品牌：



移动阅读APP：

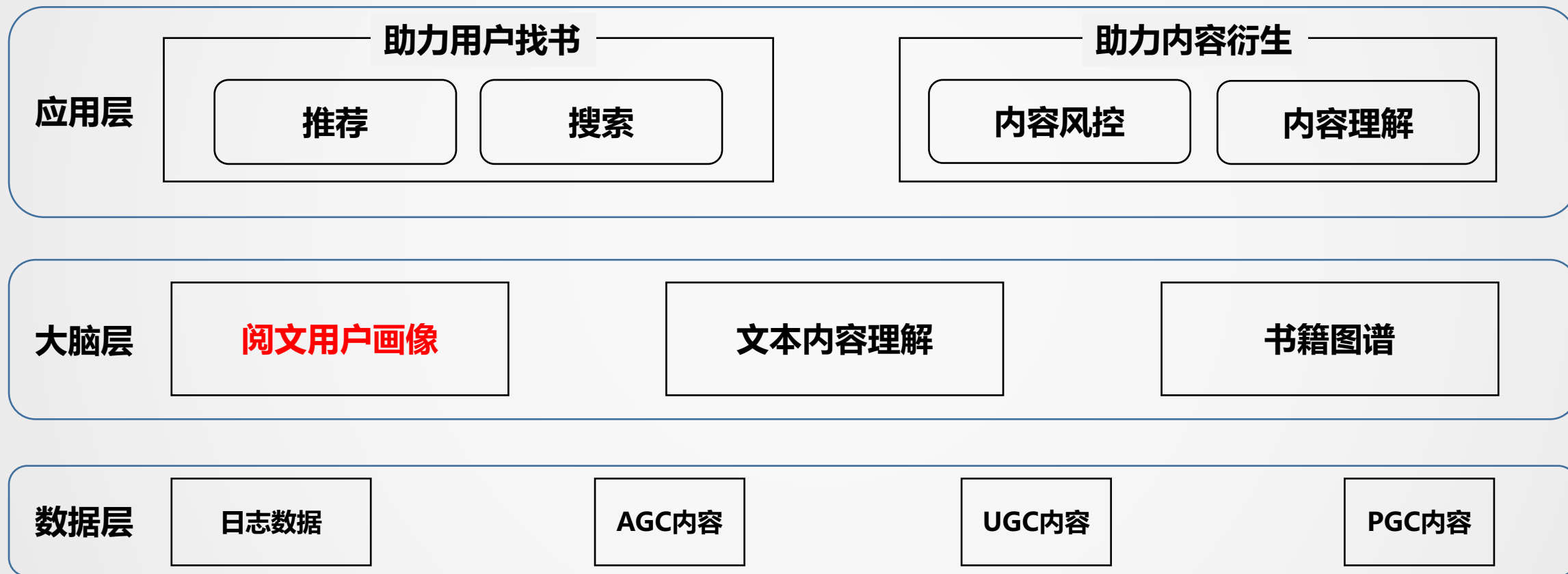


移动创作APP：



移动阅读硬件：口袋阅

阅文智能业务中心



分享大纲

- 思考和方法
- 实践和探索
- 展望

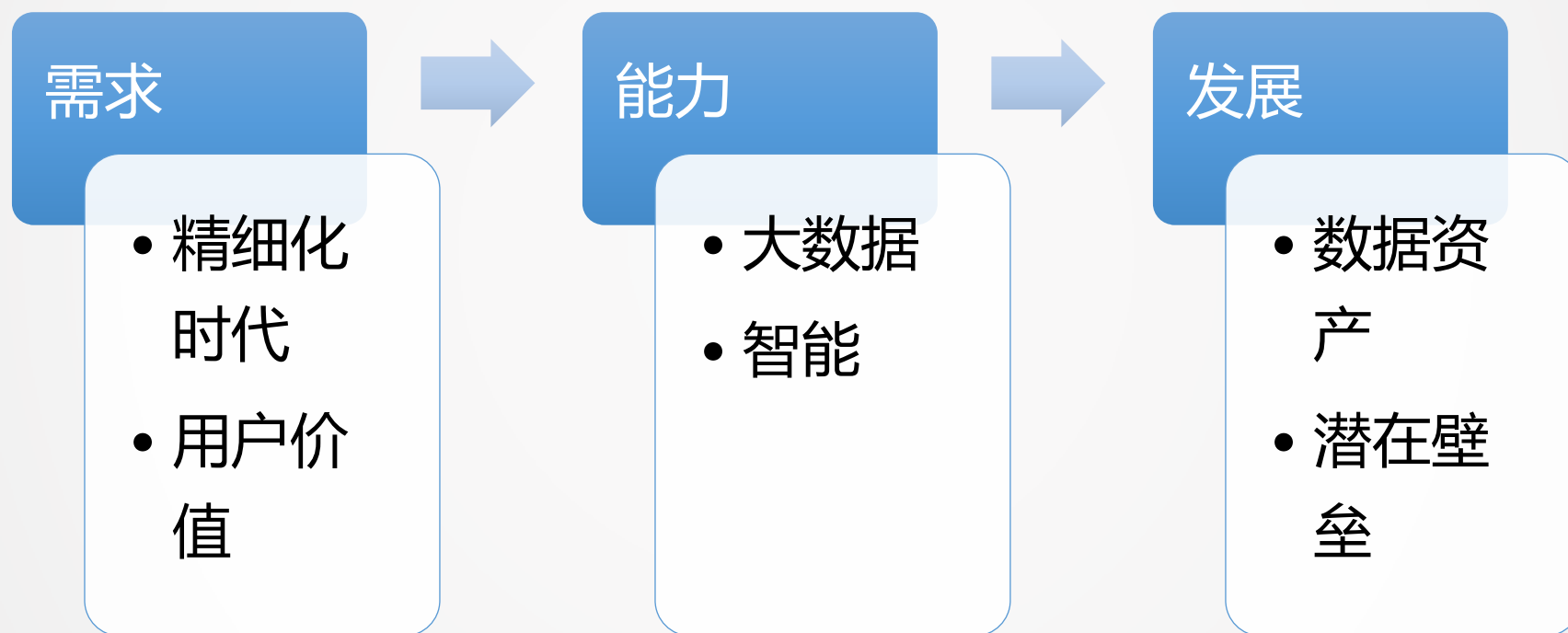
DATAFUNCON

2020大数据 AI的最新技术实践

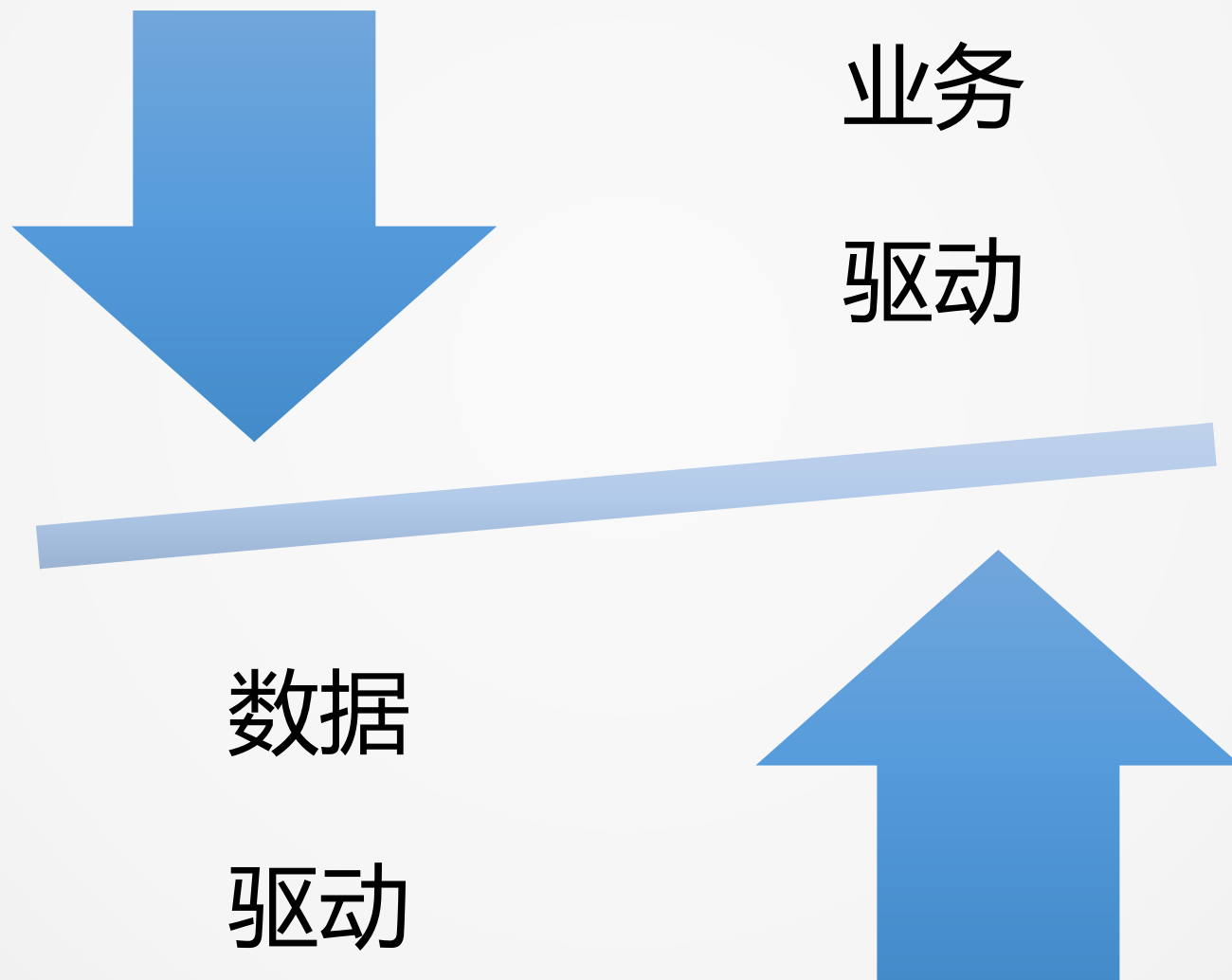
阅文集团
CHINA LITERATURE

思考和方法

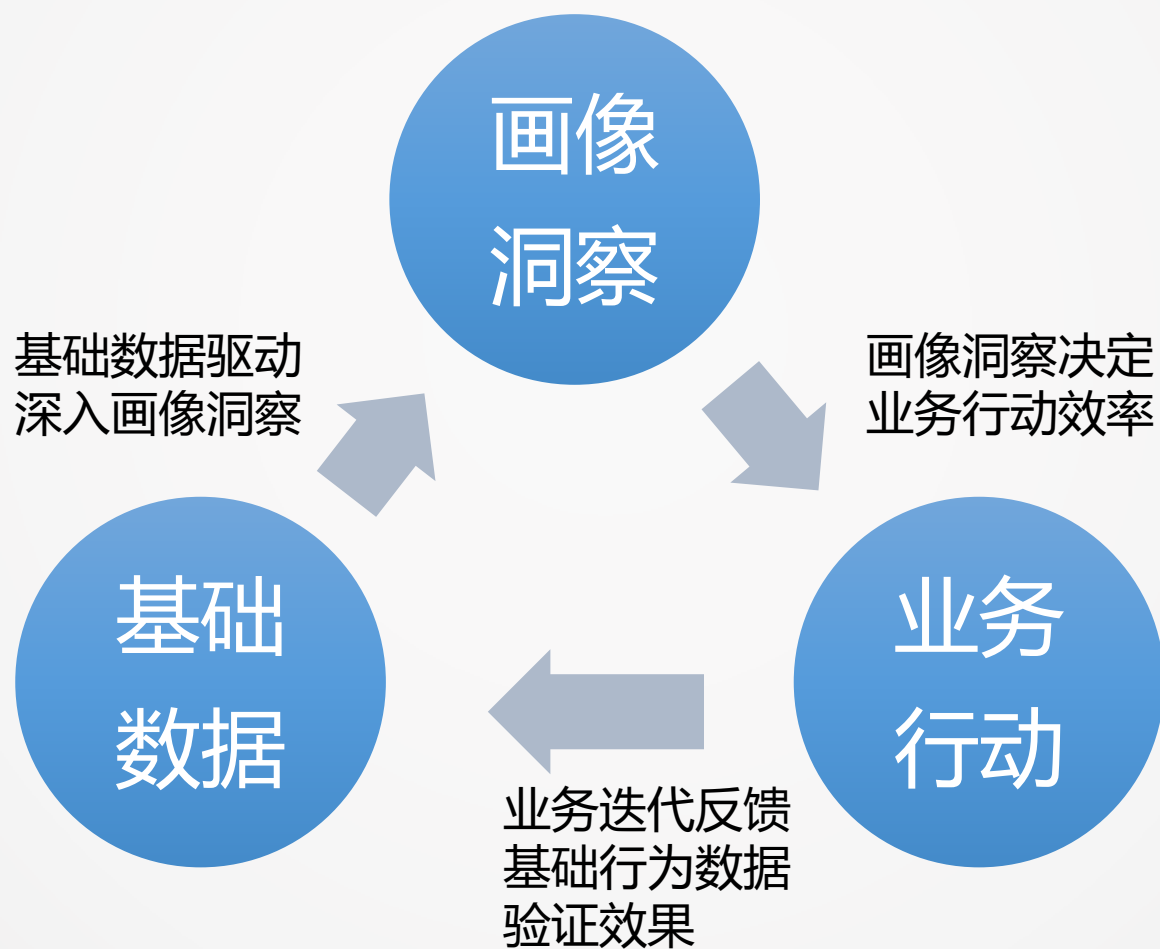
思考：画像为何存在



思考：画像如何做



方法：数据驱动+业务驱动



DATAFUNCON

2020大数据 AI的最新技术实践

阅文集团
CHINA LITERATURE

实践和探索

背景



业务标签分散



系统建设薄弱



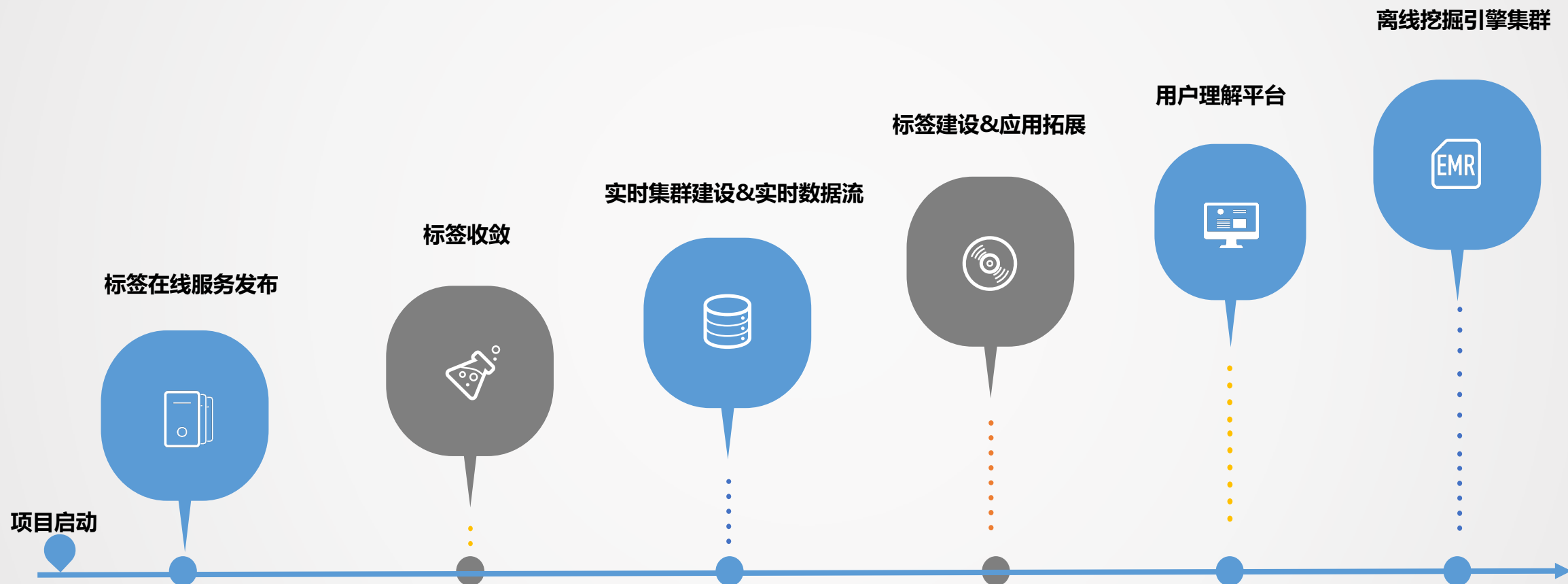
数据不完备



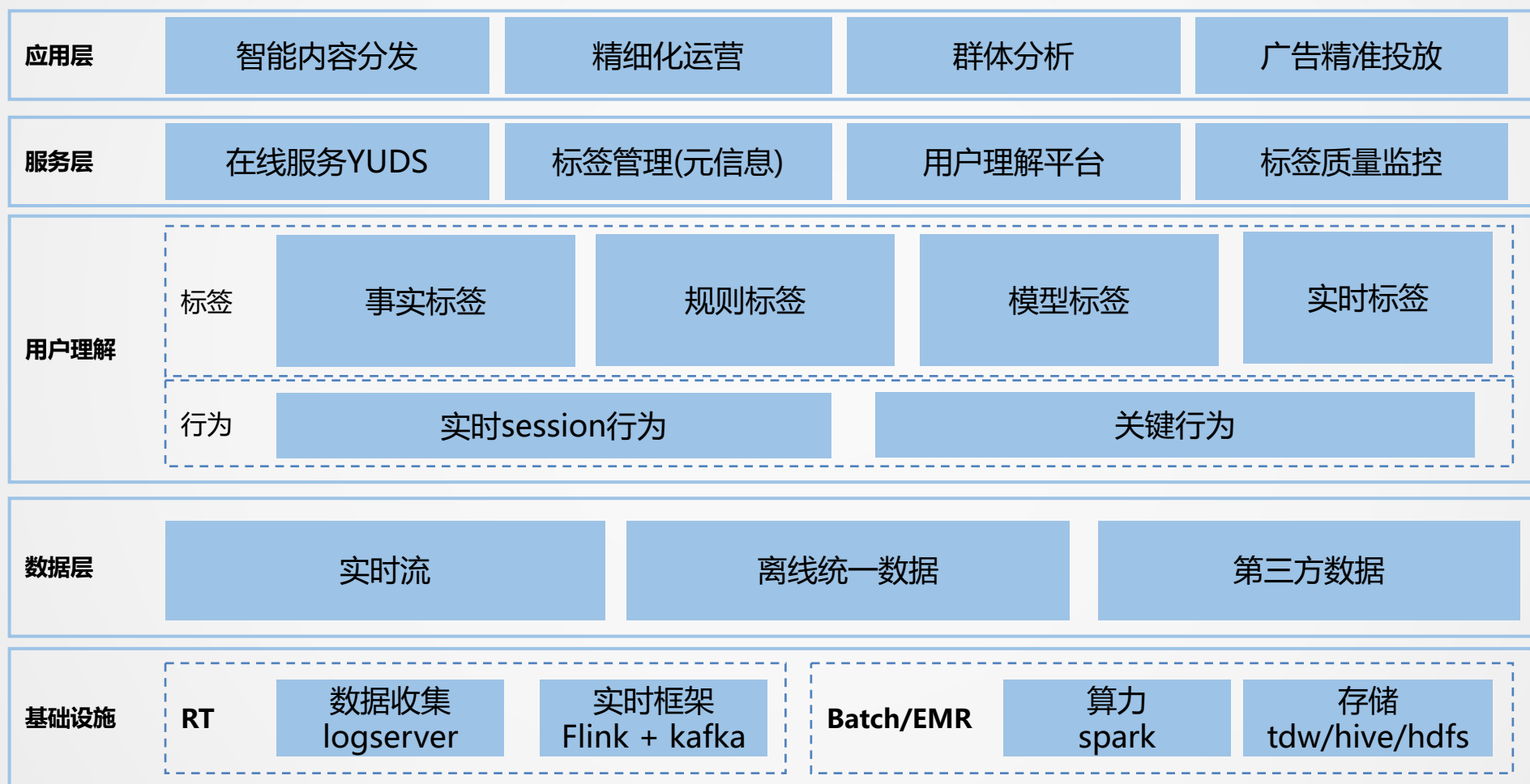
应用少，case型



方法：收敛，基建，应用，反馈



全貌



基础：不可忽视的数据建设

数据质量

- 覆盖面/准确率情况
- 数据监控体系完善

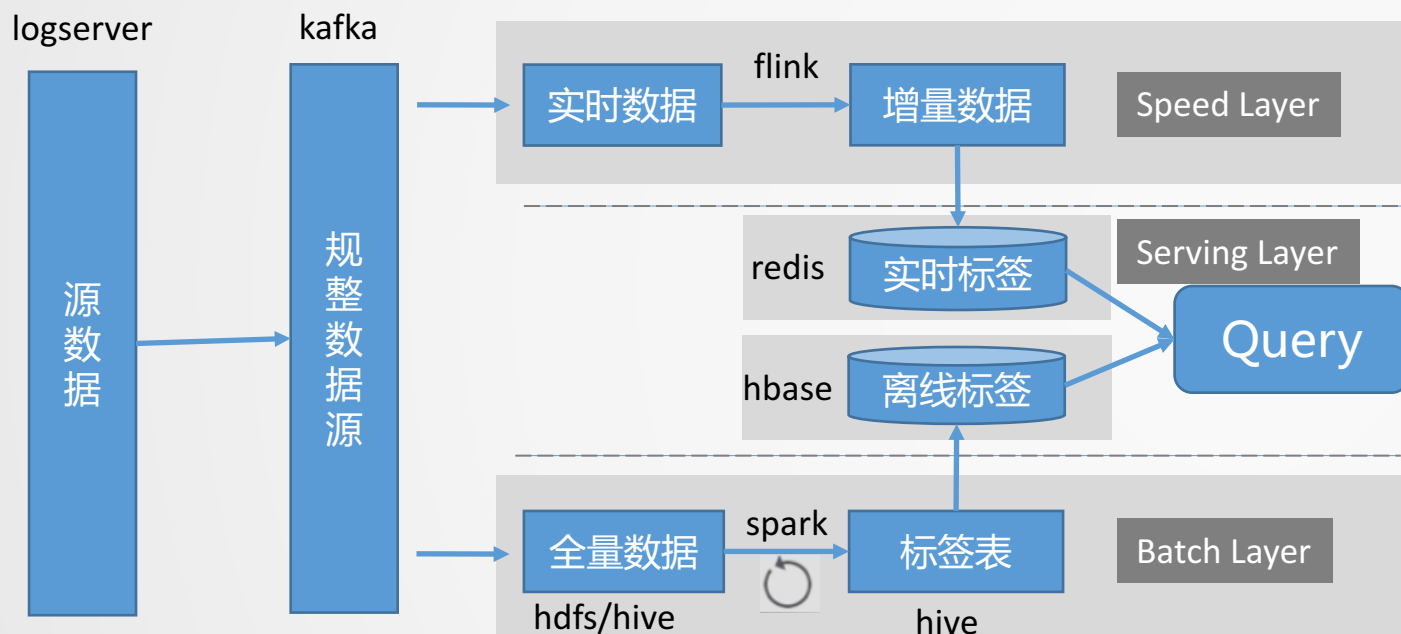
数据时效

- 线上延迟
- 实时数据流建设

数据平台化

- 数据统一/ID贯通
- 数据体系统一，埋点统一

支撑：架构建设



➤ 实时层

➤ 建立统一的logserver

➤ Kafka统一收敛数据

➤ 服务层

➤ 统一收敛服务调用

➤ 解决性能匹配，Id转化问题

➤ 离线层

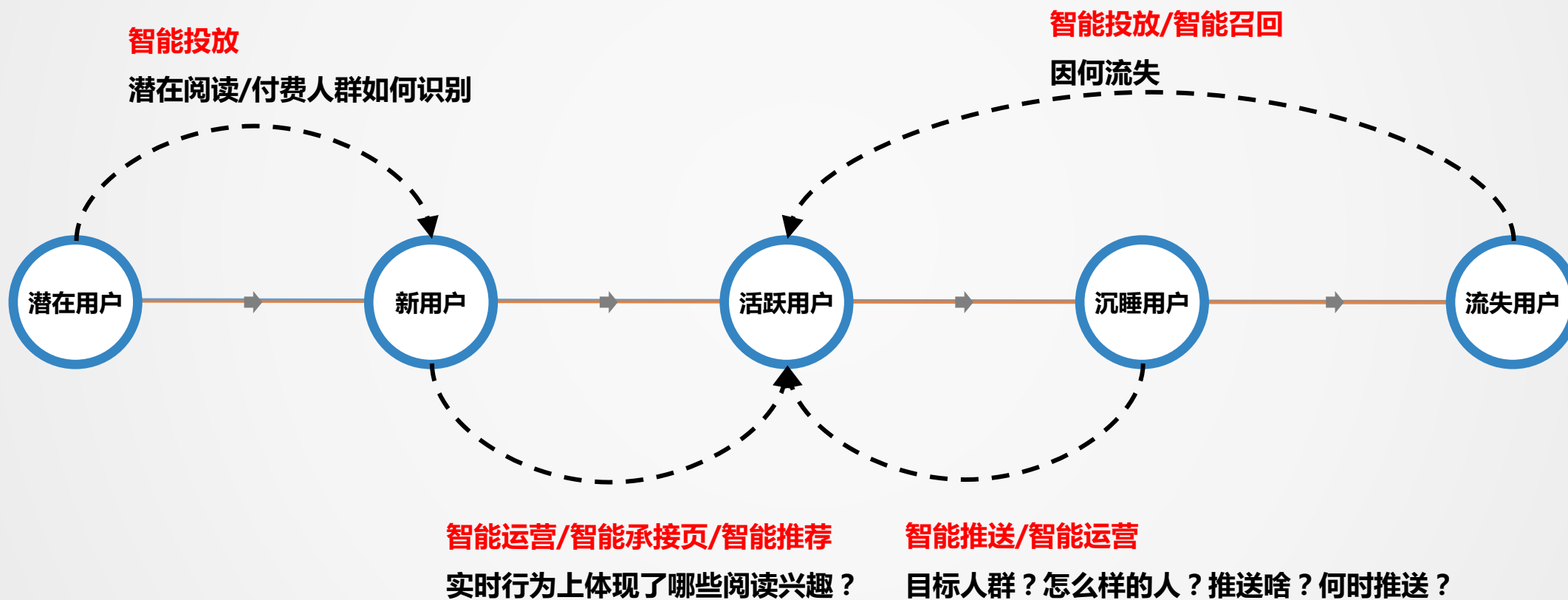
➤ 大数据挖掘算力建立

➤ 保证数据回溯

血肉：标签建设

模型标签	泛娱乐	漫画用户	漫画评分表	...
	社区偏好	UGC内容偏好	KOL	...
	消费能力	免费用户	付费流失	...
	阅读偏好	基因偏好	书籍评分表	...
	生命周期	新用户	流失预测	...
统计标签	人书标签	阅读进度	阅读日期	...
	行为RFM	付费	阅读	...
事实标签	渠道属性	渠道ID	渠道类型	...
	账户信息	余额	注册时间	...
	社会属性	国家	性别	...

价值：应用



反馈：验证+闭环

- ✓ 平台验证数据，抽样评估画像准确
- ✓ 线上AB测试，效果反馈

DATAFUNCON

2020大数据 AI的最新技术实践

阅文集团
CHINA LITERATURE

未来展望

技术突破，奠定泛娱乐的用户画像之基

业务赋能，打造用户全生命周期的赋能

DATAFUNCON

2020大数据 AI的最新技术实践

阅文集团
CHINA LITERATURE

THANKS !

阅文集团智能业务中心欢迎各位加入，开放职位（扫码联系）：

- 推荐/搜索/用户画像/自然语言处理等策略研发工程师
- 推荐/搜索/用户画像/精准运营等架构开发工程师（C++）
- 智能产品经理（2C/2B）

