

www.vip.com

唯品会大数据架构及应用实践

唯品会大数据 李创

A large pink circle on the right side of the slide, containing the Vip.com logo and tagline in white text.

唯品会
vip.com
一家专门做特卖的网站

唯品会-全球精选，正品特卖



全球精选 正品特卖

2017年Q3财报核心数据

净营收增至 

153 亿元人民币 增长 27.6%↑

总订单量增至 

7400 万 增长 23%↑

人均消费额增长至 

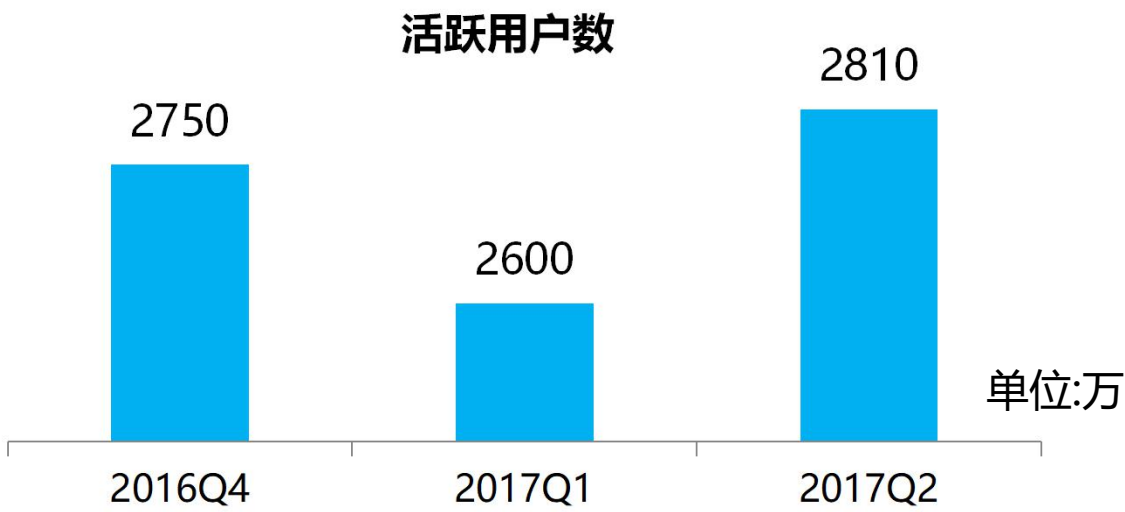
643 元

复购率增长至  复购用户销售占比

84.4%↑ 95%↑

客户粘性持续提升 

*增长百分比相较于上年同期



| 产品

数据仪表盘、数据魔方、比价系统、地图服务等

| 算法

选品、分仓与预调拨

精准推荐

基础算法库

| 数据

基础数据

用户画像

商品画像

| 系统

数据实时接入

离线计算平台

实时计算平台VRC

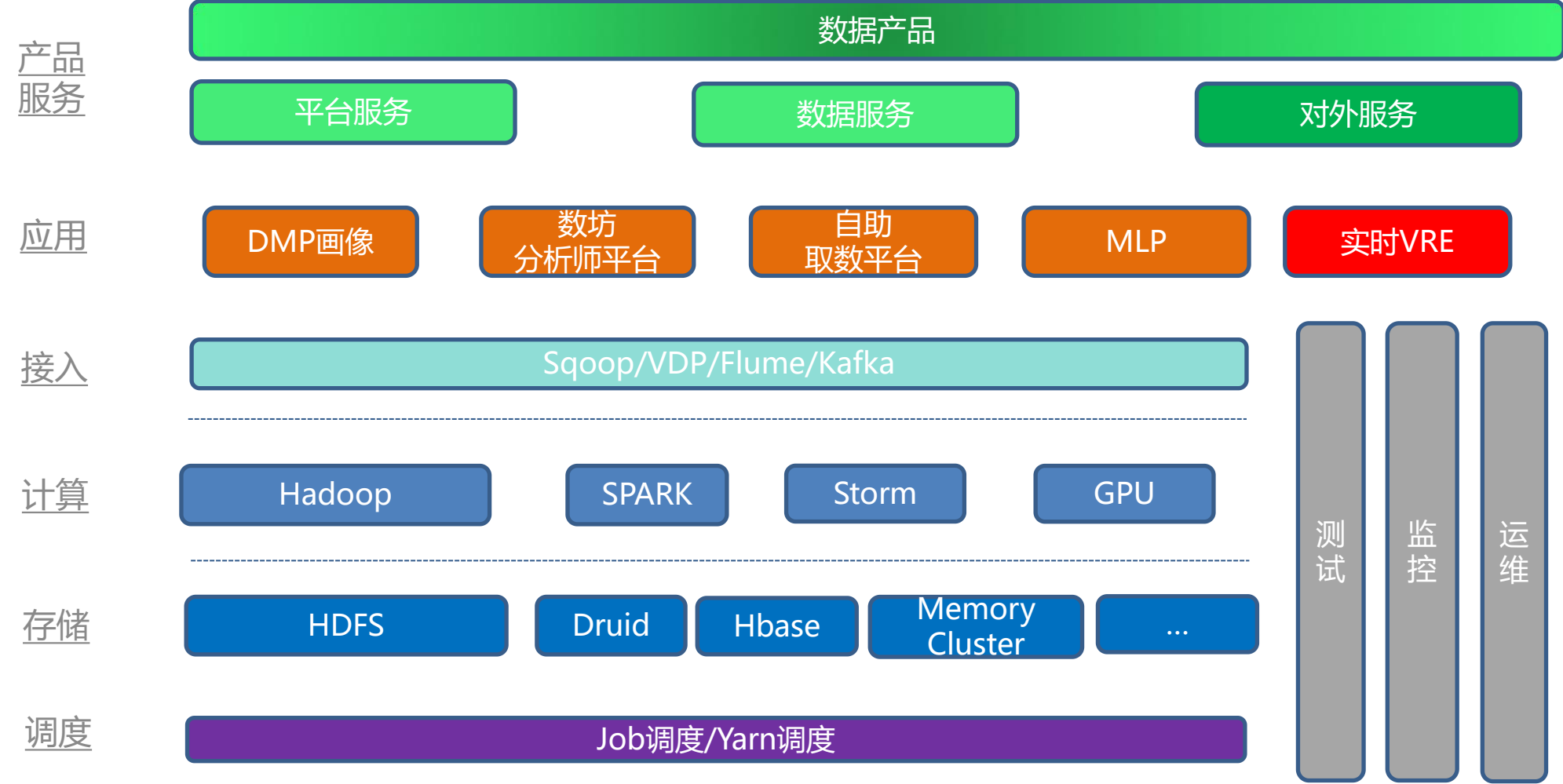
资源管理平台

运维

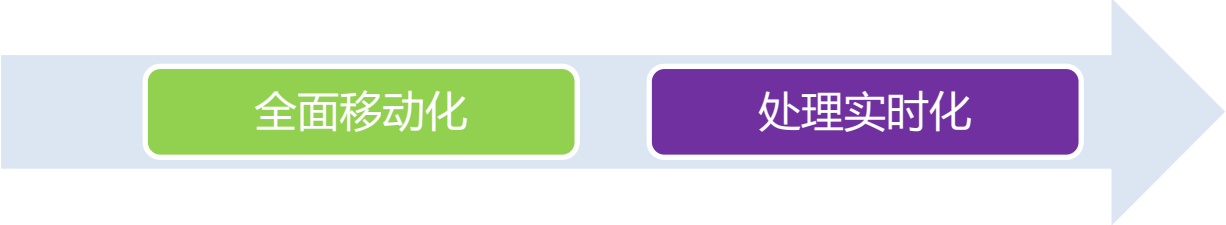
监控

测试

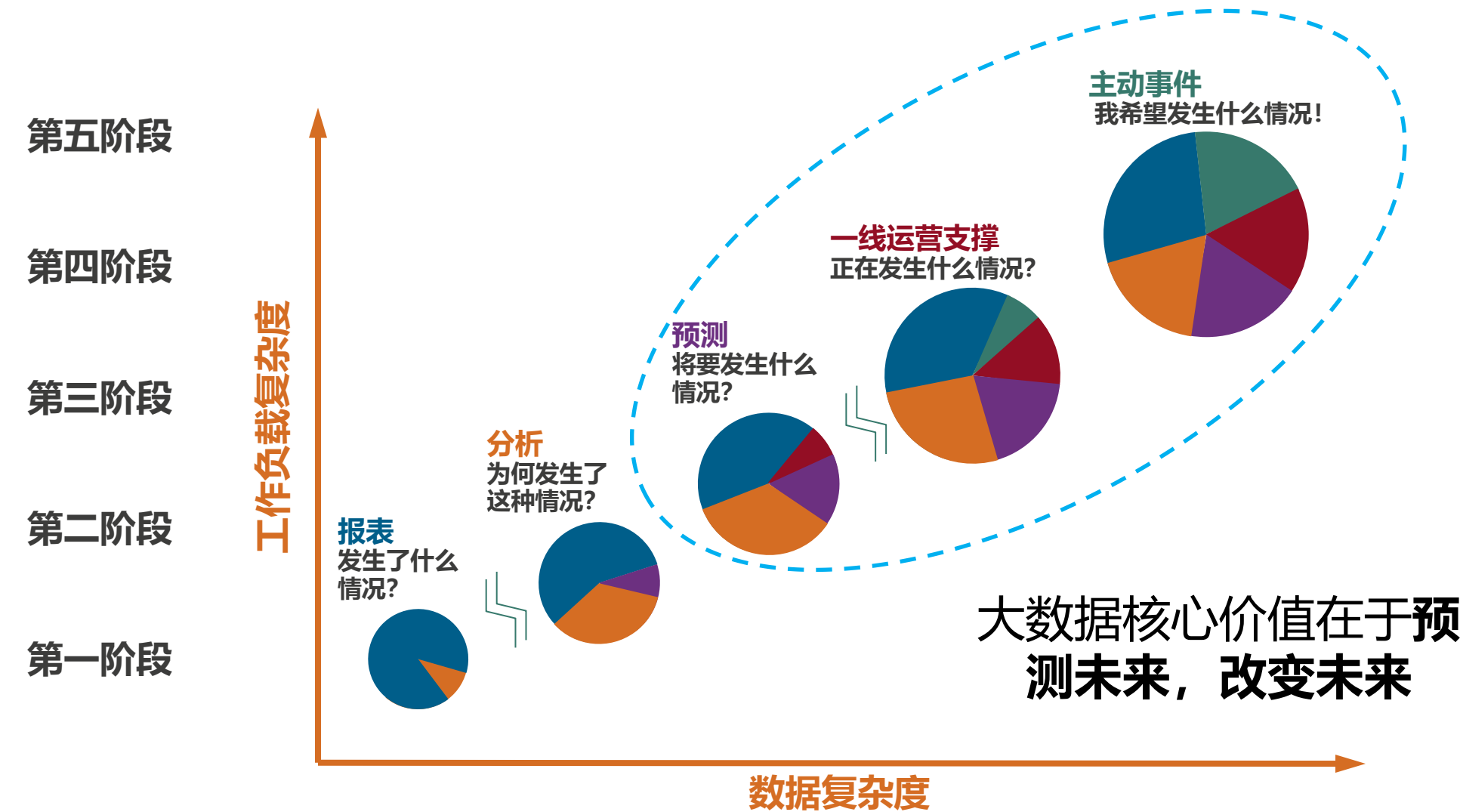
系统-实时与离线融合，关键应用全流程实时



数据-动静结合,用户画像&商品画像



数据-分析与预测融合，从分析历史到预测未来



分析与预测融合案例-寻找潜在牛股



种子
扩散

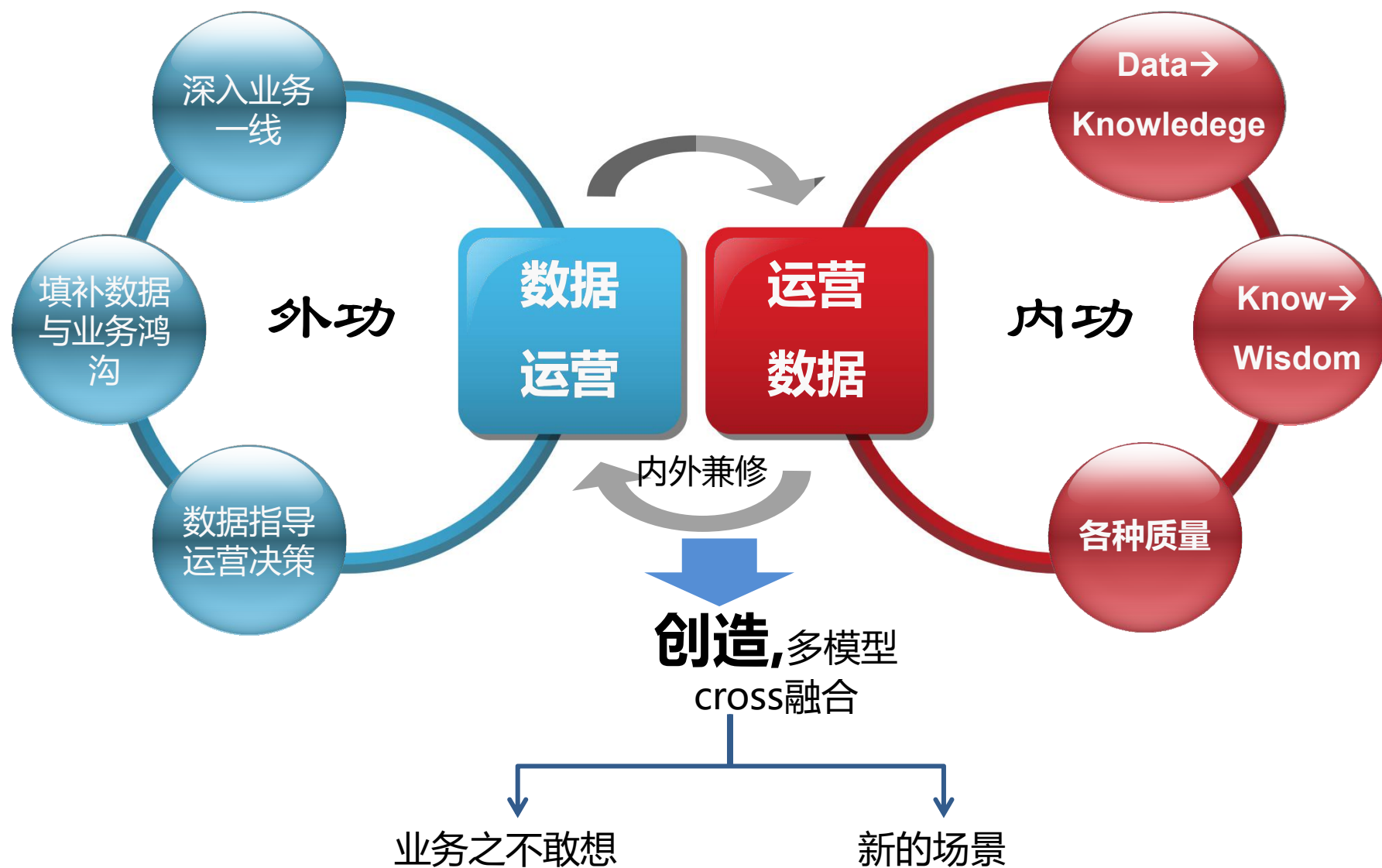
潜在
牛股
?



验证ing+期待

股市有风险，投资需谨慎

数据-从数据运营到运营数据，数据创造需求



数据创造需求案例-选品组货

老流程



组货质量如何
事后才知道

支撑商
务组货



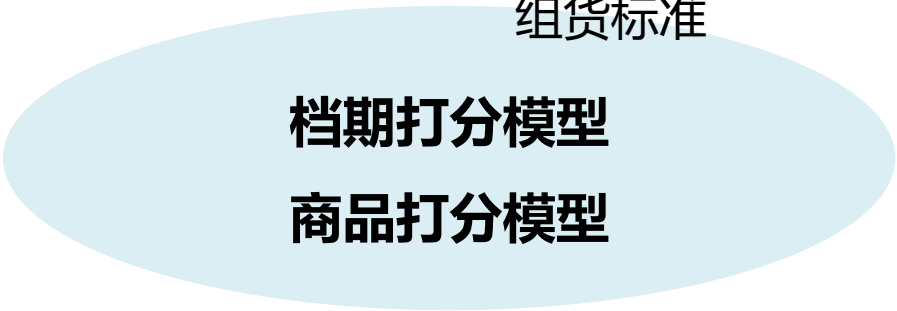
事后变事前

大数据
组货



创造新场景

建立选品
组货标准



算法-深浅结合

推荐与预测

档期实时精准推荐

商品list页实时系统

评估系统

.....

选品、分仓与预调拨

深度学习平台

CNN网络

DNN网络

并行CPU集群

并行GPU集群

数据挖掘算法库

协作推荐

内容推荐

实时混合算法模型

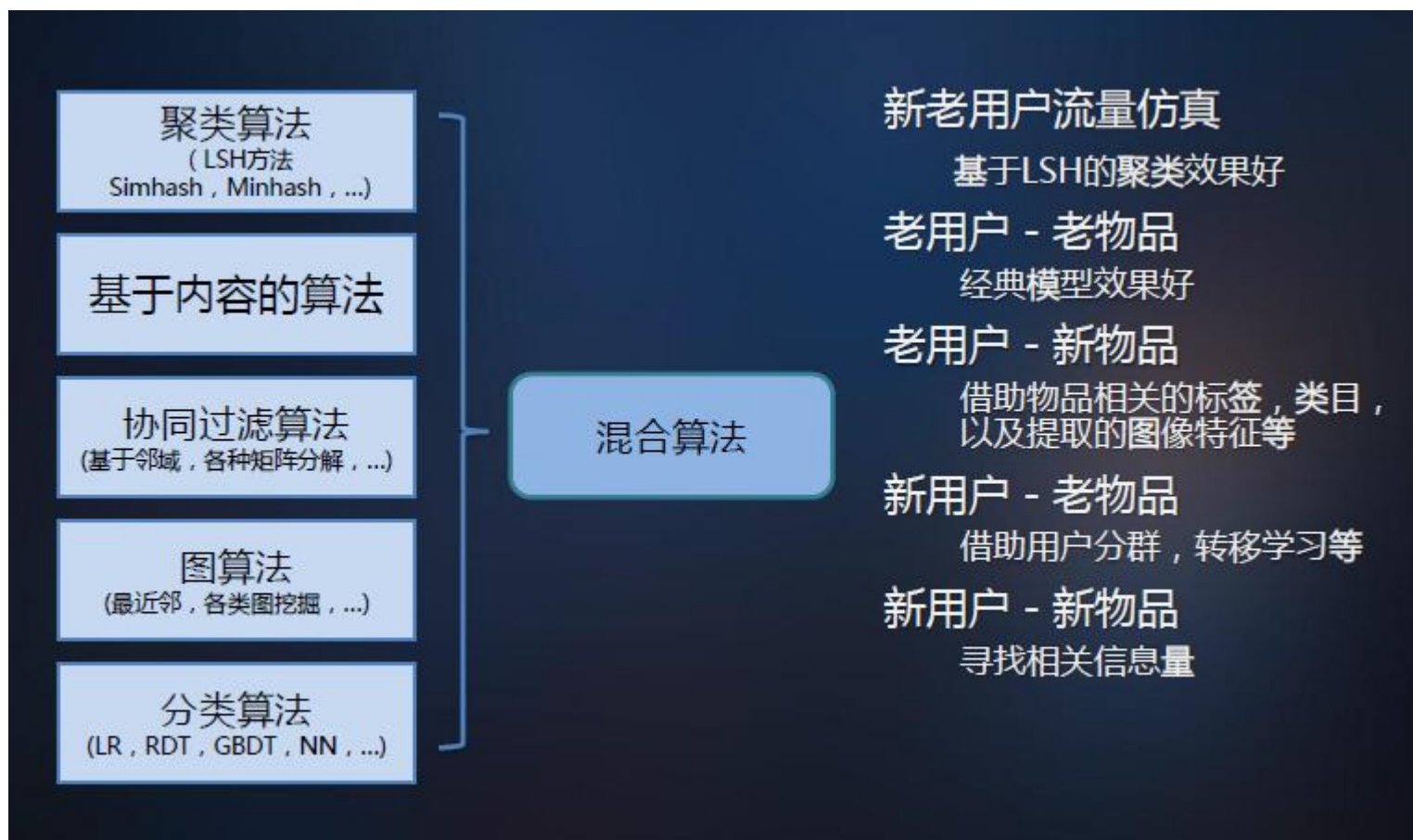
图像识别

流式挖掘算法

.....

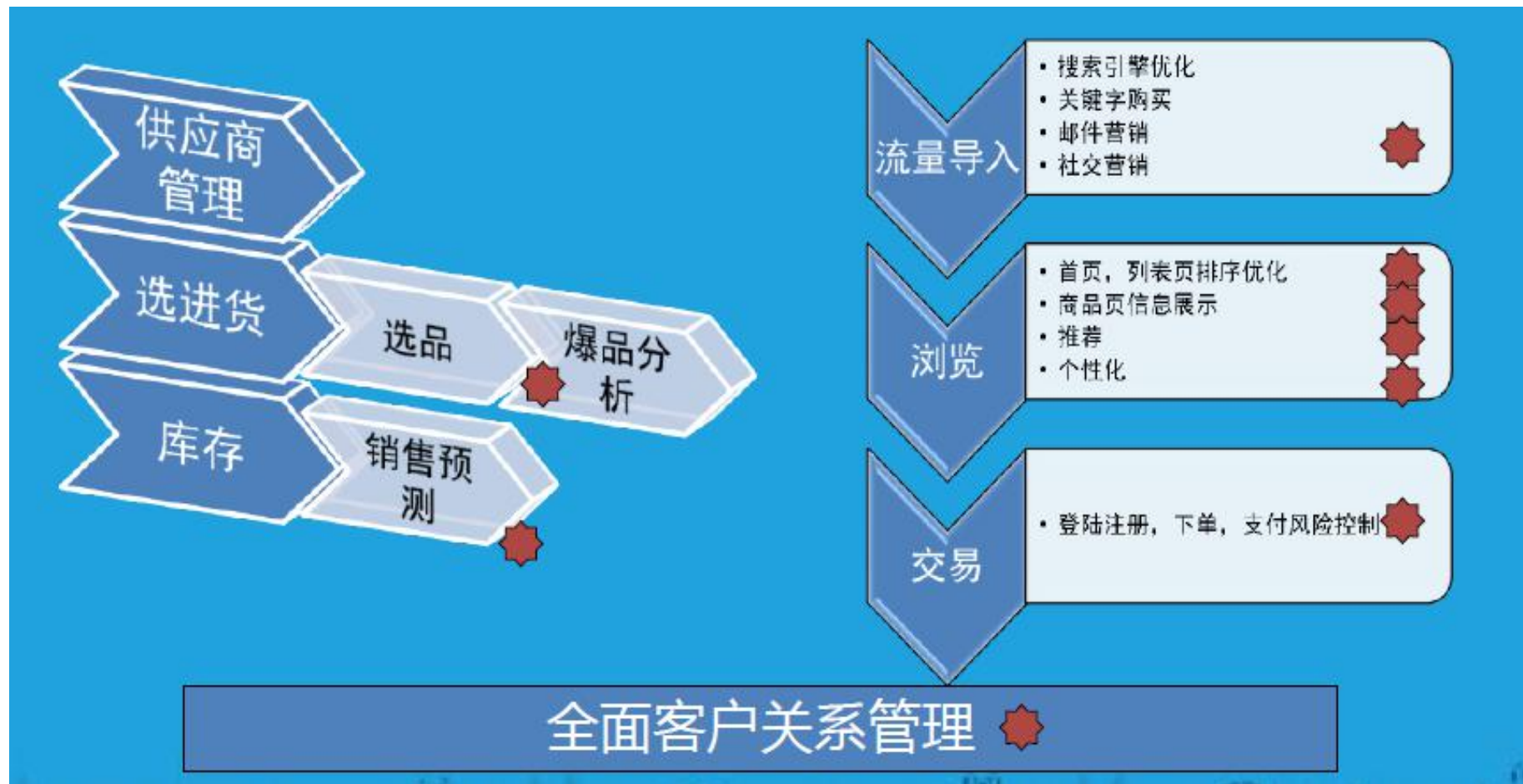
算法@VIP-谱系

针对不同的场景，基于不同的数据采用不同的算法策略

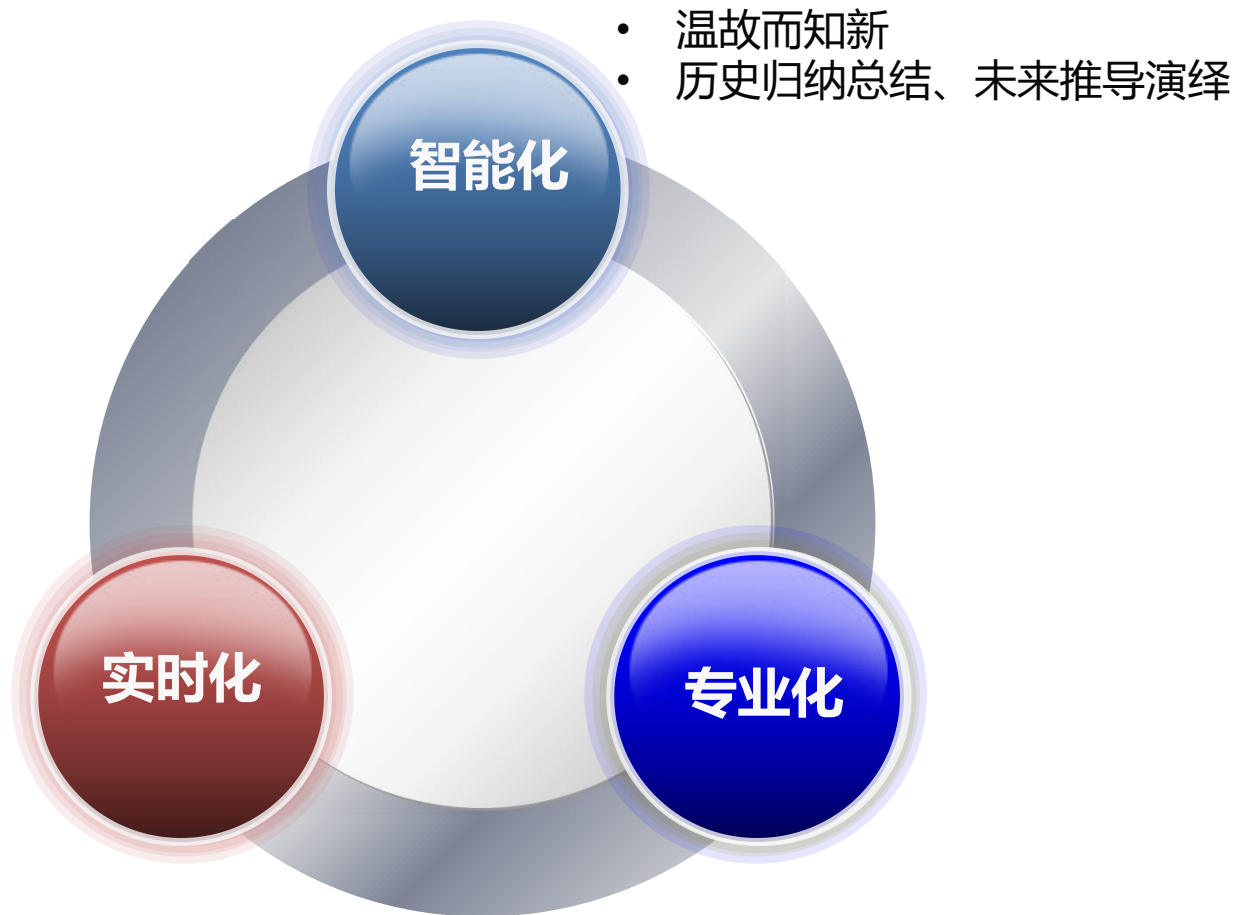


产品-前后端数据产品一体化，支撑全链条运营

- 后端数据产品与前端个性化推荐一体化运营，覆盖售前、售中、售后全流程销售场景
- 由供给决定需求向需求决定供给转变

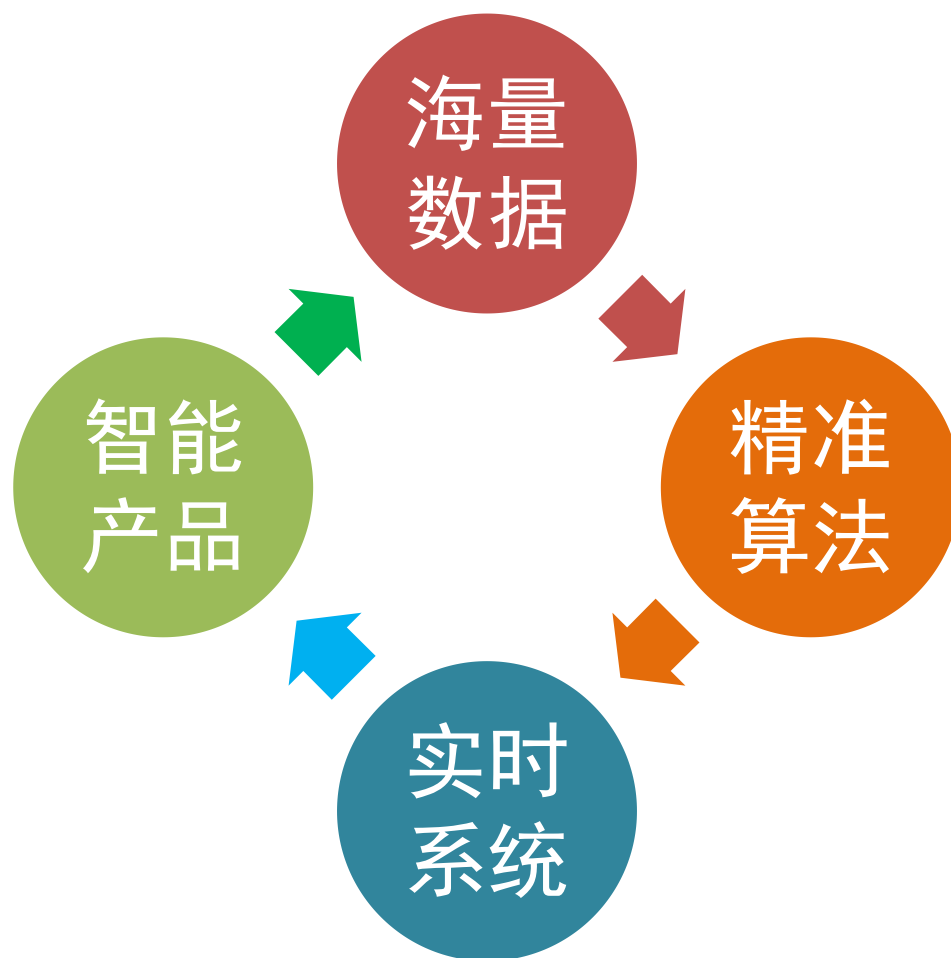


产品-智能化、实时化、专业化



解放人做创造性的工作，把高强度的工作交给机器

大数据生态链：海量数据+精准算法+实时系统+智能产品



让大数据成为唯品会增长引擎

谢谢！

