

解密姨搜

基于数据的风控系统和应用

侯松

songhou@creditease.cn

目录

1. 我们是谁
2. 基于数据如何做风控
3. 风控系统的建设
4. 应用案例
5. 大数据风控的未来



姨搜

1. 我们是谁

- 宜信：超十年的稳健经营历史，世界上规模最大的P2P金融公司之一。
- 姨搜：为整个宜信公司提供风控数据服务和模型服务，包括风控搜索引擎，大规模知识图谱，决策引擎，离线分析建模平台等。
- 本人：负责姨搜大数据风控的各种数据产品和基础设施，协助宜信各产品线和业务方提升风控水平。



姨搜

目录

1. 我们是谁
2. 基于数据如何做风控
3. 风控系统的建设
4. 应用案例
5. 大数据风控的未来



姨搜

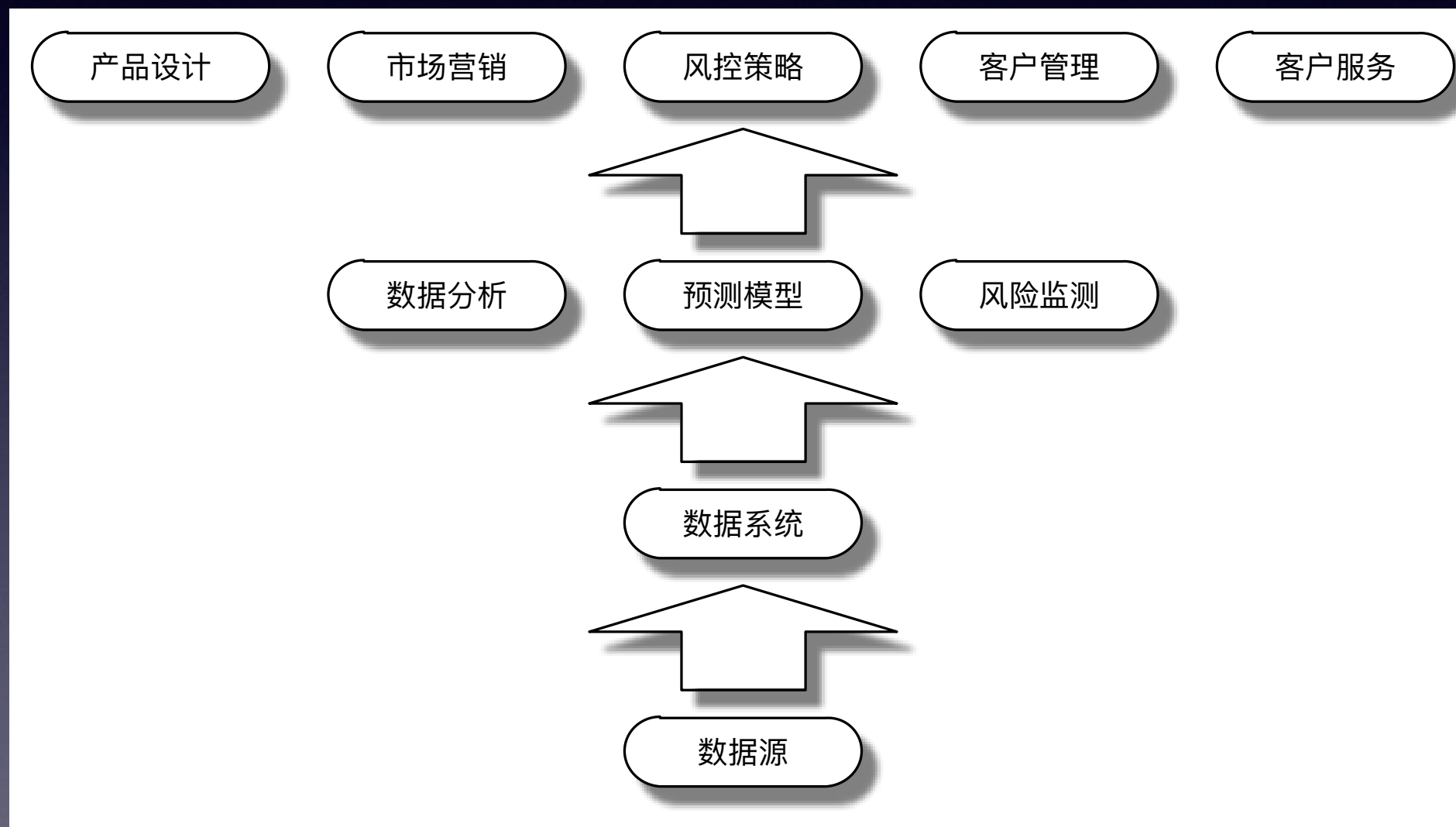
2. 基于数据如何做风控

1. 信贷产品的组成部分
2. 什么数据是有用的数据
3. 个人借款风险的原因
4. 数据与风险的关系



姨搜

2.1 信贷产品的组成部分



2.2 什么数据是有用的数据

- 可观查到并记录下来的都是数据。
- 被动收集与主动收集。
- 不同的数据有不同的价值。



2.3 个人借款风险

- 风险包括：欺诈风险和信用风险
- 欺诈风险：
 - 第三方欺诈：
 - 主动赖账：
- 信用风险：
 - 财务收入：
 - 支出管理：
 - 不良嗜好：



2.4 数据与风险的关系

- 统计模型适合个人信贷决策。
- 相关性比因果关系更重要。
- 数据整理 -> 数据整合 -> 数据特性
- 原始裸数据 -> 简单特征变量 -> 复杂特征变量
 - 信用卡原始交易记录（时间、地点、商家、金额等）
 - 简单变量（不同品类、时间段、地点的消费金额和频次）
 - 简单特征变量的组合特征（在深夜购买游戏点卡的消费金额和频次，等等）
- 模型应用。



姨搜

目录

1. 我们是谁
2. 基于数据如何做风控
3. 风控系统的建设
4. 应用案例
5. 大数据风控的未来



姨搜

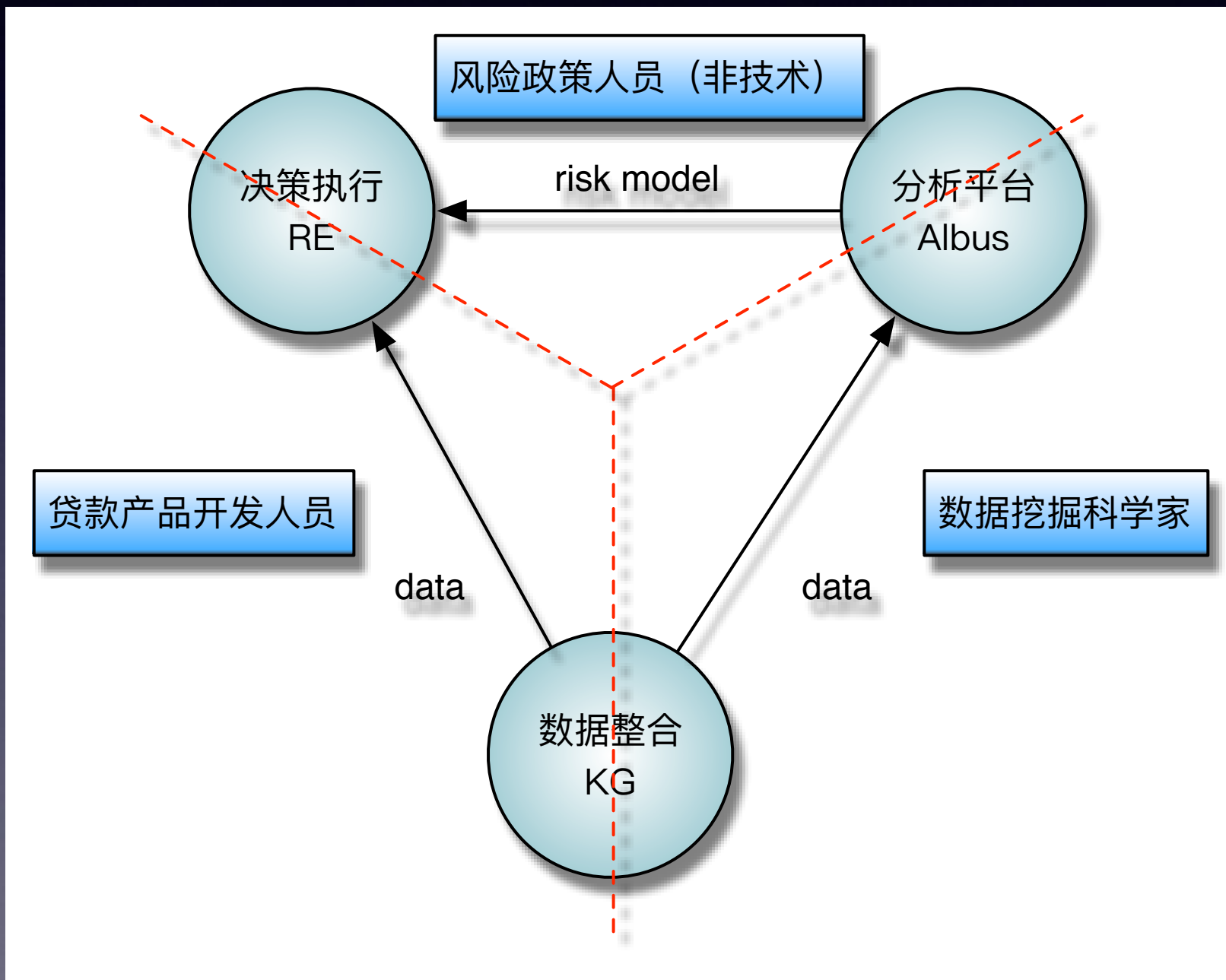
3. 风控系统的建设

1. 总体架构
2. 数据整合
3. 决策引擎
4. 分析平台

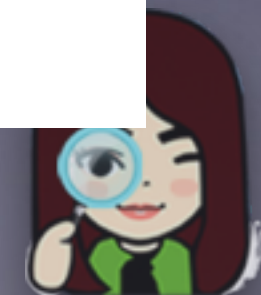
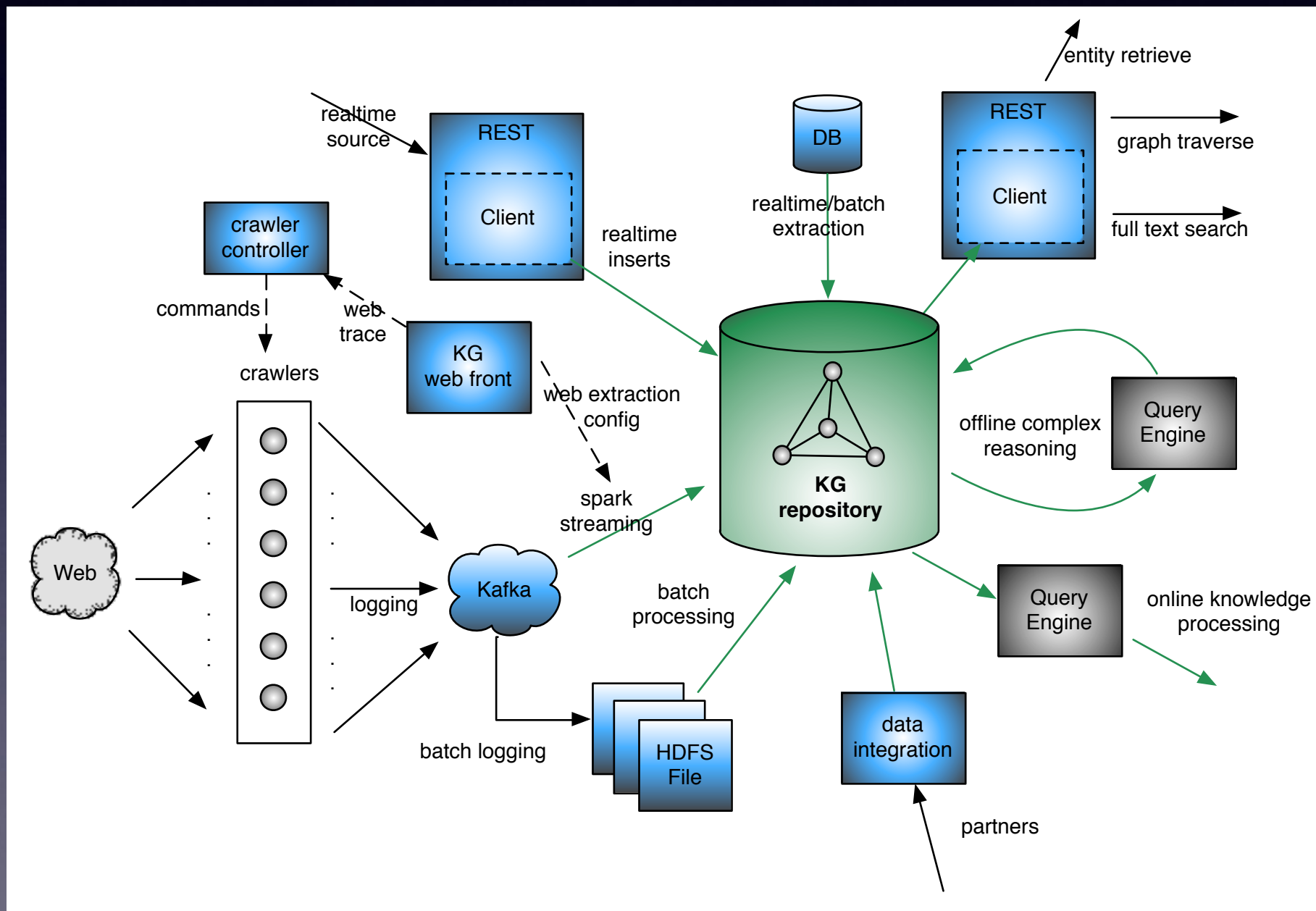


姨搜

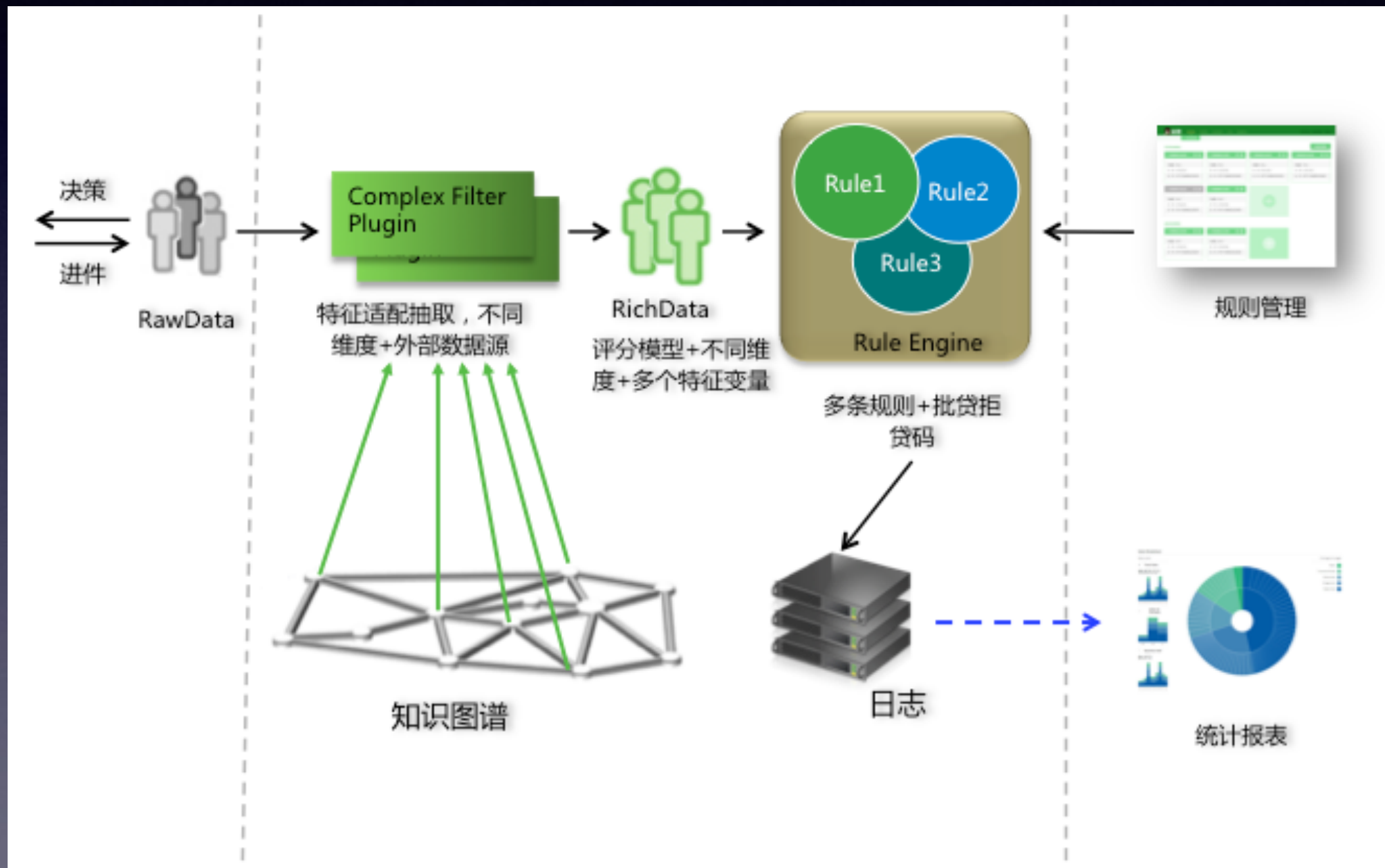
3.1 总体架构



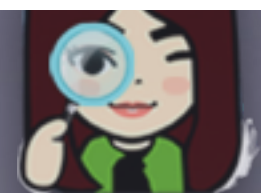
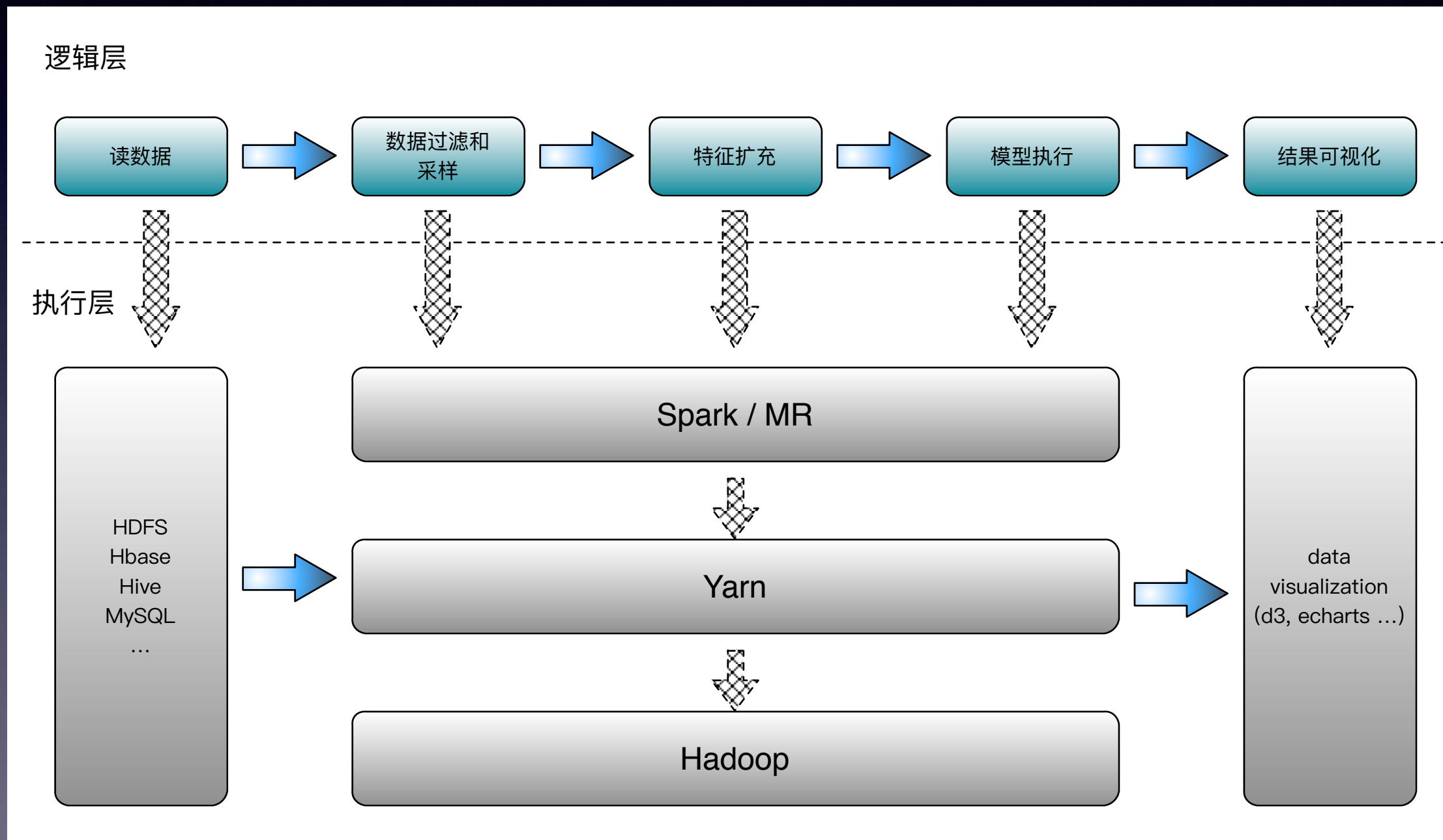
3.2 数据整合



3.3 决策引擎



3.4 分析平台



目录

1. 我们是谁
2. 基于数据如何做风控
3. 风控系统的建设
4. 应用案例
5. 大数据风控的未来



姨搜

4. 应用案例

1. 小额短期借贷场景
2. 审核流程的拆分创新



姨搜

4.1 小额短期借贷场景

- 产品特点：
 - 小额：不超过3000元。
 - 短期：30天以内。
 - 时效性强：有时甚至是一个小时以内就需要。
- 对信审流程的要求：
 - 全自动：有人工干预的话，成本无法控制，而且时效性不强。
 - 客户提供的资料有限：短时间内无法诸如房产证、收入证明等的资料。
 - 高效的试验环境：表现窗口期短，需要利用好试验环境来提高效果。
 - 顺畅的流程反馈：及时发现问题，才能及时反馈到风险政策中。



姨搜

4.1 小额短期借贷场景

- 姨搜自动化网查服务，用公网爬虫抓取各种公开数据，用NLP的方法分析和提取风险信息。
- 强授权爬虫，抓取各种用户授权数据，比如淘宝交易记录或者信用卡账单。
- 利用分析平台Albus，在数据集中快速定位人群，尝试不同的特征变量与政策模板的组合，对结果进行可视化的理解。
- 决策引擎中有对于结果的详细统计分析，并支持自定义的各类报表，各指标的有效性一目了然。



姨搜

4.2 审核流程的拆分创新

- 收集完资料的一次性审核，转变为不同批次资料的步进式审核。
 - 给用户更多的选择空间，金融服务以人为本。
 - 优质用户可以只提供很少的资料，提升体验。
 - 更快的完成审核，提升整体时效性。
 - 之前被直接拒绝的客户，现在有了更多的选择。
- 对信审系统的要求：
 - 系统更易用：将有很多的审核步骤，管理起来会比较繁琐，对易用性提出更高要求。
 - 更高性能：每一步的设计都需要尽量快，留出给用户选择的时间。
 - 过程与结果监控：需要有可定制化的报表，展示每一步的执行细节和统计。



4.2 审核流程的拆分创新

- 姨搜决策引擎支持多步骤依赖管理，在方便管理的基础上，给用户提供了最大的自由度。
- 优化决策引擎自身的性能，并整理依赖的数据服务的统计信息。
- 基于Pentaho的统计系统，可以自定义出各种报表及dashboard。



姨搜

目录

1. 我们是谁
2. 基于数据如何做风控
3. 风控系统的建设
4. 应用案例
5. 大数据风控的未来



姨搜

5. 大数据风控的未来

- 由降低风险，到控制风险，再到利润最大化。
- 更加自动化、系统化、体系化。
- 对不同数据价值的理解和使用方法进一步细化。
- 数据的价值不止体现在贷中的审批，也会在获客、预审批、贷后管理、失联修复等环节起到作用。



姨搜

广而告之

- 校园招聘即将启动，感兴趣的请关注公众号。
- 金融产品经理招聘中，详询jianmindong@creditease.cn



姨搜

Thanks!

Q & A



邮箱: songhou@creditease.cn

博客: <http://housong.github.io>