

宜信大数据实时授信平台

xuanwei@creditease.cn

传统批贷流程

- 上网学习攻略
- 准备申请材料
 - 打印、复印、盖章等等
- 到柜台，提交各种证明材料
- 补交各种证明材料
- 等待若干工作日，通知结果与额度
 - To be or not to be



- 用户提交户个人信息
 - 身份证号，联系人等
- 用户授权获取信用卡账单
 - 最近4个月信用卡账单及明细
- 实时反馈额度
- 等待放款

瞬时贷 - 用户体验

宜信
CreditEase
普惠金融

周年
2006-2014
财富管理

- 方便
 - 手机上完成，无需任何线下材料
- 快速
 - 实时授信
 - 次日放款
- 充分互动
 - 第一时间掌握额度



瞬时贷－宜信体验

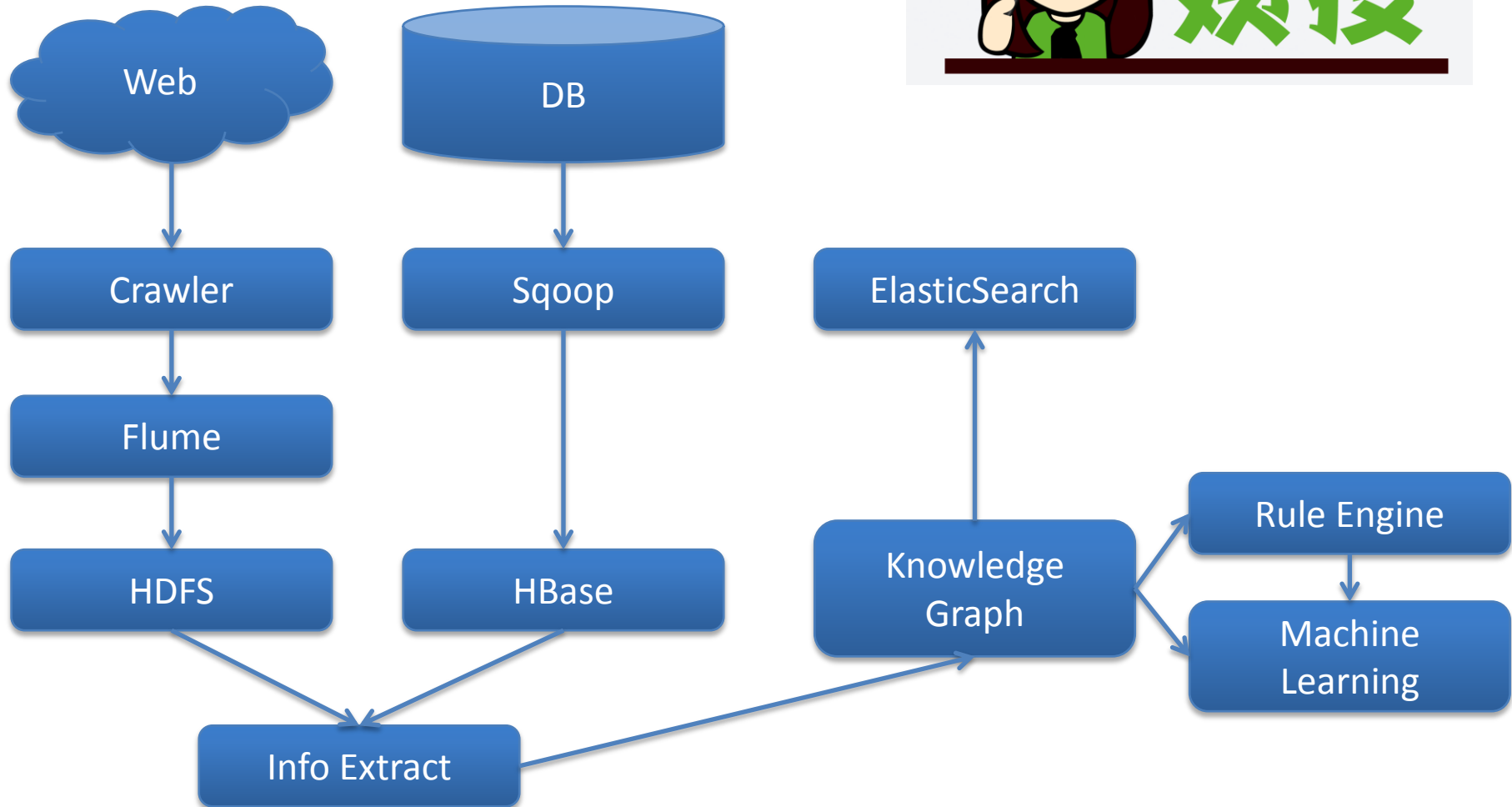
宜信
CreditEase
普惠金融



- 成本
 - － 人工成本几乎为0
- 低违约率

大数据实时授信平台

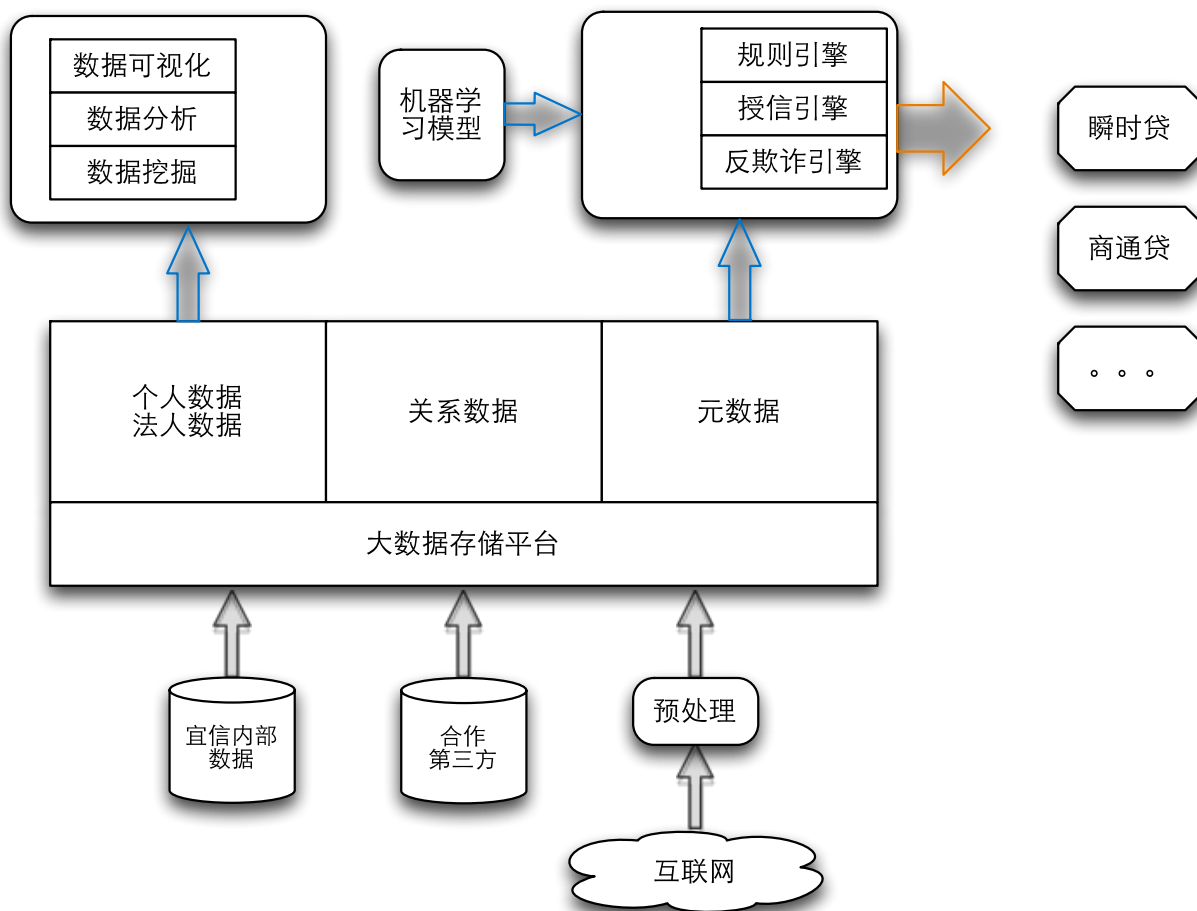
- 分析客户的信用状况和欺诈风险，实时估计授信额度、检测欺诈风险，或者辅助非实时业务进行信用及欺诈估计



大数据的边界

- 个人或者商户隐私数据一定经过用户的授权
 - 授权数据只用于宜信自己的业务
- 网络公开数据的边界以通用搜索引擎为参照

架构



- 传统授信模型
 - 作为人工的辅助手段
 - 强特征
 - 可解释
- 我们没有边界
 - 一切数据皆为信用数据
 - 维度无限制，强弱无限制
 - 一切模型我们均可使用

瞬时贷－数据

- 用户提交信息
 - － 详单信息也非常有用
- 通过姨搜获取用户的一切相关信息
 - － 用户的常用论坛，发帖记录
 - － 用户的消费水平
 - － 用户不良嗜好
 - － 用户地址房价高低

瞬时贷－信用模型

- 数据特点
 - － 少数强特征
 - 账单信息
 - 非线性
 - － 大量弱特征
 - 账单详单
 - 互联网信息
 - 非线性
- 模型
 - － GLM(LR)
 - 挖掘海量弱特征
 - － Random Forest
 - 处理强特征
 - － 模型融合

- 规则引擎
 - 专家根据对申请人群的了解，指定规则
 - 用户是否有账单逾期未还？
 - 用户身份信息是否真实？
 - 可以解决很大一部分问题
- 冷启动
 - 在产品上线前，没有分类标签
 - 专家定义什么是坏客户，模型计算什么人看起来是好客户，但是很有可能变为坏客户
 - 随着产品数据积累，不断更新模型

反欺诈模型

- 起点
 - 数据校验和黑名单库
- 利用所有数据进行校验
 - 直接发现异常、欺诈或者疑似欺诈信息
 - 生成安全问题，确认申请人身份

建模总结

- 数据很重要
 - 提取更多的特征
 - 提取更多的样本
 - 积弱为强，不要丢弃
- Domain Knowledge 很重要
 - 尽量理解数据
- 模型
 - AD领域，GLM(LR)几乎是唯一选择
 - 但是可以有变种
- 数据依然重要
 - 除了视频，我们都去研究
- Domain Knowledge 更加重要
- 模型
 - 复杂模型终于派上用场
 - 随着认识的增加，迭代更合适的模型

- 没有一种单一的算法可以解决所有的问题
 - 要通过工程将各种算法组合成强大的系统
- 单纯靠机器和单纯靠人都是不够的
 - 充分发挥专家的先验知识和机器的计算能力
- 道高一尺，魔高一丈
 - 监控，报警，更新，进化

- 应对
 - 数据充足的用户，率先享受便宜快捷的互联网服务
 - 数据不足的用户，提供信用估计供其他产品线使用
- 未来
 - 继续积累数据
 - 加强行业合作

Thanks