

# 宜信大数据金融云平台实践

郑华，宜信大数据创新中心研发总监  
huazheng@creditease.cn

## DTCC

### 2015中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2015

大数据技术探索和价值发现



# 概述

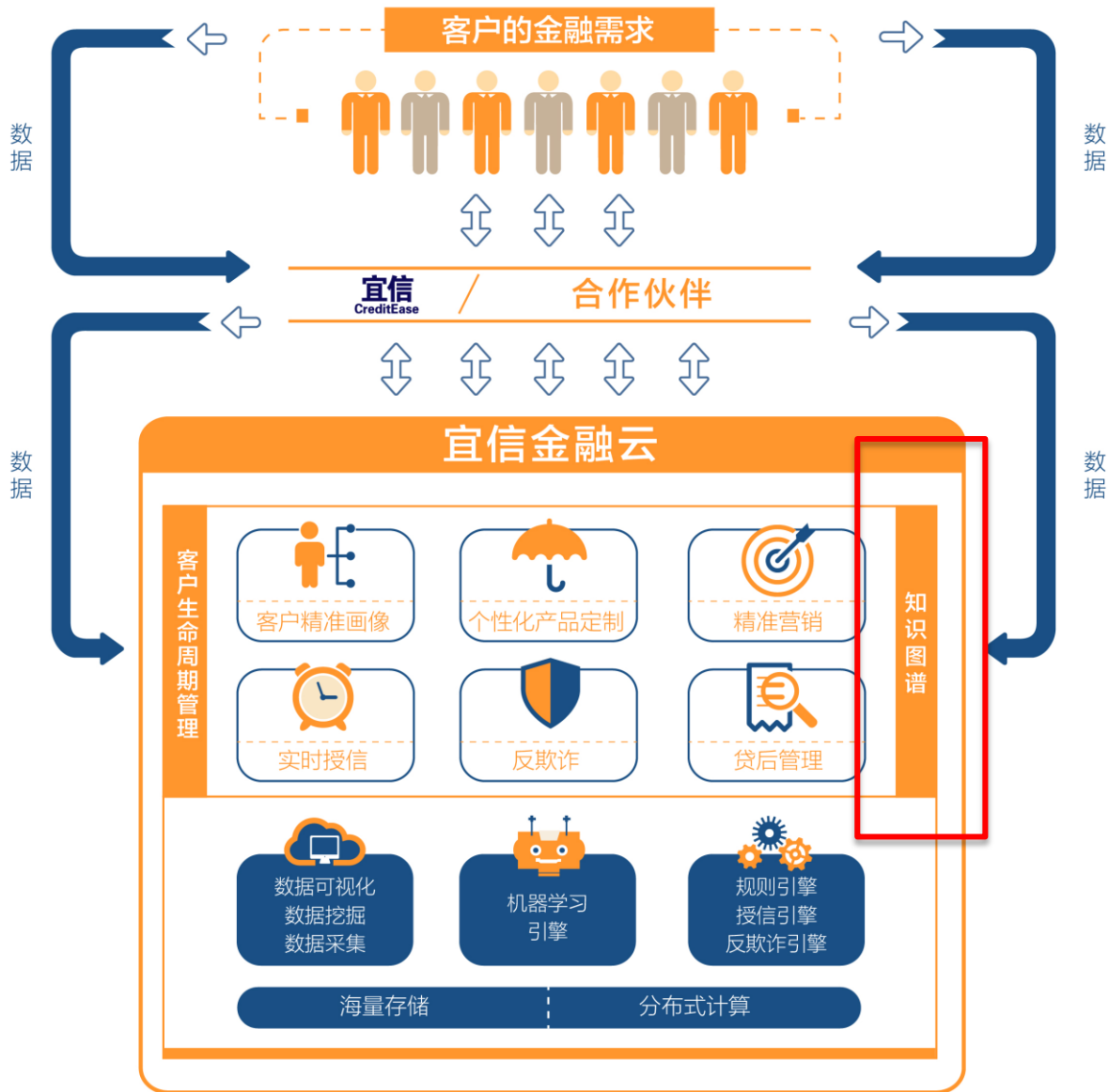
- 宜信金融云平台
- 知识图谱
- 商通贷
- 展望



# 宜信金融云平台

致力于打造世界一流的**金融云平台**，为宜信自身的各项业务以及宜信生态圈内合作伙伴的各项业务提供最好的金融云服务





# 知识图谱概念

- 知识图谱是一种数据表现形式，采用自然知识中 **entity-property-link** 的形式，拥有强大的表现能力和扩展能力，为 **knowledge-based reasoning** 提供基础。
- 以 **entity** 为中心，每个 **entity** 拥有独特的属性值，**entity** 之间有多样的关联关系
- 灵活的 **schema**，增加新的知识更加简单；
- 新应用可以无缝重用已有数据，加快开发进度



# 知识图谱的应用

- 实体查询  
黑名单查询：直接匹配 & 间接联系  
客户查重
- 实体扩展  
通过属性查询实体 => 获取实体的其他属性  
获取相似实体
- 知识查询  
例：XX公司的法人是否有失信记录  
例：XX写字楼中有多少家机构入驻
- 个性化问答  
由客户信息生成个性化问题  
根据客户回答，检测虚假身份等欺诈风险



# 知识图谱的建立

元数据  
定义

数据获  
取和转  
化

数据间  
关系确  
立

提高系  
统智能

提高系  
统性能



# 知识图谱的元数据定义

```
m:people.person rdf:type owl:Class ;  
  rdfs:label "人"^^xsd:string ;  
  rdfs:subClassOf m:common.topic ;
```

```
m:people.person.name rdf:type owl:DatatypeProperty ;  
  rdfs:domain m:people.person ;  
  rdfs:label "姓名"^^xsd:string ;  
  rdfs:range xsd:string ;
```

```
m:people.person.work_at rdf:type owl:ObjectProperty ;  
  rdfs:domain m:people.person ;  
  rdfs:label "工作单位"^^xsd:string ;  
  rdfs:range m:organization.organization .
```

```
m:organization.organization rdf:type owl:Class ;  
  rdfs:label "组织"^^xsd:string ;  
  rdfs:subClassOf m:common.topic ;
```



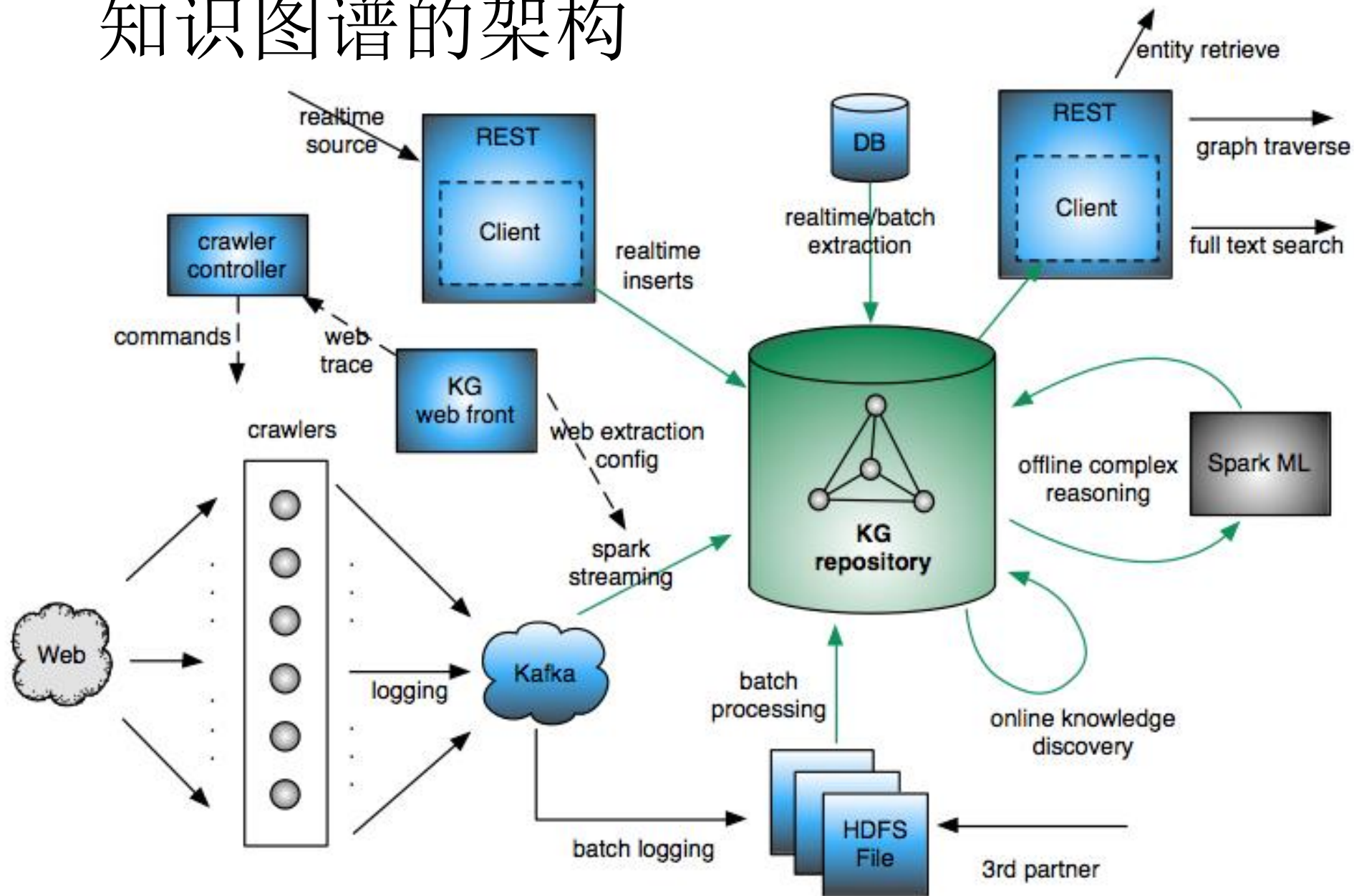


# 知识图谱的数据转化

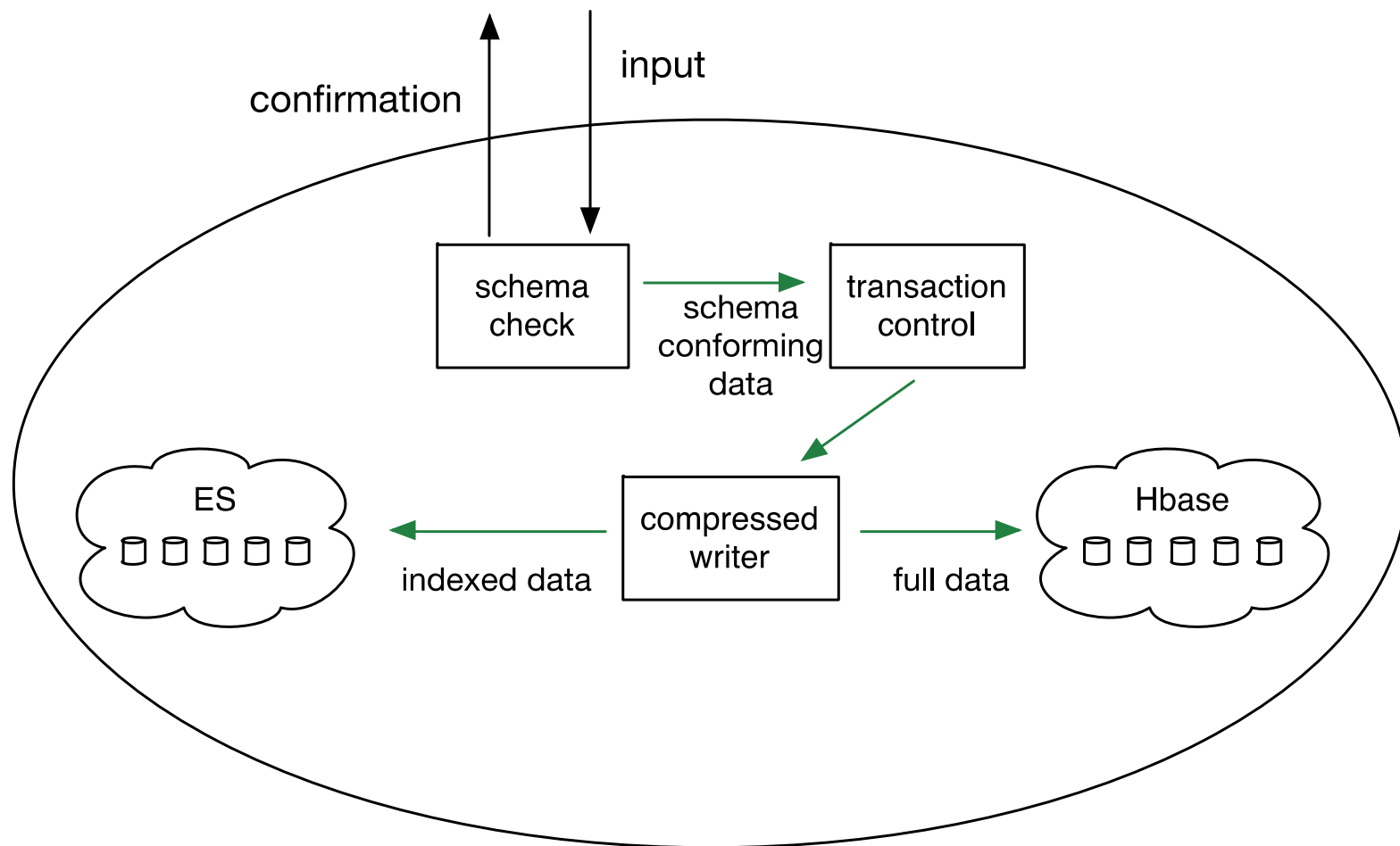
e:borrower\_1, rdf:type, m:finance.borrower  
e:borrower\_1, m:people.person.name, "张三"  
e:borrower\_1, m:people.person.phone, "13800138000"  
e:borrower\_1, m:people.person.home\_addr, e:location\_1  
e:borrower\_1, m:people.person.work\_at, e:company\_1  
e:location\_1, rdf:type, m:common.location  
e:location\_1, m:common.location.full\_addr, "北京市朝阳区温特莱中心"  
e:company\_1, rdf:type, m:finance.company  
e:company\_1, m:finance.company.name, "宜信普惠"  
e:company\_1, m:finance.company.telephone, "400-6099-800"



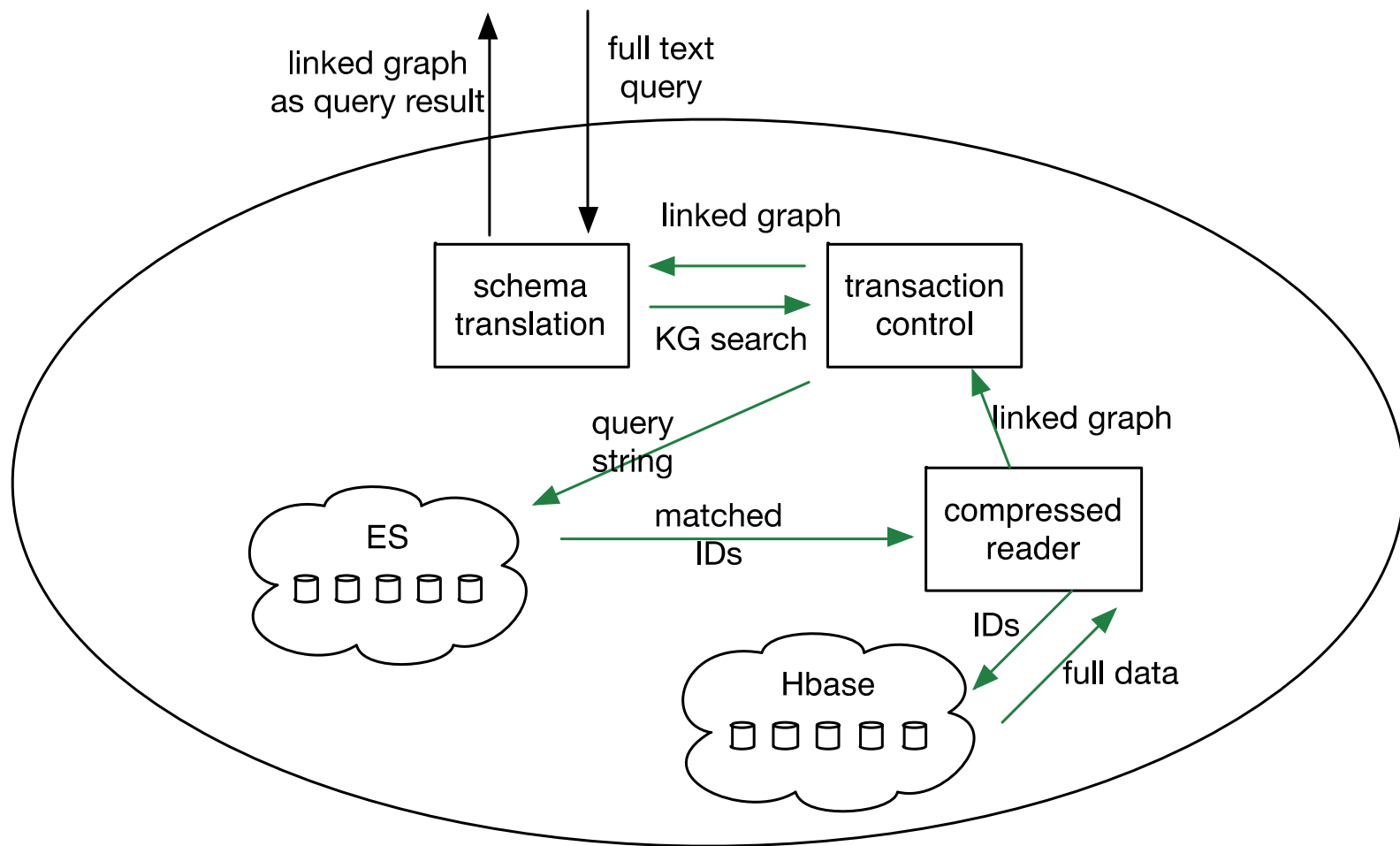
# 知识图谱的架构



# 知识图谱的写入



# 知识图谱的检索



# 知识图谱小结

- 将所有信息类别以统一的形式组织起来并建立连接关系，给了我们一个全新的视角来观察世界，开启数据管理的新模式
- 知识图谱的网状数据和人工智能与机器学习结合，能顺畅的使用所有能取得的数据，可以让大数据给互联网金融和征信带来更多机会。

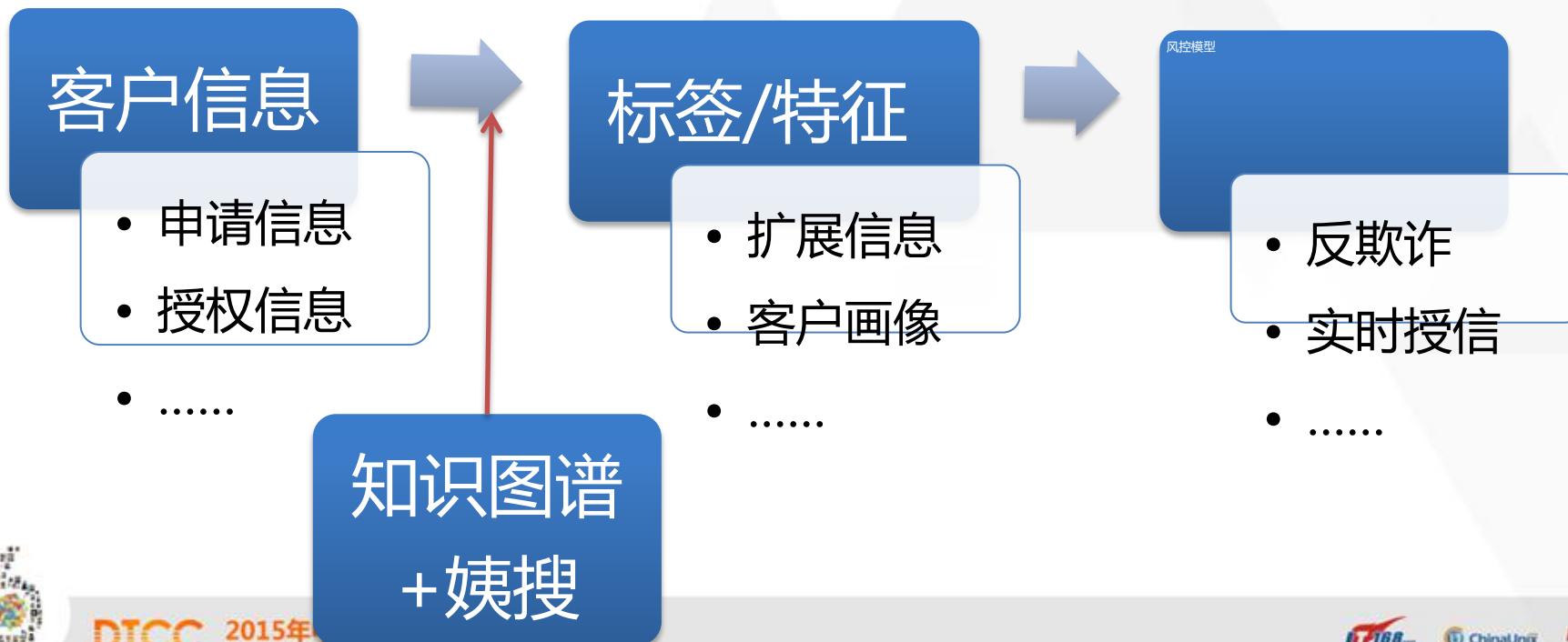
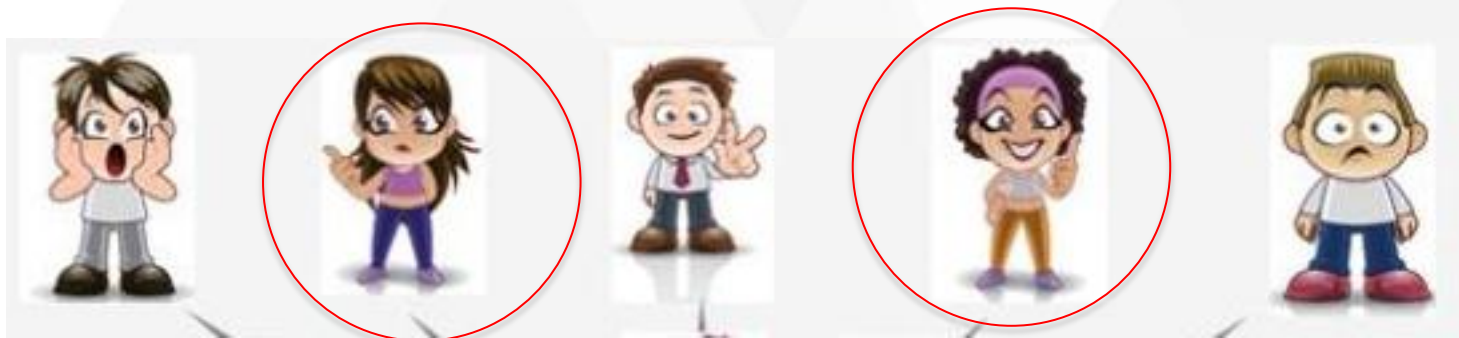


# 大数据给互联网金融带来什么

- 精准获客
- 反欺诈
- 风险评估
- 实时授信
- 交叉销售
- 贷后管理
- ...



# 贷款流程



# 商通贷

*Working  
Capital At  
Fingertips*

- 实时授信
- 全线上
- 低门槛
- 全平台支持
- 开放平台

ebay / amazon / Aliexpress 全球速卖通 / 淘宝网 Taobao.com / 天猫 TMALL.COM





# 还款

按时还款或提前还款，保持良好的还款记录，下次借款会有奖励额度哦

本期应还86,888.14元

☒ 按月还款：86,888.14元

☒ 提前还款：498,624.09元

☒ 其他金额： 元

还款

还款截止日期	本金(元)	息费(元)	逾期违约金(元)	合计款项(元)	实付款项(元)	还款状态
2015-02-12	78,264.08	8,624.06	0.00	86,888.14	0.00	未还款
2015-03-12	79,594.57	7,293.57	0.00	86,888.14	0.00	未还款
2015-04-12	80,947.68	5,940.46	0.00	86,888.14	0.00	未还款
2015-05-12	82,323.79	4,564.35	0.00	86,888.14	0.00	未还款
2015-06-12	83,723.29	3,164.85	0.00	86,888.14	0.00	未还款
2015-07-12	85,146.59	1,741.55	0.00	86,888.14	0.00	未还款

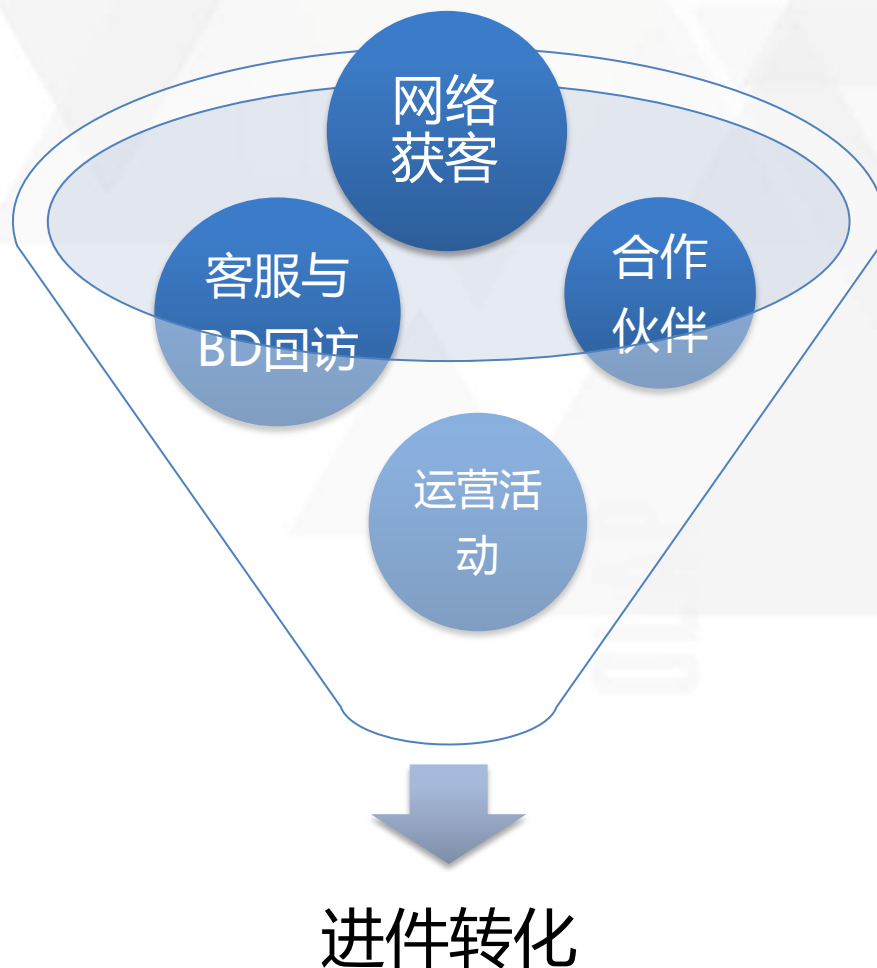


# 三大要素

- 大规模精准获客
- 优秀的风险管理
- 客户的高粘性



# 商通贷－获客



# 商通贷 – 网络获客

- 通过爬虫定向抓取平台上所有的店铺
- 根据店铺的创建时间，品类，评级，评价 (Detailed Seller Ratings)，上架货物，单价及订单数等公开信息，过滤出符合进件条件的店铺，并制定相应的营销策略
- 通过站内信，自动聊天等方式进行第一轮自动营销，筛选出感兴趣的客户后，采用人工的方式电话或者线下拜访进行转化

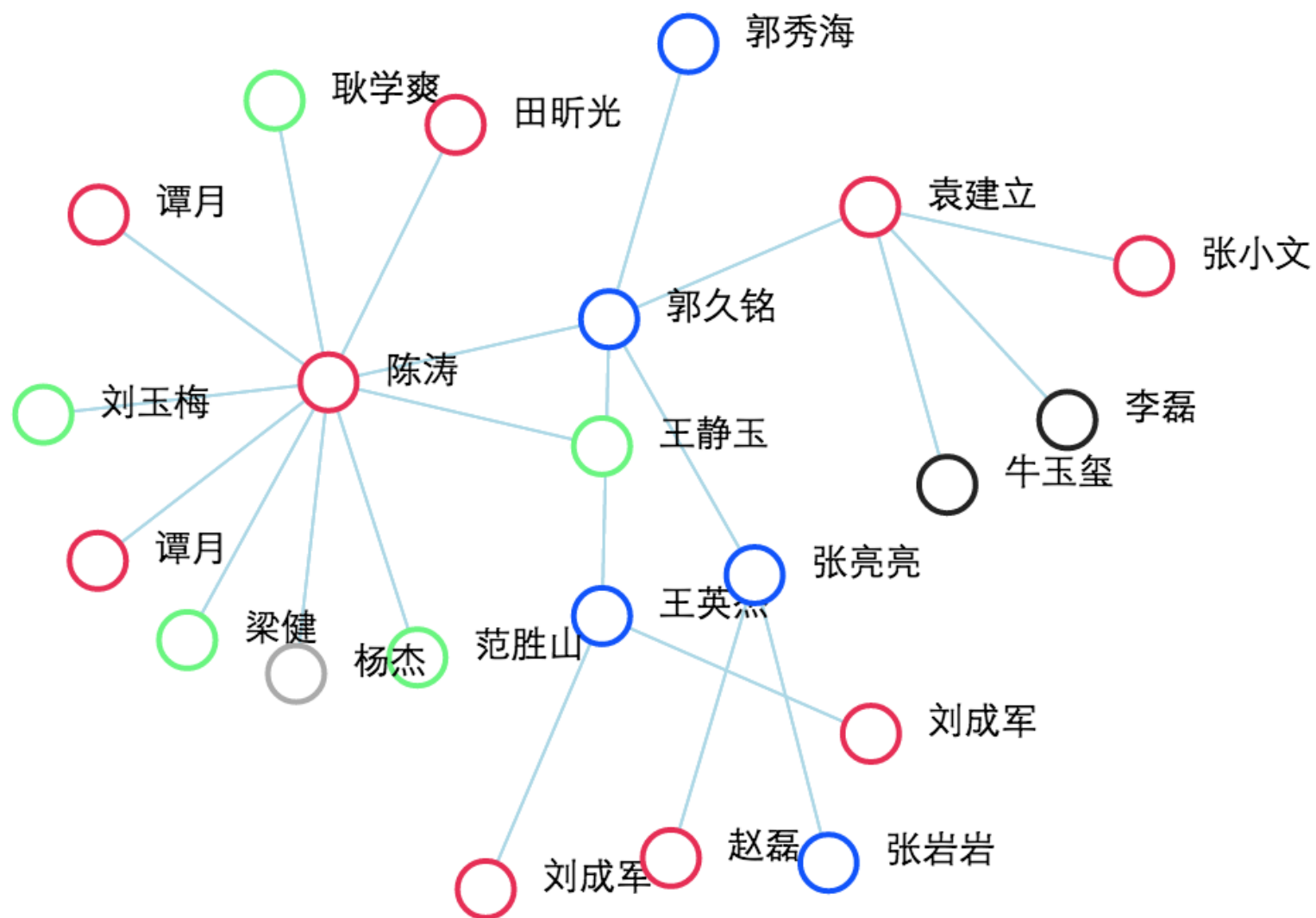


# 商通贷－欺诈风险

- 身份：平台账号，ID5，交叉验证，问答小姨
- 行为：网络活动，经营活动
- 关系：图谱搜索



# 商通贷－欺诈风险



# 商通贷－信用风险

- 还款意愿+还款能力
- 综合客户的基本信息、网店经营数据、个人信用数据、个人消费数据和互联网信息等，计算客户的综合信用评分，根据评分决定审批和风险定价
- 根据客户还款能力和实际的资金需求进行实时授信



# 商通贷 – 平台数据维度

- 经营
  - 持续经营时间
  - 经营规模
  - 经营稳定性
  - 周转速度
  - 现金流
  - 物流
- 客户
  - 客户规模
  - 客户来源 (免费, 收费, etc)
  - 客户评价
  - 网络人气和口碑
- 行业
  - 行业竞争力
  - 行业风险

...





# 商通贷 – 平台数据获取

- 合作伙伴提供;
- 采用OpenAPI的方式获取;
- 或者安装“商通贷助手”插件的方式获取数据;
- 或者QQ远程截屏征信;



# 商通贷 - 商通贷助手

商通贷

扩展程序

Alibaba Manufacturer D

datamatrix.aliexpress.com/wsDISellerAnalysis.htm

欢迎您! [Seven](#) | [站内信](#) | [退出系统](#) | [速卖通规则](#) | [卖家论坛](#) | [速卖通教程](#) | [国际版旺旺](#) | [买家频道](#) | [建议反馈](#)

**AliExpress 全球速卖通**  
阿里巴巴旗下网站

English

我的速卖通 | 产品管理 | 交易 | 站内信 | 店铺中心 | 账号设置 | 营销

经营分析

实时风暴

店铺概况

店铺流量来源

店铺装修

商品分析

商机发现

行业情报

选品专家

搜索词分析

产品支持

当前账号

开始截屏, 请稍候  
请不要关闭当前标签, 商通贷插件  
图标会显示实时抓取进度

数据纵横 BETA

在线咨询

店铺概况

店铺经营看板详情

店铺概况

**重要提示:** 目前您店铺的7日成交额, 转化率和访客数排名均使用与您同行业的卖家进行对比, 同类最佳店铺也采用此类算法, 请卖家朋友知晓!  
【平台规则】[搜索作弊行为评判标准](#)  
【重要通知】[双十一大促期间实时风暴功能修复说明!](#)

最近7日店铺概览

购买率 0.0%, 低于网站均值 0.0%;

选择时间: 最近7天 | 最近30天 | 最近90天 | 起始时间 | 到 | 结束时间

房产证! 无需银行流水! 无需存单质押! 只看店铺综合经营情况, 我们就能确

# 商通贷－个人信用数据

- 央行征信报告
- 宜信内部贷款记录
- 其他贷款记录
- 信用卡还款记录
- ...



# 商通贷 – 评分模型

- 数据特点
  - 少量强特征 + 大量弱特征（碎片化，缺失率高）
  - 与预测目标呈非线性关系
- 可解释性 vs. 预测效果
- 模型
  - L1 Logistic Regression
    - 挖掘海量弱特征
  - Random Forest
    - 处理强特征
  - 模型融合



# 商通贷－授信模型

- 考虑客户的资金需求和还款能力
- 综合行业趋势(外贸出口/进口，内贸，电商)，平台发展(ebay, amazon, wish, 移动，微商)，经营品类(服装，3C，母婴，汽配)，销售市场(欧美，发展中)，DSR，流量来源，市场能力，销售额，利润率，增长率，贷款期限，周转天数，季节因素，运营开支，外部负债等等



# 商通贷－贷后管理

- 根据客户的还款情况以及平台的最新数据，计算客户的行为评分
- 依据行为评分，采取一定的贷后管理措施，比如主动提额或者提前催收
- 对于失联客户，找到可能的联系方式，以帮助催收



# 商通贷 – 客户粘性

- 极低的坏账率
- 很高的循环贷比例
- 不错的客户口碑



# 商通贷－大数据应用

- 获取数据：内部，第三方，客户授权，网上公开信息
- 了解客户：客户画像
- 精准获客：网络抓取，自动定制化营销配合线下转化
- 反欺诈：身份，行为，社交
- 风险评估：大数据信用评分模型，收益覆盖风险
- 授信模型：客户需求，还款能力
- 贷后管理：行为评分，失联查询
- ...





# 展望

- 构建生态系统，接入更多的数据和流量  
—一切数据的变现伙伴
- 更多的创新应用  
—风控，互联网金融





THANKS