



关系图谱在贝壳找房风控体系的应用与实践





讲师简介



王学志
资深算法工程师



- 中科院博士
- 先后在华为、百度从事大数据、关系图谱相关的工作
- 现就职于贝壳找房-人工智能技术中心
- 负责关系图谱在贝壳找房风控侧的落地





目录 CONTENTS

1

贝壳的业务介绍

2

风险挑战及应对

3

图谱的落地应用

4

未来的展望





ACN (Agent Cooperation Network) 经纪人合作网络

扩展业务边界，改变过去零和博弈



人的合作网络

保障店东和经纪人在平台上的权益

房的合作网络

增加物件曝光量，提高房源成交率

客的合作网络

跨区找房全程跟随，提升客源有效利用





目录

CONTENTS

1

贝壳的业务介绍

2

风险挑战及应对

3

图谱的落地应用

4

未来的展望





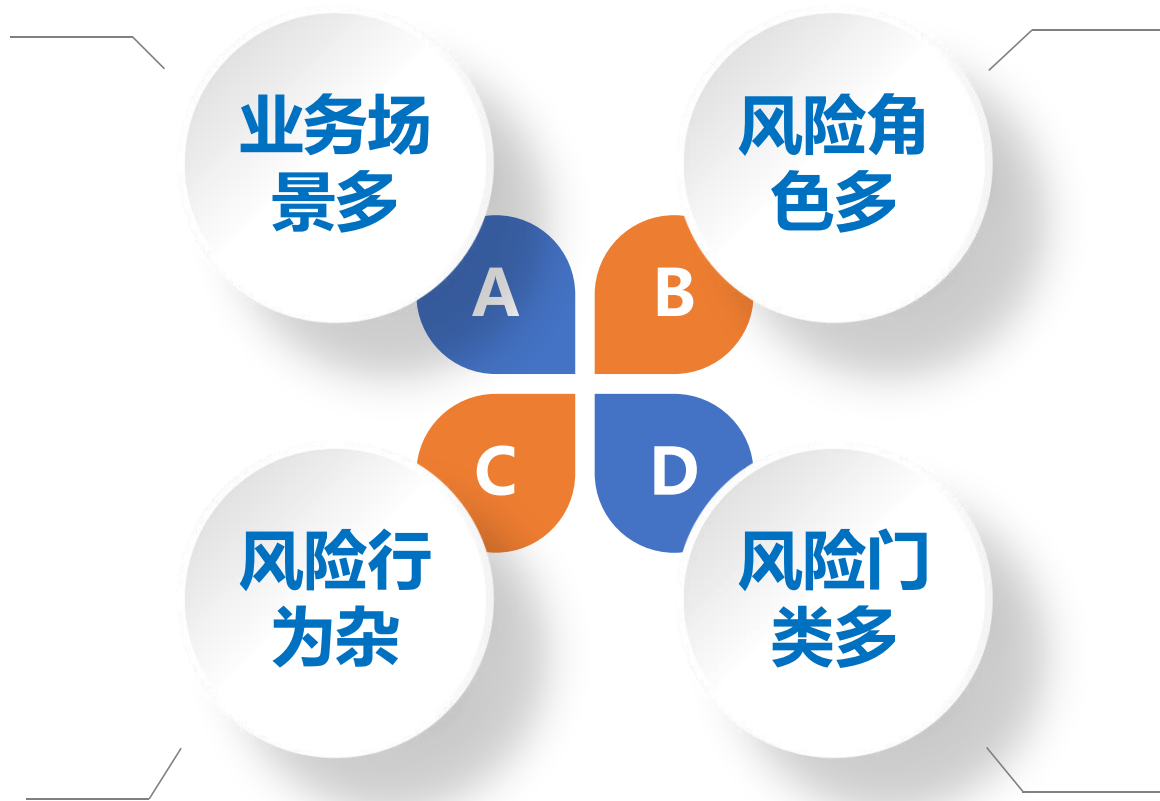
基于“房”的场景

- 新房/二手/租赁
- 装修/维修
- 家居服务
- 家居卖场
- ...

线上+线下

- 虚假流量
- 垃圾商机/IM骚扰
- 虚假房源/客源
- 藏客/藏房/虚假带看
- 影子经纪人
- 资源泄露
- 私单飞单
-

低频+大额+非标+长周期



关于“房”的人

- B端经纪人/C端客户
- 店东/圈经
- 加盟商/开发商
- 自有员工/...

关于“房”的风险

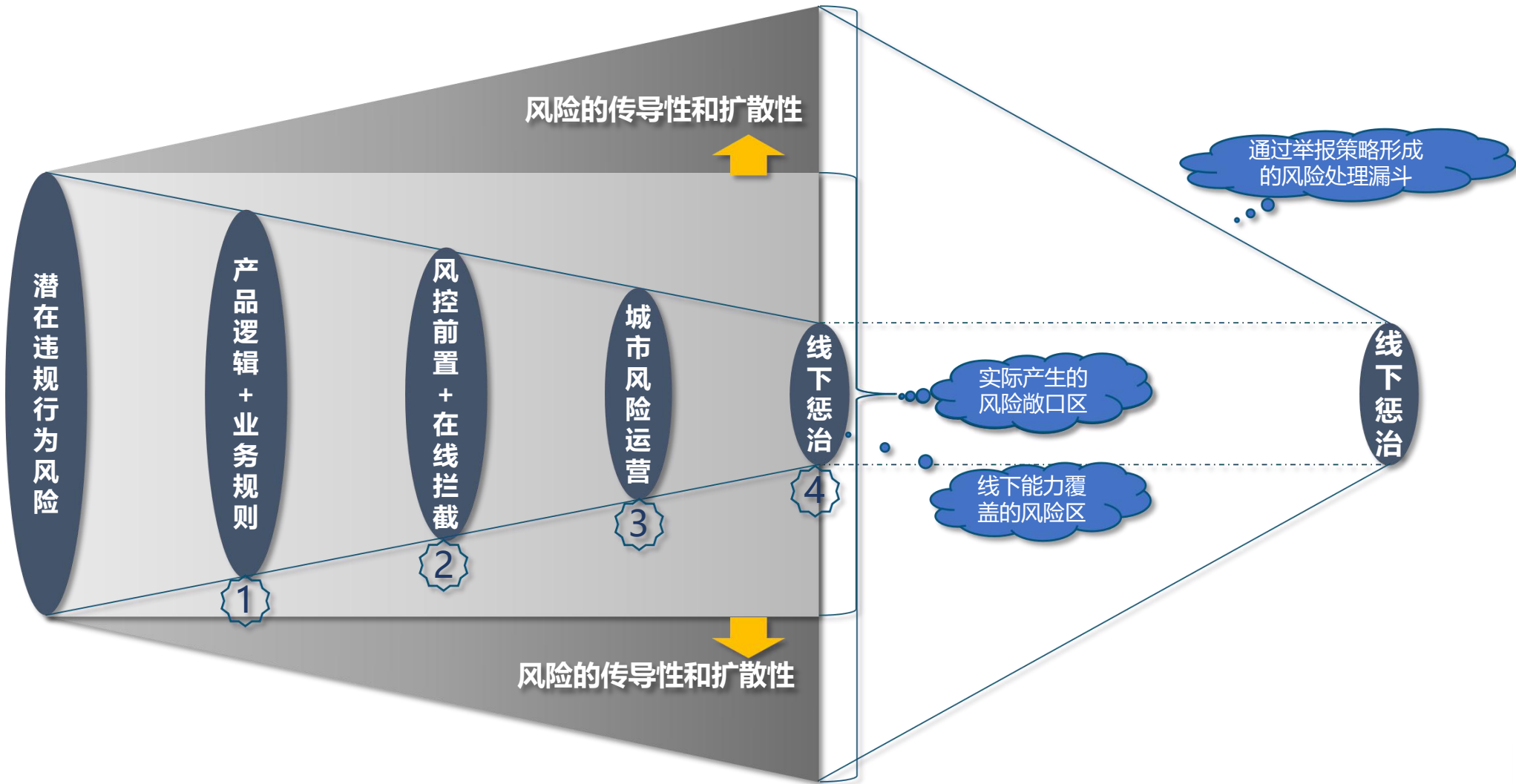
- 作业风险/营销风险
- 虚假服务承诺
- 资金监管风险
- 产权纠纷风险
- 经纪人道德风险
- 合规风险
- 职能舞弊风险
-





贝壳的风控理念

- ① 坏的行为具备“传染性”
- ② 风险管控是个系统性工程





贝壳的风控架构

msup[®]

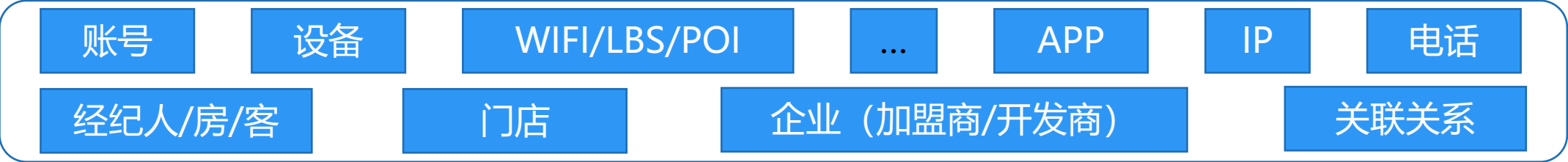
应用层



决策层



画像层



集市层



数据层





目录 CONTENTS

1

贝壳的业务介绍

2

风险挑战及应对

3

图谱的落地应用

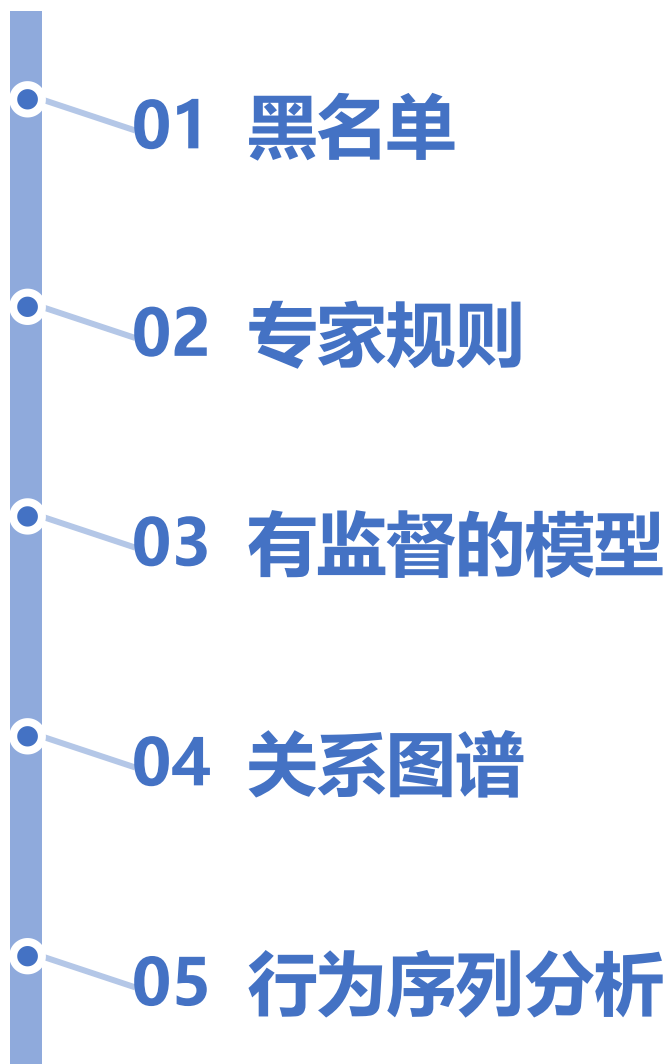
4

下一步工作展望

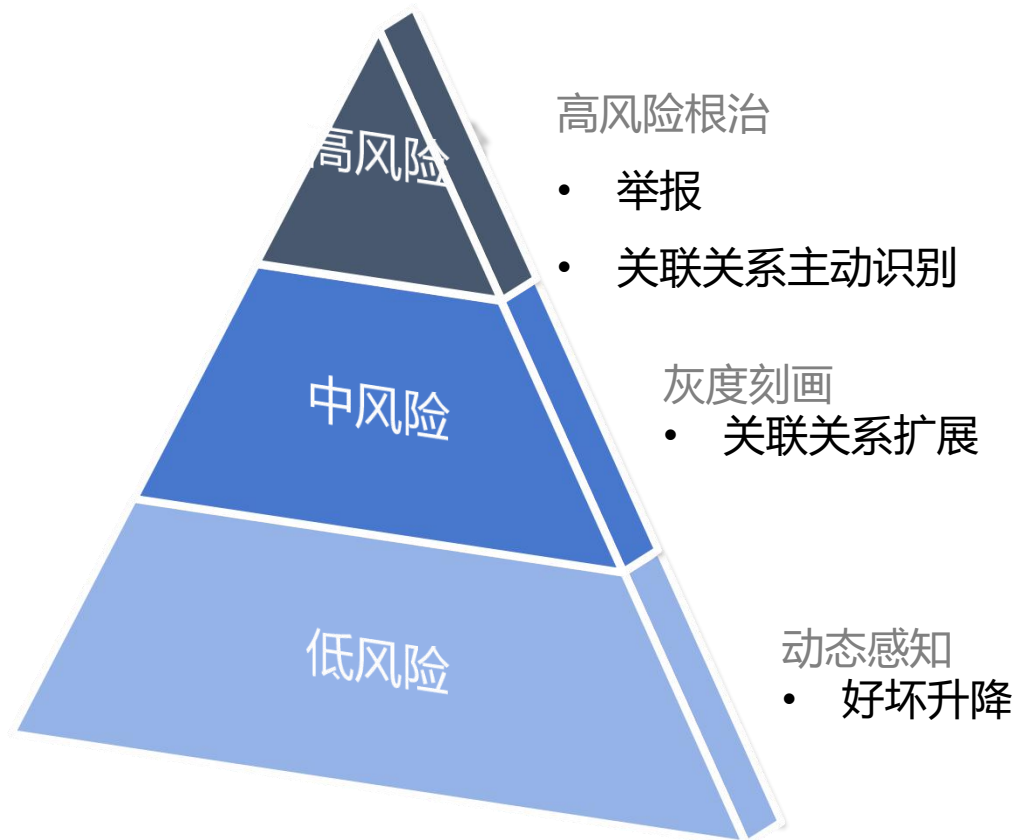




什么是关联关系



贝壳风控怎么用





业务应用层

准入防控

风险量化

品质管理

风险发现

查案溯源

知识挖掘层

最短路径

图聚类

Graph Embedding

知识推理

关联变量

标签传播

社区发现

骨干网络

关键路径

知识构建层

数据清洗

图谱设计

实体抽取

关系抽取

知识融合

基础数据层

投诉

行为数据

举报





白盒+黑盒
名单+模型
离线产出+在线服务

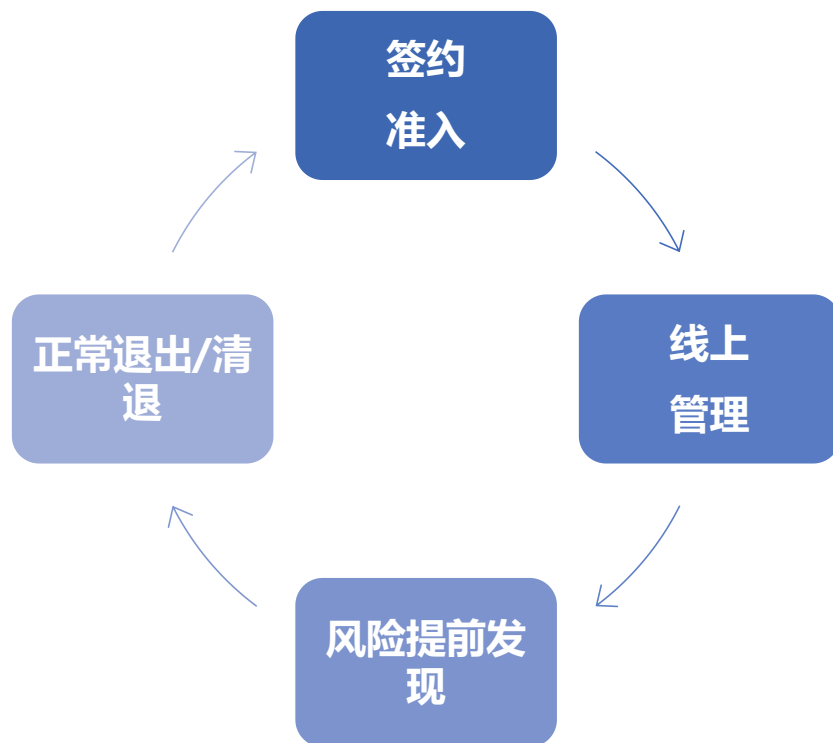






事前准入-工商风险

- 加盟商-全生命周期-线上化-风险识别



- 难点

- 企业库-张三：A公司？=B公司
- 张三：贝壳体内？=企业库？
- 外采数据不提供人的唯一id

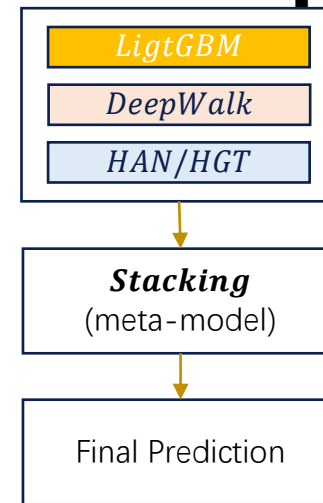
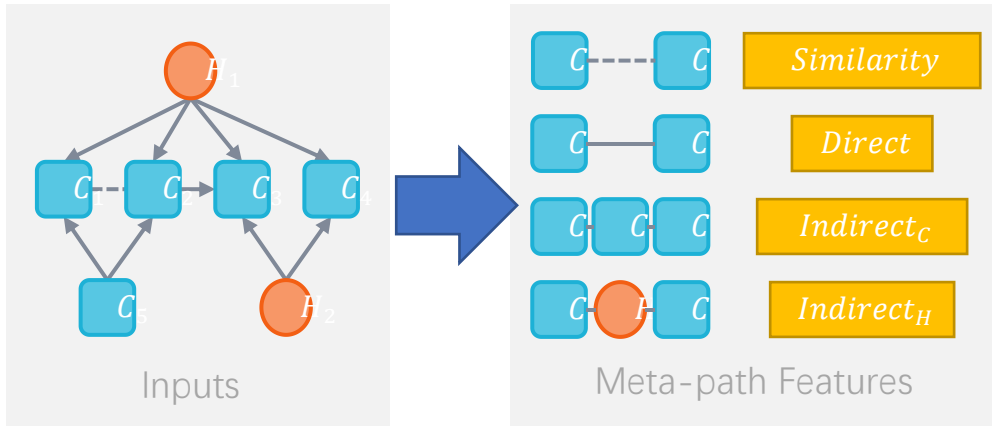




事前准入-工商风险

msup[®]

同名人员识别



V1

- 人名共现
- 总分公司



V2

- meta-path 通用特征框架
- lightGBM训练
- Spark分布式推断
- 可解释模型



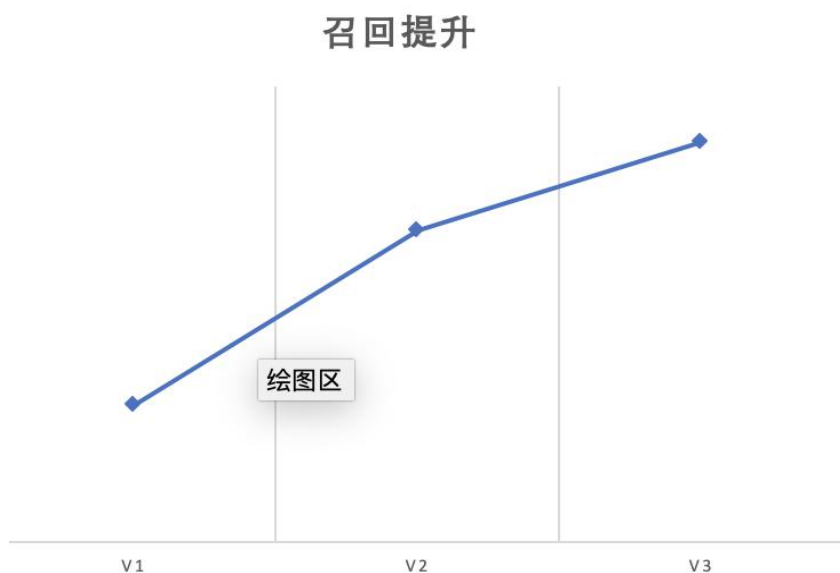
V3

- 扩大召回
- 引入图表征学习HAN
- Stacking融合





落地效果

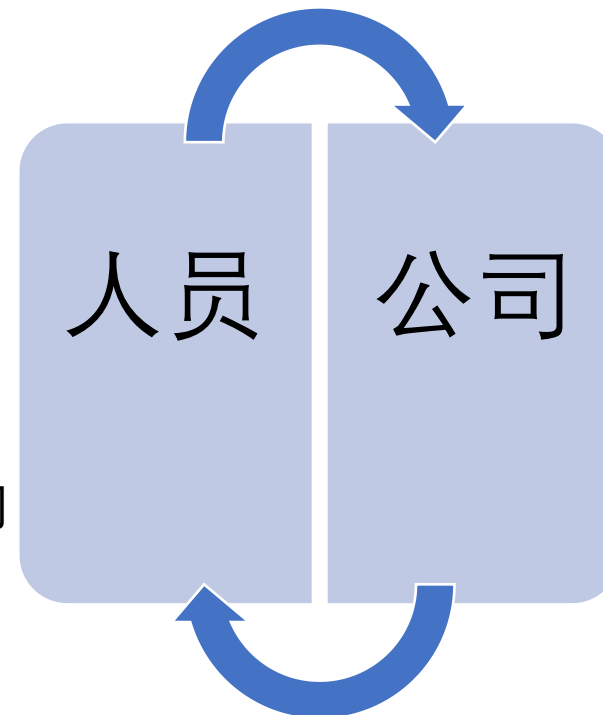


同精准率下，召回V3较V1提升**1倍+**

二手/新房生态治理

覆盖角色

- 圈经
- 店东
- 经纪人
- 置业顾问
- 客发



- 体外公司
 - 门店公司
 - 加盟商公司
 - 开发商公司
- 公司风险监测
 - 提前预警

V3较V1多识别**40%**人员风险，**25%**门店风险





事中-风险量化

事前

业务方吐槽有问题，是不是真的有问题？到底严重否吗？

事中

风控识别的是否真的解决了问题？策略当前是否失效？

事前

如何能够料敌于先，提前感知到风险的变化？

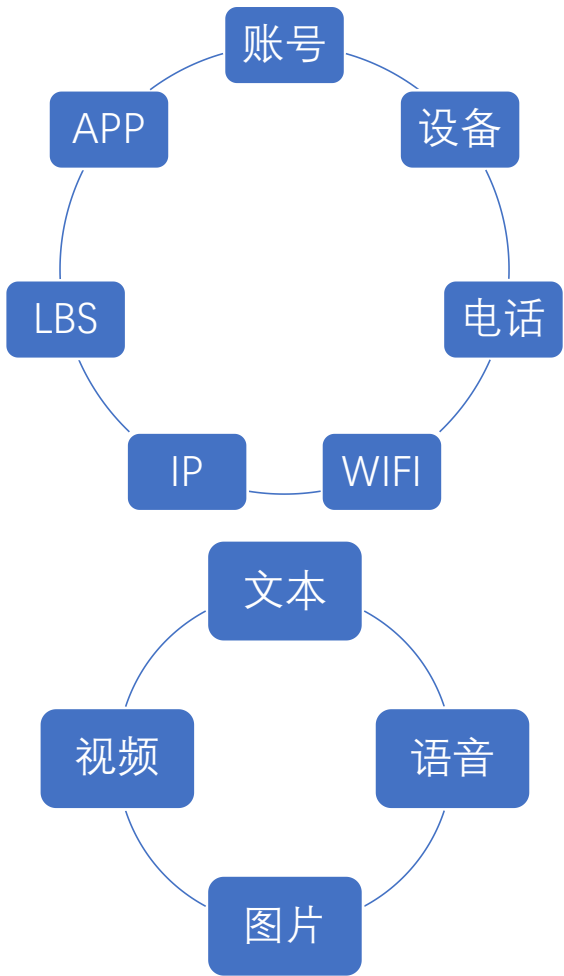
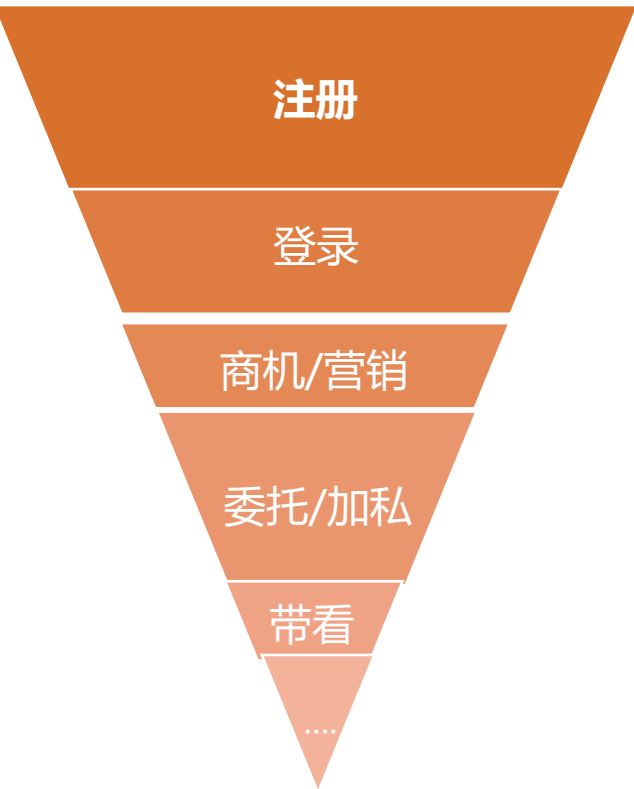
事后

还有哪些风险？
还有哪些识别盲区？





全周期-多维度-立体感知



离线+在线

集中度预警

波动预警

统计分布预警

文本/图片/...
聚类

二部图

动态网络结构

- 环境异常
- 行为异常
- 业务异常
- 团伙异常

- 流量/商机/房源/客源/渠道/门店/大区/城市





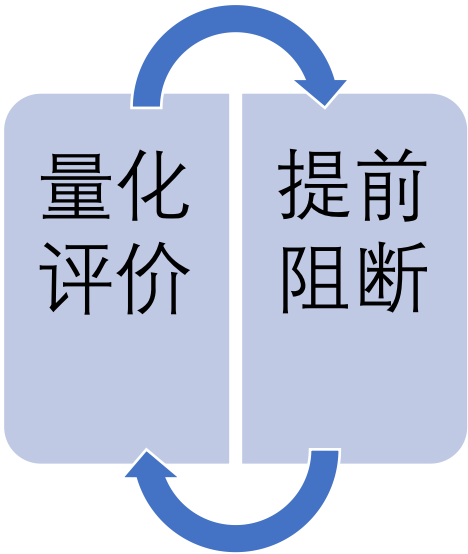
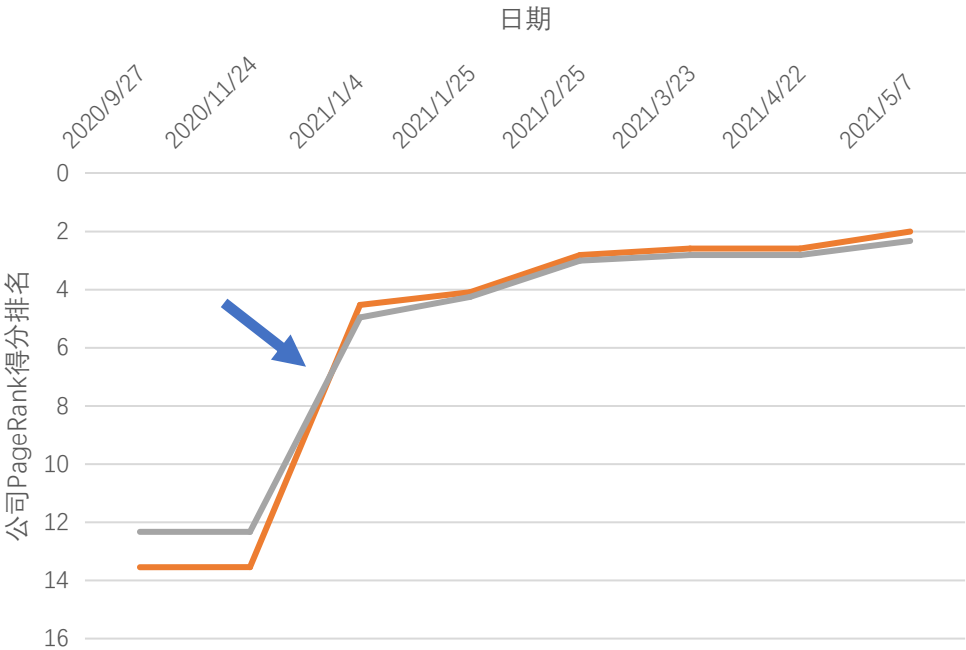
- 公司风险感知

- IM骚扰

展现形式

竞争	微商	恶意攻击
----	----	------

某公司PageRank得分排名-时间-趋势图



单点维度+群体维度

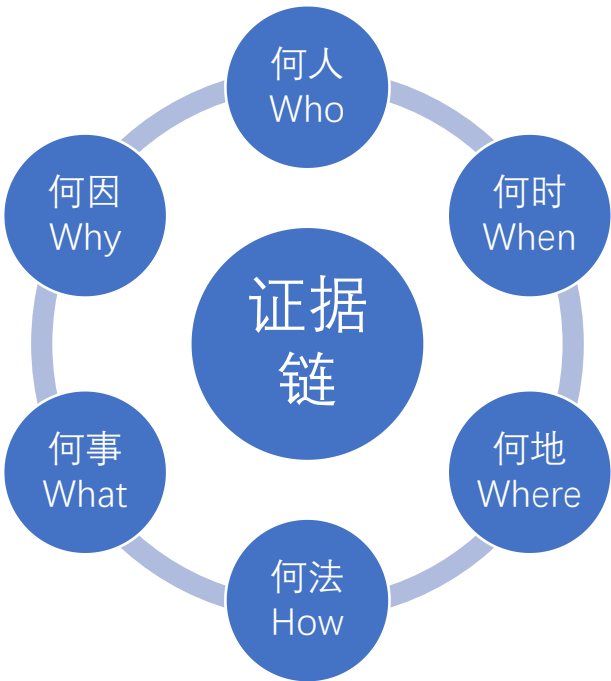
保护经纪人权益+减少经纪人骚扰+减少垃圾商机量





背景介绍

矛盾：日益增长的工单量 VS 落后的溯源效率



痛点

- 业务复杂度高
- 预案依赖经验
- 举报工单量大
- 监察人员变动
- 监察人员有限





事后-查案溯源-架构迭代

msup[®]



V1

- 快速上线



V2

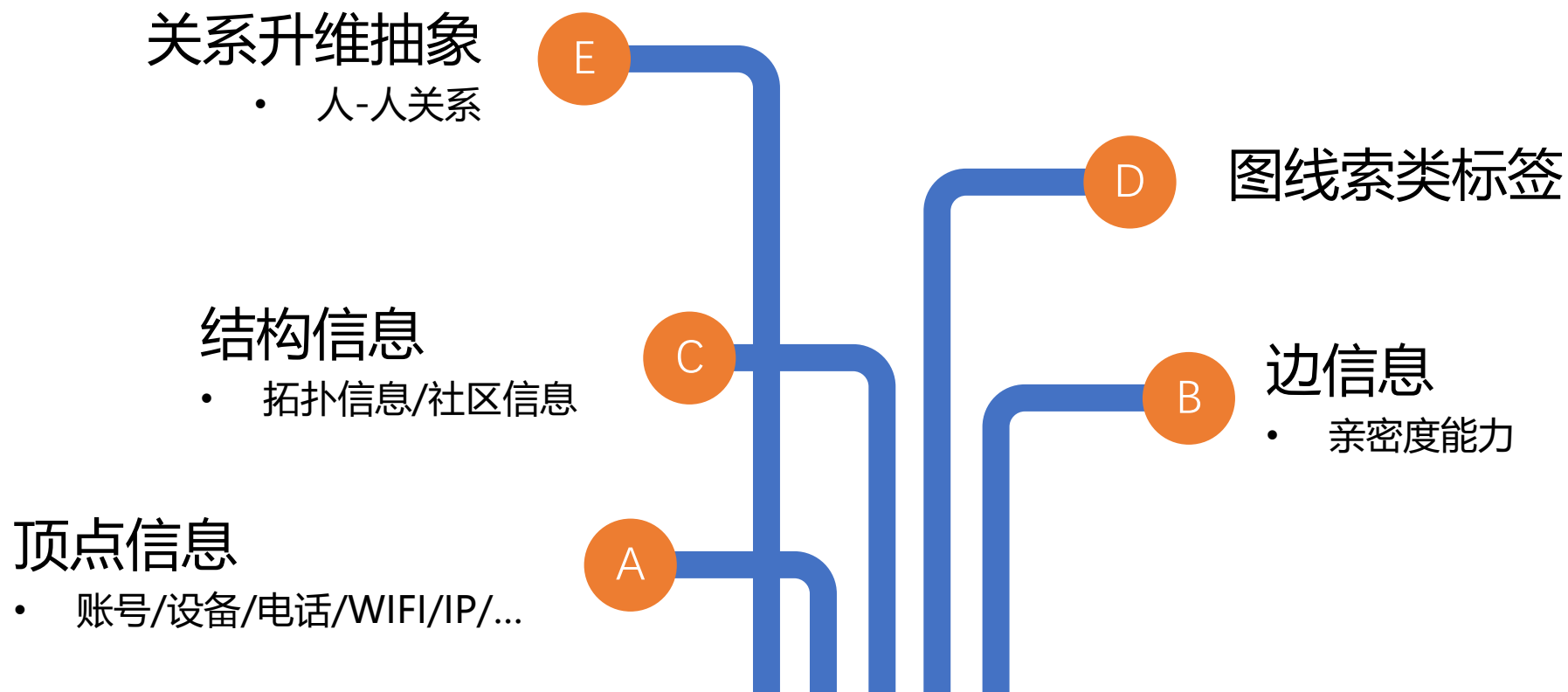
- 自研KV bulk load
- 百亿级属性，6小时内产出
- 图数据入库质量监控



V3

- 图数据bulk load
- 图数据实时化 (Flink)
- 关系发现
- 图算法的在线分析





预案盘点由零散、主观、线下 -> 体系、客观、线上+线下 标准化





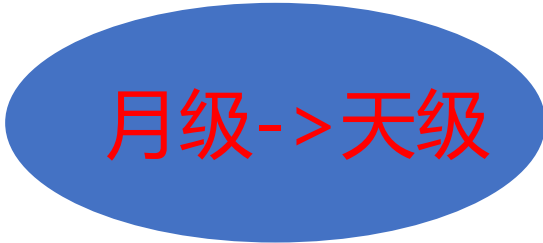
落地效果



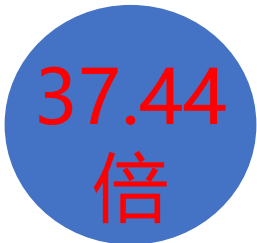
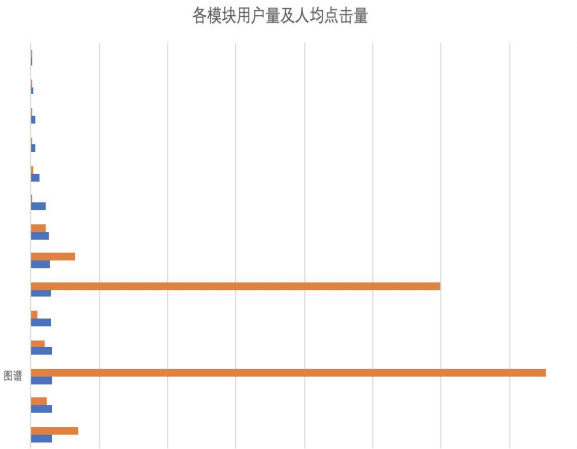
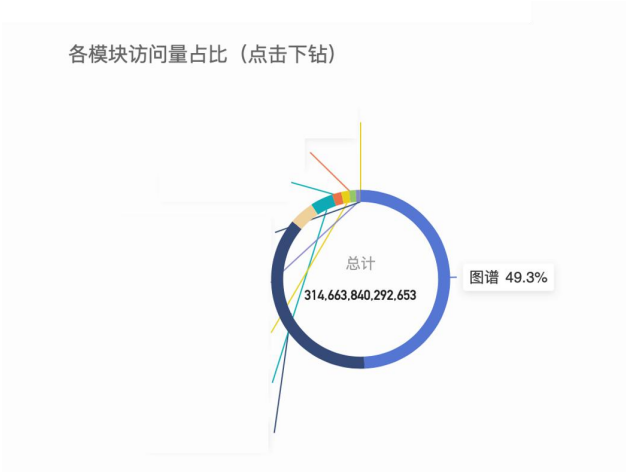
属性入库效率



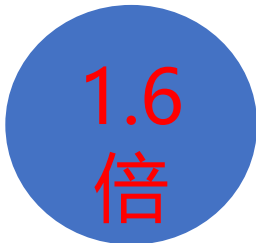
图数据入库速度



业务迭代效率



人均每日图谱
查询量



预案效率
提升





目录

CONTENTS

1

贝壳的业务介绍

2

风险挑战及应对

3

图谱的落地应用

4

未来的展望





业务应用拓展

➤ 风险治理

- 溯源智能化
- 违规范式自学习
- 风险主动感知与治理

➤ 用户增长

- 经纪人信用
- 经纪人toB/toC推荐

➤ 品质度量

- ACN健康度
- 经纪人协作
- 影响力分析

- 图的实时化
- 知识融合
- 表征学习
- 高密子图挖掘
- 图的异常检测

基础能力深耕





关注msup公众号
获取更多AI落地实践

麦思博(msup)有限公司是一家面向技术型企业的培训咨询机构，携手2000余位中外客座导师，服务于技术团队的能力提升、软件工程效能和产品创新迭代，超过3000余家企业续约学习，是科技领域占有率第1的客座导师品牌，msup以整合全球领先经验实践为己任，为中国产业快速发展提供智库。