

BIG
DATA

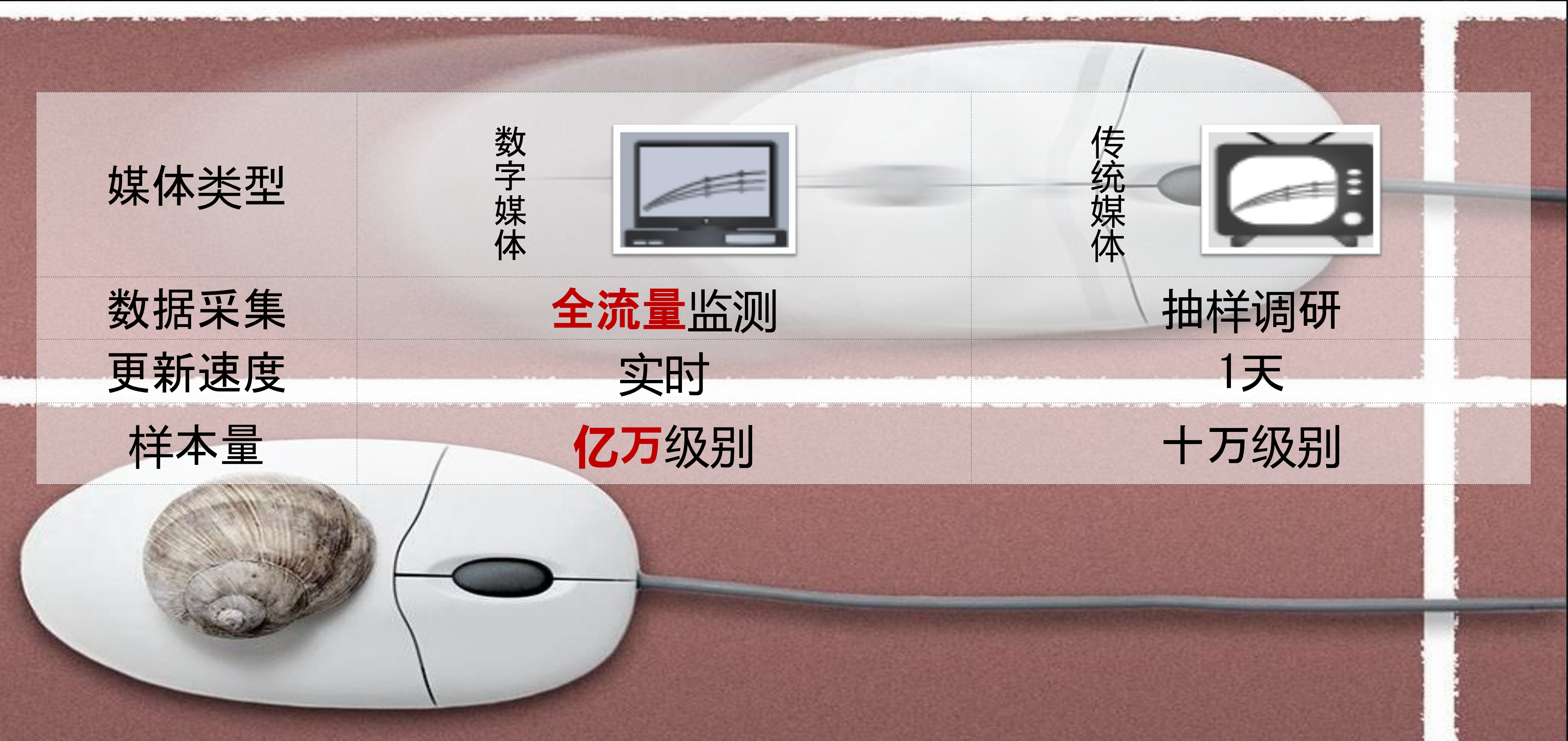
ARTIFICIAL INTELLIGENCE

数据科学在企业服务中的应用

吴明辉

北京秒针信息咨询有限公司创始人、董事长兼首席执行官
北京明略软件系统有限公司董事长

为什么说大数据时代到来了：归根结底是IT不断进步导致了数据的成本在不断降低



BIG
DATA

大数据分析的特点：从抽样到普查

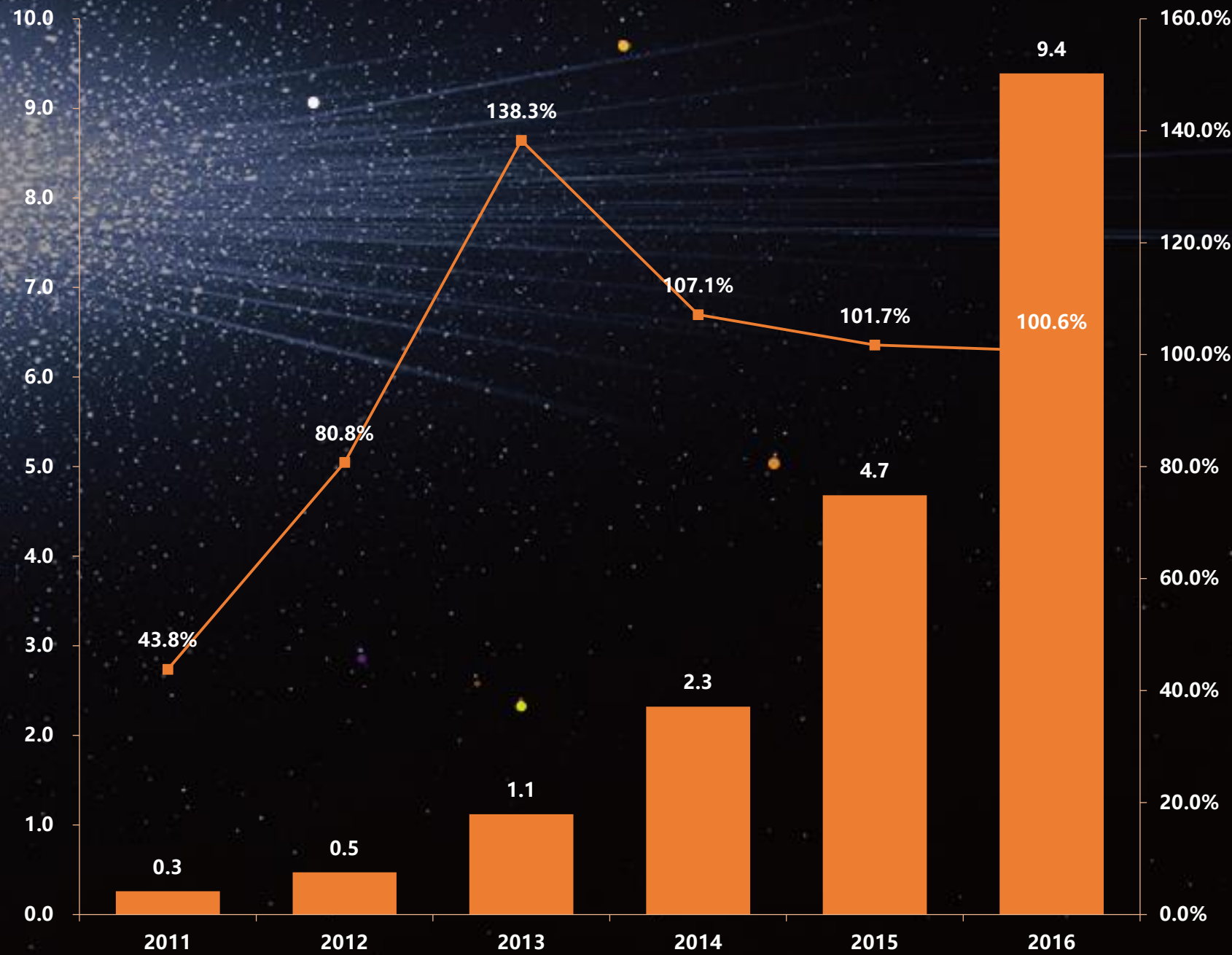
抽 样



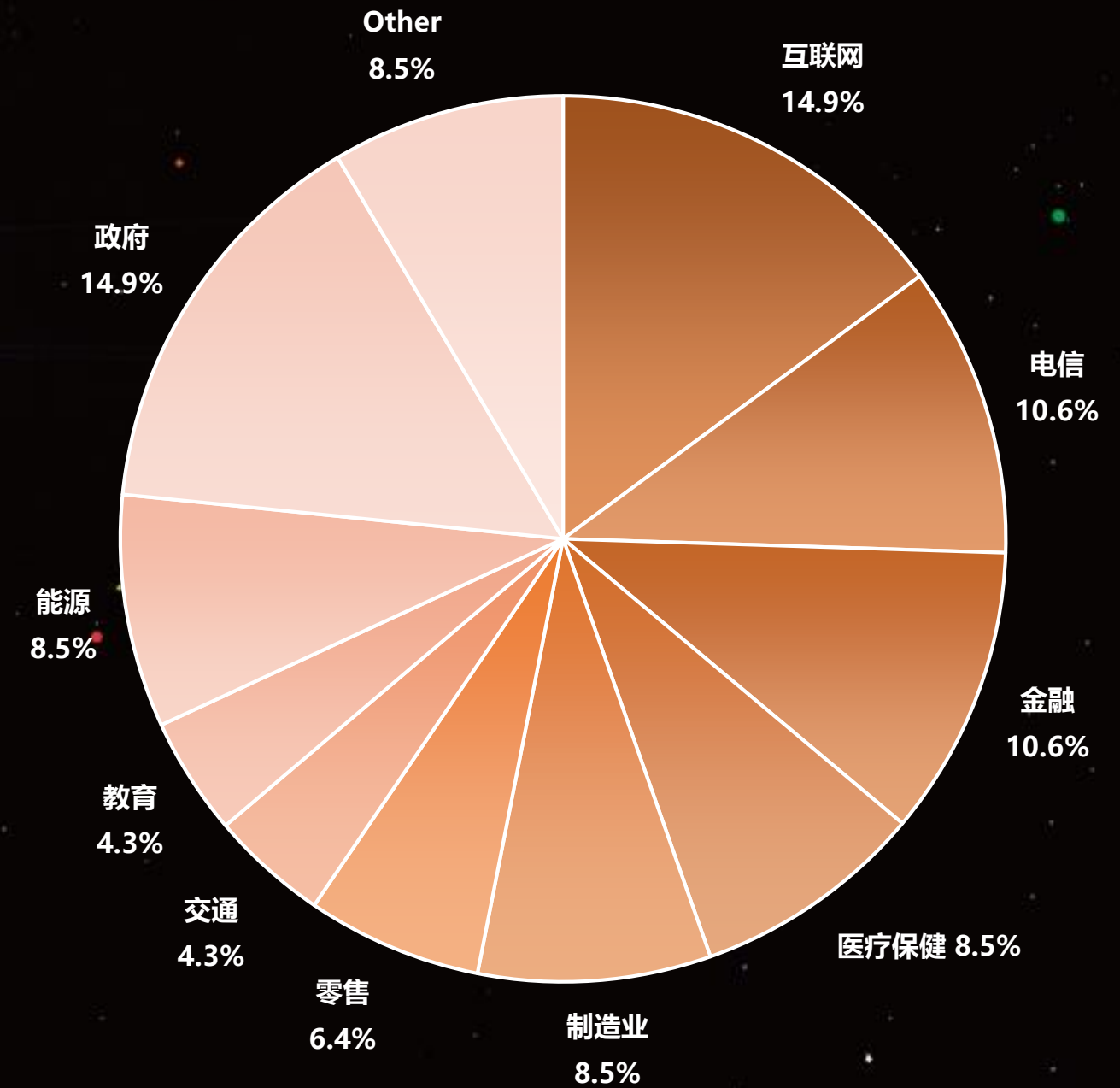
普 查



市场规模持续增长



市场份额





中国大数据产业环境-政策

“互联网+” 行动计划 ○

○ 促进大数据发展行动纲要

大数据产业 “十三五” 发展规划 ○

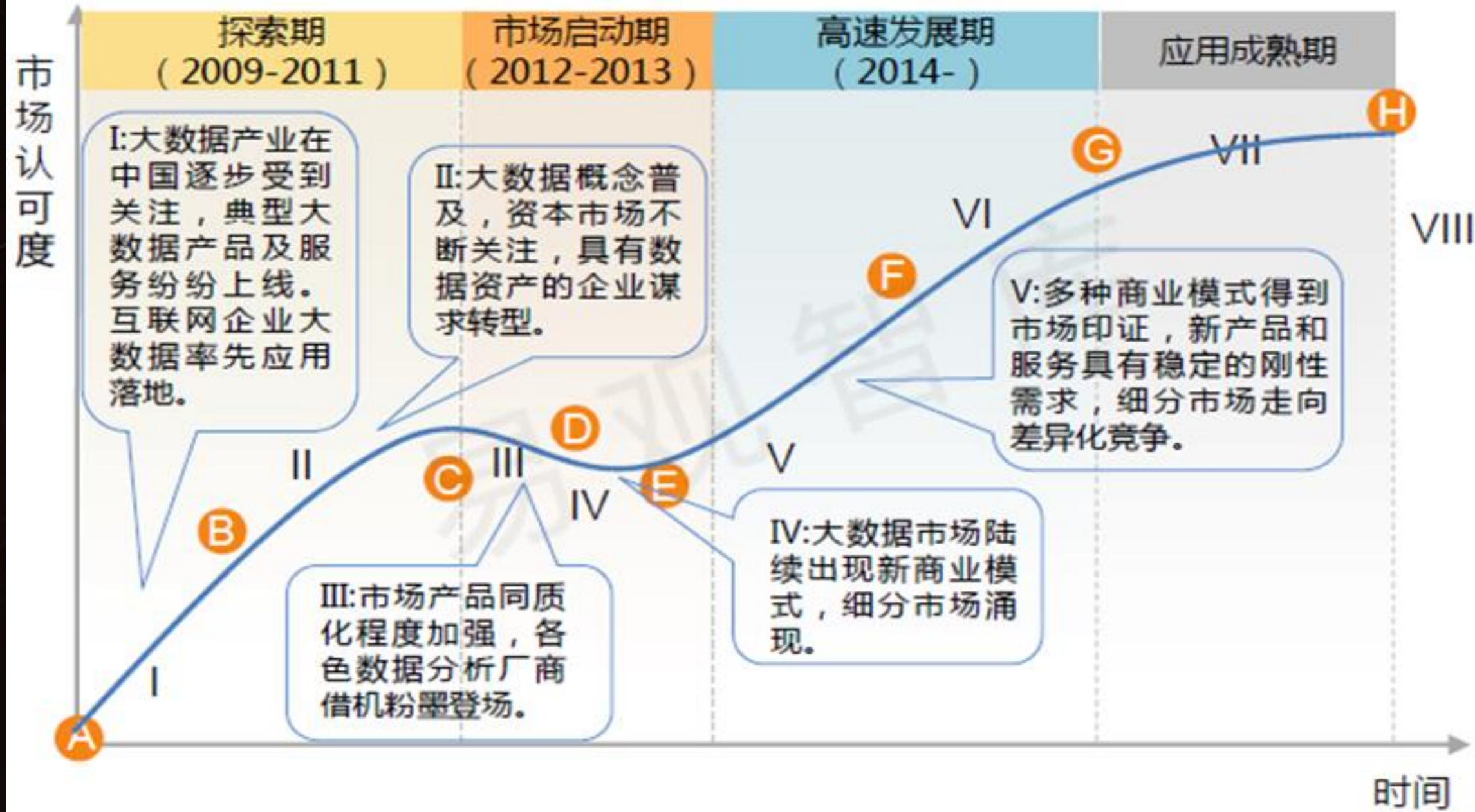
○ 中国制造2025

**BIG
DATA**

中国大数据发展虽然较西方国家晚，但是政府的推动力度是前所未有的

中国大数据市场发展历程及主要模式

中国大数据市场AMC模型



数据存储租用

通过易于使用的API，用户方便地将各种数据对象放在云端，然后再像使用水电一般按用量收费。

租售信息业务

涉及到大数据产业链的数据整理与分析环节。提供加工后的数据“半成品”。

数据增值服务

精准营销等提升企业价值链业务。
小额信贷等数据衍生新兴业务。

数据技术服务

为运营某一环节或业务问题提供方案。
针对企业系统需求，提供整体解决方案。
大数据即服务。

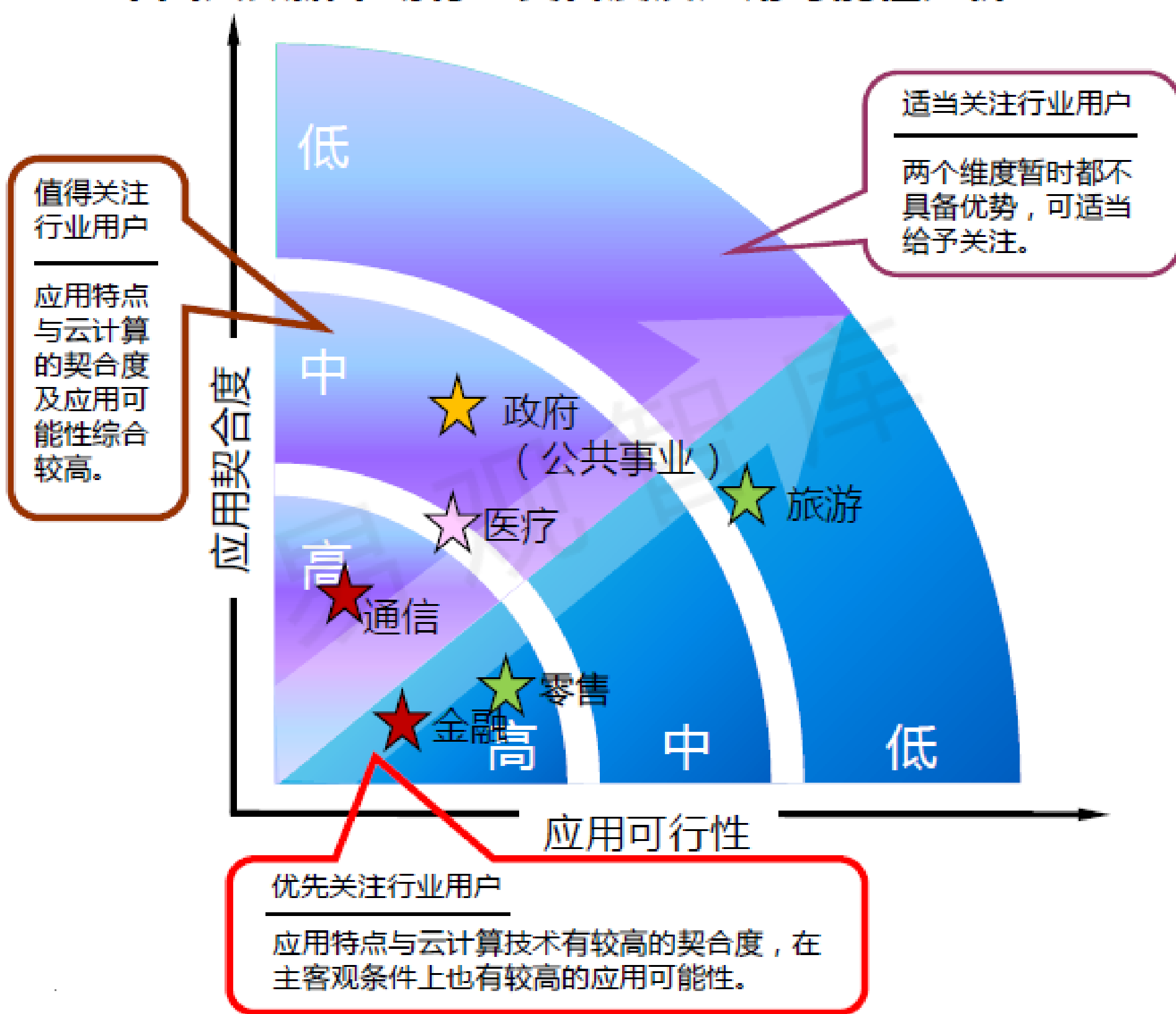
大数据时代，对数据的掌握成为国家影响力和主导权的重要体现。

大数据的挖掘、利用与保护成为各国新一轮信息技术革新以及经济发展驱动的要件。

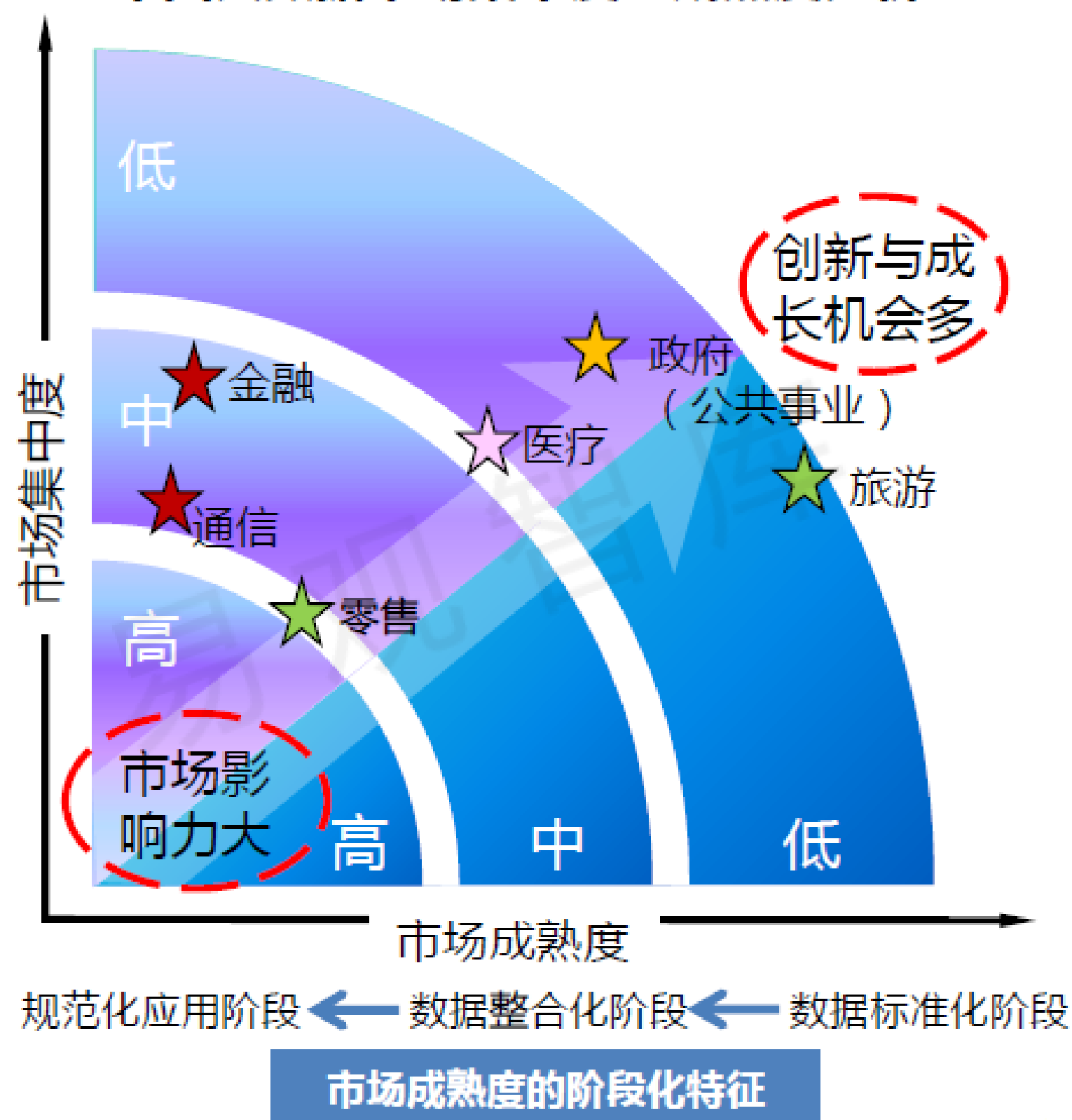
对数据资产的掌控已成为继制陆权、制海权、制空权之后的新制权——数据主权。

重点行业大数据应用与成熟度

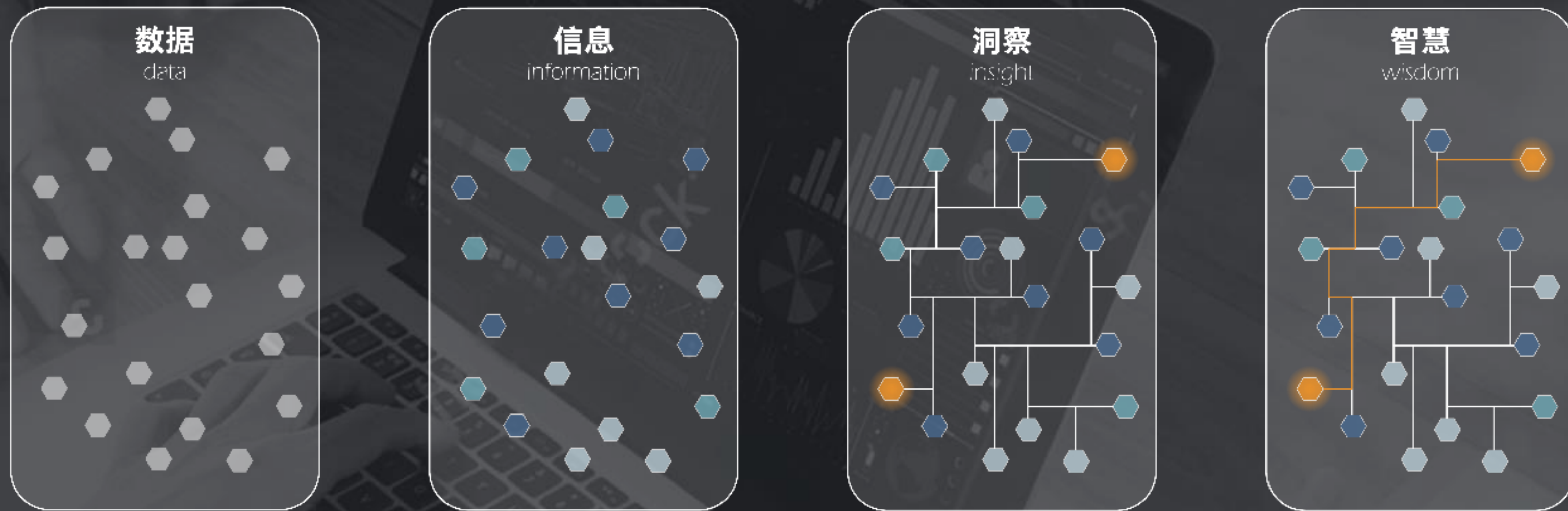
中国大数据市场行业契合度及应用可能性分析



中国大数据市场集中度与成熟度分析



企业应用数据的本质就是从数据到智慧的转换过程中让企业的员工更好地完成工作



Data Collection and Storage → Statistics and Analysis → data connection and mining → AI

Miaozhen[®] Systems



- 成立于2006年
- 北京、上海、广州、烟台、新加坡、美国6地办公室
- 拥有近400名全职员工
- 日均1000亿广告请求处理能力领先行业
- 超过70%的国际品牌广告主选择秒针
- 国家高新技术企业和双软认证企业
- 在中国内地、台湾，香港，新加坡，马来西亚，日本，印度，韩国，澳大利亚、美国等市场提供产品和服务
- 已完成四轮融资

- 成立于2014年
- 北京、上海、成都、石家庄等多地办公室
- 拥有超过300名全职员工
- 秒级处理TB级别数据，小时内部署百节点规模的大数据集群
- 国家高新技术企业和双软认证企业
- 客户横跨政府、公安、税务、金融、制造以及营销多行业
- 中关村大数据产业基地”挂牌企业
- 红杉资本中国基金在中国公共安全领域投资的第一家大数据企业

围绕知识的数据科学

行业数据科学的技术点

知识抽取

知识/本体建模

- 知识表示与本体建模
- 知识表示学习
- 知识图谱嵌入
- 知识图谱的分布式表示
- 本体重用与演化
- 本体映射、融合与对齐
- 本体评估
- 概念学习

知识图谱构建/信息抽取

- 开放知识抽取
- 众包知识工程与协同知识获取
- 人机协同知识库构建
- 维基数据的自动抽取
- 自动化知识库构建工具、语言与系统
- 基于监督学习/非监督学习的知识抽取
- 轻量监督学习与文本抽取
- 远程监督学习与文本抽取
- 机器阅读
-

知识索引

分布式知识库系统

- 知识查询与搜索
- 概率知识库存储
- 弹性知识存储与分布式计算
- 图数据库
-

知识构建

实体识别、消解、和实体链接

- 术语映射与集成
- 结构集成
- 异构知识库集成
- 跨语言知识链接与集成
- 基于本体的数据集成
- 链接数据构建与发布
- 链接数据集成链接数据挖掘与分析
-

知识推理与表达

基于逻辑的推理

- 基于概率统计的推理
- 基于自然逻辑的推理
- 基于嵌入与分布式表示的推理
- 规则学习
- 知识库补全
-

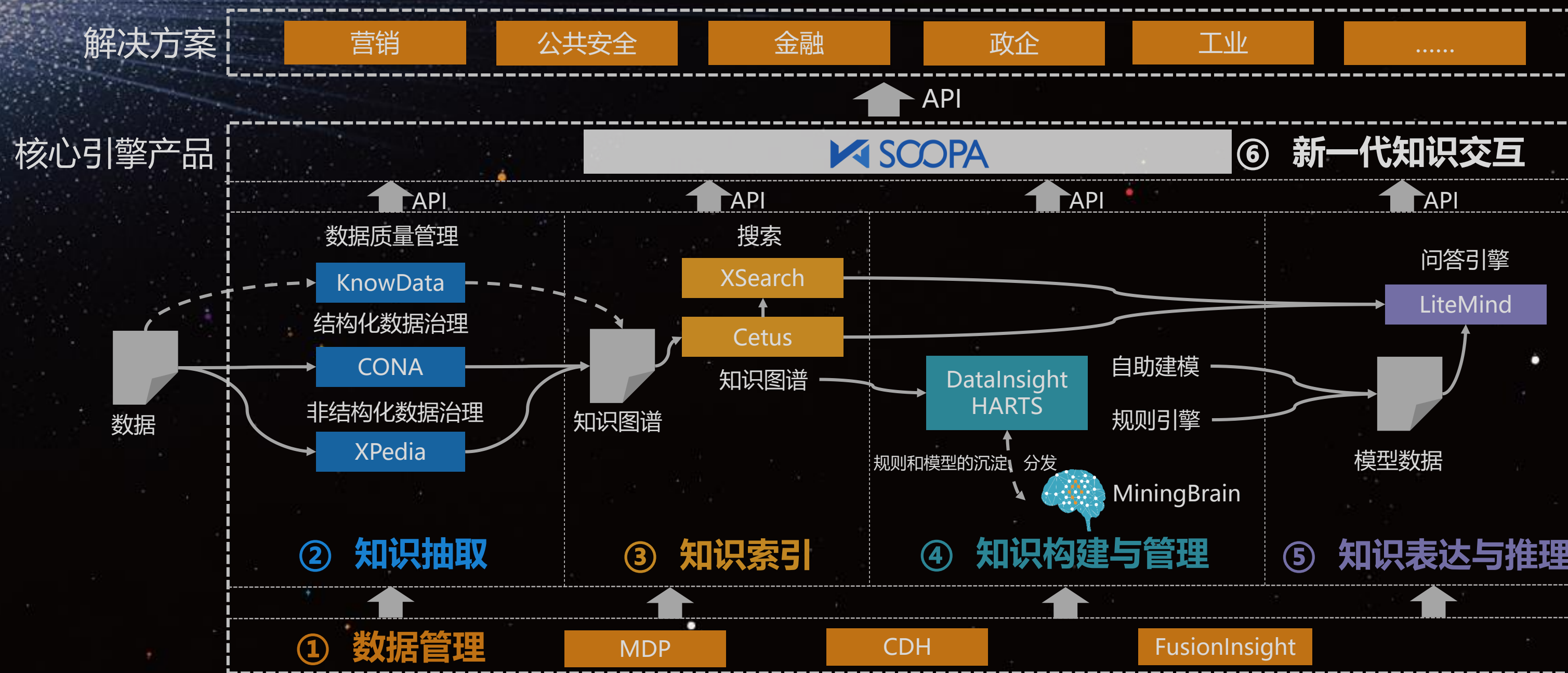
知识交互

知识图谱可视化

- 语义搜索引擎
- 基于知识的问答系统
- 智能个人助理系统
- 基于知识的语义分析
-

围绕知识的数据科学

通过不断积累形成人工智能完整技术链



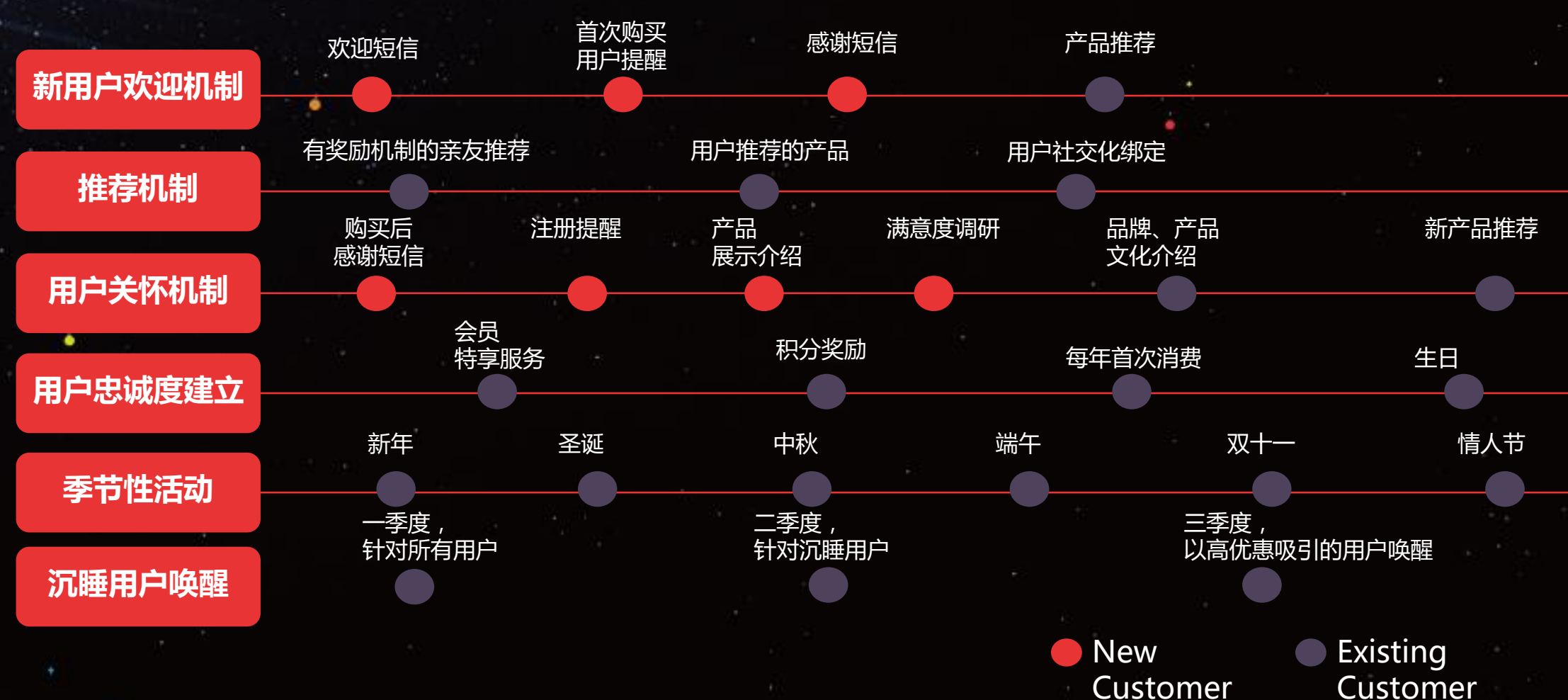
BIG
DATA

大数据的应用



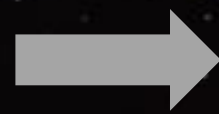
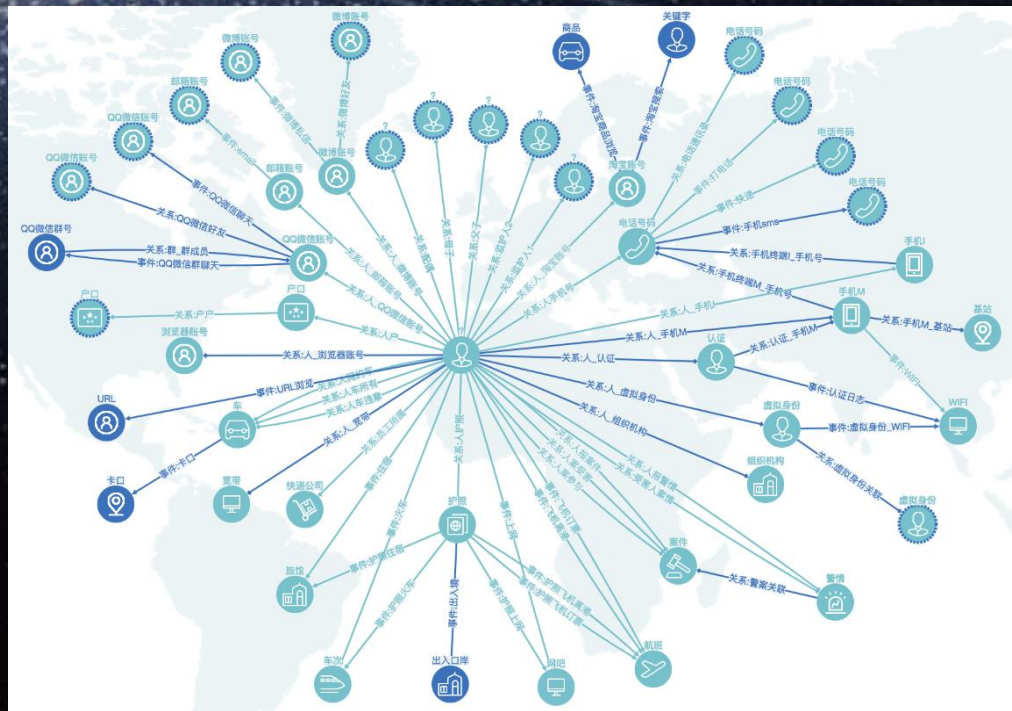
数据科学 in 营销

数据时代的营销自动化



数据科学 in 公共安全

基于公安知识图谱构建行业应用



数据科学 in 公共安全

团伙挖掘——GraphX结合规则实现重点人的社群发现



- 以重点人员为基础，扩展关系密切的隐性重点人，标注可疑团伙
- 根据侵财、涉毒等不同案件特征和特定行为数据，预测出疑似犯罪团伙
- 通过团伙活跃度、危险度、聚集度对团伙进行研判监控

数据科学 in 公共安全

重点人积分预警——机器学习挖掘特征结合规则计算积分

The screenshot displays a dashboard for '重点人积分预警' (Key Person Score Warning). It includes a navigation bar with '首页', '预警', '应用', '检索', '研判', '档案', and '协作'. The main content area is divided into two sections:

- 重点人积分列表 (Key Person Score List):** A table listing individuals with columns for '序号' (Serial Number), '身份证号' (ID Number), '姓名' (Name), '类型' (Type), '信息维度分数' (Information Dimension Score), '预警分数' (Warning Score), '置信度' (Confidence), and '操作' (Action). The table shows 23 entries, with the 11th entry (王二小) highlighted.
- 重点人积分详情 (Key Person Score Details):** A detailed view for '李雷' (Li Lei) with ID '230829198004302456'. It includes a network diagram showing relationships with other individuals and a table of '命中规则名称' (Hit Rule Names) and '积分' (Scores).

序号	身份证号	姓名	类型	信息维度分数	预警分数	置信度	操作
01	215928198604232897	张三三	涉毒	78	78	98%	
02	215928198604232897	里斯本	涉毒	98	98	40%	
03	215928198604232897	王二小	涉毒	70	70	70%	
04	215928198604232897	李雷雷	涉毒	67	67	72%	
05	215928198604232897	张三三	涉恐	86	86	36%	
06	215928198604232897	里斯本	涉恐	50	50	90%	
07	215928198604232897	王二小	涉恐	68	68	82%	
08	215928198604232897	李雷雷	涉毒	83	83	10%	
09	215928198604232897	张三三	涉毒	58	58	99%	
10	215928198604232897	里斯本	涉毒	66	66	30%	
11	215928198604232897	王二小	涉毒	40	40	60%	
12	215928198604232897	李雷雷	涉恐	55	55	66%	
13	215928198604232897	张三三	涉恐	74	74	12%	
14	215928198604232897	里斯本	涉恐	60	60	30%	
15	215928198604232897	王二小	涉毒	59	59	30%	
16	215928198604232897	王二小	涉恐	83	67	98%	
17	215928198604232897	李雷雷	涉毒	58	86	40%	
18	215928198604232897	张三三	涉毒	66	50	70%	
19	215928198604232897	里斯本	涉毒	40	68	72%	
20	215928198604232897	王二小	涉毒	55	83	36%	
21	215928198604232897	李雷雷	涉恐	74	58	90%	
22	215928198604232897	张三三	涉恐	60	66	82%	
23	215928198604232897	里斯本	涉毒	59	61	73%	

时间	类型	描述	命中规则名称	积分
2016-09-09 12:32:21	住宿	承德新环大酒店5001	住宿时间为7:00-23:00	+10
2016-09-09 12:32:21	火车	大同到北京D2118次	同日同车次同车厢有涉毒人员	+20
2016-09-09 12:32:21	订票	北京到大同D3548次	同日同车次同车厢有涉毒人员	+100
2016-09-09 12:32:21	住宿	承德新环大酒店3649	本市人在本市入住	+40
2016-09-09 12:32:21	住宿	承德新环大酒店8010	住宿时间为7:00-23:00	+10
2016-09-09 12:32:21	订票	北京到大同D3548次	同日同车次同车厢有涉毒人员	+5
2016-09-09 12:32:21	订票	北京到大同D3548次	同日同车次同车厢有涉毒人员	+10
2016-09-09 12:32:21	住宿	承德新环大酒店1320	住宿时间为7:00-23:00	+15
2016-09-09 12:32:21	住宿	承德新环大酒店8248	近180天第一次住宿	+100

- 结合重点人员特征标签、个人关系网络、活动轨迹等信息，构建重点人员积分模型，对本地重点人进行积分管理
- 当重点人员的积分达到一定预警阈值时，根据预警级别分为红、橙、蓝三色级别预警

数据科学 in 公共安全

案件和情报串并——以语义为基础实现人案关联

串并案分析-以全文检索、语义分析为基础的线索挖掘

串并案分析结果

案件编号	案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件
A9493302015396735256698	系统抽取的由自然语言组成的一些又一些的案件描述	共 20 个案件

编号	特征标签	涉案人员	案件概述	操作
0364	涉毒, 团伙作案	王天华, 刘金龙	在2015年底至2016年7月期间, 马某某多次在万寿等地向他人贩卖冰毒, 牟取暴利。经查, 以王某某为主的贩毒团伙为广东吴某某浙江高安...	打开档案 深度挖掘
164862	走私, 涉毒	刘大华, 刘金龙	2013年南京曾发生吸毒母亲哺乳女婴案, 我们今天要说的这件事, 同样发生在南京, 同样是涉毒家庭。	打开档案 深度挖掘
22648	涉毒, 盗窃, 抢劫	王天华, 刘大华, 刘金龙	2014年12月25日, 尹某某因涉毒被判处有期徒刑。在北京警方抓获, 经检验尿冰毒等, 初步确定该团伙吸毒重点涉案人员13人, 专案组也掌握...	打开档案 深度挖掘
00318	涉毒, 盗窃抢劫, 走私	王天华	北京警方表示当时搜出110余克大麻, 以平均一次大麻手卷烟使用0.4克, 房租名收藏的大麻大约能够卷250支大麻烟, 数量相当惊人...	打开档案 深度挖掘
0364	涉毒, 团伙作案	王天华, 刘金龙	在2015年底至2016年7月期间, 马某某多次在万寿等地向他人贩卖冰毒, 牟取暴利。经查, 以王某某为主的贩毒团伙为广东吴某某浙江高安...	打开档案 深度挖掘
164862	走私, 涉毒	刘大华, 刘金龙	2013年南京曾发生吸毒母亲哺乳女婴案, 我们今天要说的这件事, 同样发生在南京, 同样是涉毒家庭。	打开档案 深度挖掘
22648	涉毒, 盗窃, 抢劫	王天华, 刘大华, 刘金龙	2014年12月25日, 尹某某因涉毒被判处有期徒刑。在北京警方抓获, 经检验尿冰毒等, 初步确定该团伙吸毒重点涉案人员13人, 专案组也掌握...	打开档案 深度挖掘
00318	涉毒, 盗窃抢劫, 走私	王天华	北京警方表示当时搜出110余克大麻, 以平均一次大麻手卷烟使用0.4克, 房租名收藏的大麻大约能够卷250支大麻烟, 数量相当惊人...	打开档案 深度挖掘

通过分析已侦破的案件提取领域知识, 包括案件类型, 案发时间, 案发地点, 作案手段, 罪犯特征, 受害人特征, 作案工具等, 通过统计, 得到一系列高频案件及相关特征。

通过自然语言分析, 提取案件或者警情信息中的实体(人名、地名、组织机构名、车牌号、手机号、身份证号等)、关系提取, 并结合结构化数据构建知识图谱, 实现智能案件串并。

情报串并-情报交叉分析与骨干人员挖掘

石家庄新华路光明街附近有人闹事

官方通报称, 4月2日晚时, 晋宁街村少数群众因分子煽动部分人员, 从高铁站后山前街胡同进入高铁站站台, 造成高铁半小时停运。晋宁街官方采取强制措施, 高铁已恢复正常。

通报, 晋宁街官方此次首先通报, 对闹事事件进行了通报, 晋宁街街道办事处, 4月2日, 通报, 晋宁街街道办事处全力配合晋宁街村村委会对少数群众进行打击, 在当天上午10时, 少数煽动的分子在40多名群众在晋宁街高铁站前, 从高铁站后山前街胡同进入高铁站站台, 造成列车停运半小时, 经全力处置, 晋宁街在高铁站站台的人员于晚8时40分全部撤离现场, 晋宁街又派了4名工作人员, 对高铁站进行技术检测确保安全后, 晚8时57分列车恢复正常。

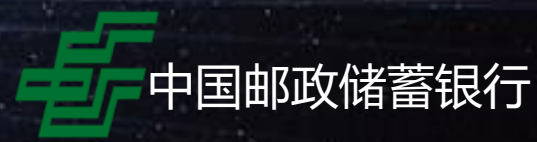
序号	事件名称	关联类别	事件规模	发生场所	处理结果	事发时间
1	事件名称事件名称事件名称事件名称	其他涉毒人员	事件规模事件规模事件规模	新市区	处理结果处理结果处理结果	2016-06-20 21:32:33
2	事件名称事件名称事件名称事件名称	其他涉毒人员	事件规模事件规模事件规模	新市区	处理结果处理结果处理结果	2016-06-20 21:32:33
3	事件名称事件名称事件名称事件名称	其他涉毒人员	事件规模事件规模事件规模	新市区	处理结果处理结果处理结果	2016-06-20 21:32:33
4	事件名称事件名称事件名称事件名称	其他涉毒人员	事件规模事件规模事件规模	新市区	处理结果处理结果处理结果	2016-06-20 21:32:33
5	事件名称事件名称事件名称事件名称	其他涉毒人员	事件规模事件规模事件规模	新市区	处理结果处理结果处理结果	2016-06-20 21:32:33

基于文本分析, 发现群体标签、重大事件情报中的关联:

- 案-案关联
- 人-案关联
- 事-事关联
- 人-事关联

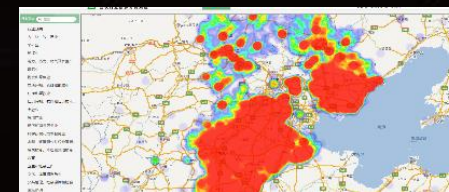
数据科学 in 金融

基于金融知识图谱构建行业风控和客户营销等应用



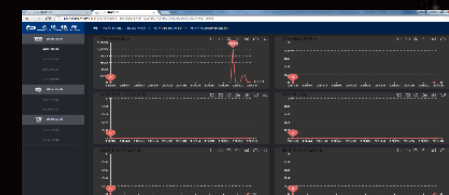
大数据风控咨询与落地完整方案

服务· 数十万信贷户 & 百亿信贷规模
架构· 20+节点 & 30+TB
业务· 不良信贷预警排查 & 行业关联挖掘



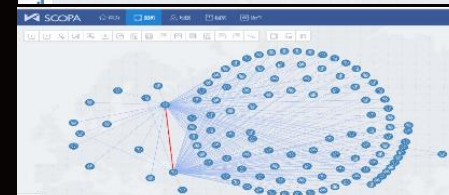
大数据实时交易风险监控方案

服务· 5000+万持卡人 & 千万日志/天
架构· 30-节点 & 60-TB
业务· 日志流式采集/计算 & 风控实时决策



大数图谱关联挖掘风险识别方案

服务· 数百万信贷客户
架构· 10-节点 & 10-TB
业务· 资金账户路径分析 & 循环担保圈识别



数据科学 in 金融

基于金融知识图谱构建行业风险管控和客户营销等应用



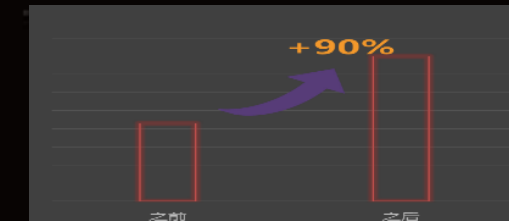
大数据平台与建模完整落地方案

服务· 600+万商户 & 6+亿流水/月
架构· 50-节点 & 200-TB
业务· 数据魔方 & 用户模型 & 中证指数



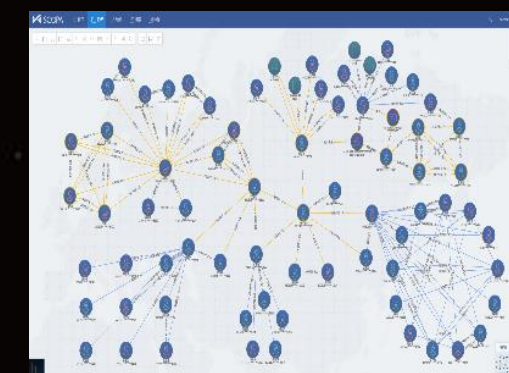
大数据客户画像与精准营销方案

服务· 500+万股民 & 千万App日志/月
架构· 20-节点 & 60-TB
业务· 标签画像 & 推荐模型 & 休眠激活



大数据实时营销与关联挖掘方案

服务· 4000+万客户 & 近亿份合同
架构· 20-节点 & 数亿实体
业务· 营销推荐 & 家庭账户体系关联挖掘

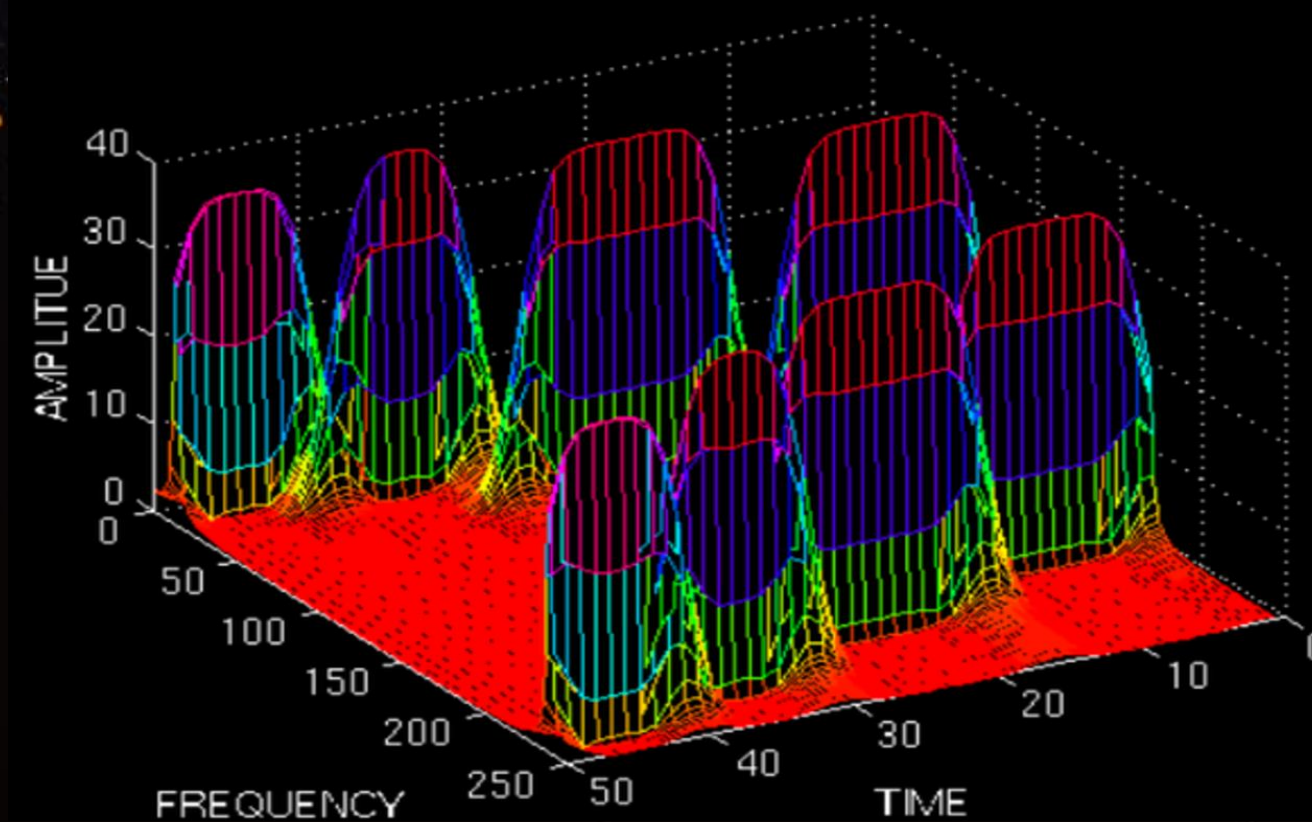


数据科学 in 工业

工业设备故障检测

- 13000+故障样本，30+种类故障，每类故障5-6个原因
- 3000+标注样本
- 10+类主要的故障占比80%+（过流、过压、接地、缺相）

时频分析



- 随机森林分类算法：逆边过流153个样本作训练，103个作预测
- 过流类故障分类准确率94%（6个故障样本错分）

数据科学 in 工业

工业设备状态修

客户核心KPI

- 安全指标（来自政府），例如：5分钟晚点
- 成本，小地铁公司相对在意

现有解决方案（计划修）

- 日检
- 月检
- 大修（5年）/架修（10年）

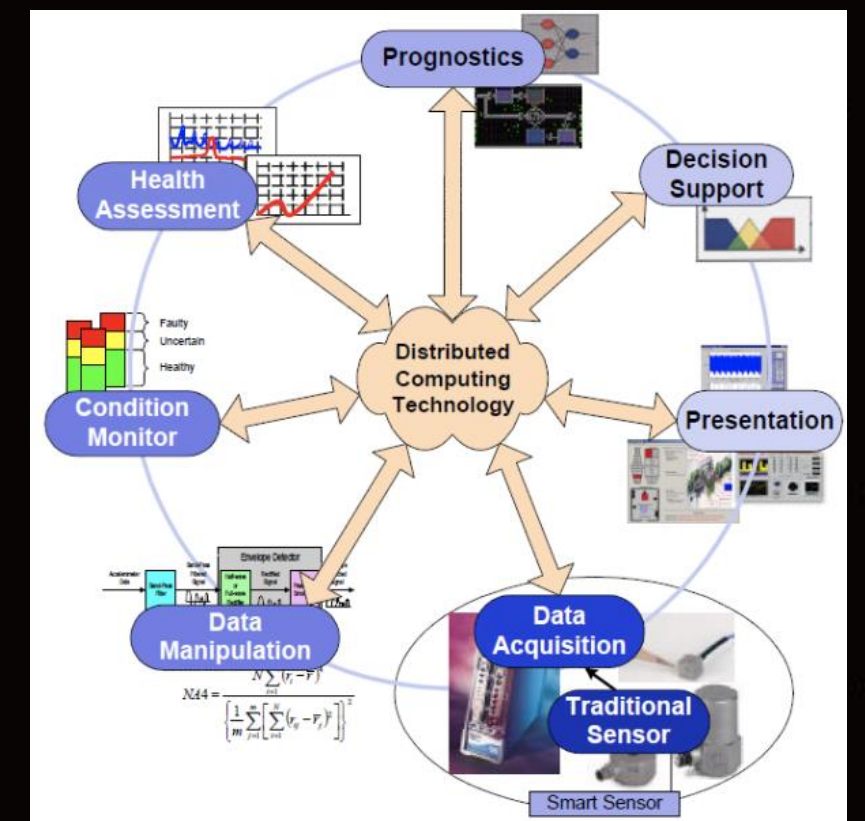
问题

- 人工效率低下
- 过修严重，造成浪费
- 指标仍然无法完成



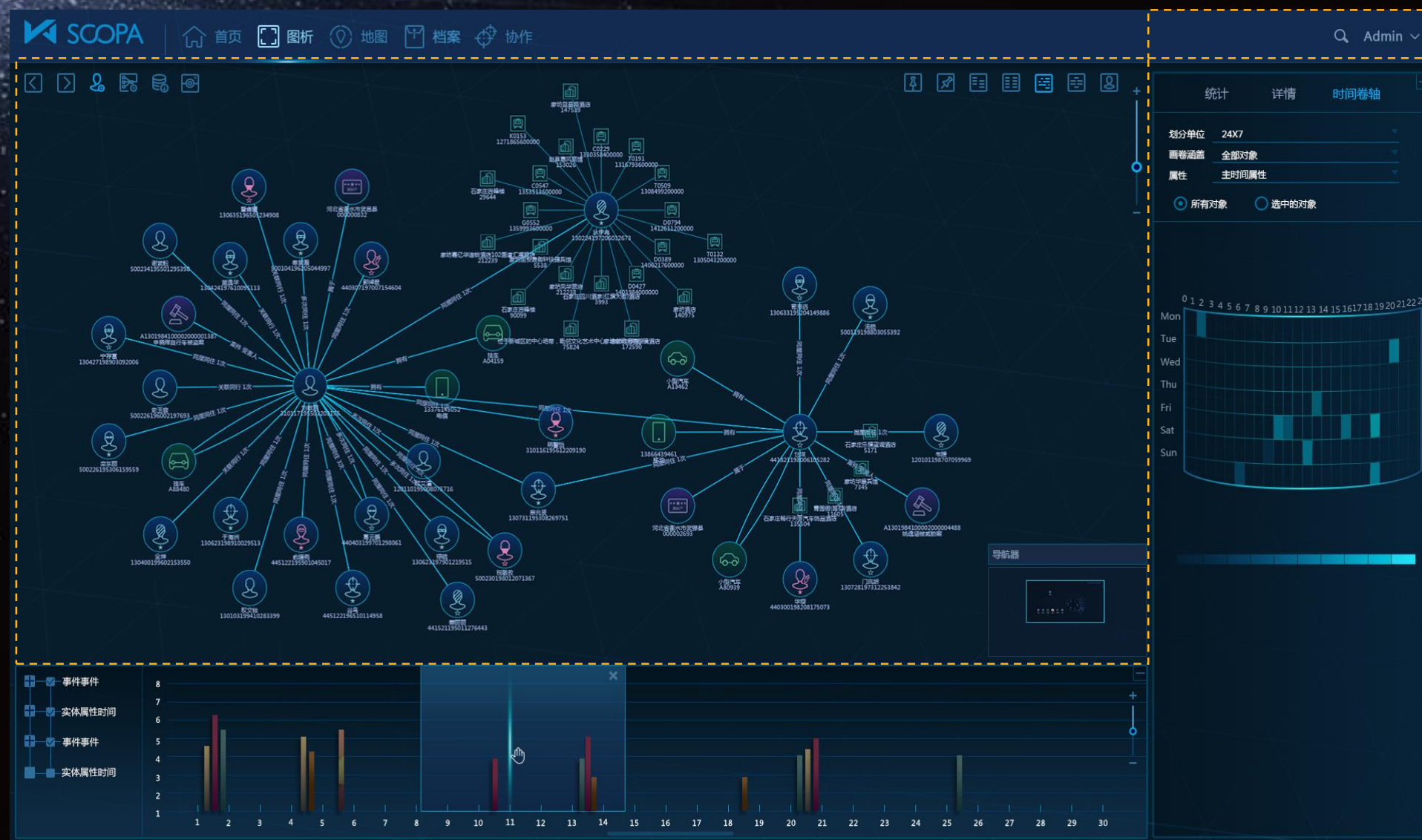
状态修的一般流程

- 数据获取
- 数据处理，原始数据解析
- 状态监控，基于阈值的报警
- 健康评估，基于趋势分析的诊断
- 故障预测，基于机器学习
- 决策辅助，提出建议
- 人机交互



数据科学 in 可视化

SCOPA使用新一代人工智能人机交互技术，结合人类大脑的发散思维与定向思维，支持可视化数据特征分析和交互能力



- 支持对数据的**发散探查(图析)**和**定向探查(搜索)**
- 基于大数据图结构的海量数据存储，支持对隐性关系的实时查询与碰撞
- 基于**实体-关系-时间-空间**的四维分析
- 基于**动态规则**的事件实时碰撞
- 海量数据的全文检索
- 革新性交互可视化
- 支持多级负载均衡和可伸缩的部署方案

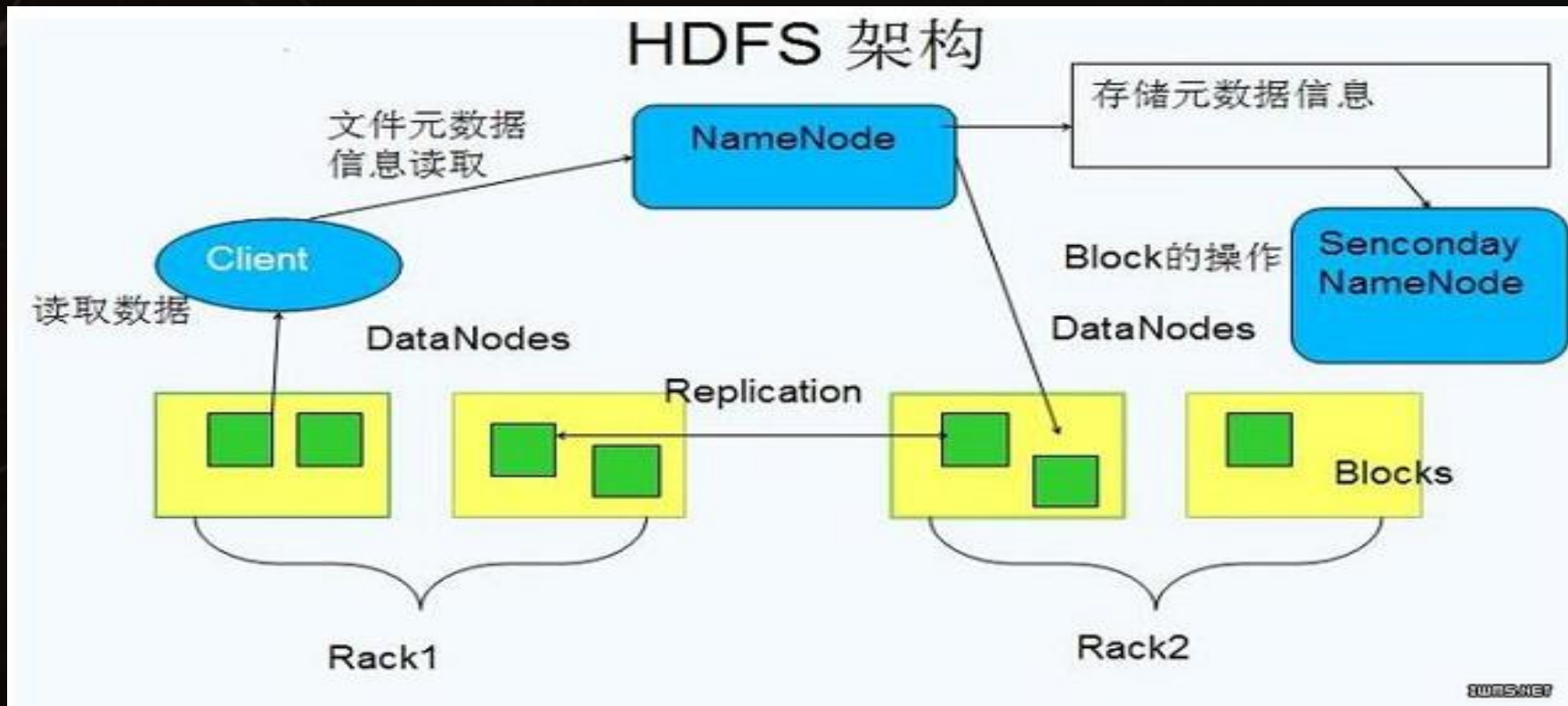
BIG
DATA

大数据的核心技术和门槛

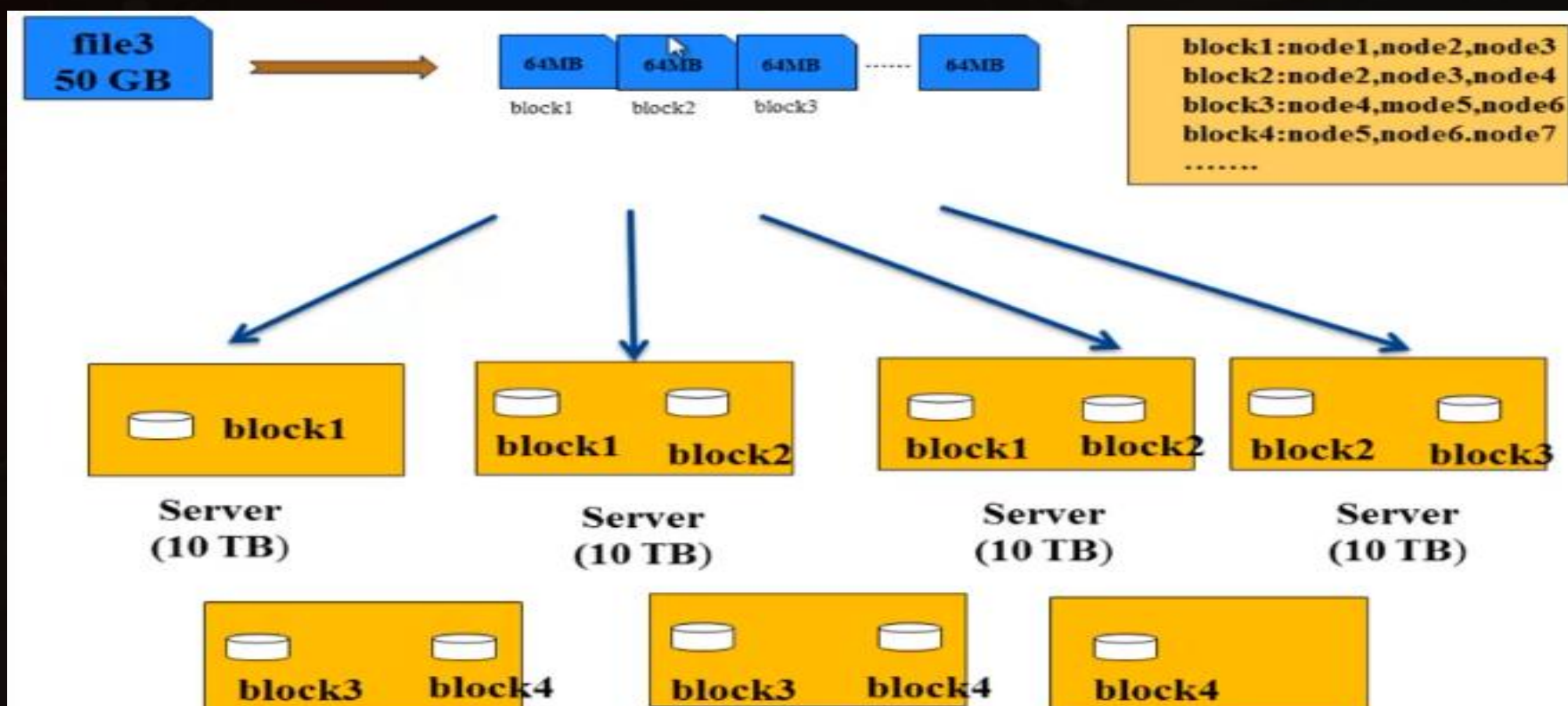
存储技术

分布式文件系统 HDFS

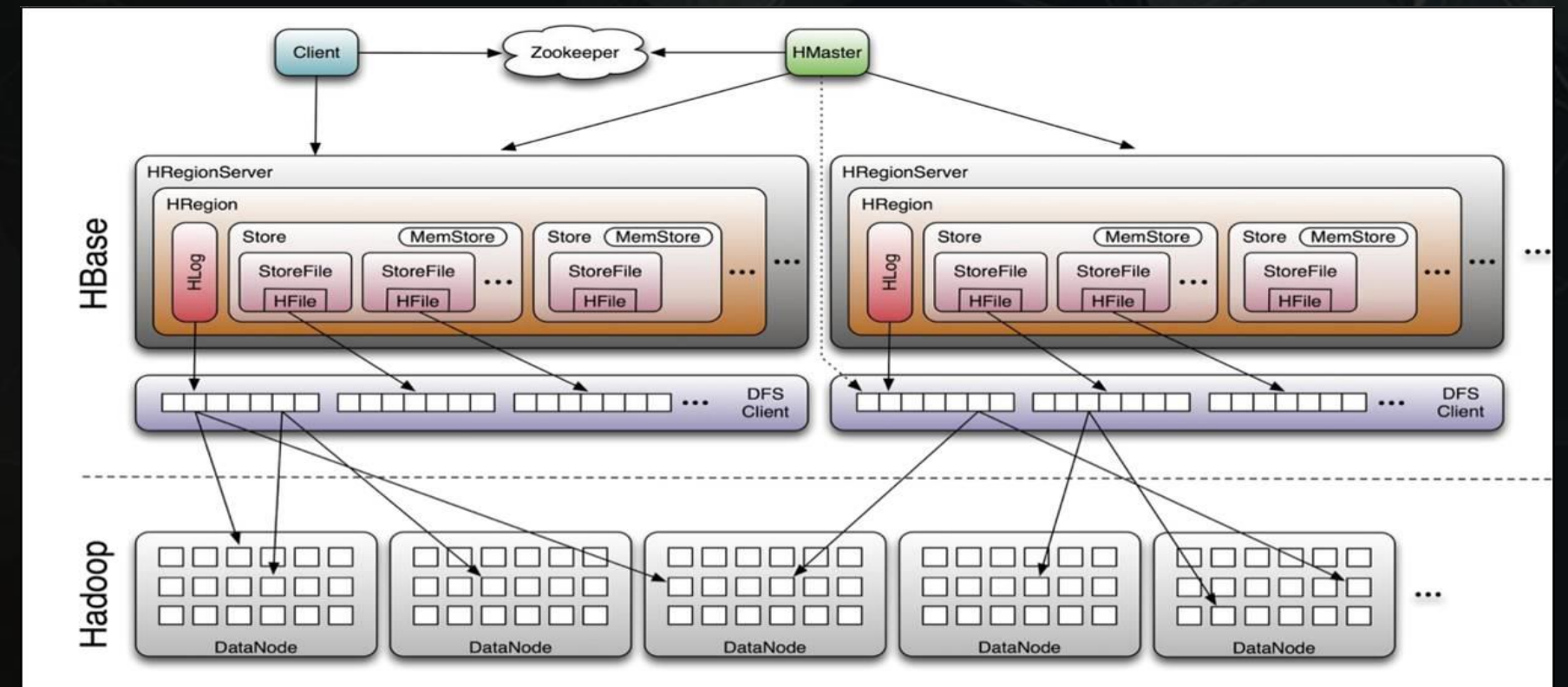
分布式文件系统 HDFS



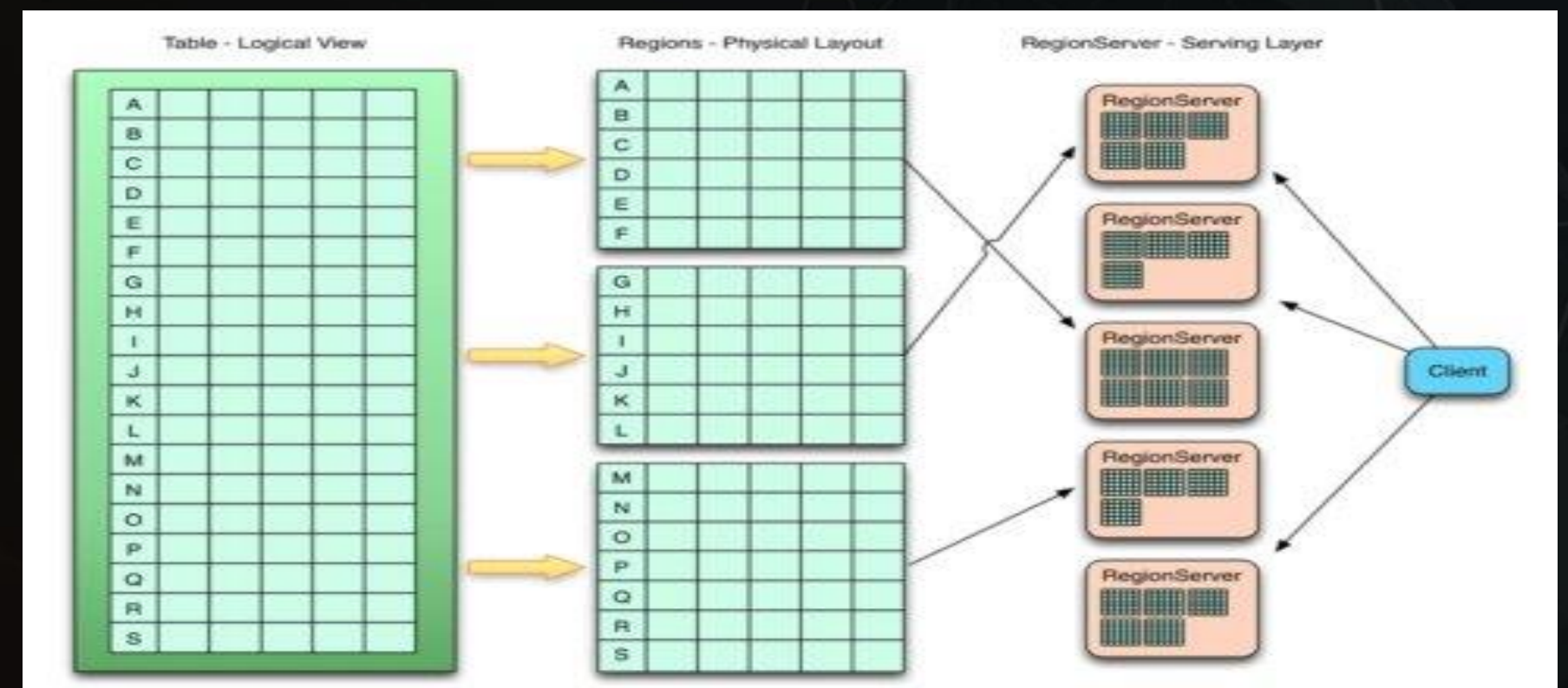
HDFS基本原理



分布式Key/Value列式存储系统 HBase



HBase基本原理





面向大数据的百亿级OLAP引擎 - Kylin

• Fast

百亿量级数据亚秒级返回SQL结果

• High Concurrency

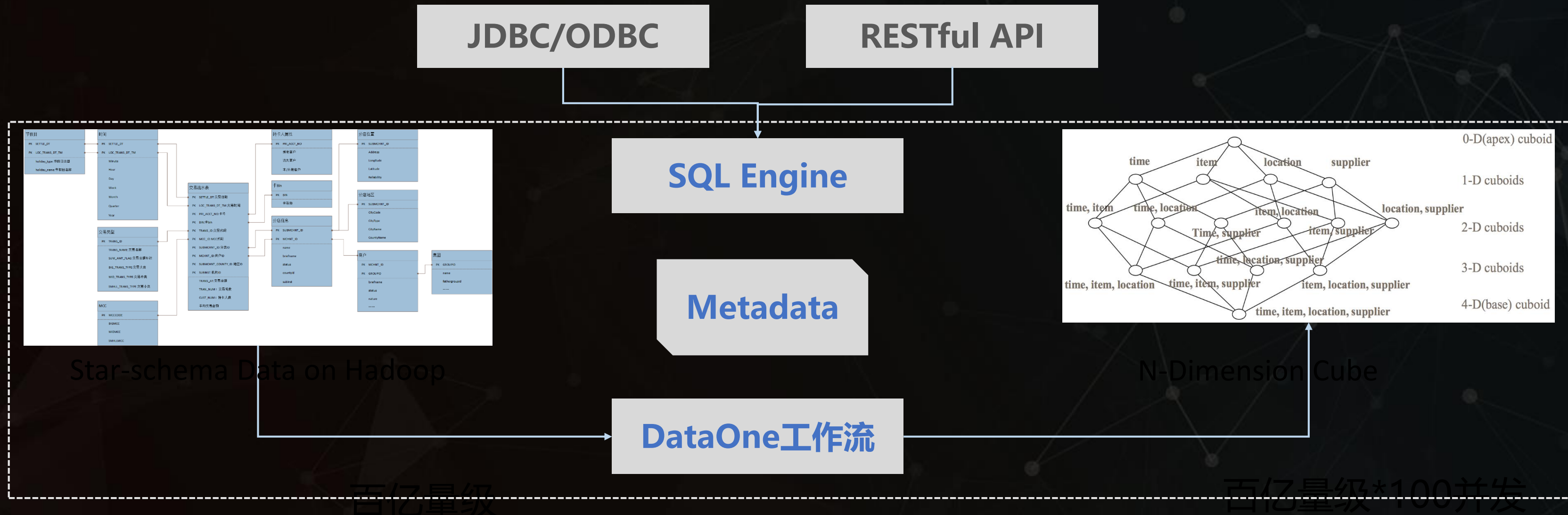
基于HBase的实现, 轻松承载高并发请求

• SQL Interface

标准SQL接口, 方便应用开发及迁移

• Easy to use

丰富的Cube管理接口, 灵活定义数据



BIG
DATA

数据挖掘



数据挖掘 —— 从**海量数据**中发现**知识**



数据连接

整合数据资源，数据挖掘平台和数据挖掘服务的一站式解决方案

明略数据拥有高质量的合规的数据供给方合作伙伴，将外部数据资源连接至明略自主研发的大数据挖掘平台DATAINSIGHT，同时明略数据科学家团队为客户提供专业的专属的数据挖掘服务。



明略数据互联解决方案

ONLINE

PC端行为数据

移动端行为数据

OFFLINE

消费行为数据

市/区政府公开数据



LINKED MOMENT
联袂科技



银联商务
China UMS



中关村数海大数据交易平台
www.shuhaidata.com



BIG
DATA

大数据安全标准



大数据相关标准及法规



中国大数据市场机制尚待完善，但真实需求决定了大量企业快速成长

BIG
DATA

大数据企业分析

中国大数据技术精锐企业图谱

CHINA BIG DATA LANDSCAPE-ELITE ENTERPRISES

DATA TECH. 数据 技术

阿里开放平台
百度数据开放平台
百分点
海量
集奥聚合
精硕科技

九次方
联袂科技
秒针系统
腾讯大数据
腾云天下
亿赞普

贵阳大数据交易所
数海科技
数据堂

国政通
华道征信
华院数据
蚂蚁金服
明略数据
数字新思
同盾科技

7G网络
东方国信
国云数据
华傲数据
华胜天成
华为
龙信数据

美林数据
明略数据
拓尔思
星图数据
亚信大数据
以萨数据

INFORMATICA
TABLEAU
百分点
博易智软
海云数据
精硕科技
久其软件

秒针系统
数字冰雹
腾云天下
永洪科技
用友软件
友盟

CLUDERA
PIVOTAL
华为
巨杉软件
浪潮
联想
星环科技
中科曙光

DATA COMMERCIALIZATION 数据商品化

搭建数据共享、交换、交易平台，为数据资源供给方和需求方提供交易环境和服务。

DATA CONSOLIDATION 数据整合应用

针对营销、风控、个性化服务等需求，整合数据资源，面向不同行业的客户提供数据应用服务，最靠近数据需求方角色。

DATA SOURCE 数据采集提供

基于自身业务及背后的监测、爬虫能力采集互联网公开数据和用户行为数据，加以清洗、整理，面向市场提供数据资源，数据资源供给方角色。

APPLICATION 应用程序供应商

应用程序供应商多是以提供产品的形式服务于客户企业。这些产品可能是BI工具，可能是数据接入或挖掘领域的细分产品。应用程序除了独立销售，很多时候也会作为整体解决方案的一部分出现。

SOLUTION 解决方案供应商

将商业产品或开源组件与业务知识和定制开发相结合，形成大数据解决方案，提供给行业客户。相比应用程序供应商和平台供应商，解决方案供应商是最直接接触行业客户的企业。

INFRASTRUCTURE 基础架构供应商

提供Hadoop发行版以及数据存储、并行计算、数据治理等周边软工具和硬件的厂商。很多基础架构供应商也提供解决方案。



国内大数据公司之一——百分点

- 百分点产品线已涵盖大数据基础层、管理层和应用层，其核心产品包括基础层的大数据操作系统，管理层的用户画像标签工场，以及应用层的推荐引擎、分析引擎和营销引擎；
- 百分点已为近2,000家互联网及实体企业提供大数据技术平台搭建和大数据驱动的SaaS应用，客户如华为、TCL、长虹、建设银行、1号店和中国电信等；
- 个性化推荐系统是百分点的过去的主打产品，公司结合大数据的应用也继续围绕着营销领域，侧重于电商、媒体、零售、家电制造、汽车、政府、电信、金融等多个行业；

BIG
DATA

数据堂——数据服务

- 深度挖掘原始数据的基础上，提供多样化的数据增值服务。
- 每天数据访问量在1亿次以上。

DATA

数据采集

众客堂 crowd.datatang.com/mobile

数据采集、数据标注、数据爬取的众包平台

数据交易

数据商城 www.datamall.com

实现大数据资源的在线共享与交易，极大的增强数据变现的空间，提高数据价值的流通性

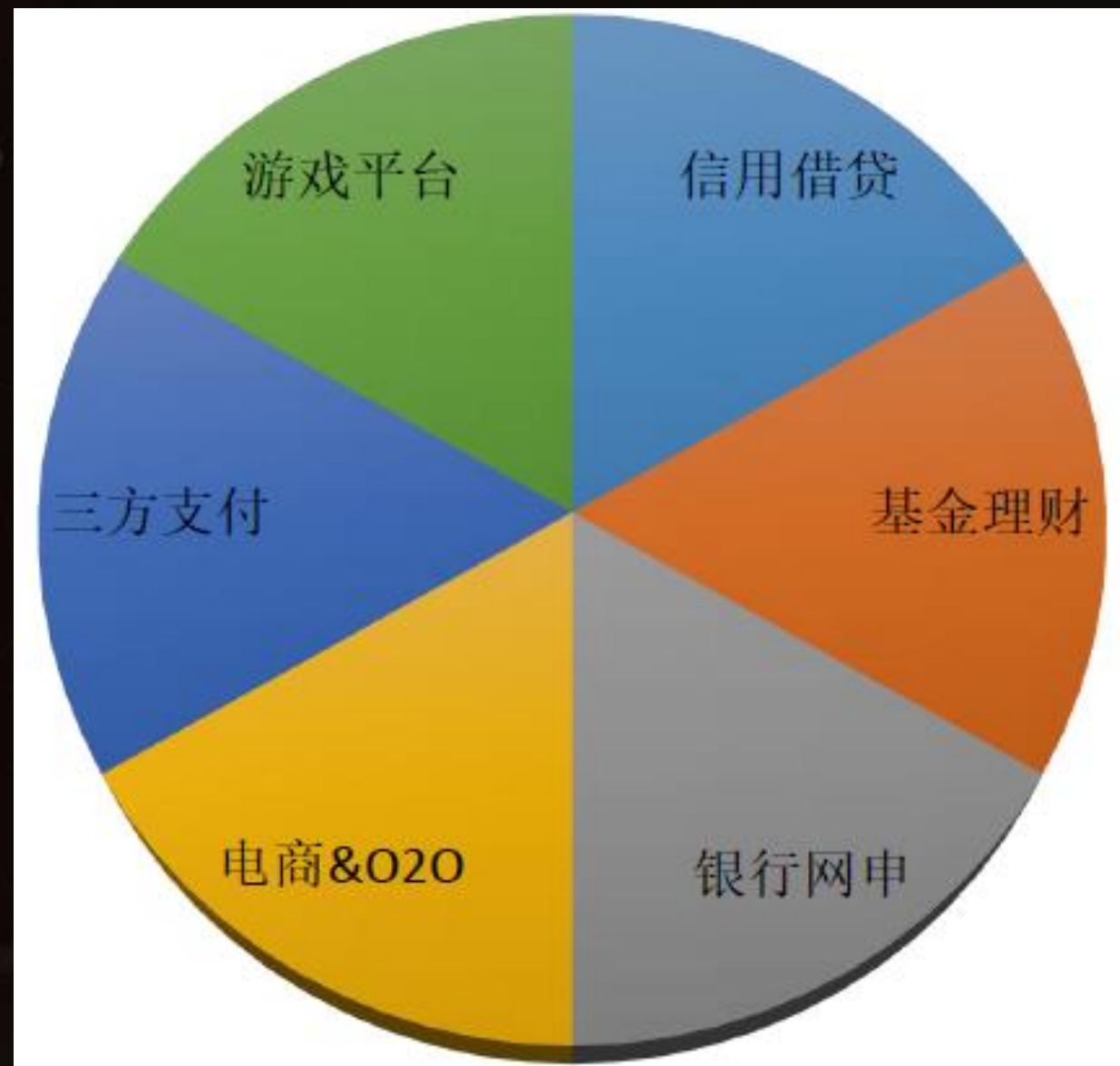
数据API

数据云服务：

深度整合各行业数据资源，全面挖掘数据价值，为客户产品和服务提供数据增值在线支撑平台

国内大数据公司之一——同盾科技

- 国内第一家大数据反欺诈服务提供商；
- 率先提出“跨行业联防联控”的风控理念；



跨行业联防联控





大数据应用工具/服务供应商 —— IBM

- IBM的大数据产品包括Hadoop平台产品与数据管理产品。包括：
 - 1.基于IBM SoftLayer全球云基础架构的“Hadoop即服务”产品；
 - 2.企业级的Hadoop发行版本；
- 推出了下一代人工智能-Watson，非常接近通过“图灵测试”；
- IBM近年来在大数据领域的投入很大；

时间	IBM 投入
2013 年	30 多起并购，包括 9 家分析解决方案中心，形成超过 6000 家商业伙伴的生态系统
2014 年	投资 240 亿美元构建 IBM 的大数据与分析能力；170 亿美元用于大数据及分析相关的并购。



国内大数据应用服务商——明略数据

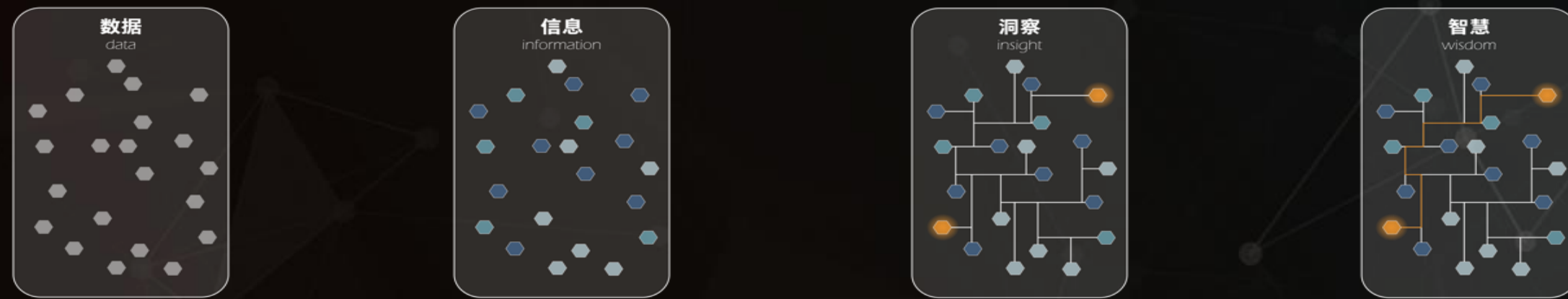
已形成多个能够给客户带来价值的行业解决方案

明略数据基于自主研发的大数据技术产品，与我们的领域专家一起结合金融、公共安全、税务等行业的普遍且重要的问题，形成了针对不同场景的一站式综合解决方案。我们相信关系到国计民生的大数据项目，需要有明略这样拥有自主知识产权和强大技术能力的中国公司来实现。



明略数据产品Landscape总结

从底层平台到上层应用



数据(MDP) → 信息(CONA) → 知识&洞察(SCOPA) → 智慧(DI)

- IT成本的不断降低造就了大数据
- 大数据的核心是数据的整合和连接
- 数据的应用比数据本身更重要
- 政府是大数据的最大用户

我们的观点

跨界为王。

麦肯锡认为未来需要更多的“translators”，能够在IT技术、数据分析和商业决策之间架起一座桥梁的复合型人才是最被人需要的。“translators”是可以驱动整个数据分析战略的设计和执行，同时连接IT、数据分析和业务部门的团队。如果缺少“translators”，即使拥有高端的数据分析策略和工具方法也是于事无补的。

BIG DATA

ARTIFICIAL INTELLIGENCE

不断沉淀的解决方案

已形成多个能够给客户带来价值的行业解决方案

明略数据基于自主研发的大数据技术产品，与我们的领域专家一起结合金融、公共安全、税务等行业的普遍且重要的问题，形成了针对不同场景的一站式综合解决方案。我们相信关系到国计民生的大数据项目，需要有明略这样拥有自主知识产权和强大技术能力的中国公司来实现。



高门槛的核心技术

明略最终形成顶级的人工智能警察、风控员、故障分析专家、审计员等基于数据工作的专家型岗位

以安全大数据平台MDP和分布式大数据挖掘系统DataInsight为基础，整合用户全网数据，构建用户360°统一视图，协助电商、旅游、新闻媒体、汽车、教育及房地产等行业深入挖掘海量数据价值，洞察客户，提高企业效益。

基于大数据的用户画像与精准营销

明略数据以第三方身份整合位置、消费等多方数据，构建统一数据分析框架，提供从数据挖掘到选址、定位、招商、运营等内容在内的市场研究整体服务，为大型商业地产企业提供商业地产大数据分析决策综合解决方案。

商业地产大数据分析决策系统

从体量庞大、类型多样的涉税数据和第三方数据中发现税源间隐藏的关联关系和固定规律，以此来更高效地解析税源各项指标，智能挖掘税源行为模式，更准确地预测税源涉税事件，支撑税务机关在税源精细化管理、风险防控、纳税服务以及法规宣传等方面的决策。

税源分析平台

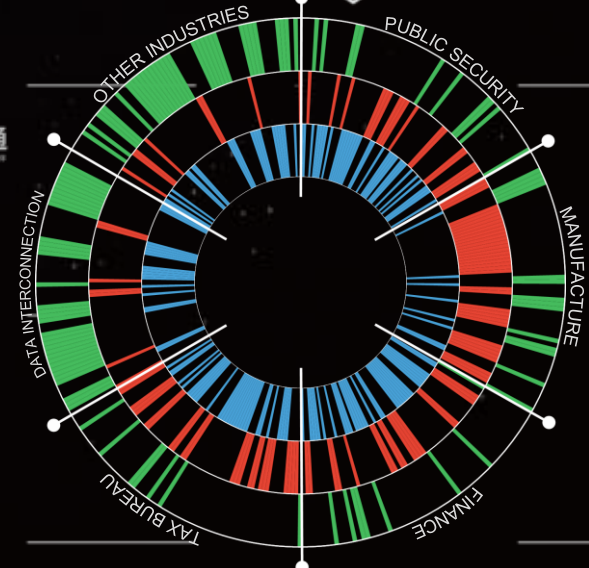
数据资产管理平台

集中管理量化税务机关数据资产总体情况，从数据量、数据种类、数据源等多维度揭露涉税数据的业务含义和关联关系，提供数据资产全景图，帮助税务机关更加系统、清晰、深入地理解自身数据资产情况，指导税务人员完成基于数据的各类应用操作。

MININGLAMP BIGDATA SOLUTION MATRIX

明略大数据行业解决方案矩阵

● SCOPA ● MDP ● Datainsight



精准营销解决方案



数据互联行业解决方案

税务行业解决方案



公安行业解决方案



制造行业解决方案



金融行业解决方案



明略官网二维码

明略微信二维码

将大数据平台MDP和关系挖掘产品SCOPA相结合，以情报业务为基础，为全警种提供综合情报分析，构建警务、情报与实战联动的作战平台。

情报综合研判实战平台

明略数据基于大数据的设备故障诊断与预测解决方案将成为制造业在数据矿井中的一盏明灯。明略数据凭借自身复杂建模能力、全线自主研发大数据产品，为企业提供产品预测性维护等服务，提升企业综合实力。

基于大数据的设备故障诊断与预测

通过整合数据和机器学习构建内控模型，结合不同业务有效识别和预警违规行为；利用关联关系挖掘产品SCOPA，构建实体关联网络，提供强大的分析功能，帮助法律合规、审计部门成功掌控内部操作风险。

大数据金融机构内控管理

金融大数据平台

基于MDP的金融大数据平台解决方案，帮助客户治理、整合、打通存量数据，进而进行针对业务需求的查询、报表、基础用户画像等应用建设。

大数据客户价值挖掘

针对营销场景需求，整合传统数据、线上线下数据及外部数据，以数据挖掘和机器学习为方法，结合金融业务特征，提供个性化服务，最大化客户价值。

以关系挖掘产品SCOPA为核心，以图数据库和战法模型为支撑，挖掘多源异构的数据关系，完成全维度情报分析，覆盖重大事件预警、重点人员管控和情报研判等场景。

大数据情报分析系统

通过整合、打通数据源，在海量关联数据的基础上针对历史的不良客户和欺诈行为进行深度机器学习，准确地识别和预警欺诈行为，帮助保险公司有效提升反欺诈效果。

大数据保险反欺诈

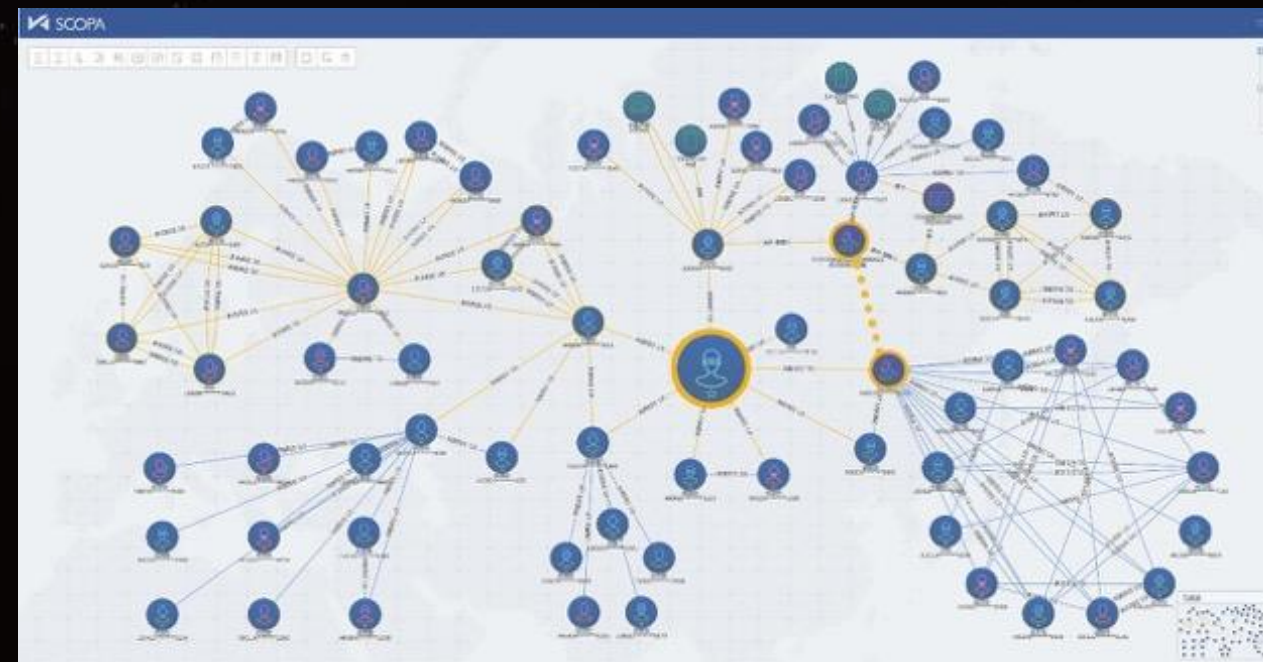
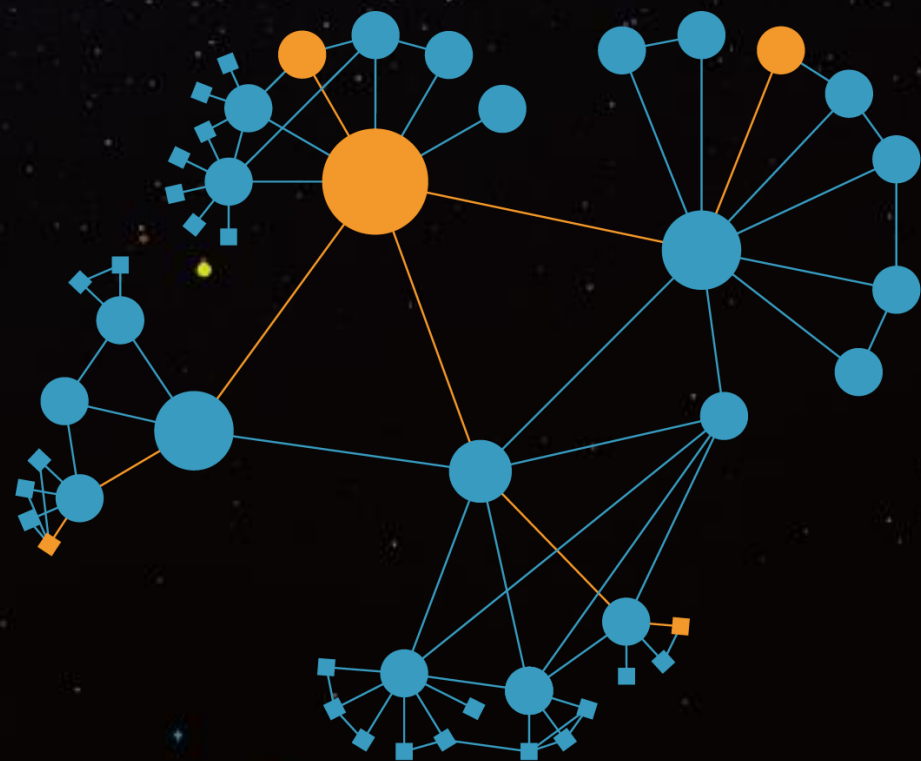
大数据信贷风险管理

利用强大的关联分析和数据挖掘技术，整合异构数据，挖掘特征并构建数据关联网络，识别新客风险，预警存量风险，帮助金融客户有效管理信贷风险。

不断沉淀的核心引擎产品

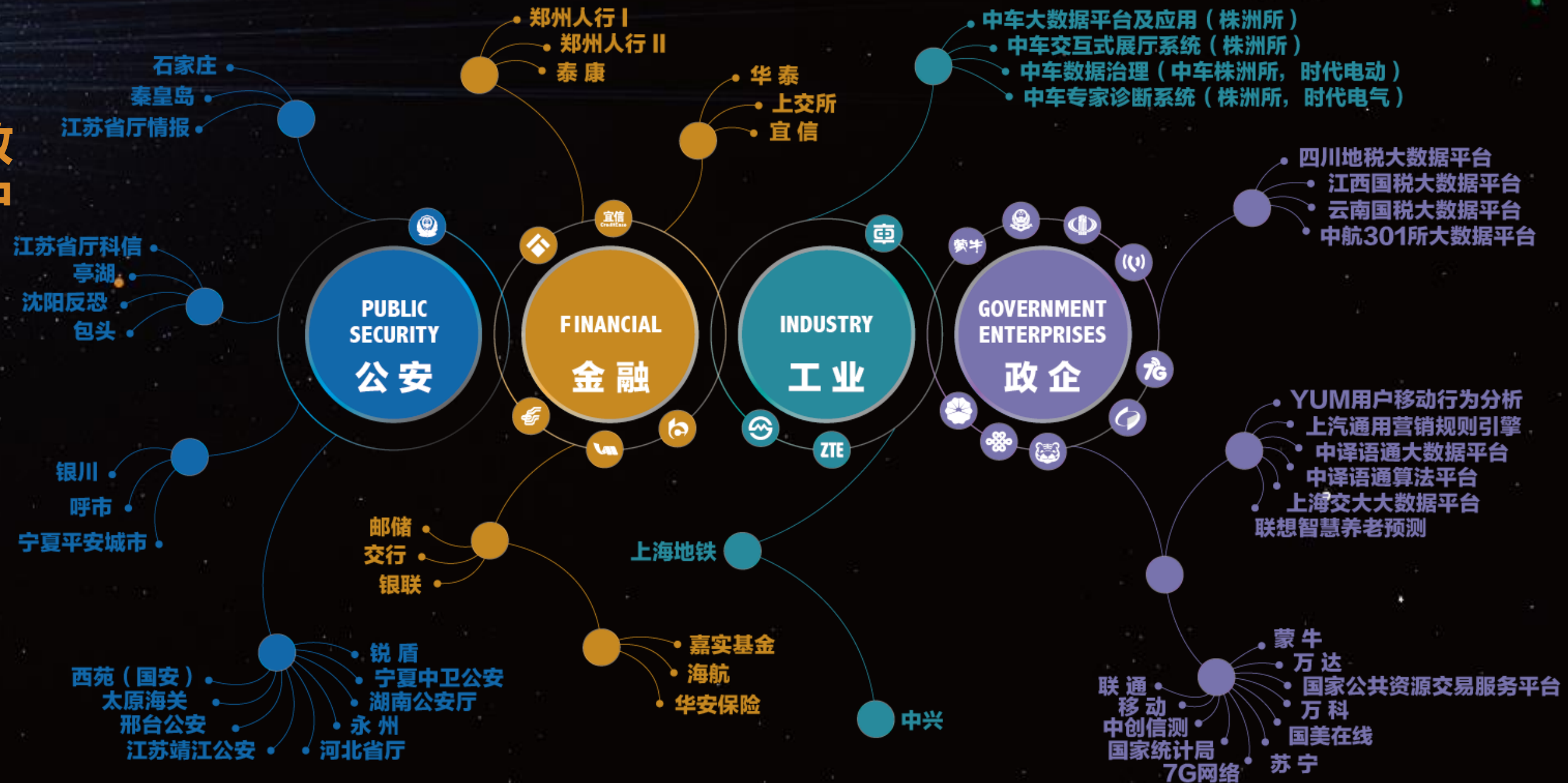
明略大数据关联关系挖掘产品SCOPA

明略数据SCOPA是基于PB级全量数据进行秒级运算的大数据分析和应用综合平台。用数据还原真实世界，基于创新的大数据技术组合，SCOPA颠覆了传统的数据查询和计算方法，打造了多维多层关系挖掘的全新数据分析模式。从数据端到用户端，SCOPA构建了一套高度智能化的人机智慧协作体系，可以广泛应用于多行业多领域，如公安、金融、工业、政企等。



明略客户

短短三年内拿下公安、金融、工业、政企等行业的顶级客户

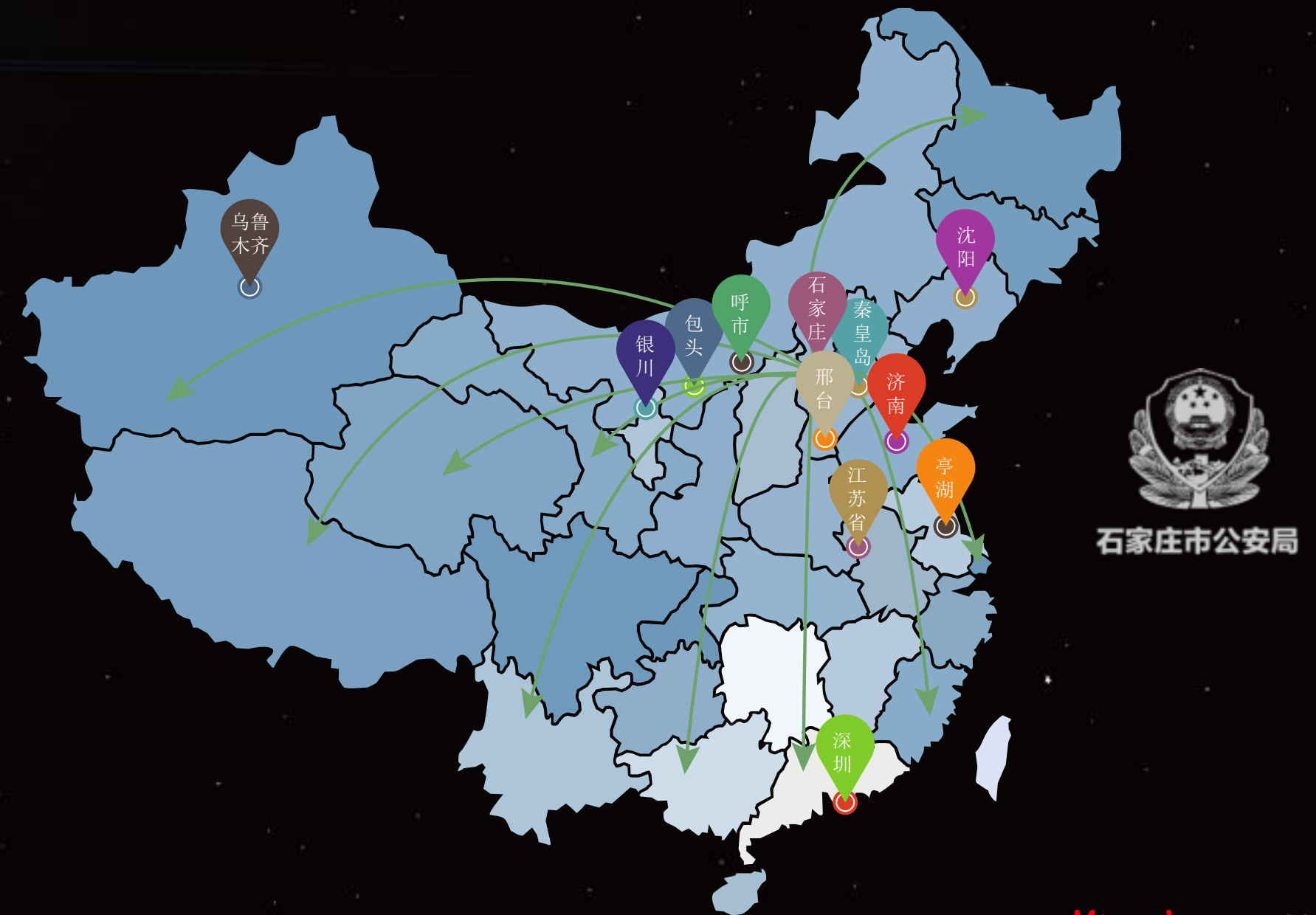


从标杆客户到行业可复制场景的快速沉淀

在公安领域里从石家庄走向河北，再到全国

rapid EXPANSION

应用场景	客户
大情报业务应用：知识图谱+在线战法	石家庄、秦皇岛、江苏...
科信垂直挖掘及数据综合治理	江苏、大连、四川...
科信、指挥中心综合数据分析服务集	包头、呼市、银川...
区县级数据汇聚、管控平台	盐城亭湖、辽宁营口...



人工智能的警察

新一代人机交互的情报研判系统

- 通过CONA和XPedia知识获取产品，能对大量孤立、异构的数据进行高效的处理、连接，转化为高价值情报；
- 通过HARTS和DI积累和管理了大量与犯罪相关规则、预测模型，结合数据科学家和公安专家的智慧，辅助进行案件研判和预警；
- 通过SCOPA进行全维度知识关联，快速智能的进行关联关系挖掘，绘制出反映真实世界的概念和关系网络；秉承“Human-Driven”和“Data-Answer”理念，使警务人员在最短的时间内找到关键线索。

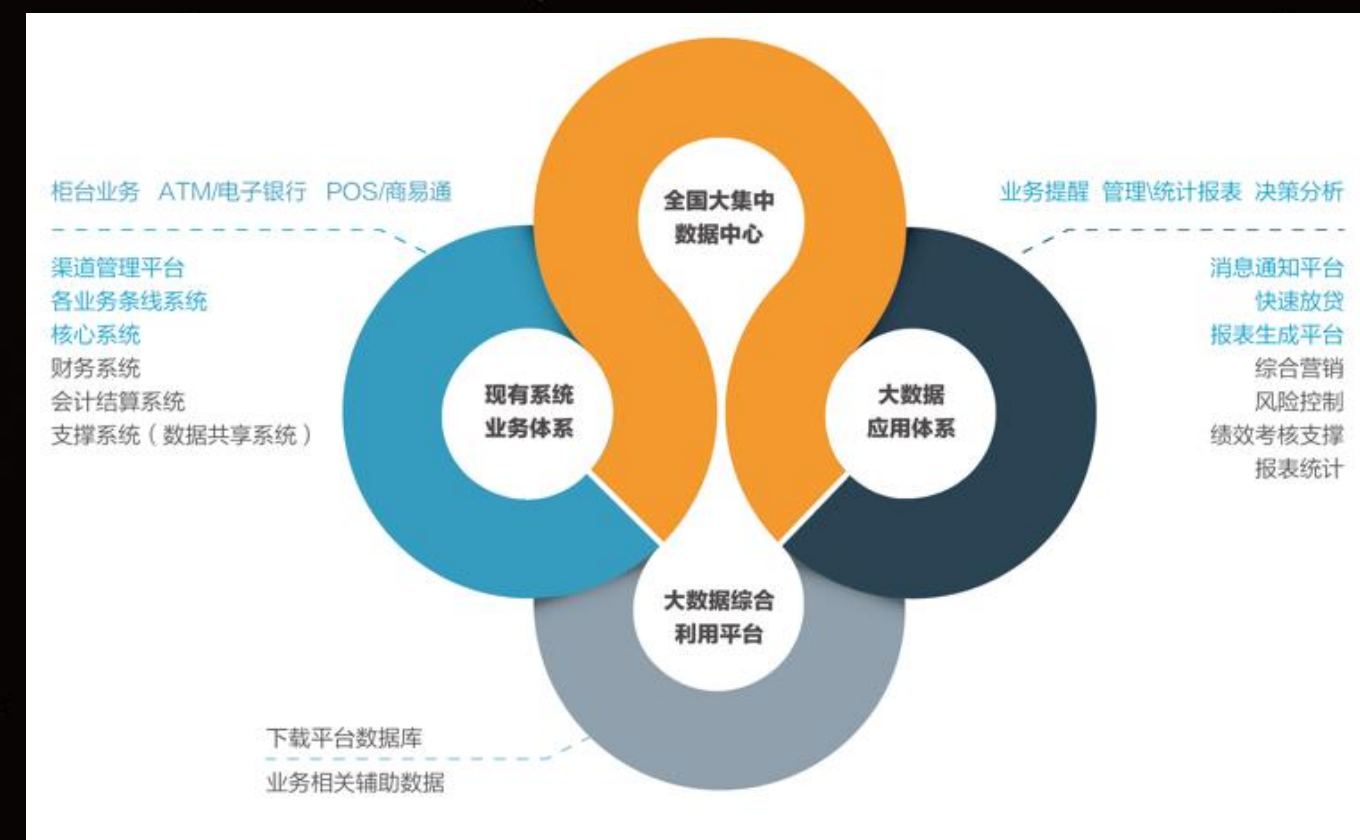


人工智能风控员

打通行内外数据，预测风险，超短期放贷

优质客户筛查、不良贷款预警、风险控制、行业信息分析等提供了多个优质的大数据应用落地，对河北分行重点发展的普惠金融业务起到了重要的支持作用。

- **对行内行外客户相关数据进行知识抽取**，形成完善的客户画像体系，支持**300+**标签。
- **通过DI积累风控等规则和模型类知识**，使得优质客户的小额业务办理时间从以前的**2-3天**缩短到**30分钟**，将不良贷款率降低**80%**、控制在**1%**以内。
- **通过与既有业务系统密切整合、及时反馈的知识交互**，让小额业务人员能随时查询优质客户和不良贷款的趋势和分布，以及预测得到的未来小额业务还款情况。

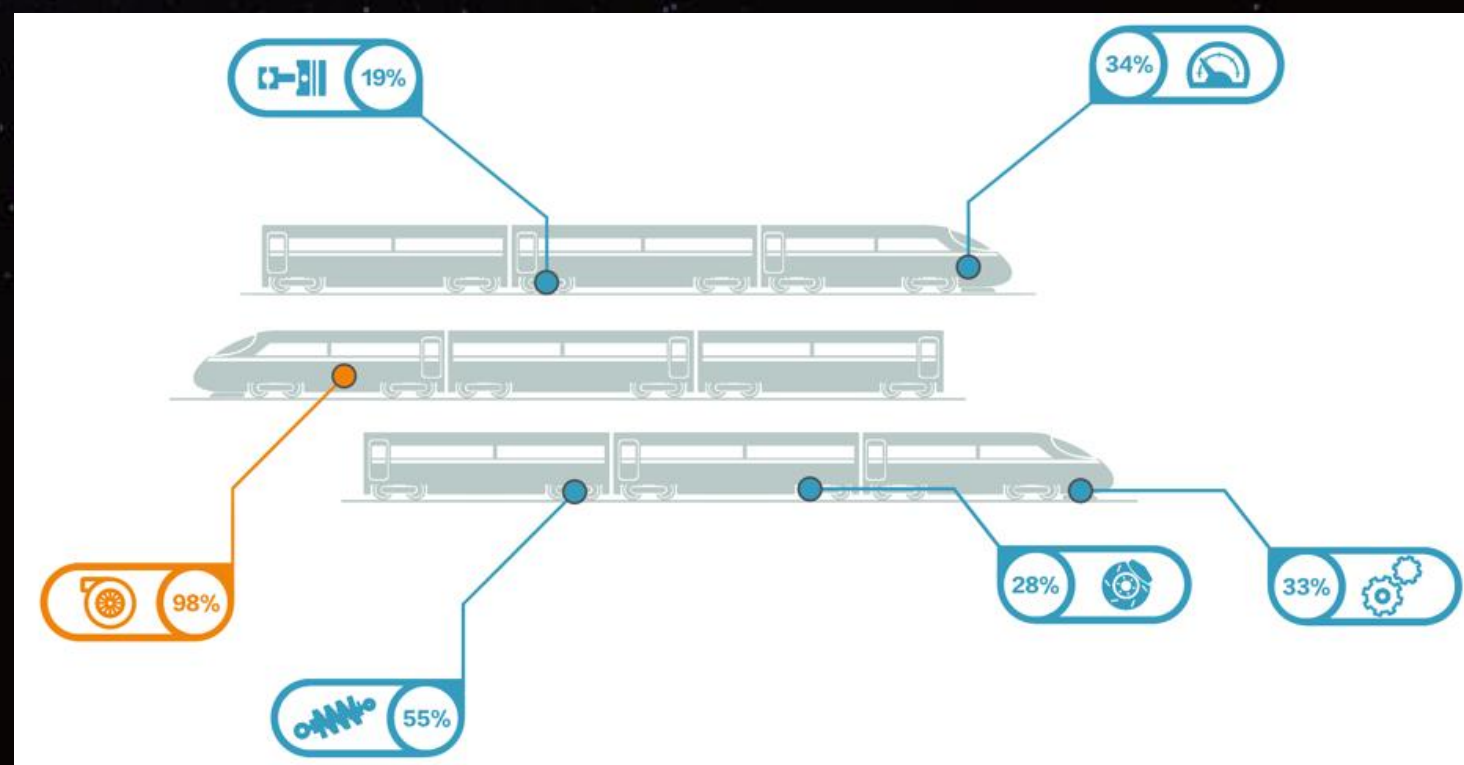




人工智能故障分析专家

机车零部件故障诊断和预测

- 依托产品实现对海量设备数据的知识抽取，进而可以进行BI分析和状态监视。
- 通过DI对故障信号的模式识别模型进行积累和管理，结合行业专家分析故障原因，并采用GPU加速深度学习。



专业委员会

公共安全大数据专业委员会

中关村大数据产业联盟公共安全大数据专业委员会（简称“安委会”），于2016年8月2日在北京成立，作为公共安全大数据领域产、学、研、用沟通的平台。

发起单位北京明略软件系统有限公司，顺应大数据产业在公共安全领域快速发展的趋势，助推大数据技术在公共安全落地实践中发挥更大价值，共同探索大数据技术在公共安全领域如何发挥最佳效用，为政府打造立体的打防管控体系，实现真正的“平安中国”助一臂之力。

荣誉顾问：吴绍忠教授，博士。中国人民公安大学公安情报研究中心主任，首席研究员；侦查与反恐怖学院情报技术教研室主任；网络空间安全与法治协同创新中心副秘书长。



公共安全大数据专业委员会
China Public Security BigData Council

高校合作——梦想、技术，成就社会责任

与北京大学计算机系共同研究人工智能算法，并将直接应用到产品中去。

与清华大学自动化系合作研究基于人工智能技术的机械设备寿命预测。

与中国公安大学反恐学院成立联合实验室，致力于基于人工智能的反恐技术研究。

.....

BIG
DATA

ARTIFICIAL INTELLIGENCE

THANKS

