


博彦科技大数据解决方案

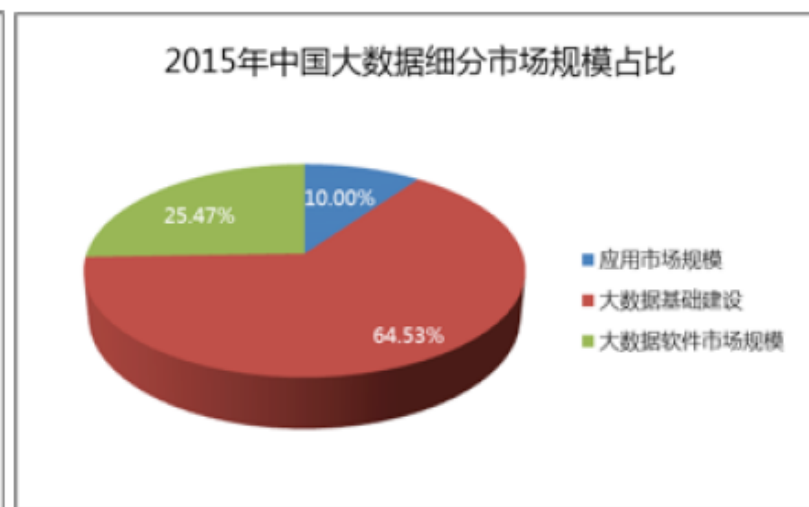
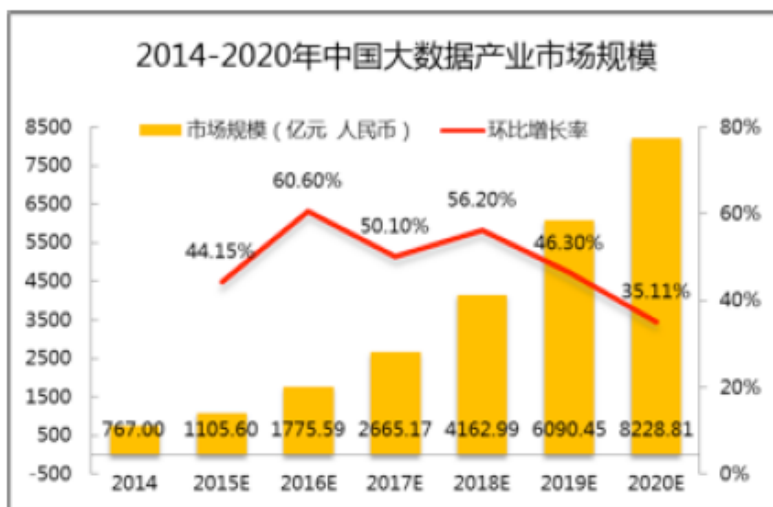
全球客户信赖的IT综合服务提供商和战略合作伙伴



目录

1. 行业现状、公司简介
 2. 专注行业、技术优势
 3. 服务领域、解决方案
- 

1 国内大数据产业



预计未来5年大数据产业市场规模增长强劲

2014年中国大数据市场规模达到767亿元，同比增长27.83%。预计到2020年，中国大数据产业市场规模将达到8228.81亿元

2 国内大数据应用市场



大数据应用市场规模同样强劲

2014年中国大数据应用市场规模为80.54%，同比增长3.23%，预计2015年市场规模增长37.28%，至110.56亿，预计到2020年，中国大数据应用市场规模将增长至5019.58亿。



更多应用，更多行业领域

互联网的迅速发展，使数据产生速度加快，规模加大，迫切需要运用大数据手段进行分析处理，提炼生成有效信息。大数据应用不仅仅停留在IT领域，在医药，科学，制造以及气象等行业，都出现海量的数据应用。

3 博彦，致力成就卓越



20

20年专业服务外包实践



17

17年离岸交付成功运营



12

12个垂直行业的定制化解决方案与服务



32

32个值得信赖的全球交付中心



8000+

8,000+名杰出专注的人才团队



100%

100% 专注卓越服务的核心团队

4 博彦，大数据足迹



中文搜索反垃圾项目

团队：20+人



咨询、图片、搜索、视频
电子商务等数据编辑处理运营

团队：70+人



社区数据运营分析

团队：300+人



舆情大数据、商业智能、
金融大数据、数据分析



团队：成立大数据业务部



收购PDL-拓展商业智能业务领域
收购红麦聚信-拓展大数据业务板块

2006

2010

2012

2015

2008-2009

2011

2013-2014



社区数据运营

团队：40+人



网页搜索、舆情分析、视频审核
信息挖掘、产品体验、地图数据

团队：150+人



网页搜索、舆情分析、视频审核
信息挖掘、产品体验、地图数据



数据运营：400+人



收购上海泓智-金融大数据迈出第一步
收购TPG-拓展大数据分析和商业智能业务



目录

1. 行业现状、公司简介
- 2. 专注行业、技术优势**
3. 服务领域、解决方案

1

专注行业领域

金融

- 智慧银行、智能现金流解决方案
- 征信大数据、风险预警
- 理财产品推荐分析
- 市场竞争产品分析
- 舆情监测、危机公关
- 客户心声分析

互联网

- 舆情大数据
- 数据运营
- 彩虹盾
- 消费者洞察
- 网站分析

高科技

- 舆情大数据
- 数据运营分析
- 智能运维
- 数字化市场
- 商业智能

零售

- 数字市场
- 商业智能
- 消费者洞察
- 舆情分析

政府

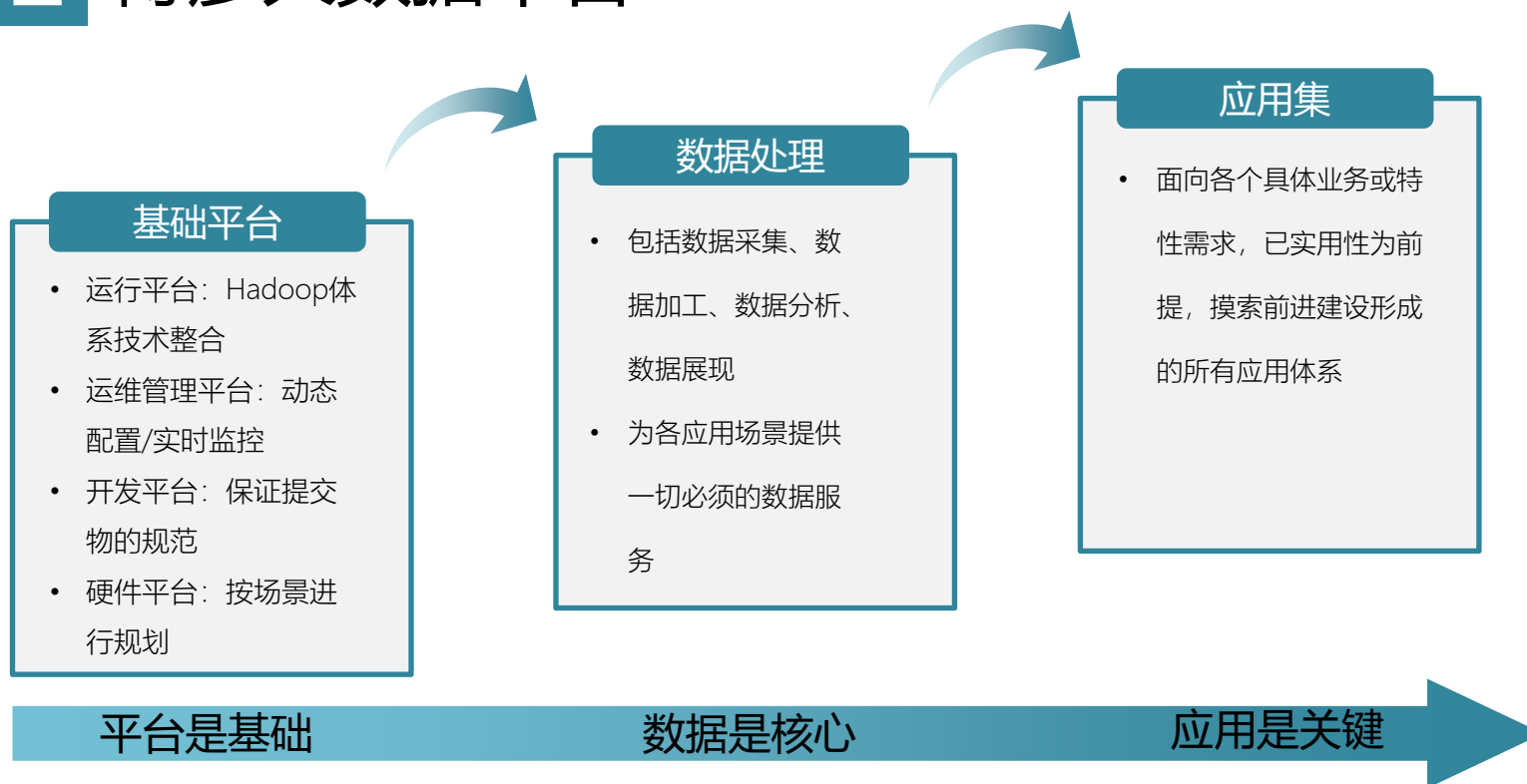
- 舆情分析
- 数据运营
- 商业智能
- 智能运维

其他

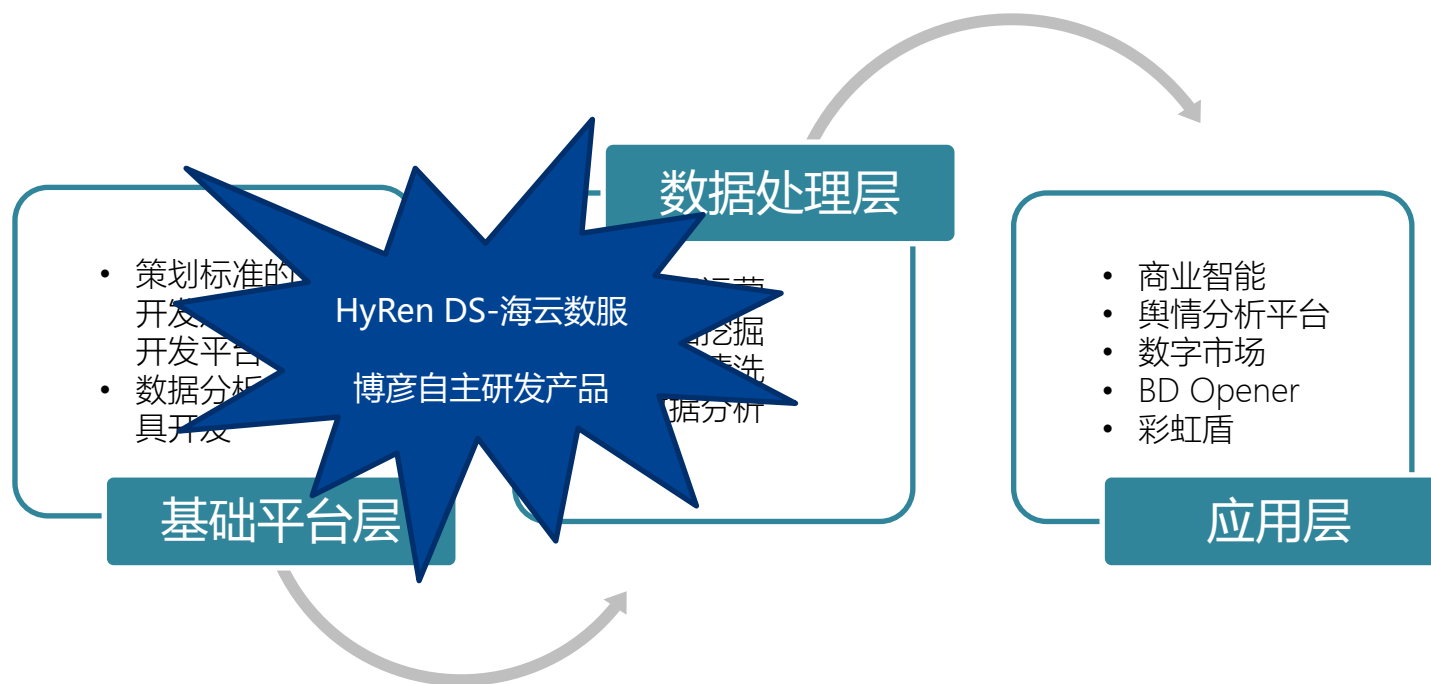
- 数字化市场
- 商业智能
- 舆情大数据
- 金融大数据
- 消费者洞察
- 数据运营
- 机器学习



2 博彦大数据平台



3 我们能做的是...



4 技术亮点

分布式爬虫技术

- 采用完全模块化设计，功能覆盖整个爬虫的生命周期：链接提取、页面下载、内容抽取和持久化，支持多线程抓取，分布式抓取，并支持自动重试



图像检索技术

- 在1秒钟之内，从1亿件的图像中筛选相似特征图像的高速检索技术
- 仅用50ms即可更新1张图像索引的数据库管理技术



非结构数据服务技术

- 网页、Word/Excel/PDF、邮件等
- 语义库管理
- 内容提取及语义分析
- 基于业务场景的中文分词及全文检索



数据分析技术

- 基于Mahour进行模型设计
- 为业务和决策提供支持

4.1 分布式爬虫技术

通用爬虫

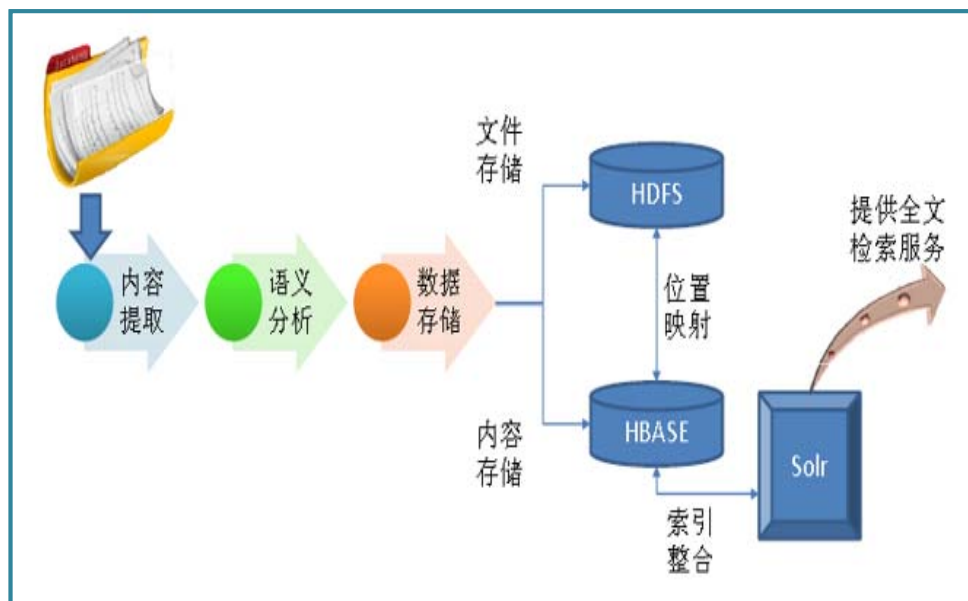
能够从网页正文中自动抽取主要内容，抽取目标可配置定义

智能网络爬虫

只抓取新的文章内容，对网站内容进行自动识别；HTML网页内容萃取，自动提取网页价值信息

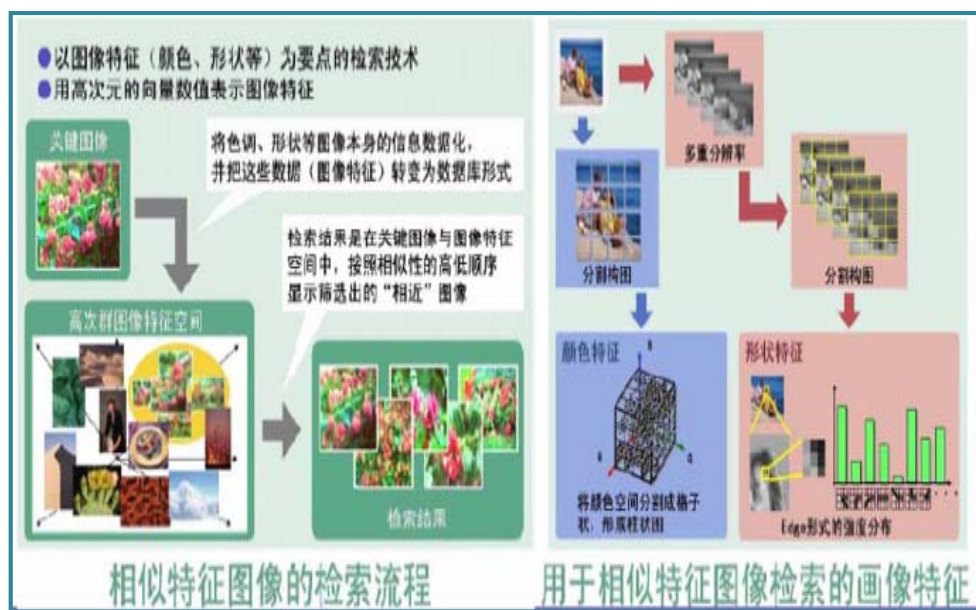
- 采用完全模块化设计，功能覆盖整个爬虫的生命周期：链接提取、页面下载、内容抽取和持久化，支持多线程抓取，分布式抓取，并支持自动重试

4.2 非结构数据服务技术



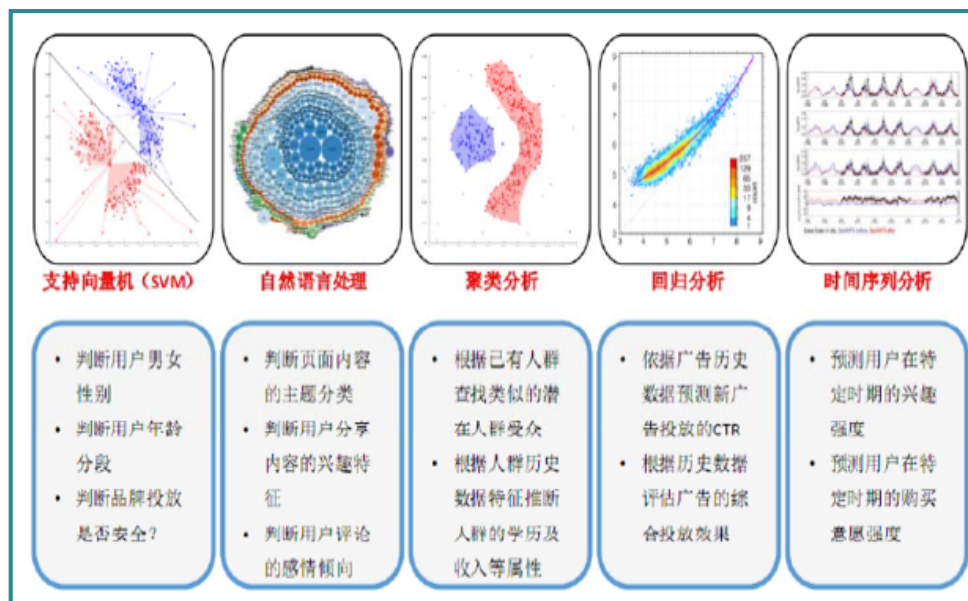
- 网页、Word/Excel/PDF、邮件等
- 语义库管理
- 内容提取及语义分析
- 基于业务场景的中文分词及全文检索

4.3 图像检索技术



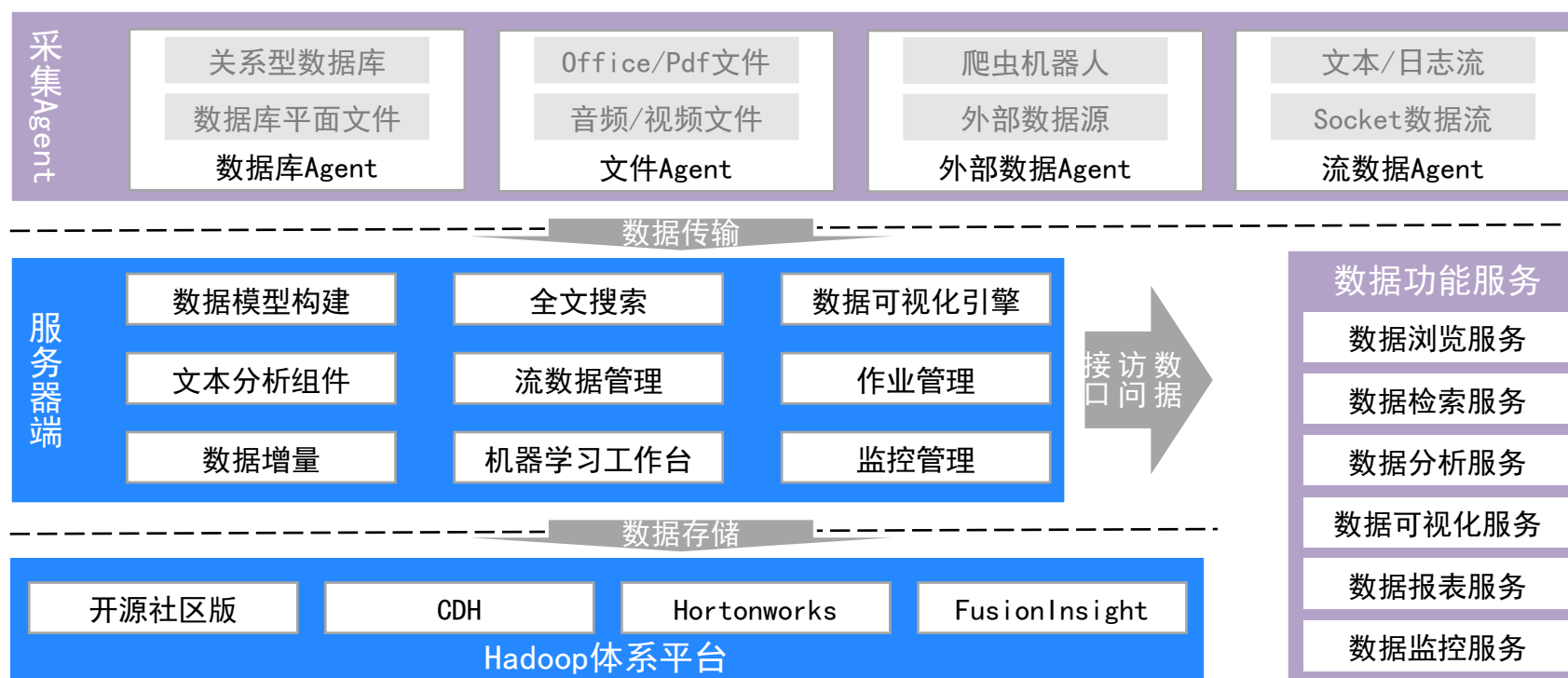
- 在1秒钟之内，从1亿件的图像中筛选相似特征图像的高速检索技术
- 仅用50ms即可更新1张图像索引的数据库管理技术

4.4 大数据分析技术

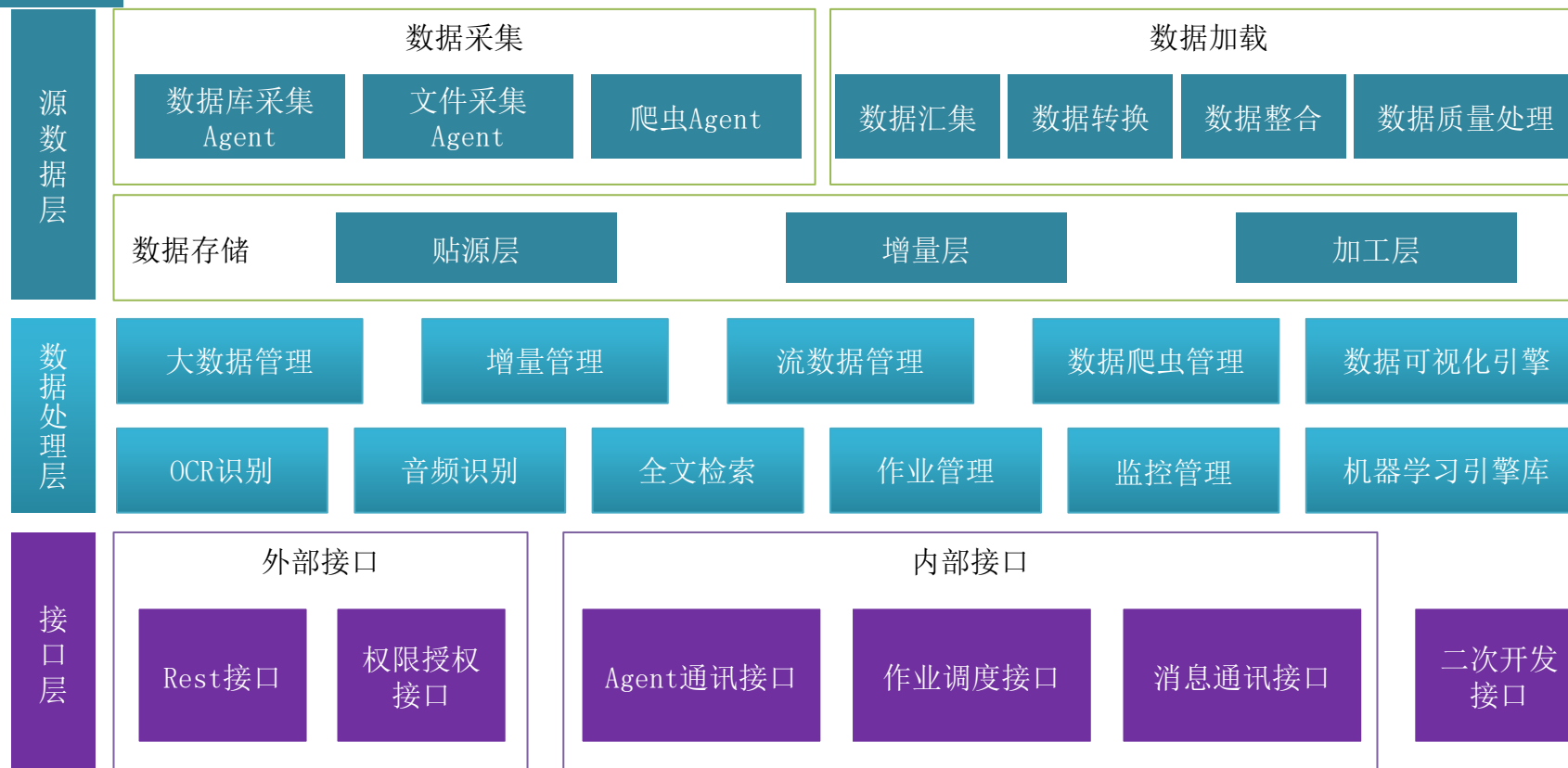


- 基于Mahour进行模型设计, 为业务和决策提供支持

4.5 HyRen DS: 企业级, 面向大数据管理的基础平台




4.6 HyRen DS-系统架构





目录

1. 行业现状、公司简介
 2. 专注行业、技术优势
 - 3. 服务领域、解决方案**
- 

1 博彦大数据领域

金融大数据解决方案

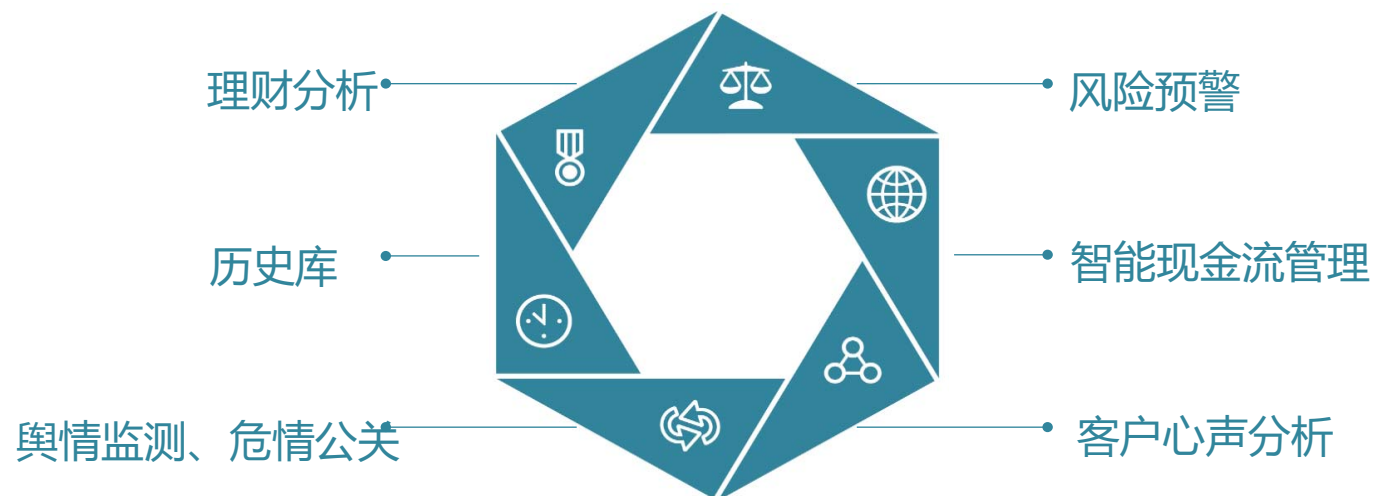
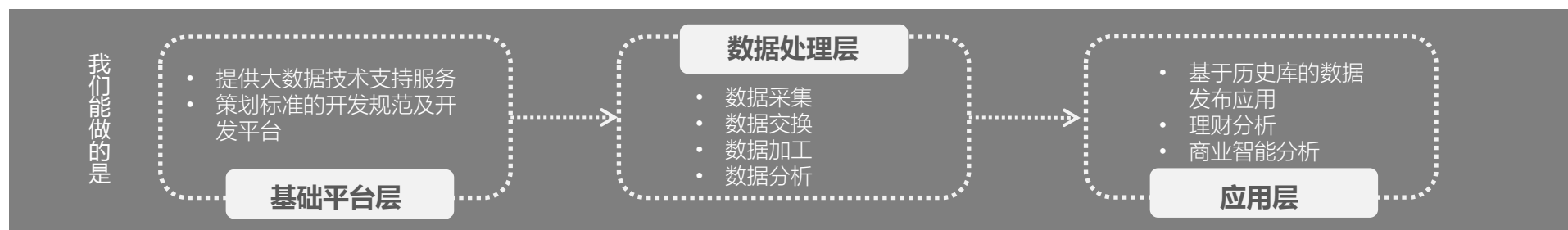
互联网大数据解决方案

数字市场&商业智能

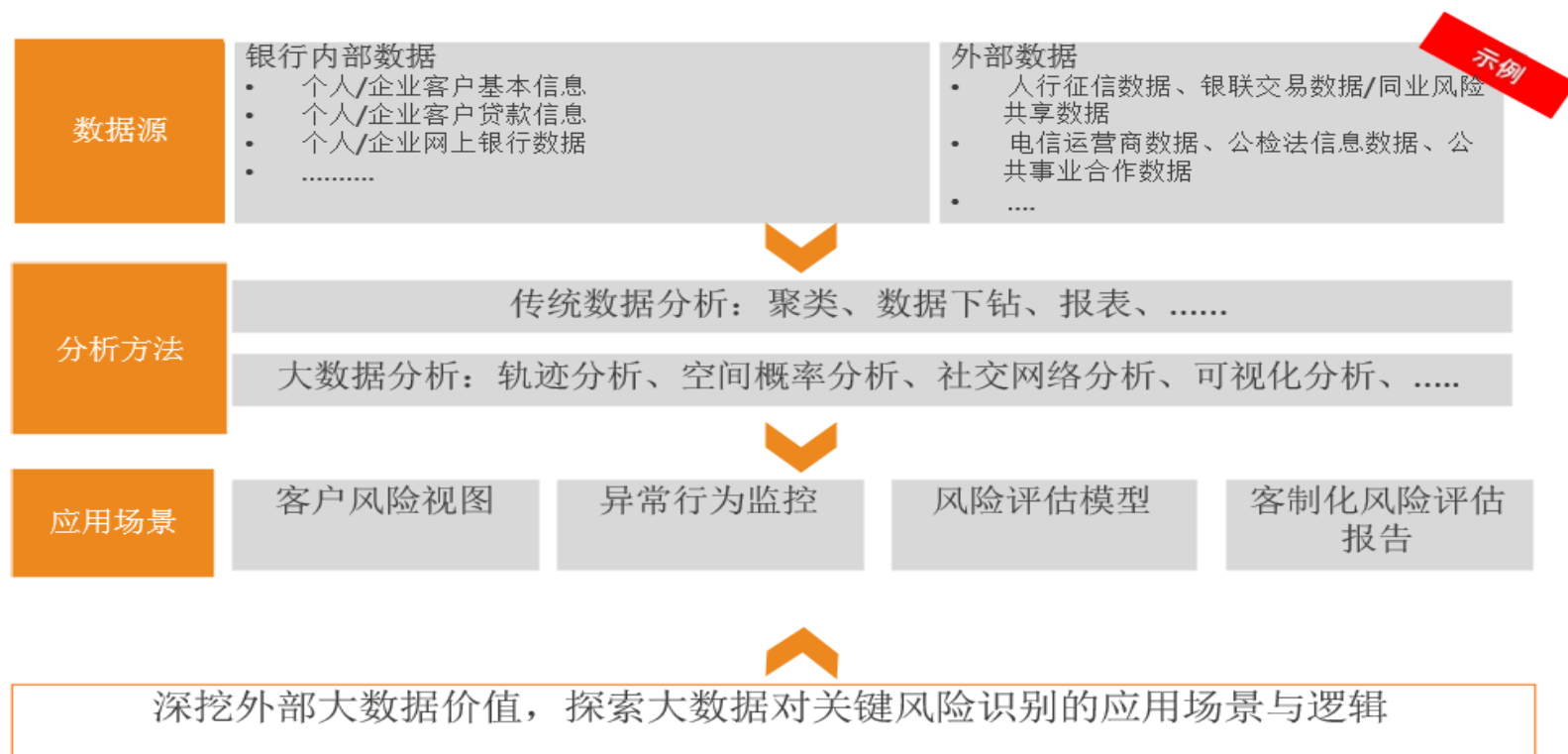
智能运维

机器学习

2 金融大数据解决方案



2.1 风险预警



示例

2.2 理财分析

■ 目标： 对当前投放市场的所有竞争产品进行信息汇总、挖掘，辅助产品开发

■ 场景：



投资理财一览：洞悉其他银行、机构开发的理财产品。了解市场动态。借鉴并开发出有竞争力的产品

问题：

理财产品千千万，如何开发出有差异化有潜力的理财产品？



■ 提案：



网络爬虫



银行网站

sina 新浪财经
finance.sina.com.cn

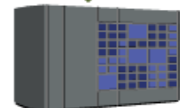


财经门户

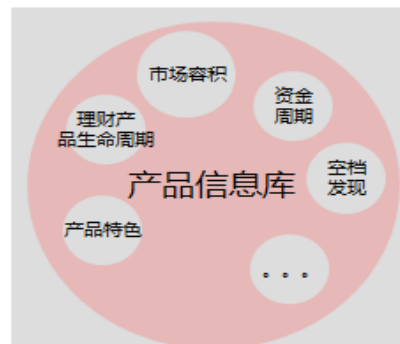
产品获取

内容存储

数据转化



全文检索



知识可见



爬虫技术可实时抓取各大网站、理财信息。



大规模实时存储系统



产品特征提取



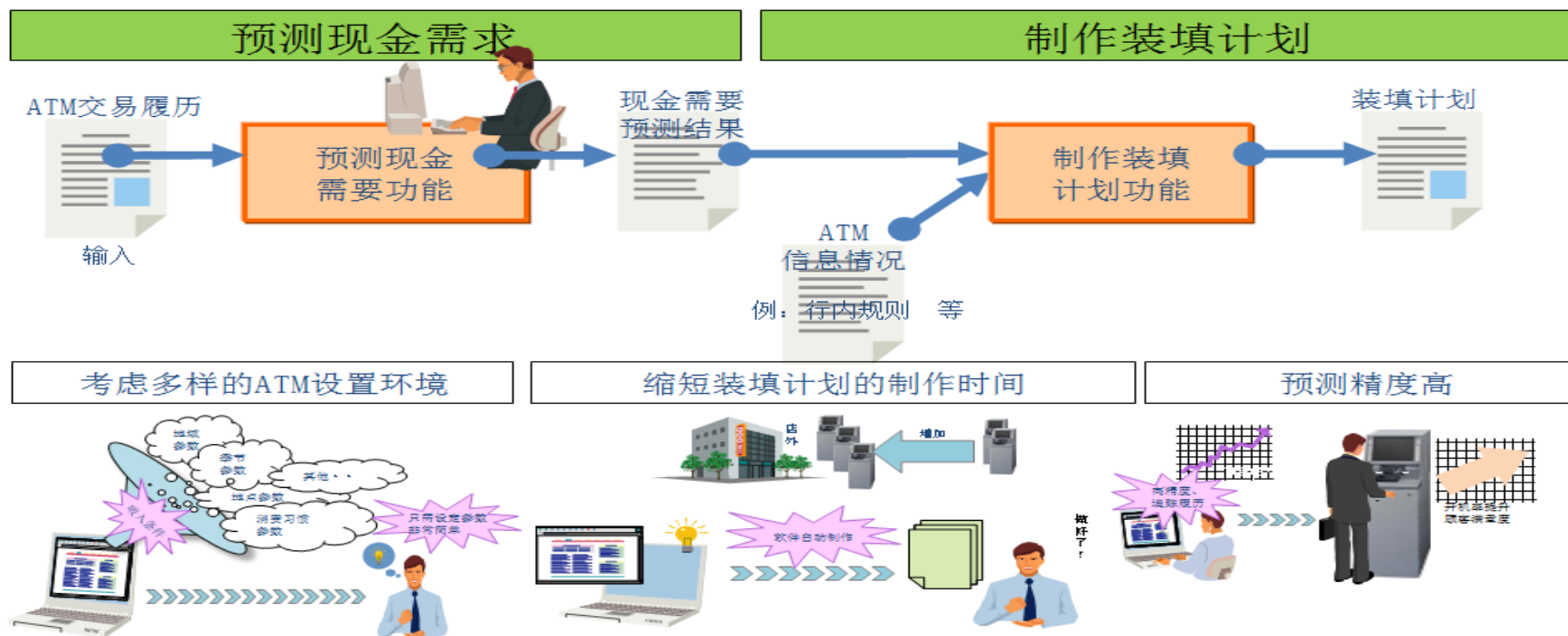
产品信息库构建



策略推荐

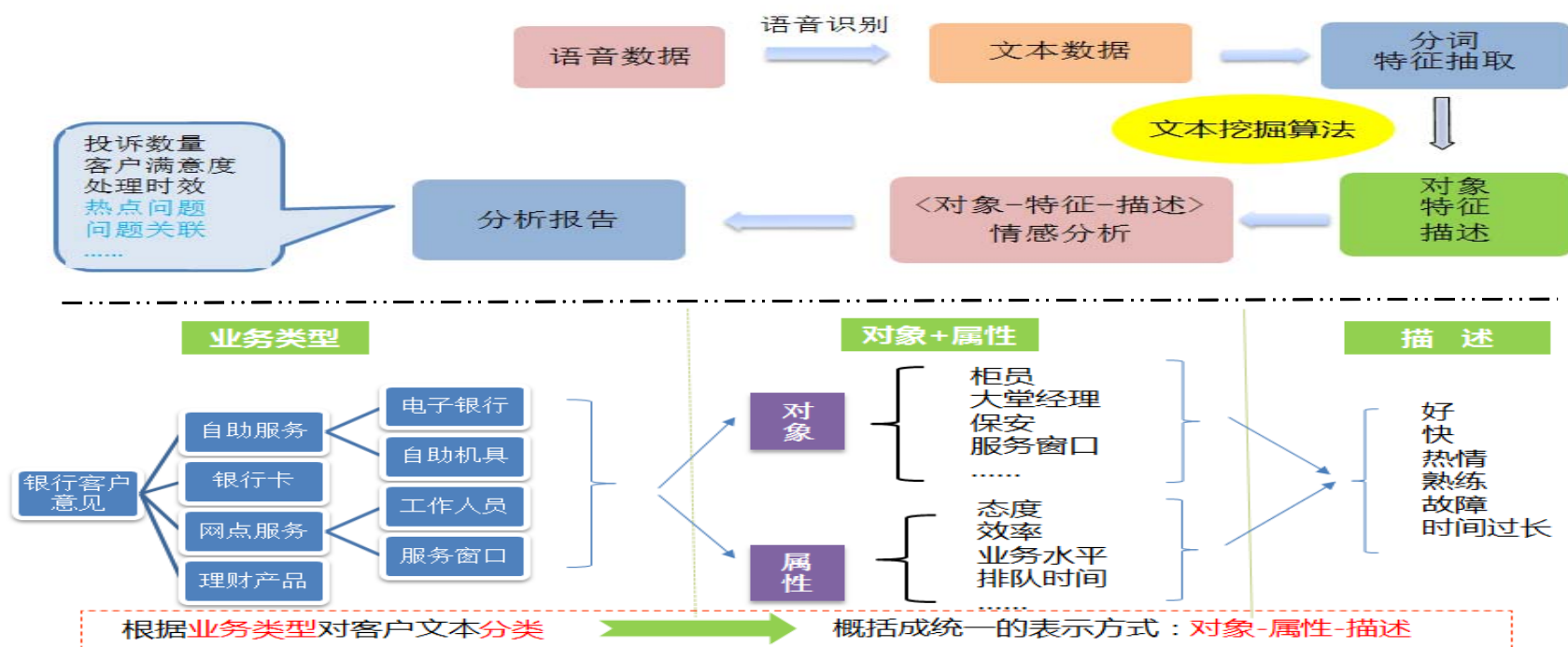
2.3 智能现金流管理

综合分析过往每一台ATM机器的交易信息，在保证ATM现金不空箱的情况下预测现金最小需求量，以达到ATM全体能在适当的现金量下运用。

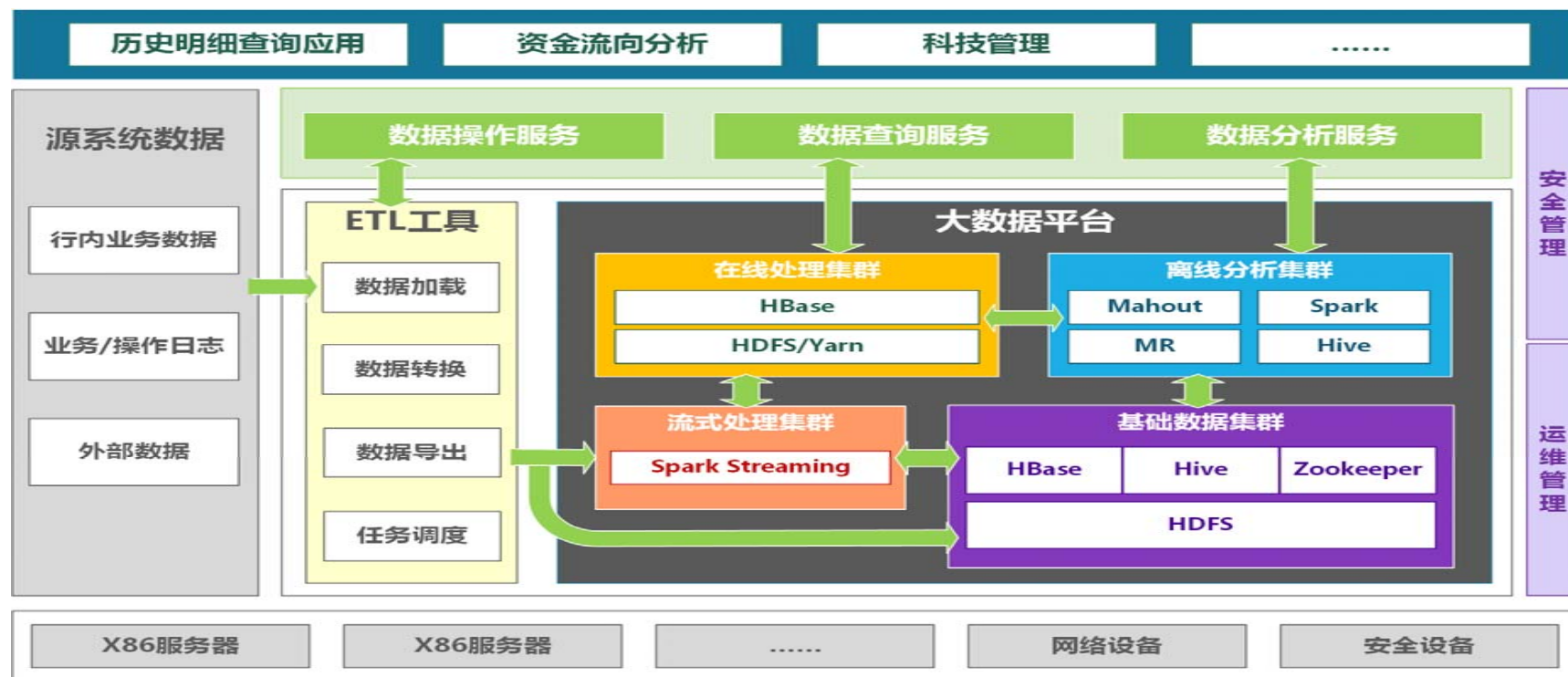


2.4 客户心声分析

结合语音识别和机器学习技术，获得电话客户的热点意见和情感倾向等分析成果，由原来人工阅读工单变为机器自动分析，更有效的提升客户满意度和忠诚度



2.5 历史库



3 互联网大数据解决方案

电子商务

- 站内搜索相关性评测
- 产品ID标注
- 推荐商品用户体验
- 产品相关性评测
- 电子商务分词标注
- 导购页面标注
- 商家认证审核
- 商品类目识别标注



搜索引擎

- 网页搜索结果排序
- 页面质量标注
- Query意图分析
- 网页搜索对比评测
- 视频搜索
- 图片搜索
- 地图搜索



舆情大数据

- 舆情信息判断
- 舆情信息挖掘
- 舆情价值分析
- 舆情信息监测



社交网络

- 社区内容编辑
- 用户审核
- 内容审核
- 评论审核
- 热门话题推荐
- 处理用户举报
- 社区版面管理



产品运营

- 分析用户需求
- 互联网营销
- 数据整理, 报表分析
- 产品策划
- 产品优化
- 语料挖掘



3.1 舆情数据分析



日/周/月/季/、年报专题事件报告、市场效果评估

3.1 輿情数据分析

全新云爬架构

- 自定义界面
- 自定义组合
- 自定义功能



多级权限管理

- 用户权限管理
- 輿情待办工作
- 操作权限控制



模块化前端

- 本地化部署
- 系统一键更新
- 智能安全防护



多终端应用

- 用户权限管理
- 輿情待办工作
- 操作权限控制

3.2 BD Opener

国内首款消费者评价分析工具



准确
Accurate

开启中文文本挖掘的智能分词技术，准确率高达99.6%



全面
Comprehensive

覆盖全网络主流平台，开启信源上各品类的消费者评价采集



快速
Fast

24小时无休，实时数据及时呈现，快速满足客户需求



易用
Easy

数据分析可视化，轻松掌握数据秘密，开启商机

3.2 BD Opener

国内首款消费者评价分析工具



行业分析

- 声量统计
- 关键词(正/负)
- 评价数据&健康指数



评价搜索

- 搜索框
- [区分平台]搜索结果
- 被搜索关键词数据量



Dashboard

- 各型号评价数量
- 声量统计
- 活跃度
- 健康度
- 关键词(正/负)



购买后评价

- 声量统计
- 正负面指数
- 产品/物流/服务等多维度
- 数据(带评价)

3.3 彩虹盾

平台对接

彩虹盾提供通用接口，与用户系统轻松对接

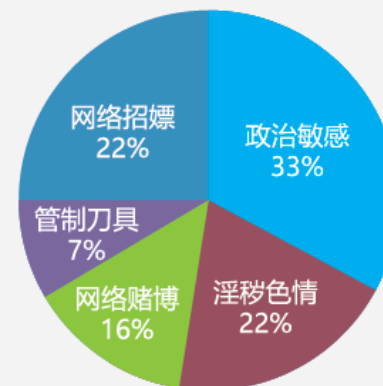
平台监控

添加用户监测平台，UGC内容与彩虹盾词库进行匹配

百万词库过滤

- 12大主类
- 35个副类
- 300多万衍生词
- 10万基础词库

过滤结果
分析



3.3 彩虹盾



3.4 数据运营

网页搜索

- 网页搜索结果排序
- 页面质量标注
- 分词标注
- query意图分析
- 网页搜索对比评测
- 页面结构/内容分类

图片搜索

- 图片相关性评测
- 图片分类
- 图片采集
- 图片质量评测
- 图片搜索结果对比评测
- 图片内容识别（色情）

视频搜索

- 视频内容审核
- 标签分类
- 视频相关性评测
- 视频补码
- 视频评论审核
- 视频信息搜集整理
- 色情视频审核

地图搜索

- POI标注
- 地图信息审核
- 地址词分类
- 地铁/公交线路标注
- 地址词去重

3.4 数据运营

电子商务

- 站内搜索相关性评测
- 产品ID标注
- 产品相关性评测
- 电子商务分词标注
- 导购页面标注
- 商家认证审核
- 商品类目识别标注

产品运营

- 数据整理，报表分析
- 产品优化
- 语料挖掘
- 语料采集
- 数据分析

音频服务

- 音频专写文本
- 音频校验
- 人工语音识别

信息采集

- 词库建立及整理
- 定向信息采集

3.4 数据运营

社区运营

- 社区内容编辑
- 用户审核
- UGC内容审核
- 评论审核
- 热门话题推荐
- 处理用户举报

资讯

- 资讯推荐
- 新闻分类
- 内容抽取
- 文章编辑
- 资讯结构分类

舆情监测

- 舆情信息判断
- 舆情信息挖掘
- 舆情价值分析
- 舆情信息监测

产品体验

- 产品用户意见收集
- 游戏类用户体验
- 不同群体用户体验

4 数字化市场分析



4.1 数字化市场分析

战略性计划实现

Strategic Planning

提供整体分析策略来支持大数据框架驾驶企业营销战略性计划

数据收集

Data Collection

使用先进的云计算，全面收集数据，为数据分析做足准备

数字化报告

Data Reporting

建立有效的可视性可实施计划，为企业营销计划提供可视化报告

数据化分析

Data Analysis

通过深入的数据分析和挖掘，全面分析消费者行为，为企业营销做战略性准备

优化商业模式

Site Optimization

消费者行为目标经验洞察分析，由消费者驱动优化业务商业目标

技术堆栈

Technology Stack

管理和提高分析平台和技术支持大数据的数字和离线营销。

供应商管理

Vendor Management

建立和维护关键供应商关系,不断提供专业知识和咨询服务

4.2 商业智能



数据管理

Data Management

 云迁移和可视化
Cloud Migration & Visualization


 技术支持
Technology Agnostic

 可扩展应用
Scalable



解决方案

Solution

 SharePoint在线
SharePoint Online

 优质BI解决方案
Power Business Intelligence

 云迁移
Cloud Migration



可视化报告

Reporting

 可视化仪表盘
Dashboards & Reporting

 多维数据集
Data Cubes / OLAP

 数据仓库
Data Warehousing & Data Marts



智能分析

Business Analytics

 可操作
Operations

 可统计
Statistical

 数字营销
Sales & Marketing

 金融分析
Finance

 关键绩效指标
Key Performance Indicators

5 智能运维解决方案

业务洞察

通过对多源异构数据的整合和挖掘，消除信息孤岛，使企业对资源状态和分配具有深度洞察，为企业运维管理提供可靠依据。

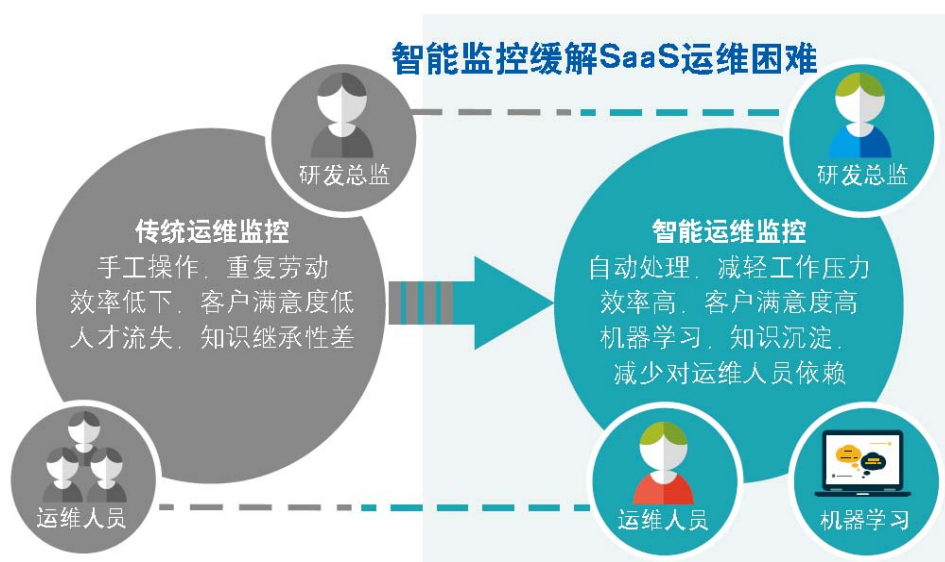
智能检测

通过对异常特征的提取和追踪，大幅提升系统检测的质量和效果；利用机器学习技术自动实现运维知识体系的梳理

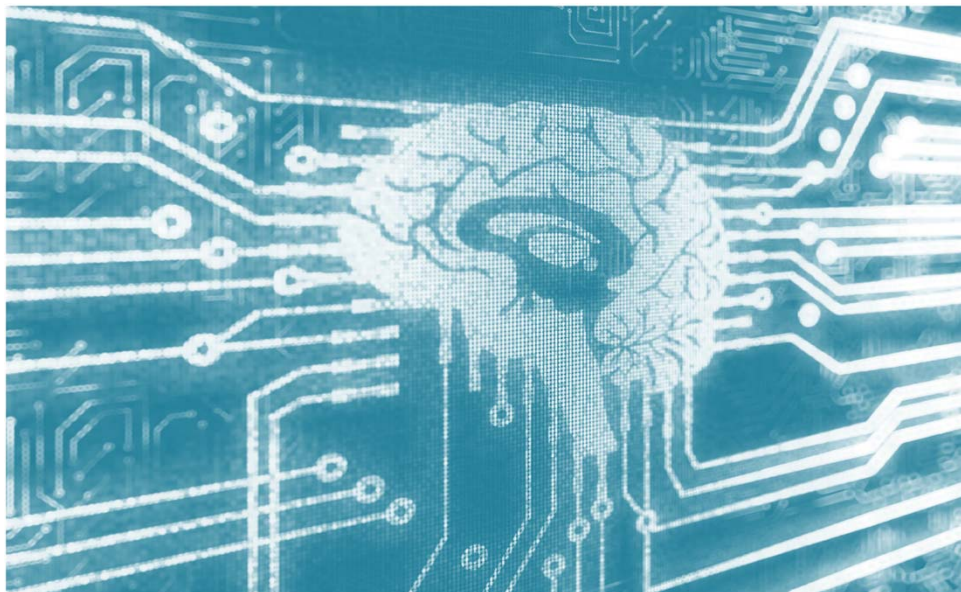
精细运营

通过对系统资源各项数据的采集和分析，为企业运维的优化和资源使用效率的提高提供数据支持；通过对业务特征进行系统分析，为精细化服务运营决策提供保障。

5.1 智能运维解决方案



6 机器学习



- 预测操作服务：
主动预测服务，成为企业真正的核心优势
- 低风险决策：
自动积累历史数据
- 自动调节服务：
增强学习反馈确保持续改进服务质量。



谢谢!

网址: www.beyondsoft.com

联系电话:

中国: 400 000 6282

海外: +1-877-896-5859

邮箱: info@beyondsoft.com



2016年08月

www.beyondsoft.com

© Beyondsoft Corporation All Rights Reserved