

智能化零售的数据引擎

石颖

京东万象资深架构师

京东

全球最佳**金融科技创新者**

将中国社会化物流成本降低**70%**，社会化物流效率提高**2倍**以上

金融子模型**100+**，模型变量**3万+**

中国首家入选**财富500强**的互联网公司

中国电商领域**价值链最长**的数据，**12年数据**积累，日处理能力**15P+**

活跃用户**2.58亿**，用户画像**2亿+**

全球**第一** 低空 通航**无人机** 物流网络

中国线上线下**最大的零售企业**，2016营收**9392亿**

用数据让智能化变成可能

用数据让智能化变成可能

数据



让精准营销成为企业变革最佳路径

数据化精准营销可以帮助企业提升客户服务水平、扩大客户范围，实现决策自动化，为企业带来生产模式和合作模式的质的变革。

技术



让科技成为企业创新的驱动力

人工智能等技术不断发展，带来了新的生产力。科技服务将从商业模式和业务价值链两个方向引领企业创新发展。

整合



让制造变为“智造”

通过数据决定我们要造什么；
通过数据发现生产中的问题；
通过数据监测提高生产效率。



数据和AI支撑智能化零售场景

基于京东在商业和技术上积累的优势能力，
在多终端、多场景、全渠道提供最优零售体验



京东金融



东优势能力
供应链



物流



AI 大数据

智能化零售-入口



AI技术

语音识别、语义识别
家庭智能助手



数据支撑

用户基本信息、地理信息、
生物信息

智能化零售-推荐



推荐应用

基于用户偏好品牌的热卖商品推荐

基于用户最近浏览的的相似商品满减信息推荐

基于秒杀商品选择最感兴趣的

AI技术

传统/DL算法
推荐特征工程
知识图谱

数据支撑

2亿+用户画像、行为链条
千万级商品画像

智能化零售-搜索



用户意图识别



统计筛选热力图



AI技术

自然语言处理：分词、词频统计、自动纠错
语音识别、以图搜图

大数据支撑

实时搜索、智能搜索

智能化零售-浏览



AI技术

智能OCR：关键词提取、全文识别、智能审核

自然语言处理：关键词过滤、文本生成

数据支撑

中文语料库、图片库、外部

智能化零售-下单



1. 用户注册和订单地址填写环节改进

--银行卡识别、身份证识别、智能地址分析

2. 反欺诈、防羊毛党



北京市朝阳区天辰东路国家会议中心 5 层

province: 北京市 | town: 朝阳区
| district: 天辰东路 | building: 国家会议
中心 | more: 5 层

AI技术

智能OCR：文本提取、号码识别、恶劣环境处理
地址分析：地址分词、地址分类、相似度算法、自动补齐

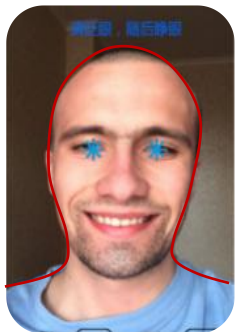
数据支撑

地址数据、POI 画像

智能化零售-支付



1.生物特征识别



2.风控、信贷

AI技术

人脸特征提取
卷积神经网络
异常检测

数据支撑

人脸图像库、自动标记

智能化零售-配送



机器人智能配送

商场



写字楼



园区



AI技术

机器人双目视觉、3D相机
深度学习



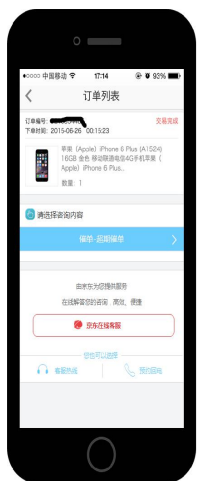
硬件支撑

无人车、无人机

智能化零售-售后



JIMI



电商知识库 社交、舆情

售前咨询
售后服务
商家智能助手
社交媒体
闲聊、百科
专属服务养成



AI技术

自然语言处理：中文语义分析、用户意图识别、文本分类
用户画像、个性化应答

数据支撑

电商知识库、用户画像、外部

智能数据引擎

赋能企业智能化转型

云计算

开放的数据服务
开放的人工智能、大数据、基础云平台
1000+数据产品
400+优质合作伙伴

大数据

人工
智能

驱动集团业务创新

数据总量210PB
日计算数据量 15PB+
日运行job数量 40万+
子模型100+
模型变量3万+
集群规模2万+

万象数据服务-汇聚数据的力量

能力输出

- 京东自有数据能力输出

多维数据

- 海量API、数据包及报告流通服务

数据质保

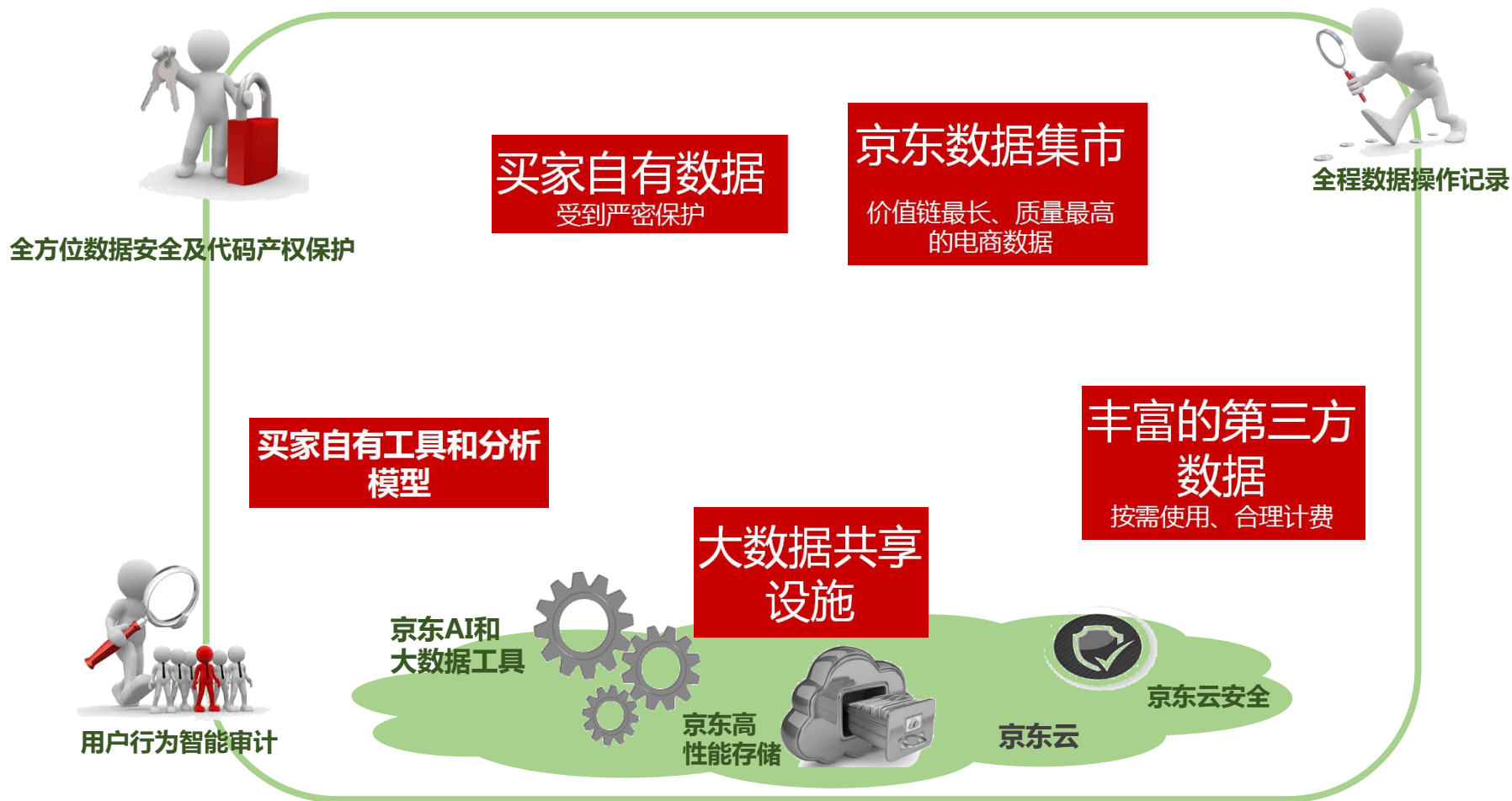
- 区块链溯源+数据DNA



万象数据服务目录

类目	标签
金融征信	身份核验 银行卡核验 反欺诈 黑名单 贷款 蓝鲸企业征信 京东小白分 授权查询 职业信息 出行信息
企业管理	企业工商 税务 司法 知识产权 招投标 资质备案 行政监管 投融资 企业综合 经营管理 人力资源
数字营销	京东数据报告 精准营销 用户画像 电商 舆情监测 商品信息 市场调研 新闻资讯
交通地理	违章查询 车辆信息 位置地图 尾号限行 油价查询 汽车产业数据 公路铁路 市内交通 航空航班
人工智能	证件牌照识别 OCR 人像识别 图片识别 语音语义 情绪判断 语料对话
应用开发	IP地址 语言翻译 网站监测 app应用 条形码查询 格式转换
移动通讯	手机号码验证 手机消费查询 手机在网时长 手机号码状态 手机号码归属 短信API 基站 设备信息
生活服务	天气查询 快递查询 环境质量 自然灾害 万年历 彩票信息 区号查询 邮政编码 星座运势 医疗健康
经济贸易	京东指数 腾讯指数 海关进出口 产业经济 行业趋势 汇率 银行 股票

万象数据共享数据使用权



登月机器学习平台



完美结合大数据能力与人工智能能力，高效训练模型，挖掘数据价值

登月机器学习平台

一站式

- 集数据导入、数据预处理、模型训练、模型评估、模型发布等功能于一体，一站式建模流程。

高效性

- 高性能CPU+GPU，百G Infiniband连接，分布式训练，硬件并行加速。

易用性

- 将复杂的机器学习算法简化为可视化拖拽组件，方便的模型调参，极大降低了使用门槛。

组件化

- 丰富的组件库，包含回归、分类、聚类、SVM、CNN、DNN、CF等多种算法模型组件。

数知全链路大数据服务



自助式的
数据分析



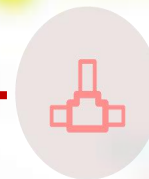
智能化的
机器学习



开箱即用的
计算资源



数据服务的
接口



多样化的
数据接入



开发环境的
集成



数据集成



数据地图



数据探索



作业调度



计算资源



数据模型



BI报表

领先的云计算服务平台

京东全产业链**能力输出**：成熟的解决方案



可信赖的**基础**服务



弹性计算



网络



云安全



存储与CDN



数据库与缓存

一站式**云数据**服务



数据集成



计算建模



数据分析



报表和发布

Thanks

