



百度：让大数据开口说话

文 / 国网能源院研究院 马云高 何琬

当大数据说话时会发生什么？

不少人也许还记得 2014 年春节前中央电视台晚间新闻推出的“据说春运”特别节目，节目首次采用了百度提供的可视化大数据，以充满科技感的酷炫的画面播报国内春节人口迁徙情况。大数据就这样一种接地气的方式走进了普通人的生活：哦，原来你是这样的大数据！

作为国内大数据资源巨头之

一，百度公司一直在探索通过大数据“索引真实世界，连接 3600 行”，让大数据成为商业新能源，为更多企业科学地解决实际问题提供支持。

百度公司总裁张亚勤曾提出打造基于人工智能（AI）、大数据（Big data）和云计算（Cloud computing）“ABC”三位一体的发展战略。在他看来，今后，人工智能会变得像电力一样重要，主导着每个人的数

字生活体验；大数据就像是新能源，能让服务商更了解用户需求，让用户更轻松地获得自己喜欢的商品和服务；云则提供了各种应用和服务运营的基础。

这一切正如维克托·迈尔-舍恩伯格教授在《大数据时代》一书中所言，“能置身于信息流中央并且能收集数据的公司通常会繁荣兴旺，而有效利用大数据需要专业的

◀ 百度 2015 年春运期间中国人口迁徙地图。

技术和丰富的想象力。”

开发众多工具 从大数据挖掘商业价值

随着大数据时代的到来，百度为满足更多企业对大数据营销分析的需求，于 2010 年推出了一款面向内部大客户和营销策划人员的大数据分析工具——百度司南。

百度司南包含了百度舆情、百度代言人、百度精算等多款产品。通过分析百度搜索引擎所覆盖的中国 95% 以上的网民行为，百度司南在消费者洞察、媒体投放规划、市场格局研究等方面，提供了成本更低、效率更高的分析方法。

其中，百度舆情以社会舆论分析思路和方法论为基础，依托百度强大的中文语义分析和网页内容挖掘能力，帮助用户进行传播分析、人群分析和情感提炼等。在传播分析方面，百度舆情通过分析判断舆情传播节点、路径，研究相关事件在互联网上的关注度、关注量级与趋势；人群分析方面，可以为舆情关注者识别人群属性、兴趣爱好等，并列举相关人群的代表性观点和媒体偏好，帮助舆情关注者有针对性地与相关人群沟通；情感提炼方面，可以将网络上的主流评论分为正反两类，帮助舆情关注者了解网民态度，以及对相关事件的关注点与侧重点。

由于百度司南侧重于营销决策分析，尽管在业内叫得响，但离公

众日常生活还有一段距离。百度大数据应用真正进入老百姓的视野，则是 2014 年春运期间百度通过手机用户定位信息数据推出的中国人口迁徙地图。

作为一项创新项目，“百度人口迁徙”不仅可以服务于政府部门帮助其科学决策，还赋予社会学等科学研究以新的观察视角和方法。长期研究定位服务的北京邮电大学教授邓中亮说：“北京市人口现在到底是什么状态，通过以往的方法，我们很难准确研判。但大数据解读就可以分析到人口流动趋势、交通拥堵情况等细节，为城市管理提供支撑。”

大数据的核心不仅在于分析、挖掘，还在于预测。百度充分利用其统计的各年度、季度及月份的数据，开发了中小企业景气指数预测、宏观经济指数预测等模型，前者可实时分析我国中小企业运行发展状况，后者可预测宏观经济发展情况。此外，百度还建立了景点预测、疾病预测、高考预测，甚至一些重要赛事的预测模型等，为政府、企业及用户提供多方面的数据分析服务。

大数据为传统行业 转型升级注入新动能

数据驱动正成为现代企业业务驱动以外的新动力。传统行业发展经过一定周期后，都会面临发展持续动力不足、创新不够等问题。科学合理利用大数据可打通产业链壁

垒，突破传统经营模式、管理模式的瓶颈，为行业转型升级注入新活力、新动能。

正如百度资深大数据专家张琪所说，不少企业时刻被业务将走向何方、如何提高产品竞争力、怎样才能更好地服务客户等问题困扰，但答案往往已经在业务系统中，只需要收集数据进行分析，便可实现业务转型发展。在无法确定因果关系的时候，数据为我们提供了解决问题的新方法，数据中所包含的信息可以帮助我们消除不确定性，而数据之间的相关性在某种程度上可以取代原来的因果关系，帮助我们得到答案，这便是大数据思维的核心之一。

辅助精准营销，是百度大数据在传统行业开展商业化应用的主要方式。依托百度搜索大数据，企业可以通过对海量用户行为进行分析锁定精准客户，进一步明确营销目标，避免造成非重点区域的投放资源浪费，还可以更有针对性地运用互联网思维定向落地人群，积聚更广泛的用户互动资源。借助百度司南等商业应用产品，以及天算等开放式大数据平台，百度帮助在线广告客户实现了产品投放前的预期评估、投放中的精准推送和投放后的数据统计分析等。

百度还以商圈实时客流分析为基础，评估店铺当前竞争力，深度挖掘全网数据，研判网民情感倾向，分析店铺口碑健康度，及时嗅探消



百度大数据和北京朝阳大悦城深度合作，多维刻画和细分消费者群体，融合打通特定消费者群体线下消费行为特征与线上全局行为数据，掌握消费者群体品牌偏好和消费特征，并对其会员进行精准个性化优惠信息短信触达，提升会员到店消费率 11%，其中提升不活跃会员到店消费率 53%，拉动朝阳大悦城 19 天内会员消费总金额增长 12%。



费者反馈并锁定问题实时优化，帮助店铺改善服务质量，降低经营成本；通过获取的实时到店客流情况，分析预测线下客流，帮助店铺提前安排产品仓储、停车排队、引流导购等工作，提升线下到店的消费体验及顾客粘性等。

此外，在金融领域，百度可以根据贷款客户历史信用情况及其信用特征，多维度挖掘用户特征，帮助金融机构综合判断客户的还款能力，有效识别具有欺诈风险的客户，并根据不同客户还款能力给予不同额度的授信，提高金融机构整体风控质量及风险识别效率。

在旅游领域，百度利用区域人群分布热力图、客流量预测及风险预警，可以制定不同客流量下的景区安全运营人力及运力安排方案，辅助景区、政府合理安排人力、运力，降低安全风险，提升景区运营效率及游客游览体验满意度。

布局智能平台生态系统 引领大数据时代潮流

目前，人工智能和大数据已成为创新制高点，开源正在引领大数据和人工智能，云端数据仓库正在升温，传统批量计算已无法满足实时要求。基于这样的趋势，百度提出了云计算、大数据和人工智能三位一体的发展战略。

百度近期正式对外发布了其人工智能平台天智，这也是百度继推出智能大数据平台天算、智能多媒体平台天像和智能物联网平台天工后，发布的第四大平台级解决方案。

“借助天智平台，企业和开发者可以获得百度的人工智能技术能力，实现业务创新、提升用户体验等。百度天智将保持开放，推动各个行业转型，进入 ABC 时代。”百度官方称。

依托百度智能平台服务，百度与中国海事局、太原铁路局、南方

航空公司、福田汽车达成深度合作，成立了覆盖陆海空车的智能交通生态联盟。百度将实时抓取散落在各个路面交通、地下交通、空中航线的海量数据，有效归类、提取、利用，打破海陆空以及行政区域的限制，最终实现多系统的配合协调。百度距离打造在技术、产业和应用等领域最有影响力的大数据、云计算、人工智能平台生态系统的目标又更近了一步。

传承开放共享的互联网精神内核，百度将大数据存储、分析和智能化处理等一整套核心能力通过平台化、接口化的方式开放，并借助智能平台生态系统开展大数据的大规模商业化应用。合作伙伴和第三方将在线使用这一系统，通过大数据的挖掘处理来改造和优化企业管理、产品服务设计、市场拓展、商业模式等环节。

对百度而言，开放大数据分析处理平台，采用统一、标准的接口，导入传统企业的信息数据和需求，将有利于百度自身的数据交叉、积累及数据工具的验证分析，进一步提高其数据分析处理的精确性。

对传统行业而言，从百度大数据引擎中获得一流的数据处理和分析能力，等于采用了技术升级手段，实现了企业运营技术水平、流程管理水平以及市场预测、分析能力的全面提升。

对数据生态而言，目前大部分企业正面临大数据应用困境，不仅

数据孤岛现象严重，数据存储与管理的规模、数据分析挖掘以及智能化能力也还存在瓶颈，处在从数据积累到数据智能应用转变的临界点

上。而百度拥有完整、领先的大数据技术，通过开放大数据接口获得流量入口，既进一步强化了自身的数据优势，也在一定程度上增强了

用户黏性，更重要的是通过技术化的平台获得了大数据迁徙、运动的通道，实现了对大数据生态模式的促进。

传统企业能从中获得哪些启示？

基于丰富的数据资源库，百度将大数据与用户需求科学匹配，充分发掘了数据资源的商业价值。传统行业企业，尤其是具备大数据资源的企业能从中获得哪些启示？

厘清大数据应用思路或模式。

企业首先要想清楚究竟用大数据做什么，而不应一窝蜂地为了用大数据而用大数据，然后才能有针对性地考虑怎么用好大数据。百度大数据开发应用思路是以行业关键诉求为牵引，以突出应用为导向，前瞻性挖掘用户千万级标签数据，打造开放共享的大数据平台生态系统，形成合作共赢的大数据商业化应用机制。对传统企业而言，如电网企业，应加快推进数据资产战略布局，推动数据与企业业务、管理、服务、技术的深度融合，从而达到提升电网安全水平、强化精益管理、创新增值业务、拓展市场空间、增大品牌价值的成效，最终实现数据治企与数据强企的目标。

做好大数据应用的顶层设计和基础管理。非结构化数据比例的增加使得大数据的数据结构越来越复杂。对于传统企业来说，面对数量庞大的历史数据、爆炸式增长的实时数据，如何做好数据存储、分析、维护以及风险管控等工作，都是不小的挑战，需要按照统一规划、统一设计的原则加强数据基础管理，同时积极与具有技术、资源、市场优势的先进数据管理企业开展合作，借鉴其先进理念和经验，实现强强联合，共同挖掘企业数据价值。

整合资源建设统一的大数据服务平台。百度充分发挥其人力、技术和数据资源等各方面优势，通过嫁接整合数据和商业资源，从数据背后挖掘盈利模式，建立了开放的大数据服务平台，培育了具有平台特征的应用场景，为各个机构、企业提供了多元化服务，提高了企业的品牌影响力和在平台生态系统中的控制力、带动力。对于传统企业

而言，由于数据一般都是按照业务分散管理、多头存放，更需要统一规范不同来源、不同类型、不同维度的数据，从基础数据层面解决各业务数据资源分散、接口不统一、标准不一致的问题，为提高企业数据使用效率、提升大数据应用服务水平奠定基础。

面向不同需求开发不同的工具。百度科学匹配大数据与用户需求，针对不同用户开发了个性化的大数据产品，并注重大数据产品商业化应用过程中的可视性效果，为提升用户体验、改善服务质量发挥了积极作用。对于传统企业而言，可以通过企业上下互动相结合的方式，一方面加强顶层引导，确保大数据开发应用符合企业管控方向；另一方面，鼓励企业各部门立足自身实际生产需求，重视用户体验，从用户需求出发建立数据开发工具库，为数据开发利用提供多类型、多用途的模型和工具。