

大数据“润物无声”大产业

本报记者 陈静

来自市场研究机构IDC的数据显示, 2011年全球数据总量已经达到1.8ZB, 这个数值还在以每两年翻一番的速度增长, 预计到2020年全球将总共拥有35ZB的数据量, 增长近20倍。而每个ZB的数据, 都相当于装满10亿个1TB的移动硬盘。

镜头一
清华—人人社会化媒体研究中心在今年高考刚结束时发布了“人人指数”, 其中不光有报考的热门区域和专业, 甚至还有高校生活费用和毕业薪酬统计; 在校大学生平均每月花销1152元, 校内每餐平均成本7.5元, 计算机专业毕业生年薪起薪超11万, 居各行业首位。这统计结果正来自大数据。

镜头二
宝洁曾经统计过玉兰油品牌在百度上的搜索结果, 发现与竞争对手相比, 消费者在“玉兰油”关键词后加“适合几岁”来搜索的比例明显偏高, 这意味着玉兰油的年龄定位比较模糊, 于是宝洁及时调整了营销策略, 还顺势推出了一款针对25岁人群的细分产品。这样的决策也来自大数据。

镜头三
自动售货机运营商友宝在广东收购了一家自动售货机公司, 销售额马上就提升了50%。原因就在于新机器连上了互联网, 可以统计出每分钟的热销产品。而传统自动售货机却是2天补货一次, 根本不会知道到底是哪些商品, 第二天就卖空了。这样的变化还是来自大数据。

数据显示, 从现在起直到2018年, 全球大数据市场的年均复合增长率将达到26%, 从148.7亿增至463.4亿美元。中国工程院院士邬贺铨也表示, 继云计算后, 大数据成为信息技术领域最为热门的概念之一。在工信部发布的物联网“十二五”规划中, 作为四项关键技术创新工程之一的信息处理技术中包括海量数据存储、数据挖掘、图像视频智能分析, 这些都是大数据的重要组成部分。

然而, 大数据究竟是怎么回事? 在现实场景中又有哪些应用? 这个“时髦”的新产业, 将给我们的生活和工作带来哪些巨大的变化?

取之不尽的“数据金矿”

什么是大数据? 《大数据时代》的作者之一舍恩伯格将之定义为“不用抽样调查的捷径, 而采用所有数据的分析方法”, 简单来说, 就是利用互联网和计算机对海量数据进行抓取、管理和处理, 并且从中提取出有意义的讯息。

大数据有什么用? 一方面大数据技术可以扎扎实实地用数据说话, 打破直觉和思维定式。舍恩伯格曾经举例说, 人们常常会说冬天要记得戴帽子和手套, 避免着凉感冒, 但大数据会告诉你, 感冒来自于病毒, 跟戴帽子手套没有直接关系。二是大数据技术可以让人们对未来进行“预测”, 通过对所有可能的数据进行分析, 而找到数据之间的关联性, 并由此预测出趋势。

日本SS制药公司运营的网站Kazemiru正是大数据技术应用的一个生动写照。这个网站通过对数千万推特信息的抓取, 找到日本各地可能正在感冒的人, 并通过与昨天的数据进行对比, 并结合气温和湿度变化来预测感冒在不同地区的流行情况, 为每个登录该网站的用户提供“感冒日历”, 通过与日本国立感染症研究所对全国5000个医疗机构的感冒患者的统计数据对比, Kazemiru的预测与实际情况基本一致。

“真金白银”领跑行业应用

去年11月份, 北京公务员胡晓璐摇到了车牌号, 在网上四处查询和对比各种汽车型号之后, 她惊讶地发现, 自己看到的汽车网络广告变多了。

胡晓璐所不知道的是, 在她点开某个网页的瞬间, 一场“拍卖”正在悄无声息地进行着。广告交易平台迅速地向各大需求方平台发出“这儿有个广告位”的信息, 需求方平台则根据她之前的浏览行为判断出“这个人可能要买车”, 然后迅速出价, 谁出价最高, 胡晓璐就会在点开网页上看到谁推送的广告, 一切都在0.1秒内完成。

这就是RTB(实时竞价)技术, 利用大数据, 它可以根据每个用户的具体偏好来精准地投放互联网广告。和这项技术一样, 与营销、电子商务、互联网金融等等“真金白银”相关的大数据技术最早被投入到对商业模式的变革之中。

1号店负责生鲜业务的副总裁郭冬东告诉本报记者, 通过对海量数据的分析、判断和整合能力, 1号店可以精确地预测第二天生鲜产品的订货量, 然后向供货商下订单, 产品第二天早上8点进货。“综合折算下来, 仅在仓储环节, 损耗率可以控制在个位数, 而传统线下渠道的损耗

智慧城市释放数据“能量”

对于广东佛山的“12345”市民热线来说, 每一通来电都是宝贵的“数据财富”, 不光是话务量分析, 通过大数据技术, 这些音频中的核心数据被提取出来, 并被统一组织、存储、管理和应用, 市民们集中的诉求会被直接向各有关部门反映, 成为政府决策分析的数据支撑。

尽管目前大数据相关的商业模式还集中于企业应用, 但与电子政务相适应的服务则被不少人当做大数据技术未来的又一重要的应用场景, 大数据也因此被视为智慧城市的“智慧引擎”。邬贺铨告诉记者, 大数据将遍布智慧城市的方方面面, 是智慧城市的智慧之源。大数据能够使政府的决策与服务、人们的生活方式、城市的产业布局和规划及城市的运营与管理方式, 实现“智慧化”或“智能化”。

然而, 将这部引擎发动起来, 却首先需要数据的开放与整合。国务院发展研究中心技术经济研究部副调研员李广乾表示, 数据开放为大数据应用提供创新“源头”。大量数据的可获得是大数据价值实现和最大化的前提。

大数据技术在今时今日的“井喷”也事出有因。大数据, 首先要有足够多的数据产生。随着互联网的高速发展和成熟, 互联网产生内

容和数据的能力不断增强, 百度技术委员会理事长陈尚义曾透露, 百度每天处理的数据量接近100PB, 而1PB相当于5000个国家图书馆的信息量总和, 此外在每分钟内, 新浪和腾讯微博发送出数万条微博, 淘宝卖出了近10万件商品……来自市场研究机构IDC的

数据显示, 2011年全球数据总量已经达到1.8ZB, 这个数值还在以每两年翻一番的速度增长, 预计到2020年全球将总共拥有35ZB的数据量, 增长近20倍。而每个ZB的数据, 都相当于装满10亿个1TB的移动硬盘。

其次, 还要有储存和分析这些数据的能力。从硬件设备来看, 计算机性价比的提高, 磁盘价格的下降, 大幅降低了大数据存储和处理的门槛, 数据显示, 2000年的硬盘驱动器平均每GB容量的单价约为16美元到19美元, 而现在却只有7美分。在重量上, 1982年世界第一台超GB存储器容量为1.2GB, 重113千克, 而现在, 32GB的微型SD卡仅有0.5克。从处理能力来看, 云计算成为最重要的推手, 数据被搬到了“云”上之后, 更容易被收集和获得, 而云计算强大的处理能力也使更多行业和企业有机会对数据进行整理和分析。

率一般是30%左右。”借助大数据, 电子商务平台得以进军传统零售业的“腹地”。

数据甚至成为网络卖家和中小企业的“无形资产”。阿里巴巴旗下的两家小额贷款公司, 截至2012年6月底, 已为超过12.9万家小微企业和个人创业者提供贷款, 贷款总额累计超过260亿元。他们能拿到贷款的关键之一, 正是此前网络交易的诚信记录。全国农信银资金清算中心运行保障部总经理王永刚对此感慨万千: “阿里巴巴的小额贷款几小时就能发放, 而我们最快也要7天。他们靠的就是开放平台的渠道优势, 和大数据的分析和处理能力, 从而大大降低了信贷的成本。”

不过这些都还是基于自己平台内部产生的数据, 如今互联网巨头们还通过频繁的并购与合作, 将外部数据纳入大数据处理的范畴。比如: 阿里巴巴在收购新浪微博18%的股权后, 宣布打通淘宝与新浪账号, 让微博产生的隐形消费需求 and 淘宝的显性消费数据得以整合; 百度将19亿美元砸向91手机助手, 从而获得了1.27亿手机用户的使用习惯和相关数据; 腾讯则通过微信公众平台直接向线下企业提供大数据技术下的客户管理增值服务。用户在哪里? 他们喜欢什么? 想要什么? 互联网巨头们正在拼凑一份涵盖方方面面生活场景的数据图。

在美国, 2009年奥巴马入主白宫后, 所做的第一件事就是要求联邦政府各部门通过“一站式”政府数据下载网站向社会公开各类非保密的数据库。如今在这个网站上, 有超过40万种各类原始数据文件, 涵盖了农业、气象、金融、就业、人口等近50个门类。

在我国, 各地也在探索数据开放的机制和方式。作为“大数据”惠民的一项重要探索, 北京市政务数据资源网预计今年年底之前正式开通, 为政府信息资源的社会化开发利用提供数据支撑。北京市经济和信息化委员会副主任董腾飞表示: “政府部门大量信息资源, 如果能开放给社会利用, 会创造更多价值。”目前, 北京市已有29个部门公布了400余个数据包, 点击量最高的“土地用途分区”已被下载684次。在上海、广东、浙江、重庆等地, 不同形式的数据开放也在探索之中。

随着数据公开, 大数据使用的监管也被提上日程, 有关专家建议, 应该尽快出台大数据相关的顶层设计, 划清保密信息、商业公开信息和免费公开信息、义务公开信息之间的界限, 并且明确信息公开和数据使用的责任与义务。

立足国内市场 推进分布式光伏发电应用

本报记者 齐慧

近年来, 在全球光伏市场需求增速减缓、国际贸易保护主义抬头、光伏产业发展不协调等多重因素作用下, 光伏产品出口受阻。而我国光伏产品90%以上依赖出口, 在欧美“双反”和国际市场需求增速放缓后, 产能面临过剩, 企业面临生存压力, 迫切需要国内市场支撑。

“推进分布式光伏发电应用是拓展国内光伏市场的有效途径。”国家能源局局长吴新雄近日表示。

按照每年新增光伏应用规模1000万千瓦测算, 40%以上的光伏产品可在国内市场消化, 能有效抵御国际市场变化, 特别是美欧“双反”带来的风险。

如何能提高这40%以上的国内市场? 今年7月, 国务院正式印发《关于促进光伏产业健康发展的若干意见》。其中, 大力推进分布式光伏发电应用是其中重要内容之一。

据了解, 我国西部地区光照条件好, 未利用土地辽阔, 适宜发展集中式大型光伏电站, 但是度电补贴需求高, 约0.6元左右, 且当地用电需求小, 大规模开发就地消纳困难, 电力需长距离外送, 变损、线损高。

然而, 在东、中部地区发展分布式光伏发电, 虽然平均利用小时数稍低, 但电力易于就地消纳, 且网购电价高, 度电补贴需求少, 约0.4至0.45元, 与西部集中式光伏电站相比, 用同样的补贴资金能够多支持30%至50%的光伏发电。如果政策措施得当, 有效推动产业技术进步、企业成本下降, 还可逐步减少补贴。

专家表示, 坚持集中式与分布式并举, 大力推进分布式光伏发电应用, 将能更为有效地拓展国内光伏市场, 缓解光伏制造企业面临的困难。

据测算, 全国建筑物可安装光伏发电约3亿千瓦, 仅省级以上工业园区就可安装8000万千瓦, 潜力巨大。

吴新雄指出, 《关于促进光伏产业健康发展的若干意见》明确, 2013至2015年, 年均新增光伏发电装机容量1000万千瓦左右, 到2015年总装机容量达到3500万千瓦以上。初步考虑, 2015年分布式光伏发电装机容量达到2000万千瓦。

为完成规划, 首先要抓紧组织推进示范区建设。示范区是分布式光伏发电应用的试点和创新区, 也是当前工作的重点。8月初, 国家能源局印发了《关于开展分布式光伏发电应用示范区建设的通知》, 确定了北京海淀区中关村海淀园等18个工业园区, 作为分布式光伏发电应用第一批示范区。这批次示范区到2015年总规模将达到182万千瓦, 2013年建设规模75万千瓦。

“2014年全国拟新增分布式光伏发电600万千瓦, 国家能源局将给各省(区、市)下达分布式光伏发电的指导性规模指标。初步考虑, 2014年重点在用价格水平较高、电力负荷较大、控制能源消费总量任务较重的长三角、珠三角、京津冀及周边地区建设分布式光伏应用, 上述地区新增规模约占全国分布式光伏总规模的80%以上。”吴新雄说。

龙源电力中标南非风电项目

本报讯 记者李予阳报道: 近日, 龙源电力集团股份有限公司在南非能源部组织的第三轮可再生能源项目招标中, 成功中标两个风电项目, 总容量为24.4万千瓦。这是继加拿大10万千瓦风电项目后, 龙源电力拓展海外风电市场取得的又一重大突破。

2011年8月, 南非能源部启动了可再生能源招标工作。前两轮风电项目总容量约为350万千瓦, 最终中标项目容量为119.65万千瓦。第三轮风电招标参加投标的企业除龙源电力外, 还有法电、意电、葡电和Mainstream等多家国际知名风电开发商和投资者。第三轮风电计划招标容量为65.34万千瓦, 最终中标项目容量调整为78.7万千瓦。

在国电集团的大力支持和指导下, 龙源电力携手国电联合动力, 共同实现了南非风电项目的成功中标, 积极落实“走出去”战略。这对于加快龙源电力国际化战略步伐, 促进国产风机出口, 以及加强中两国可再生能源技术合作与交流等方面都具有重要意义。

360发布儿童卫士手环

本报讯 记者黄鑫报道: 360公司日前推出可穿戴设备“360儿童卫士”手环, 这款电子手环能随时定位, 出安全区域预警, 并能建立通话连接, 将全方位保障儿童安全。

据360公司副总裁沈海寅介绍, 这款手环内部装置了专业级的GPS芯片, 有随时定位的功能。家长在智能手机上安装一个软件, 只要通过手机向手环发送一个命令, 就能知道孩子的位置, 并能查找孩子某段时间内的行动轨迹。手环的基站辅助定位系统增大了其接受信号的能力, 室内等封闭环境也不会对定位造成影响。同时, 系统能够自动标识孩子佩戴手环期间去的安全位置, 一旦孩子的位置超出这些安全区域, 系统就会自动向家长的手机发出报警, 以提示孩子可能正面临危险。

善用大数据

安之

美国奈飞公司利用大数据技术拍摄《纸牌屋》的案例几乎已成为行业中最著名的励志故事。作为全美最大的DVD和网络视频租赁网站, 奈飞通过数据分析技术判断出观众喜欢导演大卫·芬奇和



演员凯文·史派西, 甚至通过他们按下暂停截图的时间点, 来判断观众更喜欢哪些布景和画面。

然而, 《纸牌屋》的拍摄方法也引发了人们对大数据技术的思考。如果仅仅按照大众对某种文化产品的兴趣和取向来制作商品, 也许在市场上更容易取得成功, 但另一些更有创意的产品可能一开始就会被抛弃, 因为它们尚不为人所知。有人将之精确地概括为“大数据偏爱潮流, 但忽视杰作”。

大数据技术中的能量让人目眩, 但它也并非万能灵药, 过分迷信大数据, 则很可能过犹不及。

比如数据噪音问题, 被分析的数据越多, 它们在统计上能够呈现出的相关关系也必然越多, 但在这些相关关系中, 许多并没有实际意义, 很有可能会将人们引入歧途, 这种欺骗性有可能使真正重要的线索被隐藏在数据扩张的噪音之中。

此外, 大数据的使用还有个边界问题。对大数据的迷恋还可能带来对“数字隐私”的侵犯。毫无疑问, 可以获得的个人数据越多, 其中蕴含的信息量就越大, 从这个角度出发, 一方面要从技术层面来保障大数据时代的隐私安全, 完善用户保障体系; 而在另一方面, 对大数据采集的监管必须提上日程, 个人信息保护的相关法律条文还应在操作层面上进一步细化, 在确保用户知情权的前提下避免个人数据的滥用和非法售卖。

大数据是把“双刃剑”, 《大数据时代》的作者之一舍恩伯格最近表示, 人们必须以一种不仅欣赏其力量, 而且承认其局限的态度来接纳大数据技术。使用者对大数据技术的态度和认识程度, 将直接决定这一技术对未来的影响方向。



10月31日, 一名工作人员在第九届北京国际金融博览会上向参观者介绍自动售货手机移动支付。

当日, 第九届北京国际金融博览会在北京展览馆开幕。本届金博会为期四天, 以“金融服务实体经济、金融服务市民百姓”为主题, 20余家商业银行、50余家金融机构及诸多互联网金融企业参展, 期间举行了高峰论坛及“金融惠民”系列活动, 并首次设置了“互联网金融”展区。
新华社发