



大数据时代来了 你还在等什么?

■ 本版撰文 本报记者 杨颖 霍玉茵 徐森

编者按:早在1980年,著名未来学家阿尔文·托夫勒便在《第三次浪潮》一书中,将大数据称为“第三次浪潮的华彩乐章”。不过,大约从2009年开始,大数据才成为互联网信息技术行业的流行词汇。

如今,大数据已经成为时下最火热的词汇。悄无声息地,它已经来到我们身边。当你上搜索网站的时候,一些搜索引擎公司会记录下你的搜索习惯,并根据这些信息自动向你推荐相关广告;当你在电子商务网站购物时,一些电商公司会依据你的购物喜好,自动向你推荐有关产品;当你驾车上路时,电子导航系统会自动定位并向你指示方向……当一切皆可“量化”时,大数据潮流。

时候,一些搜索引擎公司会记录下你的搜索习惯,并根据这些信息自动向你推荐相关广告;当你在电子商务网站购物时,一些电商公司会依据你的购物喜好,自动向你推荐有关产品;当你驾车上路时,电子导航系统会自动定位并向你指示方向……当一切皆可“量化”时,大数据潮流。

开启商业变革

时下,维克托·迈尔·舍恩伯格的《大数据时代》被国人奉为经典。书中提到:大数据开启了一次重大的时代转型。

是的,大数据浪潮与互联网的发明一样,绝不仅仅是信息技术领域的革命,更是在全球范围启动透明政府、加速企业创新、引领社会变革的利器。现代管理学之父德鲁克有言,预测未来最好的方法,就是去创造未来。而“大数据战略”,则是当下领航全球的先机。

如今,大数据技术创造价值的已经显现出来。国际数据公司IDC报告显示,2012年至2016年,全球大数据技术及服务市场复合年增长率(CAGR)将达到31.7%,2016年收入将达238亿美元,其增速约为信息通信技术市场整体增速的7倍。

大数据之所以能创造巨大的价值是因为在商业、经济、政府管理等领域中,决策行为越来越取决于数据和信息,而不再是经验和直觉。一份行业报告显示,英国政府通过高效使用公共大数据技术,每年可节省约330亿英镑,相当于英国每人每年节省约500英镑。

“大数据会让我们的生产、销售、营销更加便捷。大数据分析强调模式性而不是精准性,所以它进行的是相关因素的分析,凸显的是一种趋势。比如,下雨天的时候,不仅雨伞畅销,面包的销量也会上涨,因为大家因为天气原因不再有出门购物的计划,所以要提前购买生活必需品。这就是通过数据分析,让所有的营销工作简单化,并且能够让商家与库存挂钩。”季清明告诉记者,“当有事情发生的时候,所有的营销模式以及货品的准备都会随之变化,这就是大数据在营销中的作用。”

当然,大数据往往能够带给我们意想不到的收获。

“虽然羽·泉是《我是歌手》最后的赢家,但从来自百度贴吧、搜索、新闻等平台的数据分析结果来看,黄绮珊

则是最大的赢家,她的关注度提升了321倍,而在歌曲方面,最大的赢家则是林志炫,其70%的歌曲被关注和热搜。”百度全国营销服务总经理陈潇这样介绍。在他看来,大数据可以告诉我们更有趣的故事,让我们看到更多的事情。在不远的将来,数据挖掘将会成为产品生产和市场营销等各个环节最重要的基础工作。

落地中国

据了解,大数据的应用分3个阶段来实施:数据的整合、数据的挖掘,以及把大数据变成可操作的行动,从而指导决策。由于大数据的庞大复杂程度超过传统数据库系统处理能力,因此需要特定的投资、设计、管理、维护的产业来解决,这就给了大数据产业以广阔的市场空间。IDC报告称,随着大数据技术和服务的不断升级,仅存储市场而言,2011年至2016年间将实现53%的复合增长率。诸多的调研数据都在显示,大数据将成为下一个金矿。

在国内,“大数据”作为一个较新的概念,目前尚未直接以专有名词被中国政府提出来给予政策支持。不过,在工信部此前发布的物联网“十二五”规划中,信息处理技术作为4项关键技术创新工程之一被提了出来,其中包括海量数据存储、数据挖掘、图像视频智能分析,这都是大数据的重要组成部分。而另外3项关键技术创新工程,包括信息感知技术、信息传输技术、信息安全技术,也都与“大数据”密切相关。

易观国际高级分析师黄萌对记者表示,中国大数据行业的发展比国外媒体报道得乐观,虽然有些人认为很多项目还没有落地,但据易观国际了解,中国很多大数据项目的进展还是不错的。“中国的大数据行业与国外的差距并不大,可以说是在同一起跑线上,尤其是做垂直行业数据的项目或者企业,还是很有发展前景的,有可能在将来发展成类似于现在的互联网平台这种量级的机构。”他说。

大数据的天花板

随着云时代的来临,大数据应用呈现爆发性增长,不仅衍生出了自己独特的架构,也直接推动了存储、网络以及计算技术的发展。

目前,数据几乎已经渗透到了每个行业及业务职能领域,逐渐成为重要的生产因素。不过,大数据的实现需要从海量数据中“提纯”出有用的信息,而这其中也催生了很多新问题。从事多年数据分析工作的闫先生告诉《中国贸易报》记者,现在,互联网的变化模式更新很快,虽然有能力强积累海量数据,但这其中的很多涉及隐私,而且数据整合的难度也比较大。这些都制约着大数据的发展。

伴随着各种随身设备、物联网和云计算云存储等技术的发展,人和物的所有轨迹都可以被记录。面对大数据,我们毫无隐私可言。以移动运营商为例,他们所掌握的用户信息是十分精确的,不仅包括用户的身份信息、手机号码等等,甚至连用户所处的位置、通话的状态等等都能够轻松获取,而这些数据信息对于每一个人来说都是敏感的。这些数据信息的拓展和使用如果得不到规范,后果很难想象。

更令人感到恐惧的是,很多归类分析后的数据正在被做成形形色色的商品,待价而沽;而越来越多的数据也正作为企业的战略资产被长期

保存下来。在大数据与资本营造的所谓“监察经济”(surveillance economy)之下,如何合法、智慧、有效地实现这些数据的商业智能化应用,商家如何界定和规避用户隐私,如何解决用户数据安全等问题非常关键。

好在,各界已经开始重视相关问题,我们也欣喜地看到,一些通信运营企业已经开始提高用户数据的隐私保护级别,严格用户信息保护工的规章制度和操作流程、提高客户资料管理系统的安全保护能力,希望为广大用户提供更安全的应用环境。但高度个人化的大数据的商业价值注定它极易成为黑客或泄密者觊觎的重要目标。这些风险将继续伴随着大数据的发展,不容小觑。

专家提示,在商业利益面前,我们不能指望商家自觉遵守规范,还是要通过立法来解决相关问题,以做到既鼓励信息挖掘,又有效防止权益侵害;既强调信息公开,又明确保护个人隐私;既提倡信息共享,又严密维护信息安全(包括滥用、篡改、损坏、窃取、失泄密等)。只有做到有法可依、依法办事,才能适应大数据时代的到来,并利用大数据服务经济发展,引导经济升级。

就事论事

数据浪潮来袭,企业准备好了吗?

“数据是新的石油”——现在,很多企业已经达成了共识,数据已经成为了推动企业发展的原动力。牛津大学与IBM在今年年初发布的统计数据显示,28%的全球企业已经开始进行大数据实践。而在中国,四分之一的企业也正在积极投入大数据业务,中国亦成为极具代表性的大数据实践市场。大数据大踏步的发展,在于多数人看到了它的价值。而大数据价值的体现,主要在于大数据本身的特性。

海量化是大数据最明显的特征。在信息爆炸的时代,信息的增长量超过了任何一个行业。IDC此前发布的Digital Universe Study报告显示,全球信息总量每过两年,就会增长一倍,到2020年全球被创建和被复制的数据总量将增长到35ZB,这个速度正在超越摩尔定律。而这就给了企业一个巨大的信息池,发展到今天,再没有一种方式比大数据提供的机会更多。当然,海量化的数据也奠定了多样化的基础,大数据能为企业提供种类繁多的信息。如今,企业面临的竞争环境越来越严峻,如果能够及时把握市场动态,并能快速制定出适当的运营策略,成为企业提高竞争力的关键。而反应在数据层面,则是数据的实时洞察,这也就是大数据的快速化。有了以上这些特征,大数据还并不能成气候,还需要大数据进行分析和处理,这些处理结果将给企业带来巨大的经济价值,而价值才是大数据的终极意义所在。

华尔街“德温特资本市场”公司在分析了全球3.4亿微博账户留言后作出如下判断:人们高兴的时候会买股票,而焦虑的时候会抛售股票。该公司依此决定公司股票的买入或卖出,获得7%的收益率。

而马云在卸任阿里巴巴CEO后,开始打造智能物流网,他所借助的也是大数据支撑下的订单、物流数据以及在此数据基础上形成的产品。依靠阿里巴巴多年的数据积累,马云可根据数据选择在何地建设仓储物流中心。至于哪些货品放在哪个仓库、每个仓库放多少、使用哪个物流公司,这些都是由阿里巴巴的数据计算后提供给商户。

大数据就是这样的神奇、迅速、精准,以至于大家一涌而至,对于大数据的炒作已经到达了高峰。但需要认识到的是,并非所有企业都适合大数据。

尽管大数据的作用相当了得,但是真正具备分析能力的分析师却凤毛麟角。如果企业没有这样的人才,那么在海量的数据面前,也只能是一筹莫展。另外,在成本与收益的考虑下,企业更应精打细算,因为大数据收费较高。若应用大数据,企业要选择成熟的大数据平台,且要和数据仓库有高性能的连接,这样便于让不同人员应用,根据业务需求让技术人员利用平台去快速提升数据的价值。这是企业需要资金、技术投入的“大头”,也是企业值得做功课的地方。

不能被小看的职业:数据分析师

曾经有一个故事广为流传,告知大众远离数据分析师,因为他们实在太无孔不入了。

故事很有意思,做人力资源管理的默默负责招聘数据分析师,在面试要进入最后一轮面试的人时,她只说了一句话:“我叫默默,有什么事,可以问我。”当时,她还不知道,就是这一句话,让她对数据分析师有了新的认识。

几天过后,最后一轮面试开始了。一个看起来斯文的男生拿着一叠材料,非常认真地介绍自己如何适合数据分析师这个职位。他交上来的材料,让默默大吃一惊。封面上赫然写着“默默女士的2.0洞察报告”。报告不厚,只有十几页,但关于默默的兴趣爱好、常去的地方、最关注的人(经常@的人)、亲密圈子(互相关注的博主)、经常谈论的话题以及网络口头禅等等,在报告中被一一用数据、图表展现无遗。

默默呆了几秒钟,内心翻江倒海。因为,这份报告描绘出的她,既熟悉又陌生。“难道我最爱吃的是麻辣香锅,3个月内提及这个词10次?难道我如此渴望得到某个人的回应?3个月内@了他12次?难道,一个素未谋面的小伙子,比自己更加了解自己?”默默陷入深思。

这就是一个好的数据分析师的能量值——比你更了解自己。

很多时候,许多隐藏在电脑背后的庞大数据就像一个麻烦的大线团,只需要找到那个线头一拎,就能再现一个个的网络人格,一个个连你都不知道的自己。而数据分析师就是找到那个线头,再现网络人格的那个人。而企业只要借此了解用户所需,有针对性地推送广告或其他服务,那么这是无往不利的。

北京品真艺力广告有限公司副总裁季清明告诉记者,现在大数据的技术支持不是难题,行业中最缺的恰恰是数据分析师。因为,即便有海量的数据,找不到线头、无法穿针引线,那么数据也只能一无是处。好的数据分析师需要具备不同的分析模式,能够掌握公众的思维变化。

有多年数据分析工作经验的闫先生也认为,大数据发展使其对人才的需求十分迫切。“现在,带数据分析师头衔的人很多,但多是业余选手,职业选手很少”。他认为,作为数据分析师,需要3点能力:一是业务能力,能够理解业务模式,有悟性;二是信息技术能力,能够操纵数据,能够编程、写脚本等;三是分析能力,须掌握统计学等方面知识。

目前,不少企业都已认识到若想打破大数据人才的僵局,高端大数据人才补充刻不容缓。这对于毕业后难找工作的大学生来说,可谓一条更好的选择。对于大数据职位的基础岗位,通常应届研究生毕业即可胜任,月薪在1万元左右,但是前提是对HDFS、Hadoop、Mapreduce、Hbase、Database、NoSQL等技术有扎实的基础,能完成团队领导规定的要求。而高端岗位具象化为大数据系统架构师、大数据应用研发总监及数据科学家,年薪一般在20万元至50万元,数据科学家的薪资甚至可达100万元以上。