

第3章

大数据应用价值

大数据正在催生以数据资产为核心的多种商业模式,产生巨大的应用价值。数据的生成、分析、存储、分享、检索、消费构成了大数据的生态系统,每一个环节产生了不同的需求,新的需求又驱动技术创新和方法创新,通过大数据技术融合社会应用,让数据参与决策,发掘找到大数据真正有效的价值,进而改变人们未来,革新生活模式,产生社会变化,引发积极影响。近年,伴随着物联网膨胀,移动因特网流行,社交媒体发达,交互式媒体快速发展,大数据展现其独有的时代特性,广泛应用在客户群体细分、数据搜索、虚拟现实、个性推荐、客户关系管理等方面,展现出巨大的延伸价值,越来越成为时代焦点,引起人们关注,如图 3-1 所示。

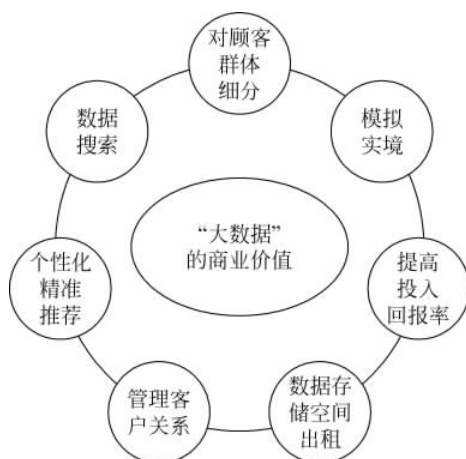


图 3-1 大数据应用价值

3.1 大数据在电子商务中的应用

电子商务发展最关键的是信息流、物流及资金流,由此,电子商务的发展也带动了许多行业的发展,以及新兴业态的兴起,例如支付宝、微信钱包、快钱等第三方支付及快递行业。

3.1.1 大数据是电子商务发展要素

随着企业处理的数据量迅速变大,处理速度飞快提升,数据处理工具的智能化程度提高,价格日益实惠,许多大型电子商务公司已经把大数据分析看成一项必不可少的工作内容,灵活运用各项数据分析手段,提炼商业智能已经成为电子商务企业在大数据时代背景下的一项必修课。

电子商务企业需要分析其核心业务数据及其不断增大的规模,不能凭直觉制定关键决策,最好分析所有与客户相关的业务数据以留住现有客户,吸引他们,同时赢得更多新客户购买更多的商品。企业应对基础设施和软件进行投资,运用相应算法处理大数据,了解消费者情绪,优化供应链,去除虚假数据。为此,聘请数据科学家完成相应工作,只有对数据进行压缩处理并智能地展现与特定内容相关的数据,大数据才能更好地利用。

非常庞大的数据集往往很难用传统的数据库管理工具进行处理,这些数据包括访问网页、登录、在线交易等,企业应使用相应工具对数据进行压缩和筛选,实施大数据策略,仅展现与特定内容相关的数据。大数据通过捕捉、存储和分析用户在社交媒体上发表的售后体验,改变业务模式,可以提高质量,改进服务。企业不仅应捕捉、存储大数据,还应开发、利用大数据,因为只有开发、利用大数据,才能挖掘出大数据蕴藏的巨大价值,特别是应使用专门工具分析和开发杂乱的、非结构化的数据。根据个人或消费群体的喜好或者消费行为分析和细分市场,提供富有个性化的产品,营销部门收集一些有价值的信息来找出购物者的兴趣所在,然后组织一些有针对性的营销活动,从而在竞争中增加企业优势。

电商公司除了要关注大数据工具的运用之外,真正应该注意的是情报数据,从日常的工作来看,情报数据处理人员出去收集情报的工作占了多数时间。他们会跟上下游供应链进行跨部门沟通。虽然这些情报数据性不强,但价值十分高。电子商务企业在有海量数据积累的基础上,还要有一套按照公司需求定制的优秀系统才可能实现大数据,在销售记录屡创新高的同时,电子商务的利润率得到增长,实现销量与利润率双增长。

3.1.2 电子商务大数据的实用措施

经过处理的大数据分析能给企业带来效益,提供增量价值,大数据只有带来实实在在的效益才会被商家接受,效益是大数据的根本要素,商家承诺产品与服务,消费者获得好的体验和价值,商家才能最终获得可观的销量和实在的利益。数据可靠性是分析大数据的前提,大数据的价值潜力来自于机器学习,大数据在经济上的应用依赖规模效应,商家需要具有大数据理念,主动开发大数据的价值,大数据就成为竞争的主要工具,利用大数据可以提高品牌忠诚度。

商家可以利用已有品牌建立互动的网络社交平台,成立网络社区,改善商业网站,在商家的社区网站中消费者与企业员工及其他消费者互动,如果得到反馈,这种网络之间的人际沟通会带来对品牌的好感,这种行为促使消费者增进对品牌的归属感,产生信赖的感觉,社交

平台促使商家赢得市场,刺激消费者增进支出,效果超过传统模式,有利于商家进行宣传活动。

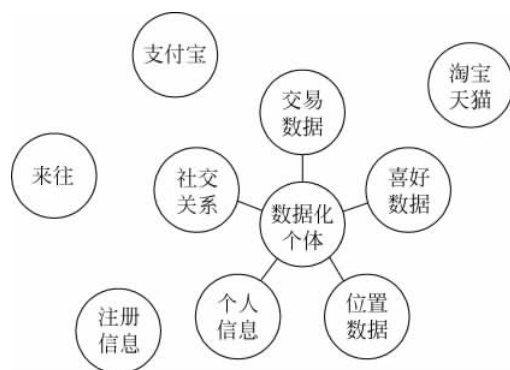


图 3-2 消费者行为数据

大数据实现市场的规模效应,反映社会性活动,消费者被网络社区聚集成一个拥有共性的消费群体,网络社交行为进一步提高,网络互动变成惯性,网络社交接近现实社会行为,如图 3-2 所示。商家通过品牌网站举行活动模拟生活或是直接将线上活动与现实中的商品关联起来,网上的潜在消费者会产生强烈的购买欲,这种感觉直接导致购买行为的发生。

大数据不是将商品打扮成大众形象,而是采集数据促使商家开发创新,找到适合的商品营销方式。帮助商家个性化生产商品,精确区分消费者群体。大数据反映消费者的需求,深入挖掘就可以更好地预测销售结果,采取有效地商业行为。

3.1.3 电子商务大数据的转型路径

1. 电子商务大数据的发展趋势

无论是传统的 IT 企业还是典型的因特网企业都在调整自己的战略,把业务延伸到大数据领域,大数据更是电子商务企业的生命线。大数据对电子商务发展的影响有如下方面:

(1) 数据将成为企业的核心竞争力和核心资产。拥有庞大的数据或者拥有分析、挖掘数据、传输数据的能力都是企业竞争力的主要来源。

(2) 跨界融合将成为普遍趋势。跨界融合既包括跨媒体、跨行业、跨地域、跨国界经营的横向跨界,也包括平台衍生内容服务的全产业链要素的纵向跨界。通信、内容、消费者、计算机具有融合趋势,IT 产业由“内容驱动硬件增长”的模式向“内容-软件-硬件”的一体化发展。苹果公司从创意设计到平台、内容、终端都是一体化运营,行业的边界日益模糊,融合成为基本趋势。

(3) 无边界娱乐成为基本生活样态。大数据大大地改变了人们的生活方式,扩展了娱乐的范围,造就了互动性强的无边界娱乐。无边界娱乐使人们不受时间和空间的限制娱乐,移动因特网时代时间和空间完全被打破,交互性超强,更满足了娱乐无边界的需求。在娱乐无边界的时代,无论是网络创作、传播、消费、阅读,还是网络购物、时尚电子产品,青少年是主流消费者。

2. 电子商务转型升级的基本方向

电子商务需要在这个大趋势之下转型升级,大数据、社交媒体、云计算、文化产业的相互联动能够促进电子商务的转型升级,将会给电子商务带来新的机遇。电子商务转型升级的基本方向具体有如下几个方面:

(1) 由销量制胜到数据制胜。在发展的初级阶段电子商务强调销量、人气,到了高级阶段聚集庞大的数据成为主要方面,销量制胜转变为数据制胜。亚马逊公司市值能达到 1500 亿,一直在行业里面遥遥领先是与庞大的数据积累有密切关联的。电子商务公司可以

将数据实现规模化,更大程度地增加用户黏度。

(2) 由规模化制造走向规模化定制。制造和定制是截然相反的,定制与制造反向,是按顾客需要销售的。大数据时代通过机器学习,这些因特网企业主要研究人类生活方式的变化,包括购买习惯和购买偏好,并满足这些个性化需求。制造类企业或服务类企业解决了原来个性化需求和大规模制造之间的矛盾,提供更有效的解决方案。在大规模定制的情况下,下一步电子商务平台可能会成为所有制造类和服务类企业的整合者,也就成为标准的制定者。三流的企业做产品,二流的企业做品牌,一流的企业做标准,通过大数据电子商务平台进行分析能够预测产业的趋势和潮流,然后通过标准和设计研发及订单的方式定制产品,占据价值链高端。

(3) 由平台为主到综合内容与平台。现在人们的生活方式变为娱乐、休闲、购物一体化,电子商务企业业务发展就要考虑变化,还要提供休闲娱乐、视频、音乐、购物、金融服务等综合业务,提升方向就是形成综合体验价值。

(4) 由资产并购到数据整合。现在围绕大数据的整合并购加速,阿里入股新浪微博,百度收购 PPS,都是平台企业收购内容企业,以实现内容和平台的综合化,并购的外在形式是资产、股权的重组,实际数据资源的共享、数据规模的扩展是其内在的本质。

(5) 电子商务由 PC 端逐步转变到手机端。从中国互联网络信息中心的研究可以看到,手机网民的规模持续上升,2013 年 6 月底达到 78.5%,比 PC 要高接近 10 个百分点,手机下一步可能成为信息中心,特别是手机在 4G/5G 推动之下会有更多的创新。从总体趋势来说,手机将成为最大的媒体终端、娱乐消费终端、购物终端。

(6) 商业模式由复制到扩展。判断数据价值,实现同一组数据无限再利用是一个主要的方式,复制是手段,扩展才是目的,而扩展包含着实现企业线上、线下整合,物理平台和虚拟平台的联动,最后实现企业整体价值最大化,复制业务模式实现企业的扩展。网上银行趋势是不可逆转的,对传统银行会造成巨大的冲击,大数据分析能力、信用体系、透明度、低成本都显示了超强的竞争力。货币的生命力不在于它是否虚拟,关键在于背后有没有强有力的信用体系的支撑,电子商务平台也有可能发行虚拟货币。

3.2 大数据在金融的应用

大数据金融是指依托于海量、非机构化的数据,通过因特网、云计算等信息化方式,对数据进行专业化的挖掘和分析,并与传统金融服务相结合,开展相关资金融通工作。可划分为平台模式和供应链金融模式。大数据金融模式需要两项能力,一是大数据分析的能力;二是数据来源的合法性、持续性能力。

3.2.1 大数据金融的提出

1. 大数据金融的兴起

社交网络推动移动支付,大数据伴随云计算,因特网信息带来大数据技术的发展,第三方支付提高信誉,B2B 和网络贷款日益增加,金融机构建立线上平台模式,传统金融吸收因特网技术,电子商务融合社交网络,发展亮点不一而足,纷繁的大数据金融借助移动因特网,共同创造出新的价值,比传统金融业更加透明,具有更强的参与度,更高的协调性,更低的中问成本,更方便的操作性,成为更便捷的新兴金融模式。

大数据金融具有革命性,今后发展速度必然具有加速趋势,网络先天具有开放性,导致多边性,进而提升生产力,有效影响人类生活,影响其他产业,力量强大,信息技术不断革新,渗透人类其他领域的各种活动,不仅仅是社会的物质基础,更从根本上改变了人们对时间和空间的看法。

大数据金融离不开电子商务的发展。近年来,中国电子商务飞速发展,势头十分迅猛。据统计,2012年我国全部电子交易高达8.1万亿,增长速度达到30%。网络零售发展更快,2012年共有1.3万亿,同比提高67%。金融依赖商务需求,商业就要交易,交易通过支付,支付越来越靠大数据支撑,融资需求实现闭环发展。电子商务的高速发展促进生态链不断整合完善。电子商务生态链的发展建设日益规范,融合平台、虚拟服务、专业营销、精准支付、网络金融、服务供应链化、物流智慧化、终端移动化,最终形成大数据金融。

电子商务获得大众关注的阿里巴巴大力推进因特网金融,其他因特网企业随之跟进。2013年6月13日,阿里巴巴推出“余额宝”,不到两个月达到250亿元。2013年7月,新浪建立“微银行”;8月腾讯实现微信5.0版与“财付通”的结合。其他大小、形形色色的企业都在拥抱大数据金融。

因特网企业逐渐进入金融领域,不是仅仅依靠技术,而是利用和依赖数据。网络服务商整合已经拥有的大数据,通过必要的信息技术,准确进行预测,提前掌握消费者,把握其习惯行为,推断未来发展趋势。大数据先天具有共通金融的本质属性,设计金融产品,通过采集各种数据,计算各种组合,实现数量匹配,得出风险定价,经过网上支付,建设大数据金融的基础与核心过程。

银行应用大数据分析直接监控企业运营,掌握企业经营状况,评估企业经营风险,解决具体贷款难题,实现全程跟踪服务,最大限度降低资金风险,缓解企业面临的困难,为企业迅速实现商业价值,增进服务用户价值。余额宝的发明是一个创新,极大地促进了大数据金融飞速发展,依靠因特网,采取优势方法采集大量全面的数据,利用新的大数据的思维,进一步促使传统金融业不断革新。因特网公司采取创新方式,通过免费手段吸引关注,迅速聚集了一大批具有高度品牌黏性的新用户,一旦客户习惯获得免费服务,因特网公司就会获得收益,可以采集相关数据,不断快速储存,加强交互联系,形成一个源源不断持续可靠的大数据来源,产生新的商业模式,满足客户需求,大数据金融进行实时分析,马上获得结果,立刻设计出相应产品,对客户进行精准营销。传统金融几乎难以想象,不可能做到。大数据金融完全可以实现给客户提供完全个性化的服务,推行适合的量身订做的各类产品,让客户享受产品价值,有大幅的直接提升,从而实现利润和因特网公司的持续盈利。因特网公司不仅仅早期为企业提前提供增值服务,现在还可以从网络支付到下一步的社会融资,大胆开辟新型的大数据金融发展道路。这其中不间断的每一次进步都依靠大数据的全面应用和大量支持。图3-3展示了以客户为中心的大数据金融,图3-4所示为大数据金融用户分析。

2. 大数据金融的发展

目前来看,大数据金融的服务主要应用在广大电子商务领域,大多数应用项目的业务经验是依靠分析已有的结构化数据。但是以后金融业的前景更多面对非结构化数据,将需要处理难以想象的海量数据。信息量膨胀,信息大爆炸越来越突出,银行决策系统必须改进,

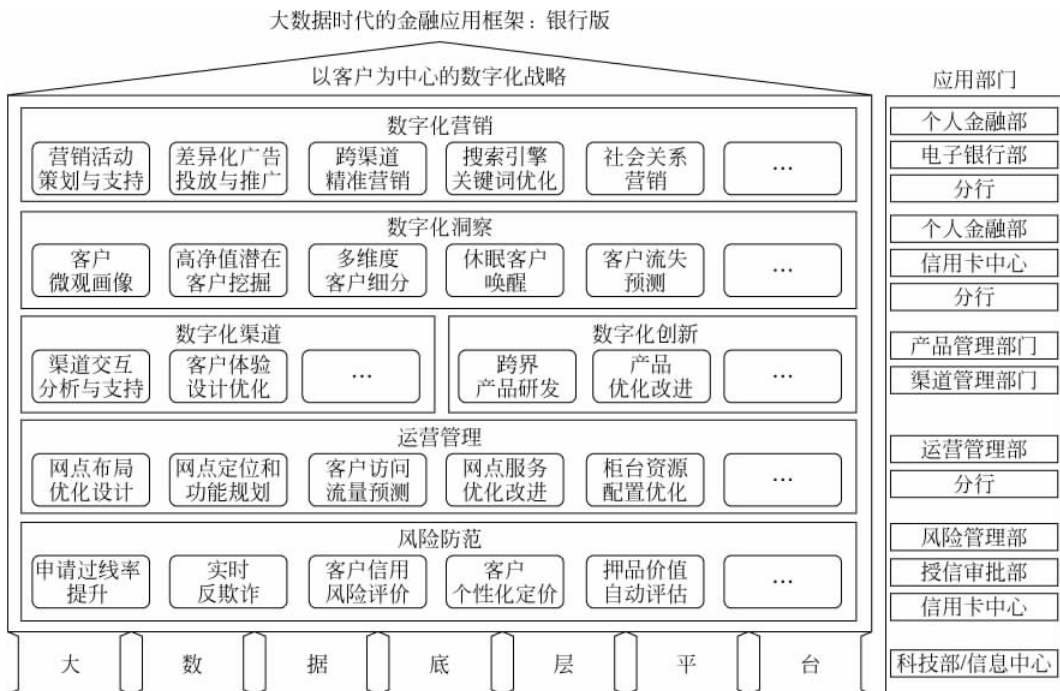


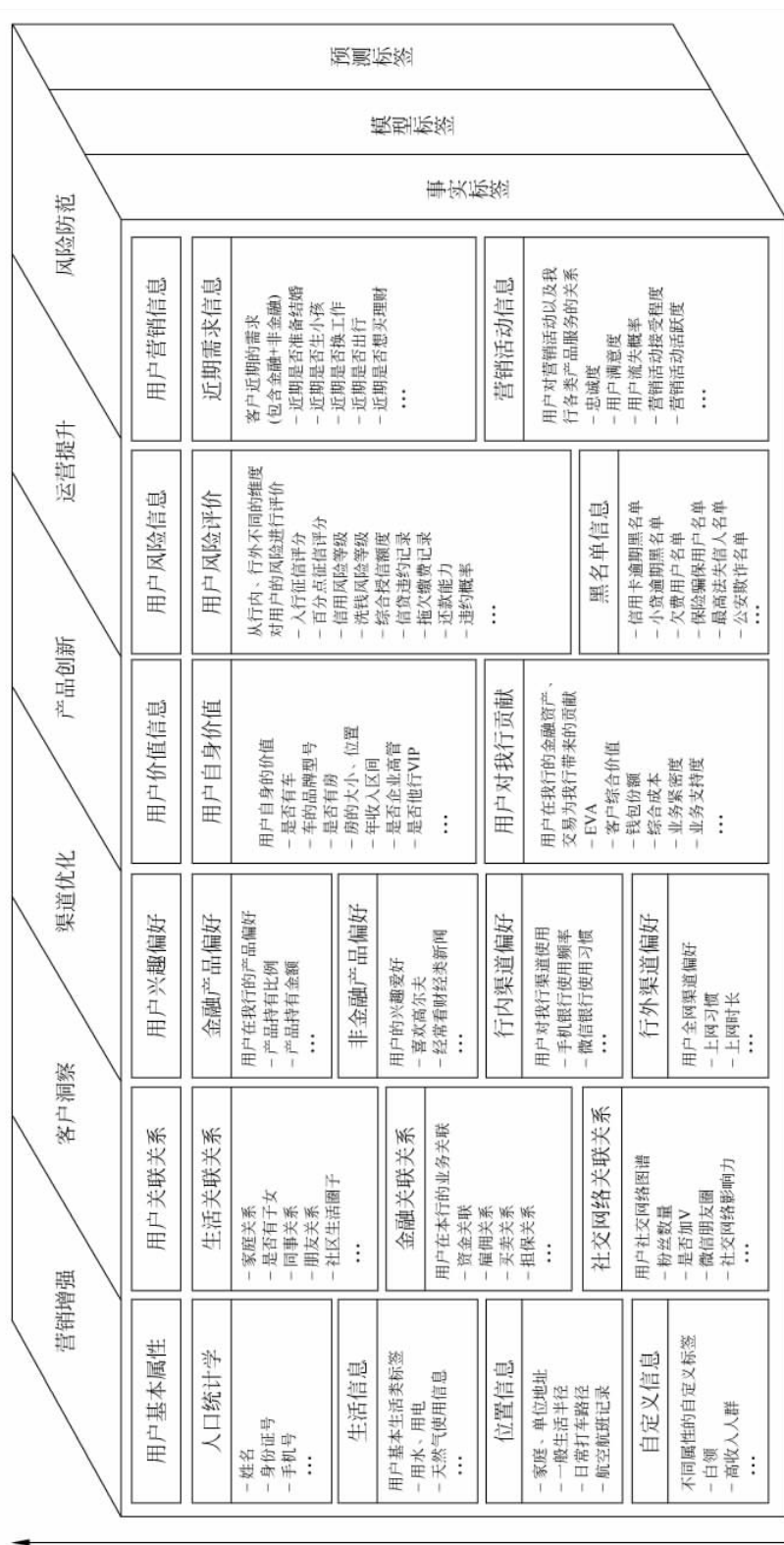
图 3-3 以客户为中心的大数据金融

才能整合处理多样化业务。银行业建设大数据平台要易用,提供系统方案,结合原系统,满足可扩展性,需要全面集成已有沉淀数据,延伸多种功能,有效解决问题,完善原有数据仓库。建立银行大数据中心平台,不仅解决数据集成,还能系统集成大数据,提供企业业务方案,实现高性能,解决高速度,实现大数据分析平台,建立智能决策体系,实时整合所有功能,组织所有系统。

金融行业不断发展,用户量持续扩大,银行业务多元,数据规模增长,为了实现数据监控,需要跨系统数据传输,跨行业整合数据,并进行快速分析,得出有效方案。银行业必须保证大数据的真实性,才能进一步应用数据挖掘功能,进行高级分析,提供决策支持体系,实时快速处理银行各种各样的静态及动态数据,通过可视化工具为企业展示,进行评估,提供决策帮助,预测未来前景。这是大数据金融的力量源泉。电子商务、电信、金融三大领域是目前应用大数据比较快的,不断优化增长,处理交易行为,时时监测用户心理,及时证券营销。电子商务拥有成功经验,银行大数据项目可以借鉴学习。而今,阿里网络金融再度启动创新模式,应用大数据技术,改变金融业传统理念。如阿里提出,针对买家发行虚拟信用卡,基于海量用户交易记录,建立有效信用机制,用于网上购物与支付,已经超过 8000 万用户,带来巨大市场利润,潜力非常雄厚。行业不同,需求有差异,关键在于构建大数据平台,实时适应行业应用。目前,金融信息化迫在眉睫,关键在于快速处理海量数据的技术性能,建设系统,注重用户体验,重视业务系统,抓好业务办理,实时传递各种数据,提高社会效率。目前急需解决金融行业的数据隔膜,信息分散就会导致计算资源浪费,利用率极低。快速扩展大数据业务,需要马上实现平台交付,需要实现客户数据集成平台,应用决策分析系统,处理分析图像及视频等非结构性数据,必将成为大数据应用的前进方向。

金融业用户画像体系：银行版

支持应用场景



第二维：标签应用深度

第一维：业务层次划分

图 3-4 大数据金融用户分析

3. 大数据金融的应用

面对海量数据做出分析是大数据平台的目标,金融行业需要大数据,核心价值在于共享,数据可视化的发展应用扩展了传统商业的视野,应用图形分析增加直观性,更加方便地发现数据特征,进而帮助其他数据分析人员抓住时机,及时操作。

过去银行里的客户经理是被动的,盲目等待客户上门,其模式难以为继。现在银行业大不一样,开始主动发掘用户的不同偏好,有针对性地积极提供各种营销服务,例如中信银行主动采用最新的 Green plum 系统,实现实时营销,已经降低数千万成本。大数据拥有巨大商业价值,体现在如下方面:一是快速定位,找到高价值客户群体,挖掘高潜力客户集群,实现对金融产品的准确营销;二是利用新型的高性能挖掘技术,进行反欺诈商业分析,避免企业各种运营风险。银行业历史产生的数据巨大,采集、存储、管理数据需要分析,应用大数据工具可以解决金融行业用户的特有需要,控制种种风险。

实现大数据应用,关键在于实时获取各种非结构化数据流,持续采集大数据,并汇总集中到数据中心,使用有效技术技能和相关数学工具来分析思考大数据,实现实时利用共享,支持业务决策。大数据平台一般采用分布式处理,以便快速定位。结合相应数据知识,推动行业开发应用是驱动大数据发展的最大力量。

大数据迫使银行和电信业提升现有业务能力,实现应用目标,还需要利用新的技术,规划需求,接受数据体系,并开发相应的战略处理海量的流数据,捕捉服务信息流数据,进行实时分析,提高服务质量。除了技术创新外,善于利用行业经验是金融 IT 企业解决问题的关键。各行各业同步发展共享数据,健全完善国家法律法规,构建合理的商业模式都同样重要,会产生无比巨大的社会价值。中国银监会设立金融消费者保护局,有力地保障大数据金融的发展。在国外,消费者金融(Consumer Financial)可以帮助客户,提供丰富便利的大数据应用服务,如对客户交易日志实施实时检测,进行债权现状分析,据此实现客户分类,提供系统评分,预测客户未来行为,实现个性精准营销,避免出现坏账。而金融管理部门及时把握交易状态,提供有效监督,做出预测分析。

4. 大数据金融的理念

大数据的核心精神是公开、透明,因特网金融将彻底改变传统金融业,从封闭僵化走向合作开放。因特网支付行业已经出现支付宝等各种不同的公司,各方选择不一样的服务对象,提供形形色色的不同服务模式。大数据金融对行业进行高度细分,提供差异服务,从业者需要设法突出特色,提升自身效率,大力改善服务质量,才能占领大量市场。

传统金融业系统封闭,思维僵化,只能服务少数客户。目前中国金融体系机构单一,过于庞大,无法及时满足百姓需求。大数据金融通过采集数据实时分析,有效降低金融服务成本,创造了金融服务的新模式。

现在银行业面临发展困境,大企业风险较小,但机制呆板;小企业利率较高,但风险太大。大数据金融充分掌握信息,公开透明及时披露,挖掘数据评估风险,合理定价约束双方,解决信贷难题。P2P 网贷的欣欣向荣很好地体现了大数据金融的发展前景。大数据金融潜力巨大,前景广阔,政府应该全力支持大数据金融的发展,大力培育良好的社会环境,创造有力的因特网金融生态环境;规划发展环节,呵护因特网金融集聚期,扶持一批因特网小额信贷公司、在线保险公司、发展因特网担保公司和第三方支付机构,支持电子保单、因特网

P2P、因特网征信等。与此同时,也要做好风险控制与安全管理,打击因特网金融欺诈,防范因特网金融犯罪,推动大数据金融发展,搞好金融建设,打造一个经济发展的新亮点和增长点。

大数据金融繁荣需要各方面努力,第一要充分重视、密切关注移动互联高速发展的明显趋势;第二是加大强化大数据的社会价值与社会应用;第三务必要加大建设力度,整合因特网产业链平台,因为任何产业升级必然沿着全球产业链延伸。利用全球化、信息化推动信息产业升级,因特网的突出作用只会越来越明显,越来越重要。沿着产业链进行产业升级关键离不开大数据金融,所以发展产业链必将引起大数据金融的日益繁荣。大数据金融时代的产业链整合是高度细分、去中心化的,所有参与者都是主体,都能够找到自己的位置,从而可以实现更好的发展,更好的服务产业。

3.2.2 大数据金融的功能

与传统金融不同,大数据金融不仅可以带来金融服务,还直接促进产品创新,以及实现用户体验的舒适变化,不断创造新的经营管理模式和业务处理方法,明显改善金融服务提供商的组织结构,根据用户特征预测数据需求与管理模式,增加产品创新力来源,提高信用,影响风险特征等,显著丰富了金融体系的多样性,增进了金融监管和宏观调控等方面的复杂性,也提出了新的课题。

1. 重组实体经济

金融机构中,无论哪个部门都在持续不断采集、积累大量数据,如抵押贷款部门采集海量的贷款客户的数据,储存并实施处理能充分描述其特征;从全球看,债券、外汇、货币和股票及衍生品交易部门能收集各种各样影响资产价格不断变化的海量信息,并试图建立可以使用的前瞻性模型;银行零售部门实时收集和分析客户行为信息;客户交易、研发、市场开发或服务运营等各个部门也隐藏了无数数据可以挖掘出巨大价值。但是由于缺乏跨部门跨行业的沟通战略,各部门已有的大数据分析技术不足,难以深刻了解不同地区金融市场之间的复杂关系,妨碍了信息的流通和及时利用。

为了打破这样的壁垒,一些企业正主动出击,试图直接推动整合跨职能部门的数据,甚至寻求办法获得外部供应商及其客户的外部信息。例如,美国纽约市新成立的 Movenbank 移动银行通过与已有的传统商业银行谈判,合作推行移动银行新业务,帮助其解决内部机构割裂问题。英国 ERN 公司提出有计划去利用用户的交易历史和消费习惯,然后参照交易位置和时间数据,向各大银行和各种商家提供相关数据服务和交易咨询。

2. 实现信息对称

大数据有助于提升金融市场的透明度。金融客户的信用状况将实现实时动态变化,随着其资产情况、经营方式和各类交易状况的变迁而变化,传统商业银行直接大量投入人力、物力和财力,建立特有的信息平台,进行收集、储存、分析和决策,以解决长期存在的信息不对称问题。近年来,大数据金融可以通过平台直接采集、整合金融交易双方的所有信息,形成了新的来源办法及金融信息的新型模式,金融客户的交易价格信息更为精细,社会经济状况等方面的数据更加透明,更为准确地形成利率,实现市场化。

新型的大数据金融企业不仅仅是平台,更可能是在价值链中间成为中介角色。例如,在

经营全球产品运输方面的发展过程中,一家运输公司不仅仅收集海量信息,还可以专门销售这些信息产品。同时进行第三方支付的企业也发现了形形色色的海量支付信息蕴含的巨大商业价值。随着价格信息不断在网上大量扩散至线下,各类基金销售企业正在自动编辑网络上数百万种不同商品信息的比价服务,这对消费者提供了信息区,为社会创造了巨大价值。

新兴市场欠缺比较成熟的各类征信机构,因此有些公司采用申请者代发的社交网络信息,加以分析后得出信用评分。例如,德国研发 Kreditech 贷款评分公司,美国设立 Movenbank 移动银行,中国香港成立 Lenddo 网络贷款公司及 Trust Cloud、Connect. Me、Briefly、Reputate 等新型中介机构,试图设计大数据金融的信用平台,打造能反映 LinkedIn、Facebook 或其他社交网络的开放平台,整合用户活动记录,通过算法自行开发软件,分析客户的好友信用状况,建立标准化格式,归纳与收集各种信用资料,成为客户信用评分的重要依据,实时将社交网络产生的种种资料直接转化成个人的因特网信用。Movenbank 直接对客户采用风险评估,其核心参考不仅仅是个人传统信息汇总,也逐渐纳入 e-Bay 等电子平台的各种交易评价,包括整个网络汇款产生的记录等综合因素,还会全面计算 Facebook 的好友人数或 LinkedIn 的人脉对象,以及 Klout 影响力分数等网络社交参与情况。

3. 衍生更好产品

大数据金融通过物联网,借助云计算,依靠社交网络等新的数字平台产生了无数新用户和海量数据,虽然直接记录了所有用户群体的社会情绪,但大数据库不能自动运行,无法自己计算总结整个群体的行为模式和活动规律。计算机科学家需要社会科学家加强互相协作,和统计学家找到新途径,使大数据研究结合小数据策略。利用大数据,金融企业可以直接分析客户行为模式,比如不同事件关联性分析,如同对照实验,即处于不同工程条件下,观察机构投资者、测试普通消费者对金融产品的不同反应,识别客户的行为关系,提高资金转化率,改善企业服务水平,实现大数据金融的良性发展和精准营销。例如,前沿的零售企业观察客户的店内行为及活动情况,监控其与商品的整个互动,结合所有交易记录分析,开展各种实验,可以指导商品的种类选择、摆放次序、售价调整。再如,通过精细化数据分析,Progressive 保险公司不断考察客户风险及其财富变化,计算家庭资产数据,并不断采集背景资料,向客户精准服务,提供专业建议和量身定制的独特保单。未来,保险公司还将根据个人位置状况和汽车信息对不同的车险产品采取不同定价,向客户提供交通信息和天气突发状况、事故高发区和限速等实时更新的信息,互动开发,有利于安全驾驶。

3.2.3 大数据金融的挑战

1. 安全隐患

随着个人位置、行动空间、购买趋向、性格偏好、身体健康和公司财务情况的海量数据产生并被收集,再伴随金融交易风格、持有资产习惯及信用状况分布被以更微小、更精细的方式采集、储存和分析,机构投资者获得更低的金融价格,金融消费者获得更符合需要的服务,市场配置从而提高,金融资源越来越丰富。但与此同时,金融市场依赖的信息基础设施更加庞大,变得越来越复杂,整个社会管理更加一体化,更加开放外向,对隐私和数据安全更加敏感,保护知识产权更加困难。针对个人隐私,大数据时代容易产生隐私的问题,以往的常规

身份确认远远不够,风险范畴不断增加。最近科学家对欧洲 150 万用户的手机进行分析,数据显示仅需要 4 项基本因素就已经可以对其中高达 95% 的个人身份进行确认。还有,基于大数据的分析可知,人们在城市中选择的路径难以置信的存在唯一性。

2. 市场情绪

大数据金融可以通过采集、分析充斥于社交媒体上的各类内容做出市场情绪分析。如今 Twitter 日发消息已经超过 5 亿条,Facebook 日益火爆,日均用户已经超过 10 亿。英国布里斯托尔大学的一个科研团队深入研究了从 2009 年 7 月到 2012 年 1 月期间由约千万英国人产生的 4.84 亿条 Twitter 信息,得出结论为公众的情绪直接源于相关财政紧缩产生的社会压力。惠普实验室的科学家伯纳多·休伯曼进行社交计算研究,在《网页法则》里分析人们目前发布在虚拟空间的微博与现实世界有关系,将其命名为“注意力经济学”。通过分析,他发现电影的票房收入能够准确预测,与人们发布在社交网站的微博相关。

大数据研究与应用密切相关,金融投资者开始投入研究,试图将其结合起来。最近两年,对冲基金开始研究,从 Facebook、Twitter、聊天室和博客等社交媒体中发现提取市场情绪信息,开发设计交易算法。一旦发现有意外信息公布,无论恐怖袭击事件还是自然灾害等,便立即抛出,获得收益。2008 年,在美国加州圣莫尼卡与理查德·彼得森筹集了 100 万美元,建立了名为 Market Psy Capital 的对冲基金,通过考察博客、聊天室,追逐网站和微博,以发现确定市场对不同企业引发的情绪,再据此确定基金的交易策略。到 2010 年,该基金回报率最高达 40%。巴黎掌握行为金融学的三位交易员运营 IIBremans,针对法国 CAC40 指数做判断,提供情绪分析;小型对冲基金 DCM 资本位于伦敦,从 Twitter 和 Facebook 等社交媒体采取手段收集信息,通过软件分析人们对某个金融工具的情绪,进行评价打分,并向客户发布预测提供零售,辅助专业投资者,做出重大的投资决定。

3. 决策误差

大数据是人类的发展成果与设计的产物,大数据的工具(如 Hadoop 软件)还在成长,并不能立刻使人们摆脱限制思考的曲解,打破隔阂和成见,数据之间彼此的相关性也不直接等同于因果关系,大数据还存在其他技术问题,比如存在选择性覆盖问题等。

例如,社交媒体虽然是大数据分析的基础平台和重要信息源,但其中大多都是年轻人,城市人占比偏高,还存在大量不活跃的空账号或死账号。比如波士顿运用 Street Bump 程序对城市路面进行统计,其坑洼情况数据来自驾驶员的智能手机,可能少收集计算年老居民和贫困市民较多的那些区域。“谷歌流感预测”曾经过高估计了 2012 年的全球流感发病率。这说明大数据有缺陷,政府决策片面依赖大数据可能带来不实,可能造成负面影响,还可能进一步加剧社会已有的不公。

2010 年,美国股票市场从恐惧情绪趋于慢慢复苏,但 Market Psy Capital 基金未能及时判断明确,其分析模型仍建立在恐惧基础之上,没有及时调整,对趋势变化考虑不充分,结果当年该基金亏损 8%。美国印第安纳大学的约翰·博伦指出,即使整体数据的准确度高达 80% 也不能轻信,仅仅 20% 的差错率就足以造成破产。只有用社交媒体衡量整体的公共情绪才有意义。

3.2.4 大数据金融创新

1. 高频交易

高频交易(High-Frequency Trading)也叫算法交易(Algorithmic Trading),指交易者为了取得高额利润,充分利用硬件设备,依赖交易程序的优势,十分快速地收取、分析、下达和输送大量交易指令,在很短的时间节点内不断买入卖出,通常不直接持有大批未对冲的头寸过夜。根据来自许多方面统计数字的综合判断,2009年至今,美国无论是期货市场还是证券市场,或是外汇市场,高频交易所占份额持续扩大,已达40%~80%。随着运用这类操作策略的高频交易越来越密集,其负面效应开始进一步凸显,且实际利润不断大幅下降。据芝加哥Rosenblatt证券咨询公司的资料表明,2012年采取高频交易的公司的总利润下滑,比2009年下降了约74%。

如今高频交易开始改变操作方式,采取“战略顺序交易(Strategic Sequential Trading)”,即根据金融大数据分析结果,以明确识别出具体的特定市场参与者,追逐其留下的金融足迹。例如,假使一只共同基金习惯在收盘前一分钟的第一秒才突然执行大额订单,那么能够识别并判断这一模式的算法就会有空间操作,将通过预判该基金在不同交易时段的大致动向,进行相应交易。那么该基金继续执行交易时将更可能付出更高的价格,从而使使用该算法的交易商最终获得较大利润。

2. 信贷评估

大数据金融可以加强管理力度,提高风险的可审查性与可预测性,支持实施精细化管理。金融机构非常希望能够了解中小微企业用户,通过收集其大量日常交易行为的数据,可以发现其业务范围、信用水平和经营状况,判断其用户定位、资金需求及行业发展趋势,从根本上解决小微企业不透明的财务制度造成的战略不清晰,难以改善真实的经营状况的难题。

阿里小贷首次实现了全程线上借贷模式,首创了从风险审核到放贷的整个流程,将贷前、贷中与贷后结合,三个环节联动,形成有效联结,使得贷款不难,改变传统金融渠道的不足,主动为弱势群体服务,批量发放小额贷款,特点是“金额小、期限短、随借随还”。

(1) 根据阿里巴巴B2B、天猫、淘宝、支付宝等一系列电子商务平台,收集大量客户积累的原始信用数据,充分利用在线视频,全方位多角度定性调查相关客户的资信,再考虑交易平台上产生的大量客户信息(客户评价度、口碑评价、货运数据等),并量化处理后两类信息,同时引入税务、海关、电力等外部产生的数据进行再次匹配,建立有效的数据库模型。

(2) 实施交叉检验身份技术,再通过第三方验证进一步确认客户信息,确保真实性。借助电子商务网络平台,统计客户的各类行为,映射为数据,建立企业和个人的信用评价体系。应用沙盘推演技术,评级地区客户,进行分层管理,研发新技术。设置评分卡体系,规定微贷通用规则,推动决策引擎,实现风险量化分析等技术。

(3) 建立网络人际爬虫系统,实现风险监管开发,跨越地理距离的限制,捕捉人际关系信息,并通过设立规则整合相关事项,实现关联性分析,得到风险评估结论,综合风险评估结论与贷前评级系统,进行双向交叉共同验证,构成双保险控制风险。依靠因特网监控技术,阿里小贷可以明确贷款的流向。如果将贷款用于扩大经营,阿里小贷将帮助评估其广告投放的方式、店铺装修的风格及销售措施。

3. 监管方式

大数据的使用不仅仅改变金融市场,随之而来需要改变传统监管市场的方式,以最大限度确保市场参与者规范地采集大数据,安全地存储大数据,客观地分析大数据。例如,2010年5月发生的“闪电暴跌(Flash Crash)”令道琼斯工业平均指数(Dow Jones Industrial Average)大幅下跌,事后美国监管部门考察,认为是高频交易的恶果,造成了快速集中抛售,引发更多不负责的抛售。2013年4月23日突发的“无厘头暴跌(Hash Crash)”的缘由更令人愕然,居然是美联社的Twitter账号出错,误发出巴拉克·奥巴马(Barack Obama)突然遭遇恐怖袭击的虚假消息。可见,大数据金融风险加大,一个数据出错就可能直接导致“无厘头暴跌”。

但是,监管机构片面限制大数据技术,或是对其使用范围进行直接限制和干预,其潜在风险反而是更加巨大的。恰恰相反,应鼓励产业界积极应对更复杂的技术,乃至继续实施更大数据的应用。

纽约大学理工学院(NYU-Poly)召开大数据金融会议,美国商品期货交易委员会(CFTC)的代表斯科特·奥马利亚(Scott O'Malia)表示,CFTC曾考虑如何实现有效监管,主张让监管机构主动出击,对金融交易商的算法进行科学认证。在实践中,利用算法采取的鲁莽行为带来更大的破坏,甚至超过传统的操纵市场行为。劳伦斯伯克利国家实验室名声在外,拥有强大的超级计算能力和独特的雄厚分析技术,能够做到针对威胁稳定交易的行为采取实时监控。传统的停市机制只能在市场暴跌后采取措施,停止全部交易,而大数据实时监控能够精细调控,将单个不规范的参与者清除,从而继续向诚信的其他参与者敞开市场。

4. 信用文化

大数据带来思维方式的变革,从而会导致传统金融业发生思维变革,首先会扭转传统金融信贷业的物质抵押文化,直接推动信用成为价值,信用借贷成为可能,走向主流。尤其传统的中国金融行业盛行抵押文化,在贷款的行为过程中片面依赖抵押物品,往往导致急需借贷的中小企业反而得不到相应的贷款服务,这种粗暴的抵押文化对金融业发展不利,让贷款提供方难以提高服务,在考量借贷时思维简单。贷款方仅仅片面的把抵押物的价值作为考量,以此确保价值的利润空间。长期而言,抵押文化对金融健康发展具有负面的影响。要想真正发展金融,就要提高信用贷款,建立有效的信用机制。真正的保值增值不是抵押物,而是人们的长期信用。

大数据金融首先表现为思维的变革。信用是抽象的,但大数据可以建立信用体系,让个人的信用或者群体的信用变得具体。这将是金融业的根本性改变,并将持续产生巨大的深远影响。个人的信用评估不是静态的,而是取决于很多的变量,是一个动态的、连续的行为特征的长期体现——资产、消费、收入、习惯、个性、社交网络等都是有效的变量,会对个体的信用产生积极的影响。个体信用通过具体的各种行为综合决定,通过大数据的整合,可以很好地采集大量的个体或者群体的信用行为,进行储存、整理、分析,只要把海量数据糅合在一起就会显示客观规律,使得人的信用不再模糊,变得鲜明、生动、立体化,从而很好地把握个体或群体信用。IT技术的改善和发展、因特网的建立和延伸、大数据的产生和应用,让市场走向全球化,摆脱了传统地域的限制,使得市场更加集中,从而使得企业规模更大,成长速度更快。而大数据技术的不断突破也催生明显的马太效应——强者越加强大,如果局限于局

域优势,就无法形成海量的用户资源和数据资产的良好管理,那么就会削弱我们的核心竞争力。

大数据是促进金融事业高发展和广开放的关键,首先要实现数据整合,保证全时在线。现在很多公共系统都是相互孤立的,即使银行的很多领域也不例外,比如对私业务与对公业务,还有银行卡业务等处于互不交流的状态,难以形成整体综合的联动效应。不仅仅是金融数据决定信用,其他相关领域的很多数据也会产生巨大的影响,这就要求数据具有更高的开放性。因特网不仅仅有开放性,还具有天生的透明性,并且这些数据还可以共享因特网,进行互联互通,推动大数据广泛应用。传统的金融业借助大数据必然会发生根本的变化。大数据必须采集、储存足够的海量数据,这是进行一切应用实现预测的最大前提。预测需要收集足够全面及非常杂多的海量信息,这是预测得以成功的最大关键。几十年来计算机和因特网的出现,IT 技术的飞速发展使得大量事物数据化,并在加速量化,直至一切皆可“量化”。

大数据金融天生具有快速发展的技术优势。因特网结合云计算可以在广泛的范围内采集信息,储存信用,完成评估,分析个人行为,整合群体信息,并将这些纷繁杂多的海量信息实时提供给高效的大数据作业系统,进行加工处理,获得价值。从这个分析来看,P2P 对大数据金融的信用评估更有独特优势。由于 P2P 的独特市场特点,导致它可以充分覆盖更多的用户群,同时因为充分利用了自我组织人人互通的特点,可以让用户源源不断的产生数据,从而迅速实现海量数据的自我膨胀和产生及循环扩张。使得数据取之不尽,用之不竭,创新成为现实。虽然大数据引发的这场巨大变革还仅仅处于早期阶段,但大数据金融的影响已然历历在目:金融服务将持续转型,从“关注整体”的粗放式管理进一步向“关注个体”的精细化管理转型;由片面简单的抵押文化向全面长远的信用文化转变;将会建立更完善的信用体制和更全面的风险管理体制;从“以利润为中心”的自我发展向“以客户为中心”的共赢发展转型。我们还可以充满信心的预见,大数据金融能够真正引发社会产生实质的改变,并且一定是由具备大数据思维的公司所推动的。只有立足精准服务,面向海量用户,占有数据资产,具备战略眼光,符合大数据的未来趋势,才能拥有长期的核心竞争力。

3.3 大数据在媒体的应用

大数据时代,信息传播方式的改变带来了突发事件话语体系的变迁。首先,突发事件议程设置主体身份话语权的变迁导致了民间舆论崛起而主流权威消解,意见领袖作用日益显著;其次,以微博为代表的社会化媒体成为突发事件的主要话语表达载体,微博成为突发事件的舆论中心,随时掀起舆论风暴;最后,突发事件话语体系一改往日的可控局面,调控难度空前巨大。而导致上述改变发生的诸多原因中信息通信技术的发展应名列首位,大数据时代通信设备的普及和巨大的信息产能使得突发事件信息得以迅速、广泛传播,进而生成舆论。在明确大数据时代突发事件话语体系的发展趋势与其中的原因后,如何应对就成为重点,可以分别从应对思维、信息管理机制的构建、具体策略实施三个方面进行仔细的考量。首先,大数据时代的突发事件应对思维应当符合大数据规模化、高相关性等特点,并且要充分适应社会化媒体对信息开放的要求。其次,大数据时代突发事件的舆论应对离不开运用信息技术对突发事件信息和数据的管理和分析。因此,以突发事件的发生时序为基准可划

分为舆情预警、信息控制及事后评估三个步骤,对突发事件事前、事中、事后每一个阶段进行突发事件信息管理并建立舆论应对的相应制度,例如信息监测体系、信息公开制度、新闻发言人制度及事后评估制度等。最后,与信息管理机制建设相配合的是具体舆论引导策略,应当从充分利用社会化媒体和加强主流媒体话语能力两个方面入手,双管齐下实现网络平台和传统媒体中突发事件的舆论引导。

3.3.1 传统媒体的不足

大数据浪潮对媒体影响巨大,价值非常可观,大数据对传统媒体冲击很大,大多数传统媒体通常分为强内容模式、强渠道模式、强服务模式,如图 3-5 所示。传统媒体如果不转型,不具备在大数据条件颠覆创新的能力,原因如下:首先本质上传统媒体业掌握的数据资源非常有限。麦肯锡全球研究所指出,行业不同,产生的大数据内容就不相同,银行、证券、通信等服务行业拥有海量数据,公共事业单位和政府组织采集储存了大量数字化数据,规模庞大。大多数传统媒体机构拥有的数据资源有限,很难与各种社交网站相提并论。其次是传统媒体业尚不具备大数据分析能力,大数据难以靠传统工具解决,大量各种各样的数据应用传统的方法无法快速解决。而传统媒体只能产生有限的数字,难以处理与适应庞大内容的数字化,在开展新媒体业务时不够专业,需重新转化编码。再从硬件和人才来看,大数据的存储处理所需要的 IT 架构,可视化所要求的基础设施和专业人员,大多数传统媒体机构也不具备。

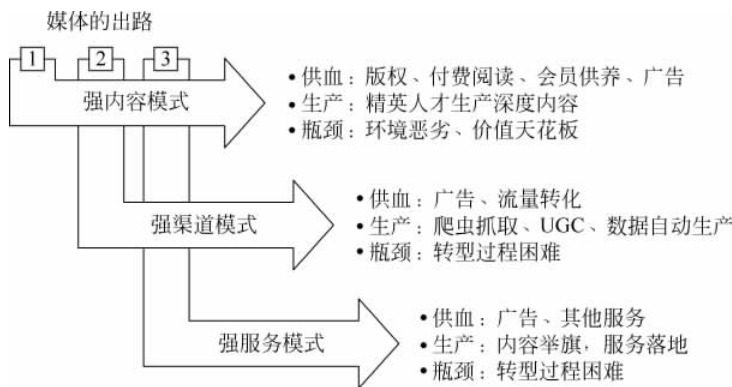


图 3-5 传统媒体的模式

3.3.2 大数据驱动传统媒体的升级

大数据浪潮之下,媒体机构应当一方面承认差距,不盲目追逐概念;另一方面避免盲目,实事求是地思考。缺少数据资源,产业不能发展;缺少数据思维,企业就要落后。数据是非常重要的生产生活资料,面对未来的发展,人人需要大数据理念,学习大数据思维,具备大数据意识,充分体会大数据价值。传统媒体机构应当更加开放,更加务实,学习大数据,掌握大数据。从大数据内容、服务、渠道等方面进行拓展(如图 3-6 所示),可分为以下 5 个方向:

1. 积累数据资产

媒体数据资产的核心是原创内容数字化和历史数据,对报业而言是全文数据库,对电视

而言是图片数据库等,建设媒体资产管理系统,扩充外部数据,通过合作,交换数据,购买因特网平台用户提供的各种内容,完善数据资产,组织数据存储,实现查询调用,提供版权管理,方便转化利用,打好应用基础。

2. 掌握数据能力

购买基础设施,技术外包合作,具备数据处理能力,掌握大数据应用能力;引进人才,培养骨干,引进数据科学家,实现商业智能,具备统计学知识,精通自然语言处理,设计新型产品,分析企业运营。

3. 用数据支持经营

传统媒体可以培养大数据意识,应用数据进行经营,增强决策的科学性,完善传统用户数据库,采集客户端网站收集的各种用户信息,精准分析,理解客户日益多元的需求,改进设计,制定符合大数据时代的营销策略,提升广告产生的效果。

4. 用数据辅助报道

通过挖掘,展示数据的背景,找到关联,建立模式,根据数据新闻学,运用可视化手段与观众互动,报道即时发生的新闻,提供新闻链接,有助于公众理解新闻,思考新闻对人们生活的影响。媒体需要学会借助搜索引擎,学习社交网站,互相合作,把握社会。

5. 真正拓展大数据业务

推出新闻产品具有社交属性,投身真正大数据的海洋,提供免费的个性化应用,采集用户行为,了解阅读内容,抓取用户数据,进而判断用户在社交网络上的个性化内容,分析客户兴趣,实现归类发送。

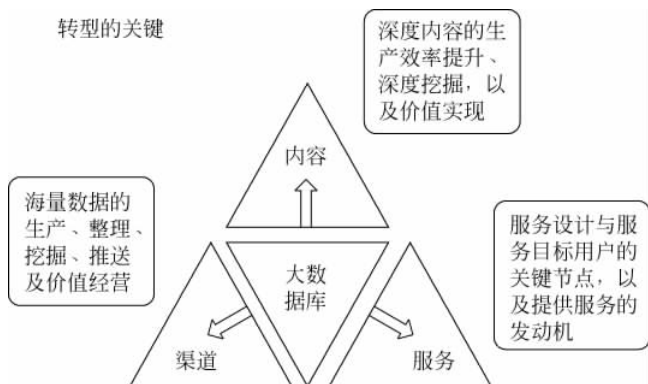


图 3-6 媒体大数据转型的关键

3.3.3 大数据引领新媒体发展

大数据结合新媒体,新的特征不仅仅是海量,而是增值并且全息可见,只有实现增值的数据才有更大的应用,只有实现全息可见数据才能被更广的传播。大数据时代呼唤大数据的融合,实现大数据的可流转才能真正发挥数据拥有的价值。大数据时代最核心的要求是数据开放,实现资源共享。如果在企业之间和社会各个方面不能做到数据的自由流动,那社会将变成一个个信息孤岛,大数据将无法发挥作用,不能得到价值最大化,所以只有实现数

据的交叉复用,达到全社会之间的自由流转,未来的商业才有可能实现繁荣。

消费者存在着信息视域过窄的问题,信息量在不断增加,但是消费者个体很难分析筛选大数据,处理信息和过滤信息的能力有待进一步提高,同时消费者出现长尾化的发展趋势。新媒体时代数据形式产生了巨大的变异,结构化数据变成半结构化甚至非结构化的数据,比如音频、视频之类。社交网络用户制造的信息也从单渠道变成多渠道。因特网和移动因特网结合,催生着跨网数据的发展。用户越来越希望利用碎片化时间,通过移动因特网获得有价值的信息,客户体验迅速下降,用户甄别信息能力与日俱减,用户兴趣数据与日俱增,所以个性化大数据是一个发展方向。

很多企业建立了日益庞大的数据平台,推动数据有效流转,在数据提供方与需求方之间自由流动。科研机构需要大量的原始数据,希望能够用来进一步做深度研究,个人用户和个人终极应用开发者也需要用户。

大数据和新媒体面临很多挑战,构建有效的消费者信息库,可以有助于企业精准出击,实现多维数据处理和实时计算,快速找到不同用户的兴趣,还有广告的信息关注怎样能够更有效、更精准地找到目标客户群。新媒体时代大数据环境下,能够非常精准预测,掌握每一个广告的投入,分析用户行为,预测广告效果,促进客户购买。

大数据时代传统媒体面临转型的难题,如何发展不仅仅是技术问题,更是战略问题,将会深刻的影响未来的媒体形态,改变现有的媒体格局。

1. 数据资源助推媒体转型

近年来,大数据时代伴随着信息爆炸式增长而来,爆发式增长的数据量带来了数据储存方式的革命,信息存储成本只是 10 年前的 1%,在 2000 年全球只有 1/4 的数据是以数字化的方式储存的,而到了 2007 年,只有 7% 的数据储存在报纸、书籍与图片等传统媒介上,其余数据全部是数字数据。新媒体的价值就体现在数据分析上,进入数字时代和智能时代,信息能够挖掘出规律,数据就是知识的基础,通过大数据分析工具帮助正确决策的数据就是最重要的资源。

大数据时代的信息不仅仅是新闻之类,而是各种各样丰富多彩的数据。媒体出现新的信息生产方式,应用新的传播方式,成为多元化媒介,不仅仅是生产数据,更要分析数据,解读信息,传播舆论,职能多元,为受众提供分众化服务,注重用户体验,实现媒体发展的大数据之路。

2. 量身打造体现发展新思路

目前门户网站互相模仿,网络媒体同质竞争,媒体和门户网站应该避免恶性竞争,利用大数据,建立关系链,为用户考虑,细分筛选,精准推荐,内容整合,通过数据分析,针对受众感受,满足不同主体的个性化要求,实现专业化发展,提供新闻资讯,重视客户体验,成为社交媒体。

大数据提供了新媒体发展的理论背景与实践手段,有助于媒体掌握大量数据源,帮助门户网站实现转型,提供了良好契机,制定了发展战略。

3. 挑战机遇赢得大数据时代的主动权

大数据时代,传统媒体需要转型,结合技术发展与客户需要找到适合自身特色的发展道路,应该思考传播规律,了解自身实际,把握受众需求,赢得机遇,接受挑战。

大数据新媒体的战略决策能力很重要,需要应对快速增长的数据,需要投入带宽,加大存储设备等基础设施方面的投入,考验媒体决策者的胆魄和智慧,转型就会赢得主动权,不然就必然被淘汰。转型就要全面变革当前的报道形式,方方面面改造现有的运行体系。只有具备数据加工能力才能应对大数据时代,大数据新媒体需要拥有专门的数据分析方法,建立全新的使用体系,招聘新型的高端专业人才,建立专门的数据管理部门和分析专家,在大数据时代的转型道路上媒体要把握方向,将既有的投资、数据和价值观整合到新的业务中去,在更高层次上得以发展,积极谋略全局,着眼长远,赢得大数据时代的主动权。新媒体的迅猛发展产生了海量数据,这些数据使我们能够更好地洞察社会各方面的细微变化。深入挖掘新媒体大数据蕴含的价值,将有力助推中国经济转型。

3.4 大数据在医疗上的应用

由于云计算产生的各种商业模式诸如商务云、物流云、医疗云等的出现,商业利益进一步推动云计算不断向前发展。2011年,麦肯锡开始应用大数据解决问题,商业互动中激增的数据量和多种多样的数据种类推动大数据存储技术和分析技术的进步,现在的大数据分析技术比20年前能处理更大更多的实时数据,产生更大的商业价值。图3-7展示了大数据应用于医疗的驱动。



图 3-7 大数据应用于医疗的驱动

实施大数据分析项目,数据企业不仅需要掌握应用何种技术,而且应该了解使用的时机和地点。医疗行业伴随因特网公司较早前就开始利用大数据并发扬大数据分析的优势,海量数据和非结构化数据的挑战带来医疗机构对于医疗信息化的需求,并迫使其投入大量资金进行大数据分析。麦肯锡研究之后指出医疗行业对大数据应用的需求不亚于银行、电信

与保险等行业,大数据分析可以帮助美国的医疗服务业排除体制障碍并创造巨额附加价值,提高医疗效率和医疗效果。

3.4.1 大数据改进临床决策支持系统

针对特定病人的最佳治疗途径可以通过全面分析病人特征数据和疗效数据进行多种干预措施的有效性比较来找到。医生确定临床上最有效和最具有成本效益的治疗方法就是通过大数据技术将医疗仪器精准分析病人体征产生的数据与疗效数据进行分析,减少过度治疗或治疗不足,因为过度治疗与治疗不足都会给病人带来不利影响。临床决策支持系统可以有效提高医疗质量与工作效率,有助于医生更好的提出方案,防止药物不良反应,医疗服务提供方可以通过部署这些系统降低医疗索赔数和事故率。

大数据分析技术主要体现在对于非结构化数据进行快速分析,加强分析技术可以提高临床决策支持系统的智能水平,挖掘医疗文献数据建立医疗专家数据库,或者使用图像分析和识别技术识别医疗影像数据,都可以帮助医生在诊疗中提高效率与质量。此外,临床决策支持系统还可以完成助理医生与护理人员在医疗流程中的大部分工作,提高医生治疗效率,避免陷入耗时过长的简单咨询工作中。

采集医疗过程中的工作数据,可以提高医疗服务的透明度与质量,进而推动医疗机构实现绩效的提升。信息透明增进医疗操作的流程优化,对整个绩效数据集实行数据分析有助于做出可视化的流程图,分析临床变异的数据,判断医疗废物的来源,提供服务质量的数据都可以实现绩效的不断提高,帮助医疗服务机构提高服务水准,带来竞争力。精简业务流程可以降低成本,找到效率更高的员工,提升护理质量,最终实现病人的更好体验,从而给医疗服务机构增加业绩,并且增长潜力。

3.4.2 大数据助推医疗产品研发

利用大数据医疗产品公司可以提高研发效率,在新药物的研发阶段,医药公司可以通过数据建模,分析投入产出比,确定最佳的资源组合,采集药物临床试验的所有数据集,建立相关模型,预测产品的安全性、有效性和潜在的副作用,评价整体的试验结果,建立模型,预测药物临床结果,选择最优药物,进行临床试验,降低研发成本。除了研发成本外,医药公司可以推出治疗成功率更高的药物,并将药物更快推向市场,使用预测模型可以帮助医药企业将研发新药推向市场的时间比原来提早3~5年。

提高临床试验设计水平,使用更好的统计工具和算法,有助于加快临床试验。评估患者,挖掘病人数据可以加快临床试验进程,找出最合适的临床试验患者,临床试验基地就能指出更有效的临床试验设计建议,在试验患者群体的规模和特征两者之间找到平衡可能是更理想的。确定药品更多的适应征和发现副作用需要分析临床试验数据和病人记录,分析病人临床试验数据,进行药物定位,检测药物不良反应,保证上市药品的安全。

发展个性化治疗需要通过例如基因组数据之类的大型数据集进行分析是另一种在研发领域有前途的大数据创新,在药物研发和用药过程中应该考虑个人的遗传变异因素,例如通过考察遗传变异、对特殊药物的反应和对特定疾病易感性的关系等。

3.4.3 大数据催生新医疗服务模式

大数据分析可以给医疗服务行业带来新的商业模式。分析患者的临床记录,处理其医疗保险数据集,将改善医疗支付方的决策能力,医疗服务提供方同样受益于医疗数据的分析处理。医药企业通过医疗数据不仅可以提高药品的疗效,而且有利于药品的销售。医疗保险数据集的整合处理有助于加快医学的发展和医疗保健行业的市场扩张。非营利性组织运营的网站网络平台是潜在的大数据启动的商业模型,大量有价值的数据已经在这些平台产生,这些网上互动信息平台就是最好的医疗数据来源。

大数据的使用可以有效地改善公众健康,通过整合全国各地的电子病历数据库,公共卫生部门可以实现全面的疫情监测,快速进行响应,控制传染病,这将减少医疗索赔支出,降低传染病感染率。卫生部门检测新传染病的速度大大提高,疫情快速得到有效控制,降低感染传染病的风险。及时提供准确的公众健康咨询,建立公众健康风险意识,可以帮助人们创造更好的生活。

由于医疗行业在服务亿万民众时必须面对海量的医疗健康数据处理需求,具有关系民生大计的特殊地位,因此成为国内率先启动大数据应用的先锋行业之一。大数据的应用首当其冲的就是智慧医疗,在医疗数字化的过程中,病历、影像、远程医疗等都会产生大量的数据,医院成了大数据产生的重要来源,把医疗大数据转换为经济价值的关键是能够提取出与诊疗有关的数据。作为实现智慧医疗的重要手段,数据分析将帮助解决医疗服务质量欠佳、医疗资源分布不合理和医疗体系效率较低这三大问题。

根据“十三五”规划中有关医疗卫生行业信息化建设规划,我国将重点建设国家级、省级和地市级三级卫生信息平台,建设电子档案和电子病历两个基础数据库,过去由于缺少统一的电子病历系统(EMR)标准,中国的电子病历系统发展比较缓慢,医院之间不能实现病患信息共享,医疗服务水平也因此受到影响。为改善这一现状,国家会逐渐加大对电子病历的投入,各级医院也将适应这一趋势,加大在数据中心、IT 外包等领域的投入。随着医疗和健康数据的急剧扩容,大数据出现几何级的增长,未来信息化工作的重要方向就是充分利用包括影像数据、检验检查结果、病历数据、诊疗费用等在内的各种数据,搭建合理先进的数据平台,服务广大医务人员,帮助患者康复,协助科研人员,支持政府决策者。

伴随大数据广泛应用于生命科学研究过程,医疗行业面临巨大挑战,其数据规模、多样化程度和增长速度都是空前的。一个基因组序列文件大小约为 750MB、一张普通 CT 图像含有大约 150MB 的数据、一个标准的病理图则接近 5GB,这些数据量乘以人口数量和平均寿命,导致一个社区医院累积的数据量就可达数 TB 甚至数 PB 之多,而且其中还附含非结构化数据,如图像、视频等。图 3-8 展示了医疗大数据的数据组成,图 3-9 展示了医疗大数据的数据来源。

医疗行业的大数据集繁杂量大,信息价值也多样且丰富,对其进行有效的存储、处理、查询和分析就可对于各层次决策服务,小到某个临床医生做出更为科学准确的诊断和用药决策,或根据患者潜在需求帮助某个医院开发全新自动服务及个性化服务,大到相关研究机构突破医疗方法和药物革新,或支持地区甚至全国医疗行业主管部门优化服务配置与医疗资源。图 3-10 展示了医疗大数据平台。

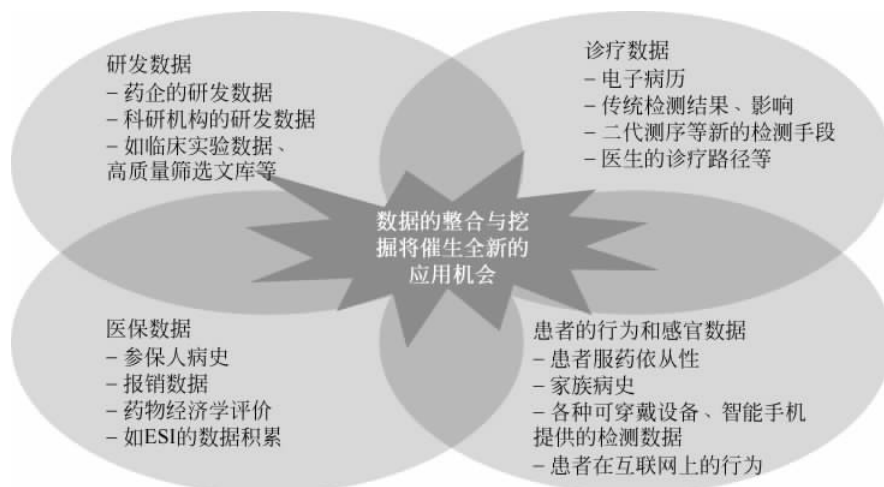


图 3-8 医疗大数据的数据组成

数据来源包括哪些？

1. 制药企业/生命科学

2. 临床决策支持&其他临床应用(包括诊断相关的影像信息)

3. 费用报销，利用率和欺诈监管

4. 患者行为/社交网络

我们如何利用大数据创造价值?(示例)

1. 个体化医疗

2. 临床决策支持

3. 欺诈监测得以加强

4. 由生活方式和行为引发的疾病分析

图 3-9 医疗大数据的数据来源



图 3-10 医疗大数据平台

3.5 大数据在教育上的应用

3.5.1 大数据教育与传统教育的优势

在教育领域中,较之于传统数据,大数据有着自己独特的优势:

传统数据主要用于辅助教育政策的宏观决策,针对宏观整体的教育状况进行分析决策。而大数据的透析可以针对个别的、微观的受教育者在课堂的状况,及时调整教学行为,实现个性化教育。

从误差大小比较看,传统数据使用阶段性评估方法,在采样中容易出现系统误差,会造成评估分析的较大误差。而大数据的采样采用即采即用或现象记录的技术性方式,系统误差较小。

数据采集的来源不同、数据应用的方向不同,这是大数据与传统数据的最本质区别。传统数据通过考试或者量表调查对学生数据进行周期性、阶段性采集,依靠数据对学生的生理和心理健康、学习状态及对学校的满意度来进行评估。信息采集具有事后性、阶段性而非实时性,并且会对被采集者(学生)造成压迫性。与之相应的,大数据采集是过程性的,关注每一个学生在上课、作业、教学互动过程的每个微观表现,采集在学生不自知的情形下开展进行,不影响学生的正常学习和生活。这些数据的获取、整理、采编、统计、分析需要经过专门的程序和专业的技术人员高效率地完成。图 3-11 展示了教育大数据系统工作流程。

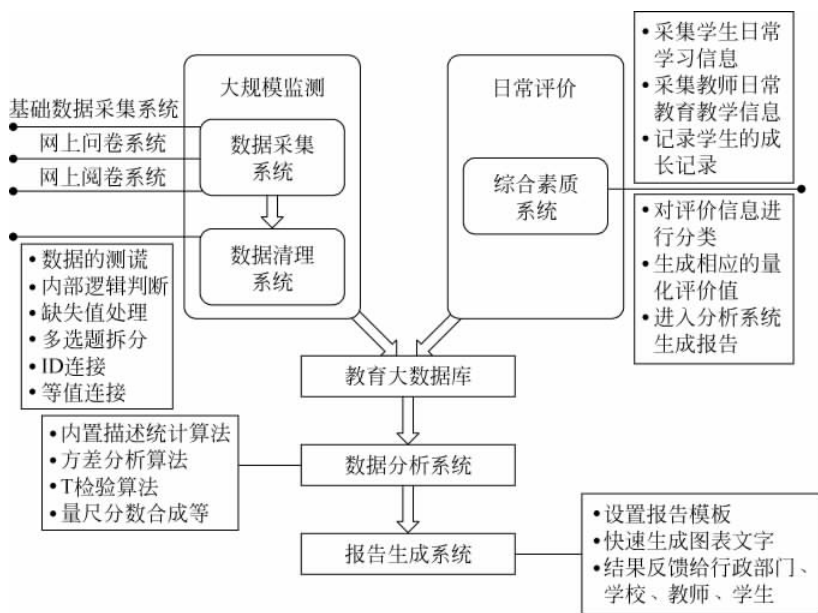


图 3-11 教育大数据系统工作流程

3.5.2 大数据教学模式的不断改善

随着因特网信息技术的高速发展,大数据成为众人瞩目的焦点,教育作为一个大数据应

用的重要领域,必将发生革命性的变化。

在教育中,大数据的运用可以改善学生的学习成绩,为学生提供个性化服务。通过大数据分析可以发现常规研究中所忽视的重要信息,革新教师的教学模式、改变学生的学习效果、优化教育政策的制定方式方法。目前,网络在线教育和大规模开放式网络课程就是大数据在教育中的典型应用。

美国国家教育统计中心等政府机构已经开始从事这项工作,他们在各教育机构收集学生行为、考试成绩、职业规划等重要信息,并利用大数据方法进行分析。通过对大数据的运用,美国教育部创建了学习分析系统,建立数据挖掘、数据模化和典型案例的联合框架,并以此向教育实施者提供更多、更好、更精确的信息,从而帮助其回答学习者如何学习等不太好回答的现实问题。

美国联邦政府教育部为了顺应并推动这一趋势,一项大数据计划于2012年在公共教育中被实施,该计划斥资2亿美元,推动大数据在改善教育中的运用。2014年4月10日,美国联邦教育部技术办公室发布了该计划的部分综述数据和案例,并宣布从财政预算中拿出2500万美元用于教育数据挖掘和学习分析。

3.5.3 教育大数据市场的广阔前景

美国的一些企业已经成功地在教育中实现了大数据处理的商业化运作。如全球最大的IT厂商IBM公司与亚拉巴马州的莫白儿县公共学区进行合作,通过对学生数据探测和干预,改善学生的学习成绩。在IBM公司的技术支持下,公司建立了跨校学习数据库,收集了100多万名学生的相关记录和700多万个课程记录的海量数据,软件分析结果不仅能够显示出学生的成绩、出勤、辍学率、入学率的趋势,还能够让用户探测性地预知导致学生辍学和学习成绩下滑的警告性信号;允许用户发现那些导致无谓消耗的特定课程,揭示何种资源和干预是最成功的;通过监控学生阅读电子材料情况、网络交流情况、电子版作业提交情况、在线测试情况,可以让老师及时诊断每个学生的问题所在,以备及时提出改进建议。

在未来教育中,高等教育的趋向将是个性化的学习。在高等教育个性化新时代中课程材料将适应性地满足每个学生的学习独特需求,电子教育、网络教育、主动教育是其显著特点。大数据分析可以应用于教育中的数据挖掘,目前教育机构已经积累了大量未结构化和结构化的数据,能使研究者有更多的新机会探究学生的学习环境。通过监测这些信息,形成教育大数据库,能够进一步总结教育规律,帮助教师理解学生、调整教育方案,掌握学生学习全过程,提供个性化的学习模式,提高学生的学习效果。

对于大数据的应用,在数据收集需要解决以下几个关键问题。一是数据收集标准化。收集数据一开始就要标准化,使用直观的方法对输入数据分类,为数据分析打好基础。二是数据获得问题。解决好海量数据的获得不仅仅是技术问题,有时还会遇到法律问题和伦理道德问题。三是数据收集者数量和质量问题。既要满足收集速度和精度要求,又需要满足数据质量保证要求。总之,大数据的教育应用可以为学生提供一个量身定做的个性化学习环境,一个教育问题早期预警系统,一个灵活调整的可控教育系统;为教师了解学生学习途径和方法提供了崭新的、可视的、可量化的新手段。

3.5.4 大数据变革教育应用的实践措施

(1) 对学生的发展进行多元评估。学生好成绩的取得有两种途径：一是依靠出色的逻辑思维能力取得，二是依靠记忆力取得。依靠记忆力可以取得好成绩，却无法培养学生的高级思维能力。依靠记忆力取得的好成绩可以掩盖学生在学习发展过程中的不足和风险。而大数据可以区分这两种差异，发现和区分这个差异有助于教育工作者及时对相应的学生提供针对性的帮助，发现成绩所反映不了的发展问题。

(2) 对学生学习实现过程性评估。教育中的大数据可以监控学生学习流程，发现学生的学习常态，通过数据流的变动分析，教育工作者可以观察到课堂流程改进的效果，促进教学的改革和发展；通过学生学习行为记录分析，捕捉学生在教室中的微观行为，实现大数据和课堂进程的结合，了解学生知识掌握度和兴趣点，促进教学效果反思。

(3) 对学生课外学习轨迹实现积累监测。学生家长使用手机可以及时获取学校通知公告，记录学生的家庭学习情况，通过区域性的数据统计，获取有研究意义、有价值的数据报告，掌握学生课外学习轨迹，加强学生学习氛围建设。

随着大数据时代的到来，社会科学领域随之受到冲击，从关注宏观整体走向关注微观个体，对于教育领域来说，大数据的应用让研究个体学习成为可能，让专门培养和针对性训练成为可能，从而比任何时候都更促进人类自身的发展和飞越。

3.6 本章小结

本章介绍了大数据的应用价值，数据的生成、分析、存储、分享、检索、消费构成了大数据的生态系统，通过大数据技术融合社会应用，让数据参与决策，发掘找到大数据真正有效的价值，进而改变人们未来，革新生活模式，产生社会变化，引发积极影响。

大数据必将在电子商务、金融、媒体、医疗、教育等行业得到广泛的应用，给人类社会带来深刻变革，产生巨大的价值。