

# 经济社会发展中的大数据应用

王钦敏

(中华全国工商业联合会, 北京 100035)

**摘要:** 通过全面系统地分析大数据特点和应用升级、大数据引发的新一轮产业革命、阐述大数据在国家基础数据库、电子政务业务、新型城镇化建设、物流公共信息服务、互联网金融、电子商务与企业信息化管理、世纪之村服务平台等行业的应用内容, 指出当前中国大数据使用面临的问题与挑战, 提出解决的方案和途径。

**关键词:** 经济社会; 大数据; 信息化; 应用

DOI: 10.11821/dlxb201505001

## 1 知识经济时代的信息化基础设施

知识经济时代是信息资源的时代。当前, 从国家到省区, 从市县到社区, 信息化工作开展得如火如荼。信息化包括政府信息化、企业信息化、社会信息化。信息化渗透于各级政府、各行各业、千家万户。如区域信息化的数字乡镇、数字城市、数字流域、数字地球等; 行业信息化的电子商务、数字农业、数字海洋、数字教育、数字文化等; 政府信息化的电子政务、公共信息服务、网络监管、社会治理信息化等; 以及综合类别的四化同步发展等。

国家信息化建设基础设施应包括网络与信息安全基础设施、信息资源基础设施和信息化应用基础设施等。目前网络与信息安全基础设施建设步伐很快, 信息资源基础设施建设还未形成体系, 信息化政策、法规、标准和规范, 特别是信息资源数据库基础设施建设需大力加强。而信息化应用基础设施建设非常薄弱, 需要引起重视。信息化应用基础设施包括: 基础信息公众服务、信息资源增值服务、数据内容的挖掘和应用服务、空间信息服务等, 缺乏统筹规划下的系统工程建设, 缺乏信息化应用服务人才和公共信息应用服务经纪人组织以有效支撑诸如企业信息化、国土资源与环境监测、智能城市发展、农村和农业信息化、数字教育、数字医疗、电子文化等信息化应用工程的实现。

当前国家信息化工程建设面临如下主要问题: 一是缺乏信息共享法规、技术标准和安全管理机制, 政务信息系统建设基本是各自为政, 数据拥有者很少使用数据, 而数据应用者难以获取数据服务。二是数据内容和质量难以保障。知识经济时代, 数据内容成为经济活动的资源, 数据质量成为发展决策的生命线。缺乏数据, 信息化将无法实施; 数据质量低劣, 直接影响决策结果。三是缺乏信息技术人才和管理人才, 特别是开展数据挖掘和信息化应用软件开发和服务的人才队伍和发展环境, 使得信息化软体体系建设严重滞后于硬件体系建设。四是缺少可共享的基础信息资源库和基于网络的信息化公众

收稿日期: 2015-02-10; 修订日期: 2014-04-12

作者简介: 王钦敏(1948-), 男, 福建福清人, 博士, 研究员, 现任十二届全国政协副主席, 全国工商联主席, 中国民间商会会长。

注: 本文根据2015年2月5日作者在中国科学院地理科学与资源研究所高端学术讲座录音整理, 经本人审阅。

服务技术平台,难以向社会、企业和公众提供基础性、公共性的信息服务,信息服务业发展艰难。五是各类信息系统更新维护体制机制不合理,重建设轻维护,造成反复立项,信息化工程招标机制与信息服务业发展需求不吻合。

因此,有必要建立公益性信息化中介服务体系和服务机制,构建联接数据生产、信息产品开发和应用服务的经纪人和服务链,满足政府内部跨部门、跨行业的业务协同和综合应用,并为社会和企业信息化的应用需求提供强有力的公共技术和信息资源支撑。

## 2 经济社会发展中的大数据应用

### 2.1 大数据特点和应用升级

大数据(Big data)是指所涉及的数据量巨大,无法在合理的时间内用常规软件工具处理与管理的数据资源。大数据的特点有数据量大(Volume)、处理时效快(Velocity)、多元与多样性强(Variety)、价值潜力大(Value)等特点,目前主要包括政务数据、各行各业数据、社会数据(包括网络)三类。1980年,著名未来学家阿尔文·托夫勒在《第三次浪潮》中,将大数据称为“第三次浪潮的华彩乐章”。由于海量信息资源的获取技术、互联网传输技术、数据存储与处理技术等跨越式发展,“大数据”成为流行词。据估计,目前互联网上的数据每年增长40%~50%。2012年和2013年全世界新产生的数据量,等于历史上所产生数据量的总和。

知识经济时代,数据是有价资产和财富资源,也是经济社会发展创新驱动的引擎。政府治理和公共服务,企业生产、经营、交易,社会中人与人、人与社会互动,自然地理、资源、环境、生态、军事等的空间定位,网络交往等信息,其数据量和实时性要求将超越现有信息技术基础设施承载能力。

大数据来源包括:一是网络用户生成数据;二是搜索引擎、运营商、电子商务服务、金融服务等所产生数据;三是政府经济社会管理、生态环境监测治理存储的数据;四是科教文卫所产生的各类海量数据;五是各类文本、音视频、图片等数据。这些数据通过网络传输与应用,并由此衍生出更多数据。大数据已经渗透到各级政府、各行各业、千家万户,成为政府治理、企业生产和民众生活的基础元素。大数据挖掘和利用已经成为新的经济增长点和质量效益的源泉。如何盘活数据资产,为国家经济、社会、生态、文化和民生发展服务,是大数据的核心与潜在价值,也是云计算的内涵与活力所在。

知识经济时代,大数据的应用必将引发新一轮经济发展热潮。中央提出的信息化与现代化、城镇化和农业现代化的同步发展,吹响了以大数据应用为代表的知识经济新时代的集结号。今后几年,谁拥有数据挖掘技术和大数据应用的专业化应用技术与人才,拥有大数据资产和信息安全保护解决方案,谁将会占领发展先机。从数据层到信息层到知识层到智慧决策层的渐进升级过程,形成了数据的科技和价值升华,其中数据层是基础,通过数据挖掘技术使基础数据升级为信息层;再通过信息分析集成与提取技术升级成知识层,最后通过专家系统决策支撑技术和验证形成智慧决策层。

### 2.2 大数据引发的新一轮产业革命

2013年3月,奥巴马政府宣布推动大数据相关产业发展,并将数据定义为“未来的新石油”。未来,对数据的占有和控制甚至将成为陆权、海权、空权之外的另一种国家战略优势和核心资产。IBM公司认为,“数据将成为一切行业当中决定胜负的根本因素,成为人类至关重要的自然资源”。它已抛弃了PC(person computer),转向了发展软件和服务,将更专注于大数据分析带来的业务增长点。德国经济部认为目前已进入了新一轮产

业革命时期。第一次产业革命由蒸汽机引领,第二次由电气化引领,第三次由信息技术引领,第四次将由虚拟—自然融合系统(Cyber Physical System, CPS)引领(例如,在3D打印技术支持下,已出现了定制生产、创新设计、动态组织、协同制造的生产模式)。新一轮产业革命的特点是:

(1) 工业制造呈现数字化、网络化、智能化、个性化、绿色化趋势。表现为:① 新机器、新材料融合;② 新生产方式、新业态、新商业融合;③ 新生产力、新产业人员结构、新岗位融合;④ 新动力、新能源融合。

(2) 产业组织出现网络化、平台化、扁平化趋势。表现为:① 网络化。一是上下游企业之间的垂直互联;二是区域内企业的水平互联;三是生产者与消费者之间的互联。② 平台化。通过研发平台、服务平台(如生产性服务业)、营销平台(如电子商务等)、合作平台和信息平台实现与中小企业的连接、“到家”服务。③ 扁平化。集中垂直式管理被分散合作式(如定制生产、众包设计、动态组织、协同制造等)替代,生产者与消费者联系更加直接;这对传统的管理体制产生挑战。

(3) 商业模式从以厂商、集中规模生产、降低成本为中心转向以消费者为中心;基于网络的个性化服务成为竞争力和利润的新源泉。

(4) 国际产业竞争格局变革将改变传统国际产业分工体系。表现为:① 新兴经济体按传统做法承接产业转移和技术转移的难度加大,产业升级和价值链环节竞争更趋激烈,谁掌握了信息,谁就将占领高地;② 数字化制造将使得一些行业的规模效益变得不明显,个性化定制、分散生产成为新特点;进而影响国际贸易格局;③ 未来各国围绕基于知识的生产要素,如人才、创新技术、品牌等方面的竞争将更加白热化。

## 2.3 大数据的行业应用

**2.3.1 国家基础数据库和信息业务处理** 目前中国已经建成了覆盖13亿人口的国家人口基础数据库,提供户籍人口查询与管理服务等;建成了完整的企业法人基础数据库和空间地理基础数据库;建成了统计信息数据库。

国家围绕宏观经济调控,市场监管、社会管理和公告服务需要,正推进部门信息共享和业务协同,扩大应用领域,提高行政效率和服务水平。例如:税务总局、人民银行、公安部,会同财政部,就利用基础信息数据开展应用服务启动了部分应用试点和应用推广。

**2.3.2 电子政务业务协同** 电子政务的部门业务协同通过分布不同部门节点上的基础数据库,开展数据交换,以达到跨部门协同和联合监管的目的。例如福建省政务信息共享平台,整合十三个涉农业务部门信息资源,开展三农信息综合服务;先后开展了省级行政事业单位人员工资编审及发放的协同业务,企业养老保险人员养老金发放的协同业务,事业养老保险人员养老金发放的协同业务等。

**2.3.3 新型城镇化建设支持** 随着国家城镇化进程的加快,对地理空间信息资源需求极大,如高分辨卫星遥感数据、航空摄影遥感数据等遥感信息;不同比例尺的地形图、土地利用图、地籍图、建筑图、道路图等基础地理信息。在此基础上,建成了政务数据库、户籍数据库、法人数据库、社区数据库、车辆数据库、地籍数据库、建筑数据库、环保数据库、社保数据库、灾害数据库、消防数据库、交通数据库等城镇信息资源库。

**2.3.4 物流公共信息服务** 中国目前物流成本约占17%,而日本仅占5%~6%。因此,发展现代物流业是提升中国物流业水平和效益的当务之急。现代物流业发展需要多源、海量、时空数据资源,以及数据综合处理信息增值产品挖掘,物联网、云计算、移动通讯、数字地球等技术支撑等。



物流公共信息服务业务主要包括：① 建设链接国内、国际海关、贸易、国检、边检、海口等口岸部门网络信息平台，实现信息共享与联网电子化作业；② 为运输的供需双方提供桥梁的平台，为物流的供需双方提供信息沟通、资格鉴证、物流金融、运输配送与监管等物联网信息和增值服务等；③ 提供物流空间信息服务，包括物流客户定位、设施选址、车辆配置等，甚至到户服务。

**2.3.5 互联网金融** 传统金融服务实体经济的模式主要是：资金供需匹配要通过两类中介服务进行，一是银行的间接融资模式，通过从储户收集资金后对用户提供的金融借贷；二是直接融资模式，通过股票和债券市场开展金融服务。

互联网金融模式下，采用移动支付、社交网络和搜索引擎等，贷款、股票、债券等的发行和交易以及券款支付直接在网上进行，将对金融模式产生根本影响。

**2.3.6 电子商务与企业信息化管理** 互联网电子商务的应用推广重点是对潜在客户的个性化特征信息发现、挖掘和商业推荐。研究证明，全球84%的顾客相信来自朋友和家人推荐的、口碑不错的、评价度高的商业信息。广告推荐可以根据观众和预期结果设计推广活动，让消费者有动力把个人见解添加到对品牌的理解中，从而推荐给朋友和读者。此外，电子商务平台对顾客网络浏览、消费行为的收集和分析，可以帮助获悉顾客消费意图，实现更符合顾客需求的广告定投，促成交易，使顾客、商家、平台网站均获利。

2014年，阿里巴巴集团连续第六年举办“双11购物节”。当天共有2.4亿网民访问淘宝，交易峰值达285万笔/分钟，交易额达571亿元，产生快递包裹2.785亿件，释放了巨大消费潜力。“双11购物节”是全社会电商“生态圈”的总动员，消费者、商家、物流、银行、超市、生产商、服务商均成为“双11购物节”的网络节点，彼此互联互通、相互协作，共同创造了这一天互联网时代的商业奇迹。

“双11购物节”所产生的巨量的在线交易与物流管理业务，需要强有力的大数据处理技术支持，都是基于阿里云的开放数据处理服务平台完成，可在6小时内处理100 PB数据，相当于1亿部高清电影。当天支付宝系统采用“云支付”构架，处理能力可达每日10亿笔以上，具备更高的服务质量、安全性、稳定性，更低的系统成本。

苏宁云商积极探索传统零售企业转型互联网零售企业的方法和路径，依托科技打造云服务，通过SAP/ERP系统，把苏宁线上线下的资源与服务融为一体，为消费者服务，也为供应商服务，打造和谐高效的供应链以实现共赢，通过1200个坐席的大型呼叫中心，为顾客提供24小时服务，将线上的便利性与线下的体验功能进行完美融合，大大提高了客户响应能力，形成可持续发展的互联网零售模式，建设互联网化的门店，将原先纯粹的零售功能，升级为集展示、体验、物流、售后服务、休闲社交、市场推广为一体的新型实体门店，为消费者提供一个融合的全渠道购物平台。现拥有18万名员工，每年创造2000多亿销售额。

恒安集团从事妇幼卫生用品、纸类用品生产，年销售额200多亿元，员工近3万人。为了实现集团对销售业务的有效管控和协同，恒安集团自1998年开始信息库建设，先后经历了单公司信息化、集中信息化、扩展式信息化和全面智能信息化四个阶段，累计投入超10亿元，实现了公司内部销售、采购、库存、人力资源、资金流业务以及外部3000多个供应商、近2000个分销商和近百万家终端门店的业务协同和信息共享。随着业务体系的要求，增加新的模块，包括人力资源、生产制造、资金管理、预算管理等，又使财务结账时间从五天缩短到一天，客户订单及时响应率达到100%，发货及时率达到90%，库存占用率下降50%，大大优化了流程，提高了运营效率。

**2.3.7 世纪之村服务平台** 福建省世纪之村农村信息化平台业务，包括了农村电子政务、

农村电子商务、农村社区信息三方面。其中农村电子政务包括村务公开系统、三资管理系统、党员廉政监管系统、农村合作社管理系统、大学生村官创业系统。农村电子商务包括农家店交易管理系统、劳务需求发布系统、企业商品展示系统、农产品ERP管理系统、便民一卡通。农村社区信息包括传递社情民意、远程教育培训、文体健身娱乐、网络舆情监控系统等。

### 3 大数据的问题与挑战

#### 3.1 大数据面临的问题

大数据改变了社会生产方式、消费方式和生活方式，但同时伴随而来的是个人隐私权受到前所未有的挑战。目前有很多互联网公司在采集公众个人信息，并保存在其各自的大数据系统中，包括个人身份信息、银行账号信息、社会互通信息、行为习惯信息、位置信息等隐私信息。这些信息的关联聚合可以准确地还原并预测个人及社会生活全貌，当数据达到一定规模时将产生巨大的经济利益。但在缺乏授权的情况下，可能导致资源利益驱动下的信息资源窃取与掠夺行为，存有巨大的隐患与威胁。

目前国内对于个人隐私信息管理还比较落后，安全技术能力方面也存在较多薄弱环节，虽然此前发布了《信息安全技术公共及商用服务信息系统个人信息保护指南》，规范了个人信息处理的全流程活动，规定了个人敏感信息在收集和利用之前须获得个人信息主体明确授权，但作为一个推荐性标准，其实施取决于相关行业主体的自愿配合。

经济发展是时代主旋律，但如何抵御数据泄露可能导致的风险，既是技术问题，也是社会问题、法律问题、国家战略问题，更是个人信息保护权益。大数据的挖掘与利用需要有法可依。中国需要尽快制定“信息保护法”与“信息公开法”，既要鼓励面向群体而且服务与社会的数据挖掘，又要防止侵犯个体隐私的行为，既要提倡数据共享又要防止数据被滥用。如何认识大数据时代隐私安全保护的症结和难点，进而探索出科学可行的保护模式，是一个体系庞大的社会工程。

#### 3.2 大数据的技术挑战

大数据技术涉及很多环节，如数据采集、数据甄别、数据存储、数据挖掘处理、数据可视化、数据智能分析、数据安全及隐私保护等。以往数据存在数据库里，一般是静态的。现在，大数据每时每刻都在产生，必须边读取、边分析。因此，大数据的分析对计算机的体系与结构提出了很高的挑战。以往的商用数据库系统采用基于行的数据存储；在大数据存储体系下，数据多采用是基于列的存储，以便实时处理数据。此外，大数据处理迫切需要实现智能化。全球大数据量年增40%，但信息技术投入年增仅5%，投入跟不上大数据挖掘和应用的需要。

现在，与发达国家相比，中国在自主可控的大数据分析技术与产品方面有不少差距，数据库、数据仓库、商业智能分析软件等领域发展基础薄弱，已经落后于国外先进企业。最后，大数据人才问题尤其关键。既了解信息技术、数据资源、又了解行业需求的人才缺口太大。麦肯锡公司预测，到2018年美国的深度数据分析人才缺口为14万~19万人，亟需熟悉行业需求和技术应用的管理者150万人。对于中国而言，能够理解与应用大数据的创新人才更是稀缺资源，需要引起高度重视。