

# 大数据中的管理问题：基于大数据的资源观<sup>①</sup>

杨善林，周开乐\*

(合肥工业大学管理学院过程优化与智能决策教育部重点实验室，合肥 230009)

**摘要：**作为重要的战略资源，大数据中包含诸多关键的管理问题。文章首先评述了基于不同视角对大数据的认识，然后，从管理的视角看大数据，指出大数据是一类重要的战略性信息资源，并从复杂性、决策有用性、高速增长性、价值稀疏性、可重复开采性和功能多样性等6个方面探究了大数据资源的管理特征。最后，提炼并探讨了大数据资源的获取问题、加工问题、应用问题、产权问题、产业问题和法规问题等6个方面的关键管理问题。

**关键词：**大数据；资源观；管理特征；管理问题

**中图分类号：**F272      **文献标识码：**A      **文章编号：**1007-9807(2015)05-0001-08

## 0 引言

随着信息技术、生物技术、新材料技术和智能制造技术等技术的快速发展，诸多领域的数据量都呈现急剧增长态势，而且数据形态日益复杂。国际数据公司（IDC）的研究报告称，2011年全球有1.8ZB的数据被创建和复制，该报告预测2020年全球将拥有35ZB（1ZB=10亿TB）的数据量<sup>[1]</sup>，这标志着大数据时代已经到来<sup>[2]</sup>。

近年来，学术界和产业界提出了“大数据”的概念，对大数据研究与应用的关注度持续升温。2008年9月，《Nature》杂志推出了“Big Data”专刊<sup>[3]</sup>，关注如何处理正在产生的洪水般的大量数据，从互联网技术、网络经济学、超级计算、环境科学、生物医药等多个方面分析了大数据带来的挑战。计算社区联盟（Computing Community Consortium）于2008年发布了大数据报告“Big-Data computing: Creating revolutionary breakthroughs in commerce, science, and society”<sup>[4]</sup>，总结了大数据的相关技术和应用及其面临的挑战，并给出了政府投资的行动建议。《Science》杂志也在2011年2

月推出“Dealing with data”专刊<sup>[5]</sup>，指出了大数据带来的挑战与机遇并存。2011年6月，麦肯锡（McKinsey）公司在其发布的大数据报告<sup>[6]</sup>中指出“大数据时代已经到来”，并详细分析了大数据的影响、关键技术和应用领域等。

2012年以来，大数据获得了全球范围内更广泛的关注。2012年1月的达沃斯世界经济论坛上发布了大数据报告“Big Data, big impact: New possibilities for international development”<sup>[7]</sup>；2012年3月29日，奥巴马政府推出了“大数据研究和发展倡议（Big Data research and development initiative）”<sup>[8]</sup>，这标志着大数据研究和发展已经成为国家的发展战略；2012年5月，联合国“Global Pulse”倡议项目也发布了专题报告<sup>[9]</sup>，旨在尽可能具体和公开地界定发展大数据面临的挑战，并给出方法建议以解决其中的问题。

然而，现有关于大数据的研究主要基于数据的获取<sup>[10]</sup>、存储<sup>[11]</sup>、分析<sup>[12-14]</sup>、处理<sup>[15-16]</sup>、管理<sup>[17]</sup>和应用<sup>[18-19]</sup>等技术层面，一些研究聚焦于大数据的潜在商业价值<sup>[20-21]</sup>，预测大数据可能对经济社会产生的深远影响<sup>[22-24]</sup>。这些研究对大数

① 收稿日期：2013-08-11；修订日期：2015-01-27。

基金项目：国家自然科学基金资助项目（71131002）；中央高校基本科研业务费专项资金资助项目（JZ2015HGBZ0093）。

通信作者：周开乐（1987—），男，江苏铜山人，讲师。Email: zhoukaile@hfut.edu.cn

据发展过程中自身面临的诸多关键管理问题关注较少。

IBM 公司认为大数据是一类新的自然资源<sup>[25]</sup> ,高德纳咨询公司( Gartner) 也认为大数据是一种信息资产<sup>[26]</sup> . 虽大数据与自然资源有着一定程度的类似 ,但二者并不完全相同 ,大数据是一类特殊的战略性信息资源. 由于其自身特性 ,这类特殊资源为管理领域带来了新机遇的同时 ,也带来了诸多挑战.

当大数据被看作一类“资源”时 ,就不可忽略这种资源的获取问题、加工问题、应用问题、产权问题、产业问题和法规问题等 ,其中每个问题都是重要的管理问题. 因此 ,本文首先评述了几种对大数据的不同认识; 并根据大数据的资

源观 ,给出了基于管理视角的大数据定义 ,接着指出了基于管理视角的大数据资源 6 个方面的重要特征; 最后 ,基于大数据资源的特征 ,提炼并分析了其 6 个方面的关键管理问题;

### 1 对大数据的不同认识

大数据来源于生产生活的诸多方面 ,其应用又服务于生产生活各个方面. 如生产制造、物流交通、电子商务、医疗卫生、社交媒体、生物科学、资源环境、教育文化和公共安全等<sup>[27]</sup> .

近年来 ,基于不同的视角 ,学术界和产业界对大数据做出了不同的定义. 本文总结了几个有代表性的定义及其定义所基于的视角 ,见表 1.

表 1 关于大数据的几种定义及其视角

Table 1 Some definitions of Big Data and corresponding perspectives

来源	定义	视角
麦肯锡公司 ( McKinsey & Company)	大数据是个大的数据池 ,其中的数据可以被采集、传递、聚集、存储和分析. 与固定资产和人力资本等其它重要的生产要素类似 ,没有数据 ,很多现代经济活动、创新和增长都不会发生 ,这正成为越来越普遍的现象 <sup>[6]</sup>	强调大数据对全球经济社会发展的重要性 ,把大数据看作是一种与固定资产和人力资本类似的重要生产要素. 这一定义指出 ,大数据可以扮演一个重要的经济角色
高德纳咨询公司 ( Gartner Group)	大数据是大容量、高速度和形式多样的信息资产 ,它需要低成本的、形式创新的信息处理 ,以增强洞察力和辅助决策 <sup>[26]</sup>	认为大数据是一类信息资产 ,从技术特征、处理方法和应用价值 3 个方面对大数据做出了界定
IBM 公司	可以用 4 个特征来描述大数据 ,即规模性( volume) 、高速性( velocity) 、多样性( variety) 和真实性( veracity) ,这些特征相结合 ,定义了 IBM 所称的“大数据” <sup>[28]</sup>	指出了大数据 4 个方面的技术特征
维基百科 ( Wikipedia)	大数据是指规模庞大且复杂的数据集合 ,很难用常规的数据库管理工具或传统数据处理应用对其进行处理. 其主要挑战包括数据抓取、策展、存储、搜索、共享、转换、分析和可视化 <sup>[29]</sup>	主要从大数据的处理方法和处理工具的视角认识大数据
美国国家科学基金会( NSF)	大数据是指由科学仪器、传感器、网上交易、电子邮件、视频、点击流和/或所有其它现在或将来可用的数字源产生的大规模、多样的、复杂的、纵向的和/或分布式的数据集 <sup>[30]</sup>	主要基于大数据的来源和大数据的技术特征 ,强调了大数据来源的多样性和特征的复杂性

从表 1 可以看出 ,目前人们认识大数据的视角是存在差异的 ,造成这种差异的原因主要是人们分析大数据的背景和应用大数据的目的不同. 除了上述具有代表性的定义外 ,还有许多关

于大数据的不同定义 ,然而对大数据的认识就像是盲人摸象 ,每个定义都是基于特定的视角 ,如大数据的技术特征、应用价值、来源和处理方法等.

## 2 从管理视角看大数据

大数据的产生和发展正在对实际的管理活动产生深刻的影响。在广度上,诸多领域的大数据对不同产业的发展以及人们的生活方式都产生了深刻影响;而在深度上,大数据正在影响人们生产和消费信息的方式,从而对管理机制和决策模式产生显著影响;在细微之处,大数据存在于人们日常生活的方方面面,并潜移默化地影响着人们的行为方式。

### 2.1 基于管理视角的大数据定义

前面给出的大数据的不同定义主要基于大数据的技术特征、处理方法和应用价值等视角。然而,为了分析和研究大数据中的管理问题,提出基于管理视角的大数据定义如下:“大数据是一类能够反映物质世界和精神世界运动状态和状态变化的信息资源,它具有复杂性、决策有用性、高速增长性、价值稀疏性和可重复开采性,一般具有多种潜在价值”。

基于大数据的资源观和管理的视角,认为大数据是一类能支持管理决策的重要资源。因此,为了有效管理这种资源并充分发挥其潜在价值,就需要研究并解决这种资源的获取、加工、应用、产权界定、产业发展和政策保障等管理问题。对大数据资源管理特征的准确认识是研究其具体管理问题的前提,下面将从复杂性、决策有用性、高速增长性、价值稀疏性、可重复开采性和功能多样性等6个方面逐一分析大数据资源的管理特征。

### 2.2 基于管理视角的大数据特征

一定意义上,大数据资源与煤、石油、天然气等自然资源有一定的相似性,至于其管理特征,则有明显不同。作为一类重要的信息资源,大数据具有以下特征:

1) 复杂性 正如很多定义所指出的,大数据的形式和特征是极其复杂的。大数据的复杂性除了表现在其数量规模之大、来源的广泛性和形态结构的多样性外,还表现在其状态变化和开发方式等方面的不确定性;

2) 决策有用性 大数据本身是客观存在的大规模数据资源,其直接功用是有限的。通过分析、挖掘和发现其中蕴藏的知识,可以为各种实际

应用提供其它资源难以提供的决策支持,大数据的价值也主要通过其决策有用性体现;

3) 高速增长性 大数据资源的这一特征与石油等自然资源是不同的。不可再生的自然资源的总存量会随着人类不断开采而逐渐减少,而大数据却具有高速增长性,即随着不断开采,大数据资源不仅不会减少,反而会迅速增加。大数据资源的增加是指数性的,甚至呈现爆发性态势。例如在互联网上,通过搜索引擎、社交媒体和电子商务等方式,每秒会产生大量的数据。据统计平均每秒有200万用户在使用谷歌搜索,Facebook用户每天共享的内容超过40亿条,Twitter每天处理的推文数量超过3.4亿<sup>[31]</sup>。由于大数据的数据量之大及其高速增长的动态性,而且大数据的内容及其数量增长容易受到偶然因素的影响,如社会突发事件在社交媒体上的传播<sup>[32]</sup>,因此,利用大数据支持管理决策面临的主要挑战就是实效性问题;

4) 价值稀疏性 大数据的数据量之大在带来了诸多机遇的同时,也带来了不少挑战。其主要挑战之一就是大数据价值的低密度问题,大数据资源的数量虽大,但其中蕴藏的有用的价值却是稀疏的,这就增加了开发和利用大数据资源的难度;

5) 可重复开采性 自然资源的开发利用过程通常是不可重复的,随着不断开采,其存量会逐渐减少。但对于大数据资源,它们可以被重复开采。对于给定的大数据资源,任何拥有该资源使用权的人或组织都可以对其进行开采和挖掘。一些人进行开采之后,该大数据资源仍可以被其他人或组织继续开采和挖掘;

6) 功能多样性 对于一些自然资源,如煤、石油和天然气等,它们的功用是有限的。而对于特定的大数据资源,基于不同的开发目的和方式,具有多样化的功能。例如:

——基于社会管理的目的<sup>[33]</sup>,大数据可以用于医疗卫生管理、舆情监控和公共安全管理等;

——基于商务管理的目的<sup>[34]</sup>,大数据可以用于社交网络分析、商业模式创新和市场营销等;

——基于企业管理的目的<sup>[35]</sup>,大数据还可以用于生产销售管理、客户关系管理和人力资源管理。

除了正面积的功能,大数据还存在着诸多

风险和潜在危害性. 因此, 在开发利用大数据资源时, 还应该防范由于方式不当或非法利用可能对个人、企业和组织甚至国家和社会带来的严重危害.

### 3 大数据中的管理问题

基于管理的视角, 当大数据被看作是一类“资源”时, 为了有效地开发、管理和利用这种资源, 就不可忽视其获取问题、加工问题、应用问题、产权问题、产业问题和法规问题等相关的管理问题.

#### 3.1 大数据的获取问题

正如自然资源开发和利用之前需要探测, 大数据资源开发和应用的前提也是有效地获取<sup>[36]</sup>. 美国 Google、Apple 和 Facebook 等大型信息技术企业已经收集并存储了大量数据, 掌握了较为成熟的大数据技术和管理机制, 并建立了自身比较完善的大数据技术体系和服务框架. 中国的相关企业和组织也已经意识到大数据资源的重要价值, 如中国的百度、阿里巴巴、腾讯等信息技术企业已经将大数据相关业务作为重要的发展战略之一, 尝试推出了相关服务. 中国要想发挥大数据在经济社会发展中的价值和作用, 政府就要采取各种有效政策措施推动和保障大数据采集和获取的相关产业发展.

大数据的获取能力一定意义上反映了对大数据的开发和利用能力, 大数据的获取是大数据研究面临的首要管理问题. 制定大数据获取的发展战略、建立大数据获取的管理机制、业务模式和服务框架等是这一方向中需要研究的重要管理问题.

#### 3.2 大数据的处理方法问题

大数据资源的开发和利用主要基于传统的计算机科学、统计学、应用数学和经济学等领域的方法和技术. 例如, Wu 等<sup>[37]</sup> 基于数据挖掘视角, 提出了数据驱动的大数据处理模型. 麦肯锡公司也在其大数据报告<sup>[6]</sup> 中总结了可用于大数据处理的关键技术.

除了大数据的基础处理方法外, 基于不同的开发和应用目的, 如市场营销、商务智能、公共安

全和舆情监控等, 还需要特定的大数据资源开采技术和处理方法, 称之为应用驱动的大数据处理方法.

针对具体的应用问题, 设计合理高效的应用驱动的大数据处理方法是这一方向需要研究的重要管理问题之一. 大数据的处理方法问题是大数据发展中重要的基础性管理问题.

#### 3.3 大数据的应用方式问题

大数据资源的应用需要考虑的重要问题是如何将大数据科学与领域科学相结合. 大数据资源的应用方式可以分为 3 大类, 首先是在领域科学的框架内来研究和应用大数据资源, 称之为嵌入式应用; 第二种方式是将大数据资源的开发和利用与领域科学相结合, 二者相互作用, 这种方式称为合作式应用; 此外, 大数据资源的开发应用还可能引起领域科学的变革, 称作主导式应用. 如图 1 所示.

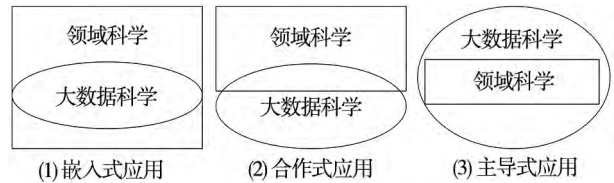


图 1 大数据的应用方式

Fig. 1 Application modes of Big Data

为了更好地发挥大数据的决策支持功能, 其应用方式问题是不可忽视的重要管理问题. 这一方向具体的研究内容包括对不同的应用方式进行选择和设计等.

#### 3.4 大数据的所有权和使用权问题

大数据在哪里、谁拥有大数据资源? 这是大数据发展过程中必须回答的问题.

目前大部分大数据资源掌握在大型企业或组织的手里, 例如:

- 互联网公司: 新浪微博、Facebook 和 Twitter 等;
- 电子商务企业: 阿里巴巴、Amazon 和 Ebay 等;
- 搜索引擎公司: 百度和 Google 等;
- 软硬件服务商: IBM、Apple 和 Microsoft 等;
- 大型企业或公共部门: 沃尔玛、国家电网等.

目前, 大数据主要掌握在大型企业或组织手中, 而个人拥有的数据则相对较小. 这就为个人利用大数据开展研究和应用带来了挑战.

然而, 这些企业或组织拥有的“大数据”是由大量“小数据”组成的, 而“小数据”是由一个个用户产生的, 如社交媒体上用户发布或交互的信息、用户网上购物的消费记录、使用搜索引擎的搜索记录和用户消费数据等. 产品和服务提供商垄断所有用户产生的这些数据, 对用户来说是不公平、不合理的. 对于无法利用这些数据开展研究的研究人员来说也是不公平的<sup>[38]</sup>.

因此, 通过有效的管理机制来界定大数据资源的所有权和使用权是至关重要的管理问题, 如 Lazer 等<sup>[39]</sup>指出, 需要建立产业界和学术界协作和数据共享的稳健模型, 从而在促进科学研究的同时保护用户的隐私. 解决大数据的产权问题需要回答以下几方面的问题: 谁应该享有大数据资源的所有权或使用权? 哪些大数据资源应该由社会公众共享? 如何有效管理共享的大数据资源, 以实现在保障安全和隐私的同时, 提高使用效率?

大数据背景下的数据所有权界定要比传统数据库环境下的产权界定问题<sup>[40]</sup>复杂得多. 对大数据进行分类是界定其所有权和使用权的重要方式之一. 基于云计算中对不同类型“云”的划分思想, 本文提出可以将大数据划分为私有大数据 (private Big Data)、公有大数据 (public Big Data) 和混合大数据 (hybrid Big Data), 如图 2 所示.



图 2 大数据资源的分类

Fig. 2 Classification of Big Data resources

各类大数据资源的简要描述和特征如表 2.

表 2 不同类型大数据资源的简要描述

Table 2 Brief introduction of different types of Big Data resources

大数据资源类型	描述
私有大数据 (private Big Data)	私有大数据是由于安全性或保密性等特殊要求限制, 仅能由某些特定企业或组织所有、开发和利用的大数据资源
公有大数据 (public Big Data)	公有大数据是可以由公众共享的大数据资源, 公有大数据可以为大数据相关科学研究的开展提供便利
混合大数据 (hybrid Big Data)	混合大数据介于私有大数据和公有大数据之间, 可以通过交易、购买或转让等方式在私有大数据和公有大数据之间转换

大数据所有权和使用权的界定问题是其面临的管理问题中的热点和难点, 也是大数据产业健康发展的重要保障.

### 3.5 大数据产业发展问题

大数据的完整产业链包括数据的采集、存储、挖掘、管理、交易、应用和服务等. 大数据资源产业链的发展会促进原有相关产业的发展, 如大数据对传统数据采集、存储和管理的软硬件设备要求更高, 会促进数据采集、存储和管理软硬件相关产业的进一步发展.

大数据资源产业链的发展还会催生新的产业, 如大数据资源的交易会促使以大数据资源经

营为主营业务的大数据资源中间商和供应商的出现. 此外, 还有可能出现以提供基于大数据的信息服务为主要经营业务的大数据信息服务提供商. 如基于服务的决策支持系统 (DSS in cloud)<sup>[41]</sup>将分析和大数据放到云端, 这种决策支持系统服务会促进大数据与云计算交叉产业的形成和发展.

在研究基于服务的决策支持系统时, Delan 和 Demirkan<sup>[42]</sup>提出了数据即服务 (data-as-a-service)、信息即服务 (information-as-a-service) 和分析即服务 (analytics-as-a-service) 的相关概念. 这里, 考虑大数据产业的发展问题, 也将大数据的服务模式分为 3 类, 分别为数据即服务 (data as a service, DaaS)、挖掘即服务 (mining as a service,

MaaS) 和分析即服务 (analytics as a service , AaaS) ,如图 3 所示.

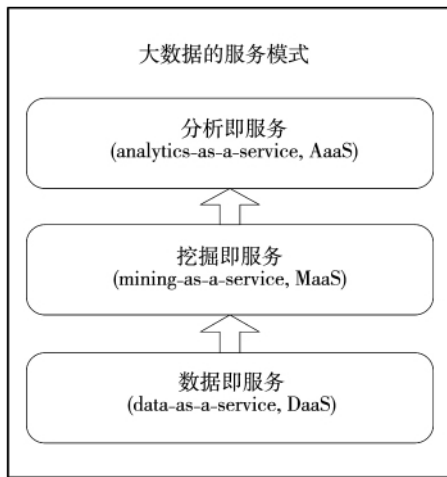


图 3 大数据的 3 种服务模式  
Fig. 3 Three service models of Big Data

DaaS 是基础性服务 ,是以数据的采集与提供为主要业务内容的服务模式; MaaS 则是在 DaaS 的基础上 ,以提供专业的数据挖掘服务为核心的服务模式; AaaS 是高级的服务模式 ,是在 DaaS 和 MaaS 的基础上 ,对大数据中挖掘出的知识进行分析和可视化展示的服务模式. 大数据的上述 3 种服务模式都会推动或催生相关产业的发展.

对大数据产业发展问题的研究是实现大数据潜在商业价值的重要环节 ,而大数据产业发展中面临着的一系列比传统商业环境下更复杂的优化问题、决策问题、预测问题和评估问题 ,这些都是大数据产业发展中需要研究的重要管理问题.

### 3.6 大数据的相关政策和法规问题

大数据资源的发展还必须有完善的政策和法规支撑. 例如通过对大数据资源的所有权界定 ,有效维护大数据所有者的权利 ,促进大数据产业的健康发展. 数据的安全与隐私保护问题<sup>[43]</sup>是大数据资源开发和利用面临的最为严峻的问题之一 ,除了在安全和隐私保护技术方面不断突破外 ,还需要相关法律法规对大数据资源的开发和利用进行严格有效的规范.

### 参考文献:

[1] IDC. The 2011 digital universe study: Extracting value from chaos [EB/OL]. <http://www.emc.com/collateral/demos/microsites/emc-digital-universe-2011/index.htm> , 2011.

[2] Brown B , Chui M , Manyika J. Are you ready for the era of 'big data'? [J]. McKinsey Quarterly , 2011 , ( 4 ) : 24 - 35.

此外 ,大数据资源开发也是一把双刃剑. 对大数据资源进行正向积极的开发和利用 ,可以给人们的生活带来更多便利、促进经济社会的发展. 而大数据也有可能被不法分子反向利用 ,从而给经济社会带来严重危害. 因此 ,就需要严格的法律法规以指导、规范和约束大数据资源的开发和利用.

大数据的研究和应用才刚刚起步 ,相关的政策法规还很不完善 ,这就需要管理学者的深入研究 ,为政府部门和行业组织制定并不断完善大数据产业发展的相关政策法规提供有力的决策支持.

## 4 结束语

大数据的研究和应用应该遵循顶层设计、实践导向和理论升华的科学路径. 顶层设计就是要制定好大数据资源产业的发展规划 ,科学引导大数据的研究工作; 实践导向就是要在大数据发展实践中发现科学问题 ,并解决实际问题 ,坚持从实践中来到实践中去. 事实上 ,在“大数据”的概念被提出以前 ,大数据已经在很多领域被研究和应用 ,已经有很多关于大数据获取、存储、处理和应用的研究和应用成果 ,这表明 ,大数据的实践应用是先于理论研究的 ,当然深入的理论研究一定能够促进大数据应用实践更好更快的发展; 理论升华就是要将解决大数据实践问题的科学方法凝练总结 ,建立并完善大数据相关的理论体系.

在云计算<sup>[44]</sup>和物联网<sup>[45]</sup>等新兴信息技术环境下 ,诸多领域的的数据正在急剧增加. 大数据产业的发展除了需要更先进的数据采集、传输、存储、分析和挖掘等相关技术 ,同样需要更加科学有效的管理技术与方法的支撑. 本文尝试从管理视角审视大数据 ,提炼了大数据资源中存在的诸多关键管理问题 ,但本文仅做了简要分析 ,其中每个问题都有待管理学者系统地研究 ,以更好地发现和利用大数据中的潜在价值 ,并推动大数据产业的健康发展.

- [3] Nature. Big data [EB/OL]. <http://www.nature.com/news/specials/bigdata/index.html>, 2008.
- [4] Bryant R E, Katz R H, Lazowska E D. Big-data computing: Creating revolutionary breakthroughs in commerce, science, and society [R]. [http://www.cra.org/ecc/files/docs/init/Big\\_Data.pdf](http://www.cra.org/ecc/files/docs/init/Big_Data.pdf), 2008.
- [5] Science. Special online collection: Dealing with data [EB/OL]. <http://www.sciencemag.org/site/special/data/>, 2011.
- [6] Manyika J, Chui M, Brown B, et al. Big data: The next frontier for innovation, competition, and productivity [R]. McKinsey Global Institute, 2011: 1–137.
- [7] World Economic Forum. Big data, big impact: New possibilities for international development [EB/OL]. [http://www3.weforum.org/docs/WEF\\_TC\\_MFS\\_BigDataBigImpact\\_Briefing\\_2012.pdf](http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf), 2012.
- [8] Obama administration unveils “Big Data” initiative: Announces \$200 million in new R&D investments [EB/OL]. [http://www.whitehouse.gov/sites/default/files/microsites/ostp/big\\_data\\_press\\_release\\_final\\_2.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf), 2012.
- [9] UN Global Pulse. Big data for development: Challenges & opportunities [R/OL]. <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>, 2012.
- [10] Brinkmann B H, Bower M R, Stengel K A, et al. Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data [J]. *Journal of Neuroscience Methods*, 2009, 180(1): 185–192.
- [11] Chang F, Dean J, Ghemawat S, et al. Bigtable: A distributed storage system for structured data [J]. *ACM Transactions on Computer Systems (TOCS)*, 2008, 26(2): 205–218.
- [12] Herodotou H, Lim H, Luo G, et al. Starfish: A self-tuning system for big data analytics [C]// *Proceedings of the Fifth CIDR Conf*, 2011: 261–272.
- [13] Prekopsák Z, Makrai G, Henk T, et al. Radoop: Analyzing big data with rapidminer and hadoop [C]// *Proceedings of the 2nd Rapid Miner Community Meeting and Conference (RCOMM 2011)*, 2011.
- [14] Dai J, Huang J, Huang S, et al. Hitune: Dataflow-based performance analysis for big data cloud [C]// *Proceedings of the 2011 USENIX conference on USENIX annual technical conference*, 2011: 87–100.
- [15] Chen Y, Alspaugh S, Katz R. Interactive analytical processing in big data systems: A cross-industry study of MapReduce workloads [C]// *Proceedings of the VLDB Endowment*, 2012, 5: 1802–1813.
- [16] Yan W, Brahmakshatriya U, Xue Y, et al. p-PIC: Parallel power iteration clustering for big data [J]. *Journal of Parallel and Distributed Computing*, 2013, 73(3): 352–359.
- [17] Borkar V, Carey M J, Li C. Inside big data management: Ogres, onions, or parfais? [C]// *Proceedings of the 15th International Conference on Extending Database Technology*, 2012: 3–14.
- [18] Schadt E E, Linderman M D, Sorenson J, et al. Cloud and heterogeneous computing solutions exist today for the emerging big data problems in biology [J]. *Nature Reviews Genetics*, 2011, 12(3): 224.
- [19] Wang F Y. A Big-Data perspective on AI: Newton, merton, and analytics intelligence [J]. *IEEE Intelligent Systems*, 2012, 27(5): 2–4.
- [20] LaValle S, Lesser E, Shockley R, et al. Big data, analytics and the path from insights to value [J]. *MIT Sloan Management Review*, 2011, 52(2): 21–32.
- [21] Miller H G, Mork P. From data to decisions: A value chain for big data [J]. *IT Professional*, 2013, 15(1): 57–59.
- [22] McAfee A, Brynjolfsson E. Big data: The management revolution [J]. *Harvard Business Review*, 2012, 90(10): 60–68.
- [23] Kirkpatrick R. Big data for development [J]. *Big Data*, 2013, 1: 3–4.
- [24] Bughin J, Chui M, Manyika J. Clouds, big data, and smart assets: Ten tech-enabled business trends to watch [J]. *McKinsey Quarterly*, 2010, 56(8): 1–14.
- [25] IBM. Big data: The new natural resource [EB/OL]. <http://www.ibmdatahub.com/infographic/big-data-new-natural-resource>, 2013.
- [26] Gartner. Big data [EB/OL]. <http://www.gartner.com/it-glossary/big-data/>, 2013.
- [27] Villars R L, Olofson C W, Eastwood M. Big data: What it is and why you should care [R]. White Paper, IDC, 2011.
- [28] IBM. What is big data? [EB/OL]. <http://www-01.ibm.com/software/data/bigdata/>, 2013.
- [29] Wikipedia. Big data [EB/OL]. [http://en.wikipedia.org/wiki/Big\\_data](http://en.wikipedia.org/wiki/Big_data), 2013.
- [30] NSF. Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA) [EB/OL]. [http://www.nsf.gov/funding/pgm\\_summ.jsp?pims\\_id=504767](http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504767), 2012.

- [31] Tencent. How much data generated per minute: 3600 Instagram photos [EB/OL]. <http://tech.qq.com/a/20120623/000018.htm>, 2012.
- [32] Oh O, Agrawal M, Rao H R. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises [J]. *MIS Quarterly*, 2013, 37(2): 407–426.
- [33] Abbasi A, Chen H, Salem A. Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums [J]. *ACM Transactions on Information Systems*, 2008, 26(3): 1–34.
- [34] 冯芷艳, 郭迅华, 曾大军, 等. 大数据背景下商务管理研究若干前沿课题 [J]. *管理科学学报*, 2013, 16(1): 1–9. Feng Zhiyan, Guo Xunhua, Zeng Dajun, et al. On the research frontiers of business management in the context of big data [J]. *Journal of Management Sciences in China*, 2013, 16(1): 1–9. (in Chinese)
- [35] Shaw M J, Subramaniam C, Tan G W, et al. Knowledge management and data mining for marketing [J]. *Decision Support Systems*, 2001, 31(1): 127–137.
- [36] Ramakrishnan T, Jones M C, Sidorova A. Factors influencing business intelligence (BI) data collection strategies: An empirical investigation [J]. *Decision Support Systems*, 2012, 52(2): 486–496.
- [37] Wu X D, Zhu X Q, Wu G Q, et al. Data mining with big data [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(1): 97–107.
- [38] Huberman B A. Sociology of science: Big data deserve a bigger audience [J]. *Nature*, 2012, 482(7385): 308.
- [39] Lazer D, Pentland A, Adamic L, et al. Computational social science [J]. *Science*, 2009, 323(5915): 721–723.
- [40] Alstyne M V, Brynjolfsson E, Madnick S. Why not one big database? Principles for data ownership [J]. *Decision Support Systems*, 1995, 15(4): 267–284.
- [41] Demirkan H, Delen D. Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud [J]. *Decision Support Systems*, 2013, 55(1): 412–421.
- [42] Delen D, Demirkan H. Data, information and analytics as services [J]. *Decision Support Systems*, 2013, 55(1): 359–363.
- [43] Clifton C, Marks D. Security and privacy implications of data mining [C]// *ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, 1996: 15–19.
- [44] 杨善林, 罗贺, 丁帅. 基于云计算的多源信息服务系统研究综述 [J]. *管理科学学报*, 2012, 15(5): 83–96. Yang Shanlin, Luo He, Ding Shuai. Survey on multi-sources information service system based on cloud computing [J]. *Journal of Management Sciences in China*, 2012, 15(5): 83–96. (in Chinese)
- [45] Atzori L, Lera A, Morabito G. The internet of things: A survey [J]. *Computer Networks*, 2010, 54(15): 2787–2805.

## Management issues in Big Data: The resource-based view of Big Data

YANG Shan-lin, ZHOU Kai-le\*

Key Laboratory of Process Optimization and Intelligent Decision-making, School of Management, Hefei University of Technology, Hefei 230009, China

**Abstract:** There are many key issues in the management of Big Data, which is an important strategic resource. We first review some definitions of Big Data from different perspectives. Then, we point out that Big Data is an important information resource. Then, the management characteristics of Big Data resource are investigated mainly from six aspects, including complexity, decision-support, rapid growth, value scarcity, the characteristic of repeatedly mining and function diversity. Finally, the key management issues of Big Data, including the acquisition issue, processing issue, application issue, property issue, industry issue and regulatory issue, are refined and briefly analyzed.

**Key words:** Big Data; resource view; management characteristics; management issues