

信息时代的大数据现象值得哲学关注

——评《大数据时代》

黎德扬

(武汉理工大学, 湖北 武汉 430070)

[摘要]《大数据时代》传达了当前 IT 技术新发展的走向及其对社会生活带来深刻影响的信息,深化了对当代世界变革的认知,加速了人类文化品位的提升。同时指出大数据是信息时代的一种现象,不宜作为时代的标识,并认为大数据不能删除现实世界中的精确性、因果性和“为什么”的追求。书中提出的思维方式变革向哲学提出了一种诉求,哲学应作出回应。

[关键词]大数据;时代;思维方式

[中图分类号]N031 **[文献标识码]**A **[文章编号]**1672-934X(2014)02-0010-04

Big data Phenomenon in Information Age Worthy of Philosophical Attention

——A comment on *The Age of Big Data*

Li De-yang

(Wuhan University of Technology, Wuhan, Hubei 430070, China)

Abstract: *The Age of Big Data* has revealed the trend of new information technology and the profound influence on social life that brought by information. It has deepened our knowledge about the reformation of contemporary society, and speeded up the promotion of our cultural taste. This article, at the same time, has pointed out that big data is just a phenomenon of information age, and it is not adequate to consider it as a landmark of the age. The author believes that big data can not wipe out the pursuit of precision, causality and "the why" in the real world. The thinking pattern brought up by this book has appealed to philosophy, to which philosophy should respond.

Key words: big data; age; thinking pattern

2013 年是一个有趣的年份,有人说,2013 年是进入大数据时代元年。大数据(Big Data)在信息时代并非什么特别新鲜的事物,Google 的搜索服务就是一个典型的大数据运用,在全球几十亿人参与的网络世界、智能系统随时都在吐纳着大数据。2013 年前后,在我国瞬间,大数据不仅成了 IT 行业中最摩登的词汇,也成了经济、政治、文化领域说事的流行语言和对未来社会预测的抢眼根据。

这种现象的出现和三件事情有关:第一,2012 年 3 月,美国奥巴马政府宣布投资 2 亿美元启动《大数据研究和发展计划》,并将大数据定义为“未来的新石油”,希望增强政府收集、分析和萃取海量数据的能力。第二,是被誉为“大数据时代的预言家”英国维克托·迈尔-舍恩伯格著《大数据时代》和《删除:大数据取舍之路》,2013 年 1 月由浙江人民出版社出版中文译本,将大数据由 IT 行业扩散到大众文化,使人们感到无比的新颖。第三,2013 年 6 月斯诺登用 41 张幻灯片,让美国大数据监控项目“棱

收稿日期:2014-02-28

作者简介:黎德扬(1938—),男,湖南浏阳人,武汉理工大学教授,主要从事科学技术的哲学与文化学研究。

镜”浮出水面,令人对网络安全不寒而栗。数字化记忆的两大威胁:一个没有安全与时间的未来,直接面对人类。“在信息权力与时间的交汇处,永久的记忆创造了空间和时间圆形监狱的幽灵”^[1],引发了因不甚了然网络生存的人群对未来的忧虑甚至恐怖。这些事件叠加在一起,风生火起,热闹非凡,大数据就成了世界特大事件。

面对大数据风潮,笔者就《大数据时代》和《删除》两本书产生如下认知和思考。

《大数据时代》从当下信息技术迅猛发展及其所引起的社会深刻变革入手,运用商业典型个案进行分析、做出解读,并提出对未来的预测及其对策。它使人们强烈感受到世界的变迁,当今工业化、信息化的推进是何等神速,催人奋进。同时从中也可获取许多关于 IT 技术新发展的走向及其对社会生活带来深刻影响的信息,对深化世界变革的认知,加速人类文化品位的提升,都大有裨益。

《删除》中关于人类从‘记忆难忘却易’走到了向‘记忆易忘却难’的转变触动很大,深感既庆幸又悲凉^{[1](P6)}。庆幸的是人类文化找到了储存的路径,历史可以凝固并在虚拟世界中重现,免除了人类在记忆上所耗费的海量劳动。悲凉的是如果不能删除,人们都将沦为数字王国的奴隶,赤裸裸地跪拜数字王面前,没有任何权利和隐私。当然也想到,今日世界只是贮存了部分上网人群的信息,是否能收集人类活动所有信息,还是一个难题。何况世界是人的世界,人是主体,人脑的智能还只开发了 20%,数字王国只是人的一种虚拟,人总会想出有效监控网络的办法。庆幸和悲凉只是一纵即逝,但还是要感激作者对网络安全的警示。

由于书的内容涉及面十分宽广,论及的又是关于时代性的人类问题,自然就涉及科学技术和社会生活诸多方面知识的综合运用,更离不开哲学的思辨。虽然作者是 IT 行业顶尖专家,但毕竟是从行业视野看世界,当刻意凸显思维方式的变革而不能不涉足哲学时,由于哲学的繁杂,不免会留下值得商讨之处。最为突出的是:在张扬 IT 技术的进步中添加了一些界定不很清晰的哲学概念。

二

大数据(或数据大)是信息科学技术发展到云计

算阶段出现的新现象,是信息时代的一个新表征,不宜用大数据来标识我们时代。

上世纪五十年代以来,在系统科学群中信息理论和信息技术领先蓬勃发展,信息通信技术按照自身技术发展的逻辑,向更高级智能阶段繁衍和演进,是技术生态的正常过程。近二、三十年以来,由于信息高速公路建成和互联网构筑,信息感知和采集终端获取了海量的数据,又以云计算为代表的计算技术提供了强大的计算能力,一个空前巨大的、日益清晰的数字世界或网络世界耸立在世人面前,对人类社会的发展和走向产生全面的、风暴式的强大影响。自然受到全球人们,特别是未来学家的关注,如托夫勒、奈斯比特、福来斯特以及当下的舍恩伯格等等。舍恩伯格说:大数据带来的信息风暴正在变革我们的生活、工作和思维,大数据开启了一次重大的时代转型,并用三个部分讲述了大数据时代的思维变革、商业变革和管理变革^[2]。确然如此,大数据的根基是 IT 系统,大数据将驱动 IT 产业新一轮的架构性变革是无疑的。但是否就由此而推论出社会走向了一个新的时代——大数据时代,却需要思量。

自从上世纪五十年代信息科学技术的产生和发展以来,就开启了世界观念的转型,从认为世界是物质、能源的变换,向世界是物质、能源和信息的变换,并确认了信息是这一变换的灵魂。物质生产方式和人们的生活方式都已转入到信息世界之中,堪称为一个新的时代——信息时代,以示其与物质和能源时代相区别。正因为如此,全球都认同人类进入了信息时代。尼古拉·尼葛洛庞帝(Nicholas Negroponte)在 1995 年出版的《数字化生存》中就说“计算不再只和计算机有关,它决定我们的生存”^{[3](P7)},“数字化生存使人挣脱了时间、空间的限制和‘原子’的束缚,得以遨游更为广阔的世界,接触更广泛的人群。”^{[3](P15)}他这些概括性的说法,已经表明信息时代的某些基本特征。舍恩伯格指出了数字化和数据化的区别,“数字化指的是把模拟数据转换成用 0 和 1 表示的二进制制,这样电脑就可以处理这些数据了。数据化是指一种把现象转变为可制表分析的量化形式的过程”^{[2](P104)},提升了数字化的内涵,反映了技术进步。比尼葛洛庞帝的数字化生存进了一步,或许可以称,当今人类处于数据化生存更为真切。虽然舍恩伯格引述了近一、二十年来计算机技

术的新进展,主要是云计算及其带来的人类社会生活、工作和思维方面的新资料,值得人们关注,但只是进一步证实了关于信息时代的预言,在总体方面却没有超出尼葛洛庞帝等的看法。最多也只能说,大数据是信息时代面临的新现象和新阶段。说 2013 年是大数据时代元年就更不符合历史事实。

数据逐渐成为重要的生产要素,人们对海量数据的运用将预示着新一波生产力增长和消费者盈余浪潮的到来。大数据将会创造一个新的经济领域,该领域的全部任务就是将信息或数据转化为经济利益。数据对于企业而言是一种重要的战略资产,谁能把握这一机遇并迅速行动起来,就能在未来的竞争站得先机至关重要。现在关于大数据的热门是 IT 行业,因为他们面临着处理现实大数据技术上的直接挑战,竞争十分激烈,关系它们的生存和发展的命运。大数据中,结构化数据只占 15% 左右,其余的 85% 都是非结构化的数据。如何处理非结构化和半结构化数据,如何把通过数据挖掘,将粗糙知识与被量化后的主观知识相结合,转化为“智能知识”;如何根据大数据复杂性、不确定性对其进行刻画并为大数据系统建模等等,都是大数据现象提出需要深入探索的深层技术问题。我们这类发展中国家 IT 技术尚处初级阶段,更是应当奋起直追,迎头赶上世界先进水平。在这一行业排忧解难中,由于商业利益而做些鼓动,使用一些过分夸张的广告语言是可以理解的。但学界更富有理性,目前人们还很难从学理上论证信息时代与大数据时代在质地上的区别,不宜于跟随商业的宣传而将大数据时代扩展到所有社会领域,随意使用时代这个相当严格严肃的词汇。如果时代的提法过多,将某一时代的某一阶段、某一现象也说成一个新时代,未免将时代太碎片化了,而模糊人们的大视野。

三

大数据不能删除精确性、因果性和‘为什么’的追求。舍恩伯格《大数据时代》中关于思维变革共有三章:第 1 章更多,不是随机样本,而是所有数据;第 2 章更杂,不是精确性,而是混杂性;第 3 章更好,不是因果关系,而是相关关系。大数据时代最大的转变就是放弃对因果关系的渴求,只要关注相关关系。也就是说只要知道“是什么”,而不需要知道“为什

么”。并认为这就颠覆了千百年来人类的思维惯例,对人类的认知和与世界交流的方式提出了全新的挑战^{[2](P29)}。如果我们面对的只是一堆大数据,这样的想法是合理的、有效的。如果我们进入到物理世界、现实生活,我们不能满足于数据本身,而是要对数据进行的分析与处理,将粗糙知识转化为智能知识,对现实生活中将要发生的事件做出预测。在这一过程中,还是不能删除精确性、因果性和‘为什么’的追求。

诚然计算机网络的建成和繁衍,构建了一个数据高速增长的世界,数据达到海量,有价值的和没有价值的数据浑然杂呈,网络数据呈现出混杂性。但是并不能因此而放弃对精确性和因果性的追求,更不能说人类的认识停留在“是什么”就够了。舍恩伯格书中明确指出:“大数据的核心就是预测”。预测不是算命,是以事实为依据的,只有依据科学规律做出的预测才能保障预测的实现。规律就是最为本质的相关性,相关性并不一定是因果关系,但因果性却必定是一种相关性。人类是有理性的,总是凭着理性追求现象背后的本质和规律,即事件之间的相关性、恒常性及其所蕴含的因果关联(线性的和非线性)。理性思维是实践中产生的以一定的‘格’固定下来的思维惯例和逻辑思维能力。信息科学和计算机技术无时无刻在运用着数理逻辑的支撑,如果放弃逻辑思维和逻辑程序的设计,大数据便不复存在。

数学是关于数量和空间关系的科学,其特点是抽象性、精确性和应用的广泛性。虽然[美] M·克莱因《数学:确定性的丧失》一书否定了关于数学确定性的‘神话’,然而绝大多数人,今天仍然认为数学是关于物质世界的不可动摇的知识体系。他强调数学不合逻辑的发展方式,应用数学反对“纯”数学的问题以及在二十世纪数学逻辑结构的连贯性遇到的挑战等事实,令人振聋发聩,从迷信数学中惊醒。然而翻阅数学发展史,人们不难发现,数学危机是多次发生的,史上就发生过三次大的危机。数学公理体系的无矛盾性始终是数学的追求,然而每一次危机都表现为悖论的出现,新的数学体系又以悖论的破解应运而生。每一次危机都带来了一次新的发展,数学发展呈现为连续和间断的统一。因此当数学遭遇危机的时候,决不是悲观地放弃精确性,而是从公理系统中解决悖论,提出新的公理系统。

人们面对的现实世界是灰色的,是混然杂呈的现象世界。生活中有许多模糊现象,应对具体的生活,使用模糊语言和大概的估算是必须的,可以节省许多不必要的精确计量,不必要饭也一粒粒数着吃,规定每顿吃多少粒。但是,人类进步,科学发展,文明的演进,就在于对这一混沌世界中的现象进行辨析,确定现象的性质和数量以及它们的关联。近代科学技术史就是实验分析和数学方法相互结合的历史。科学研究中必须做到十分精确,只有这样才能预测未来,并规范人类行为以实现生存发展的目的。地震预报就是根据大数据而做出的,人们努力的目标就是力求做到更加精准。人类探月活动中各种装置就必须设计得十分精密精准,一个数据上的错误,甚至误差,都会导致行为的失败,造成灾难性的后果。数学中有句警言:差之毫厘,失之千里。1965年美国控制论学者 L.A.扎德发表论文《模糊集合》,标志着模糊数学这门新学科的诞生。但实际运用中还是要还原于精确数学,通过常规计算才能解决实践问题。世间现实的事物都有质和量的属性,没有无质的量,也没有无量的质。如同数学中的点、线、面在现实中是不存在的,只是几何学中的虚拟,数字也是如此。数字化是计算机的必要条件,计算机是以 0 和 1 二进制的变换为其根基的。计算技术所构建的网络世界是一个虚拟世界,并非真实的现实世界。计算机行业有句名言:输入的是垃圾,输出的也是垃圾!诚然,当下我们有了多媒体、网络、云计算等等,有理由乐观地预言将来会创造出更多更新的新技术。但是,对作者所说“量化一切,数据化的核心”^{[1] (P105)},颇感疑惑。从毕达哥拉斯提出数是世界本原的哲学理念以来的数千年中,就有世间的一切是否皆可量化(或数学化)的争论,这将是人类旷日持久认识论上的论争。能否量化一切尚不确定,核心自然就难于形成。根据当下遇到的大数据现象就放弃对世界认知的精确性,用储存器中碎片的混然杂呈取而代之,只能是认识道路上的迷雾。大数据

每一种表示形式都仅呈现数据本身的侧面表现,并非全貌。任何大数据都是对一定物理事件的描述,对于赤裸裸的数据堆积,混然杂呈,人们是很难说出什么来的。所谓混杂性代替精确性,实际上不免贬低了人类的知识体系及其价值,包括 IT 技术在内。在大数据潮涌面前,数据挖掘与分析应用方面,最直接的、最大的问题是分析不准确和分析速度慢,不能满足决策的需求,制约从数据中获取价值。IT 行业应努力提高数据的可用性和准确度,最大限度地提取信息数据的价值。它深刻反映了当代人们对精确性的期待。

读完《大数据时代》,深感它所描述的信息时代思维变革正在发生,必须高度关注。同时该书也进一步证实了科学技术与哲学具有相关性:科学技术寻求哲学的支撑,哲学也只有在与科学技术的同行中才能成为时代精神的精华。对于科学和哲学的相互关联,1938 年爱因斯坦曾作过这样的描述:“哲学的推广必须以科学成果为基础。可是哲学一经建立并广泛地被人们接受以后,它们又常常促使科学思想的进一步发展,指示科学如何从许多可能的道路中选择一条道路。等到这种已经接受了观点被推翻以后,又会有一种意想不到和完全新的发展,它又成为一个新的哲学观点的源泉。”^[4]当新技术向我们走来的时候,就给哲学进步带来了新源泉和动力。《大数据时代》反映了当代科学技术前沿的哲学诉求,哲学应当作出回应,在回应中创新哲学。

[参考文献]

- [1] [英]维克托·迈尔-舍恩伯格.删除:大数据取舍之路[M].袁杰,译.杭州:浙江人民出版社,2013.
- [2] [英]维克托·迈尔-舍恩伯格,[英]肯尼斯·库克耶.大数据时代:生活、工作与思维的大变革[M].盛杨燕,周涛,译.杭州:浙江人民出版社,2013.
- [3] [美]尼古拉·尼葛洛庞帝.数字化生存[M].胡泳,范海燕,译.海口:海南出版社,1997.
- [4] [美]A.爱因斯坦,[美]L.英费尔德.物理学的进化[M].周肇威,译.上海:上海科学技术出版社,1962:59.