

大数据技术的本体分析

The Ontological Analysis of Big Data Technology

李君亮 /LI Junliang 陈艳 /CHEN Yan

(1. 广西民族师范学院马克思主义学院, 广西崇左, 532200; 2. 广西民族师范学院数学与计算机科学学院, 广西崇左, 532200)
(1. Marxism Department, Guangxi Normal University for Nationalities, Chongzuo, Guangxi, 532200;
2. Mathematics and Computer Science Department, Guangxi Normal University for Nationalities, Chongzuo, Guangxi, 532200)

摘要: 本体指的是存在或存在之存在, 又可称之为是或是之所是。数据作为给定的、已完成的事实是大数据技术“是之所是”的基本质料。技术让作为大数据技术“是之所是”基本质料的数据说话, 是大数据技术“是之所是”的形式、动力和目的。关系才是大数据技术的“是之所是”。关系作为大数据技术的“是之所是”具有客观性和确定性, 这就满足了本体论配置的可寻性原则。

关键词: 大数据技术 本体 数据 技术 关系

Abstract: Ontology refers to being or means to be. As given things, data are the basic materials of the being of big data technology. Technology makes the data talk and it is the form, motivation and purpose of the being of the big data technology. Relationship is the being of the big data technology. As the being of big data technology, relationship is objective and affirmative, which satisfies the discoverability principle of ontological configuration.

Key Words: Big data technology; Ontology; Data; Technology; Relationship

中图分类号: N0 文献标识码: A DOI:10.15994/j.1000-0763.2018.11.003

大数据技术的广泛应用给我们的思维方式和思想观念带来了最为深刻的变革, 大数据时代的到来也使学界展开了对于大数据技术认识论、知识学的深入探讨, 如吕乃基《大数据与认识论》一文对大数据认识主体和对象、认识过程、认识结果、认识论依据等的深刻阐述, 苏玉娟等对大数据知识表征的社会建构、确证问题、机制及其意义、大数据知识实现的维度、大数据知识的相关伦理问题等展开的系列讨论, 吴信东等构建的大数据生成大知识的BigKE工程模型。^{[1]-[7]}学界对大数据本体论的讨论也已经展开, 如刘红在《大数据的本体论探讨》一文中将大数据看作是本体上对于“万物源于数”的思想的回归,^[8]但该文并未对大数据或大数据技术的本己本体进行追问; 段伟文《大数据知识发现的本体论追问》一文关

于大数据时代知识发现预设的本体考察, 在该文中, 段伟文提出在知识发现视野下大数据预设的本体不过是世界的数字化表象, 在这表象下, 大数据的本体实际上自我隐匿了, 因此他提出要从多元主义、诠释学和能动实在论等进路追问大数据的真实本体。^[9]那么, 大数据技术本身的本体究竟是什么? 我们又该如何揭蔽大数据技术数字化表象背后隐匿的真实本体, 本文将从对本体论内涵的分析出发, 努力揭示出大数据技术隐匿着的真实本体。

一、本体论的内涵分析

哲学的发生是从本体论开始的, 虽然近代发生了哲学讨论的认识论转向, 但直到今天, 本体

基金项目: 教育部人文社会科学研究青年基金项目(项目编号: 15YJCZH015); 广西民族师范学院校级科研项目(项目编号: 2018YB021)。

收稿日期: 2018年6月1日

作者简介: 李君亮(1975-)男, 江西南康人, 广西民族师范学院马克思主义学院副教授, 研究方向为分析哲学等。Email: lijunliang1975@163.com

陈艳(1977-)女, 湖南湘潭人, 广西民族师范学院数学与计算机科学学院讲师, 研究方向为数学哲学、数学应用、数理逻辑等。Email: 79593434@qq.com

论依然是哲学最重要的基本论域之一，本体依然是哲学最为重要的基本概念。那么，什么是本体？什么又是本体论？

本体论（ontology）一词是由17世纪的德国经院学者郭克兰纽（R. Goclenius）首先使用的，该词由onto和ology这两个拉丁词根构成，onto相当于英语to be或being，即“是”或“存在”，ology表示“学问”或“学说”的意思。因此，本体论简单地讲就是关于存在以及存在之所以存在的学说。由于onto意为“是”或“存在”，所以又经常将本体论称为是论或存在论。

哲学与近现代科学不同——科学是通过研究现象发现事物运动变化的根本规律，即通过现象看本质；哲学则是研究本质的本质，即事物之是，to be或being。最早提出to be, being——即存在或是的是古希腊贤哲巴门尼德，他说，“……存在（或是），……是说服之路（因为有真理相随）/一条路，……因为对于思想和对于存在是同一件事。”^[10]在这段以女神口吻说出的话语中，巴门尼德提出了既可以作为实义动词理解的“存在”，也可以作为系动词解读的“是”。

“存在”或“是”之所以成为本体，即成为哲学的基本追问对象，这和古希腊语以及由古希腊语逐渐发展起来的西方拉丁语系的表达密切相关。在西方语言的表述中，其基本的语言表达结构是“主语（S）+谓语（P）”，即在表达中总是主语去“存在”或“是”，即使主语省略，to be或being却不可或缺，而to be或being实际上暗示着主语的存在，这也就是蒯因所说的“本体论的承诺”。这样的语言表达结构实际上揭示了人类对一种确定性、规定性的诉求，因此，当我们对这种确定性、规定性进行追问时，本体即to be或being就成为主语了——亚里士多德是将to be, being作为主语的肇始者。本体作为“存在”或“是”，本体论也就被称为“存在论”或“是论”了。

当然，在古代汉语中没有“本体论”一词，但有“本根论”或“道论”，“本根”或“道”指称的是一切皆从之而出又回归其中的那个东西，这指的实际上也就是古希腊哲学中的那个“本体”。

在亚里士多德看来，自然界的运动变化规律有物理学来研究，生命运动变化规律有生物学来研究，地理运动变化规律有地理学来研究，等等。而物理学背后的元在则由元物理学（metaphysics）

即形而上学来研究，因此，自亚里士多德而后，形而上学即是本体论。但是到了中世纪，to be, being的问题被交给了上帝，古希腊哲学与基督教神学的合流形成了经院哲学这一奇特的哲学形式，这个时候——直到郭克兰纽，形而上学成了哲学的代名词，本体论、宇宙论、目的论、神学世界观，这些都被囊括在了形而上学中。近代科学兴起后，天文学接管了宇宙论，目的论重新由物理学、生物学等接管，神学世界观则被拉下神坛并逐渐发展为近现代宗教哲学，于是郭克兰纽在17世纪提出ontology一词取代形而上学中最哲学之部分，即取“本体论”之意。

由上述可见，本体即存在或是，又是存在之存在或是之所是；本体论即存在论或是论，是关于存在或是，又是存在之存在或是之所是的学说。

那么，大数据技术之存在或之所是是什么？

二、数据：大数据技术“是之所是”的基本质料

大数据技术有其本体吗？当然有！因为我们必然会问：“大数据技术是什么？”回答是：“大数据技术是……”。这就表明，大数据技术定有其“存在之存在”或“是之所是”。既然如此，我们就要问，大数据技术“如何去是”？以及大数据技术“是之为何”？“如何去是”是大数据技术是之方式、方法、手段及途径方面的问题，即认识-方法论的问题；“是之为何”才是大数据技术“是之所是”的问题，即本体论的问题。没有“是之所是”，也就无所谓“如何去是”。但是由于大数据技术历史绽放的特殊性，我们必须首先考察其“如何去是”的问题，然后才好回答其“是之所是”的问题。

大数据技术“是之所是”何以可能？历史地看，大数据技术的发展历经了从数到数据再到大数据的演变发展过程，进入大数据时代，大数据被收集、存储、分析、处理、应用，大数据技术方始得以绽放。^[11]对大数据技术历史蕴演的这一过程进行考察我们会发现，“数据”成为了其生成与绽放不可或缺的逻辑项，这一逻辑项正是亚里士多德“四因说”语境下大数据技术何以可能的“质料因”回答。

数据是大数据技术“是之所是”的基本质料，数据也是我们在大数据时代深度认识世界的本源。^[12]那么数据又是什么？舍恩伯格在《大数据时代》

中指出,“‘数据’(data)这个词在拉丁文里是‘已知’的意思,也可以理解为‘事实’。”^[13]从数据的数之起源看,数据表征世界确实是表达着事实,但却不是事实本身。数才是事实本身。当然,在逻辑主义者罗素看来,数都是类或通过类构造的,而类是虚构的,那么数也就是逻辑虚构的。^[14]关于数的类构造观点,罗素与亚里士多德的看法是一致的,因为亚里士多德也认为,数是种类的相,这个相存在于为其所数的那组对象中。([14], p.63)直觉主义者则相信,数的本质是人的思维构造,([14], p.183)但在结构主义者看来,数只是其结构中的一个位置,因为独立于它们所属结构的数是没有意义的。([14], pp.254-255)卡尔·B.博耶在其《数学史》中考察了数的起源,认为数的观念是在从差异性到同一性认识发展中产生的,^[15]在这一数的观念产生过程中,数首先作为祭神时诸神的出场顺序这一事实被表达出来。([15], p.6)当被表达的事实通过量的积累仅仅依靠人脑记忆已无法有效保存时,这一事实就必须作为信息被刻入石头、兽骨、木块,最终作为数据而历时性地呈现在了我们面前。([15], p.5)

我们认为,从数-数据的历史生成考察数据,卡尔·B.博耶的假说是科学可信的,并且他也在其《数学史》中提供了人类考古学的证据。([15], p.5)因此,事实被记录为数据,数据只是蕴含或表达着事实。从词源看,数据(data)一词为拉丁文,是datum的复数形式,表示“已知的或已知的事实,即(thing) given。”“data”对应的希腊语是“δεδομε'vo”,“δεδομε'vo”也是given的意思。^[16]given是给定的,已完成的,也就是说,数据是给定的、已完成的事实。是被什么给定和完成的事实?从数-数据的运演逻辑看就已经知道,是被数给定、完成的事实。给定的、完成的事实只能被作为大数据技术的质料,而不会是大数据技术本身,即数据不会是大数据技术的“是之所是”,而是使“是之所是”成为可能的质料基础。

数据是大数据技术成为可能的质料,而不是大数据技术之本体,因此,它作为给定的、已完成的事实的表达表征着世界,就无需承担起大数据知识发现的主体责任。相反,作为给定的、已完成的事实,此时数据隐匿着的恰恰就是知识——从起源处看,其隐匿着的是关于神的排序的知识;而随着人类社会基于数据视域下从微数据到小数

据再到大数据时代,^[17]数据隐匿着的知识也从神拓展到自然世界最后到人的思想行为本身,数据蕴藏的知识量当然也就从算术级数过渡到几何级数最后到指数级数。这样,大数据知识发现需要发现的就不是作为大数据技术“是之所是”的基本质料的数据本身,大数据知识发现需要揭蔽的应是数据表达的给定的、已完成的事实及其关系,这实际上就是大数据技术的绽放,也就是说,大数据知识发现本身也作为表象之一,在数据为大数据技术提供的质料之基上,通过大数据技术被表达出来。

如舍恩伯格在《大数据时代》给出的大数据技术经典案例——谷歌流感趋势(GFT),就是谷歌公司通过对每天来自全球超过30亿条搜索指令留下的数据足迹展开数据挖掘后对冬季流感传播作出的概率预测。其它公司无法完成这样的预测,根本原因在于“他们缺乏像谷歌公司一样庞大的数据资源”,([13], p.3)当然也不具备如谷歌公司那样对庞大数据进行处理的能力和统计的技术。因此,大数据技术正如舍恩伯格所言,就是“以一种前所未有的方式,通过对海量数据进行分析,获得有巨大价值的产品和服务,或深刻的洞见。”([13], p.4)在大数据技术的应用中,没有数据,信息将失去承载;没有数据,我们的决策将丧失依据;没有数据,我们的行动将迷失方向。数据,海量的数据,已经成为现时代人们获得新的认知、创造新的价值的源泉,并将为改变市场、组织机构以及政府与公民关系服务。([13], p.9)

在GFT这一经典案例中我们可以看到,数据,海量的数据为大数据技术及其应用提供了最基本的质料,通过对大数据技术这一质料的分析处理,我们发现了数据给定的、已完成的事实,大数据知识发现实际上也就是重新发现给定的、已完成的事实的相互关系,从而使大数据技术的开发应用成为可能。

三、技术:让作为大数据技术“是之所是”基本质料的数据说话

仅仅通过作为大数据技术“是之所是”的基本质料的海量数据,大数据技术的本质并不会被完全揭示出来。海量数据给定的、已完成的事实如果作为世界的表象,那么它表征着的是已知的、

给定的事实，只是这给定的已知事实被海量数据遮蔽着，现在我们必须让它在敞亮之处敞明。大数据技术能让数据说话——确切地说，是大数据技术之技术让作为大数据技术“是之所是”基本质料的数据说话，从而敞明隐匿在海量数据之中的给定的、已完成的事实及其诸关系。

涂子沛援引美国管理学家、统计学家爱德华·戴明的话说，大数据时代，“除了上帝，任何人都必须用数据来说话。”^[18]要用数据说话，我们必须先让数据说话，让数据向我们绽放出其给定的、已完成的事实中隐匿着的诸知识——大数据技术作为技术，其承担着的正是让作为大数据技术“是之所是”基本质料的数据说话的使命。这样，技术被关键性地卷入到了大数据技术之中，并成为大数据技术“是之所是”的形式与动力，帮助实现大数据技术“是之所是”的目的。

首先，大数据技术“是之所是”以技术这一形式因表现出来，即大数据技术是作为技术的形式收集、存储、分析、处理海量的作为大数据技术“是之所是”的基本质料的数据。大数据技术并不是一个抽象的概念，它具有作为技术的实体实在性，即大数据技术离不开技术社会创造的计算机、智能手机、网络设备、存储设备等人工制品，这些人工制品也正是大数据技术在常识意义上的表现。没有计算机的发明，人类不可能进入到信息社会；没有网络设施、智能终端设备等的发明与发展，人类不可能进入大数据时代。因此，大数据技术作为技术首先就表现为人造物等实体实在，从而具有实体实在性。以技术的实体实在这一形式，大数据技术具有了收集、存储、分析、处理海量的作为大数据技术“是之所是”的基本质料的数据的可能，并且从现实上展开了收集、存储、分析、处理海量数据的一系列社会实践活动。如GFT预测，以及日本先进工业技术研究所的坐姿研究与汽车防盗系统，（[13]，pp.102-103）这正是以技术实体实在为基本形式的大数据技术收集、存储、分析、处理海量数据的大数据技术应用的典型案例。并且在这两个大数据技术应用的典型案例中，通过技术，海量的数据在说话：它告诉我们流感会在某个时期某个地域爆发；它告诉我们关于驾乘姿势与汽车防盗或安全驾驶之间关系的知识。

其次，大数据技术“是之所是”以技术作为

动力因表现出来。任何技术的出现总是为了解决人类社会发展过程中遇到的现实性问题，技术进化也必将会解决人类社会发展面对着的现实性问题。如瓦特对蒸汽机的改进是为了提高燃烧效率以解决工业化进程中遇到的能量能源问题，电子管计算机（ENIAC）的发明是为了解决弹道计算问题。大数据技术作为技术的生成正是为了解决现代社会人类在生产生活中制造出来的海量数据的收集、存储、分析、处理问题，因此我们说，技术是作为大数据技术“是之所是”的动力因绽放出来的。

技术作为大数据技术“是之所是”的动力因表现为技术的技艺实在，即它总是要具有某项能力，能解决某个或某类问题。如瓦特蒸汽机能提高燃烧效率从而有效地解决工业活动中的能量能源问题，ENIAC能有效地计算出炮弹弹道轨迹。大数据技术是运用计算机、网络设备、智能终端设备等人造物获取海量数据并对其进行储存、分析、处理的使用过程，在这一使用过程中，必须具备相关的知识和能力，懂得相应的技术与技巧，否则便会如舍恩伯格对其他公司在季节性流感传播预测所做的分析那样，这些公司不能像谷歌公司一样作出这种预测，就是因为他们不仅缺乏像谷歌公司一样庞大的数据资源，也由于他们缺乏相应的处理能力和统计技术。（[13]，p.3）

在技术作为大数据技术“是之所是”的动力因中我们可以看到，技术必须具备让作为大数据技术“是之所是”基本质料的海量数据说话的能力，如果它不能完成对海量数据的收集、存储、分析、处理，那这一技术就不是大数据技术。

再次，技术还是大数据技术“是之所是”的目的因。在亚里士多德看来，技术总是以某种善为目的。^[19]大数据技术作为技术让数据说话也必然是为了实现某种善，这就是技术作为大数据技术“是之所是”的目的因。作为大数据技术“是之所是”目的因的技术表现为技术功能实在，即它具有总是为了实现技术善的目的的功能。大数据技术被一再强调已经对我们的生活、工作、思维、政务、商业等产生了革命性变革，并使“人们获得新的认知、创造新的价值”，“为改变市场、组织机构，以及政府与公民关系服务。”（[13]，p.9）GFT预测的成功以及日本先进工业技术研究所的坐姿研究与汽车防盗系统等经典案例已对此作了

有效辩护,即大数据技术作为技术总是具有某种效应、作用、功能,总是为了实现某种善,这是大数据技术作为技术功能实在的体现,也同时说明作为功能实在的技术是大数据技术“是之所是”的目的因。

通过以上分析可知,技术作为大数据技术“是之所是”的形式因、动力因、目的因让数据说话。那么,它让作为大数据技术“是之所是”基本质料的海量数据说出的是什么?大数据技术说出的是海量数据给出的、已完成的诸事实间的关系。

四、关系:大数据技术的“是之所是”分析

维克托·迈尔-舍恩伯格和肯尼思·库克耶告诉我们,在大数据时代,“大数据的相关关系分析法更准确、更快,而且不易受偏见的影响。”([13], p.75)因此,“相关关系帮助我们更好地了解了这个世界。”([13], p.83)关系,这才是大数据技术的“是之所是”。

大数据技术以海量数据作为基本质料,舍恩伯格和库克耶把它称作“神奇的钻石矿”。([13], p.127)大数据知识发现即数据挖掘挖掘的不是数据本身,而是挖掘隐藏在数据这座“神奇的钻石矿”中的大数据技术之“是之所是”,即要发现和挖掘海量数据给定的、已完成的诸事实之间的相关关系。作为技术的大数据绽放的大数据技术也不是其“是之所是”,相反,大数据技术之“是之所是”在作为技术的大数据技术绽放之时将其带入在场,从而使大数据技术之“是之所是”在敞亮处敞明。作为技术的大数据技术敞明的也是海量数据给定的、已完成的诸事实间的相关关系。

因此,关系才是大数据技术之“是之所是”。

关系作为大数据技术之“是之所是”有其客观性。“是之所是”通常总是隐匿着的,由于其隐匿性特征,它总是容易被遗忘,这正如海德格尔所说的那样,由于存在不能作为存在者通达,所以存在就由有关存在者在存在者层次上的规定性来加以表达,^[20]这就造成日常的存在之被遗忘,我们在日常中言说的只是存在者。大数据技术在生产生活中的广泛应用让我们大数据不离口,但是海量数据给定的、已完成的诸事实及其关系却不被言及,因此,我们需要作为技术的大数据技术将其带入在场,将其敞明在敞亮之处。而作为

大数据技术“是之所是”的关系之被带入在场而敞明,根本就在于其先行于在场与敞明,即它已作为给定的、已完成的事实相关性隐匿于海量数据之中。现在,大数据技术的绽放将大数据技术“是之所是”揭示在敞亮之处。

我们仍以GFT为例以说明关系作为大数据技术“是之所是”的客观性。在GFT案例中,用户每次通过谷歌搜索引擎检索“哪些是治疗咳嗽和发热的药物”就会产生一条数据并被谷歌服务器记录和存储,将这条数据蕴含的信息释放,它给出的已完成的事实就是“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”。那么,用户为什么要在谷歌搜索引擎上检索该词条?可能是因为用户(总之是某人)患了流感。这样,“哪些是治疗咳嗽和发热的药物”这一条检索数据同时也给出了“某人患了流感”这一事实。海量检索记录产生的海量数据实际上也就给出了已完成的“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”和“X地的某人在Y时间患了流感”的海量事实。在这海量事实完成和给定之际,“X地的某人在Y时间患了流感”和“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”的相关关系(至于相关性是强或是弱,这和环境参数e有关,此处我们暂不讨论)也就生成了。正是由于给定的、已完成的“X地的某人在Y时间患了流感”事实和“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”事实的相关关系的客观存在,通过数据挖掘,谷歌公司才可以做出流感爆发的趋势预测。

关系作为大数据技术“是之所是”有其确定性。舍恩伯格和库克耶提出大数据时代我们不必非得知道现象背后的原因,而知道“是什么”——即相关关系——就够了,([13], p.67)这引起了学者们的极大争议。王天思把大数据相关关系看作因果派生关系,认为大数据相关关系根植于因果性,^[21]我们同意这种观点。其实舍恩伯格和库克耶并没有说相关关系中没有因果性的存在,他们只是说在大数据时代我们对因果性和相关性关注的焦点已发生变化:不是因果关系,而是相关关系成为了大数据时代关注的焦点。正因为相关关系根植于因果性,正是相关关系中内蕴着的因果关系,作为大数据技术“是之所是”的关系才有其确定性。在GFT案例中,我们总要问,用户为

什么要在谷歌搜索引擎上检索“哪些是治疗咳嗽和发热的药物”？正是由于给定的、已知的“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”的事实总是处于事实的因果链中，正是事实蕴含着的因果性的存在，我们确立起了“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”和“X地的某人在Y时间患了流感”这两个给定的、已知的事实间确定的相关关系。根据海量的这两组事实间的确定的相关关系的出现，我们就可以做出X地在Y时间流感爆发的趋势预测了。

需要注意的是，作为大数据技术“是之所是”的关系的确定性是相对的，这种确定性要受到事实本身所处的环境的影响。如由于广告植入导致用户对治疗咳嗽和发热药物的检索，这种检索数据给出的、已知的事实并非出于用户（或某人）患了流感这一原因。在数据挖掘中，如果不充分全面地考虑诸给定的、已知的事实间相关关系由于环境原因造成的相对确定性，这就会出现GFT夸大流感趋势等预测失准的情况。

作为大数据技术“是之所是”的关系由于具有客观性和确定性，这也就满足了段伟文教授所说的本体论配置的可寻性原则。^[9]一方面，正是由于作为大数据技术“是之所是”的关系的客观性和确定性，在数据挖掘中，大数据技术的应用有了清晰的定位。如在GFT案例中，“X地的用户在Y时间检索了‘哪些是治疗咳嗽和发热的药物’”指向的是“X地的某人在Y时间患了流感”，海量数据给定的、已完成的同类事实指示出的就是流感将在X地的Y时间爆发的趋势。另一方面，也正是由于作为大数据技术“是之所是”的关系的客观性和确定性，可寻性原则的导航清晰成为可能。

数据给大数据技术的“是之所是”提供了基本质料，技术让作为大数据技术“是之所是”基本质料的数据说着关系的话语，关系作为大数据技术“是之所是”有了质料、形式、动力、目的这“四因”的辩护，所以我们说，关系是大数据技术的“是之所是”，也就是说：大数据技术的本体是关系。

- [1] 吕乃基. 大数据与认识论[J]. 中国软科学, 2014, (9): 34-45.
- [2] 苏玉娟. 大数据知识表征的社会建构[J]. 中共山西省委党校学报, 2017, (1): 104-107.
- [3] 苏玉娟. 大数据知识表征的确证问题[J]. 晋阳学刊, 2017, (4): 140-145.
- [4] 苏玉娟、魏屹东. 大数据知识表征的机制及其意义[J]. 科学技术哲学研究, 2017, (2): 63-68.
- [5] 苏玉娟. 大数据知识实现的维度分析[J]. 理论探索, 2017, (2): 64-68.
- [6] 苏玉娟. 新时代大数据知识的伦理问题及其应对[J]. 中国井冈山干部学院学报, 2018, (1): 102-108.
- [7] Dong, W. X., Jin, H., Lu, R. Q., Zheng, N. N. 'From Big Data to Big Knowledge: HACE + BigKE'[J]. *Acta Automatica Sinica*, 2016, 42(7): 965-982.
- [8] 刘红. 大数据的本体论探讨[J]. 自然辩证法通讯, 2014, 36(6): 115-121.
- [9] 段伟文. 大数据知识发现的本体论追问[J]. 哲学研究, 2015, (11): 114-119.
- [10] 聂敏里. 论巴门尼德的“存在”[J]. 中国人民大学学报, 2002, (1): 45-52.
- [11] 陈艳、李君亮. 大数据技术：从数的解蔽到数据的遮蔽之现象学考察[J]. 江苏大学学报（社会科学版），2017, (2): 32-37.
- [12] 潘平、郑辉、兰立山. 大数据系统的本质特征及其哲学反思[J]. 系统科学学报, 2015, (3): 26-29.
- [13] 维克托·迈尔-舍恩伯格、肯尼思·库克耶. 大数据时代[M]. 杭州：浙江人民出版社, 2013, 104.
- [14] 斯图尔特·夏皮罗. 数学哲学——对数学的思考[M]. 上海：复旦大学出版社, 2014, 119-120.
- [15] 卡尔·B. 博耶. 数学史(上)[M]. 北京：中央编译出版社, 2012, 3.
- [16] 刘红、胡新和. 数据哲学构建的初步探析[J]. 哲学动态, 2012, (12): 82-88.
- [17] 苗东升. 从科学转型演化看大数据[J]. 首都师范大学学报（社会科学版），2014, (5): 48-55.
- [18] 涂子沛. 大数据：正在到来的数据革命[M]. 桂林：广西师范大学出版社, 2013, 13.
- [19] 亚里士多德. 尼各马可伦理学[M]. 北京：商务印书馆, 2003, 3.
- [20] 海德格尔. 存在与时间[M]. 北京：生活·读书·新知三联书店, 2012, 110.
- [21] 王天思. 大数据中的因果关系及其哲学内涵[J]. 中国社会科学, 2016, (5): 22-42.