

大数据知识的实现方法探析

苏玉娟

(中共山西省委党校 理论研究室,山西 太原 030006)

摘要:从知识论的发展历程看,当代知识论不仅研究知识的发现问题,还研究知识的确证和实践问题。大数据知识论作为当代知识论发展的重要成果,主要研究大数据知识的发现、确证和实践问题。从方法论视角看,大数据知识的实现方法主要包括大数据归纳方法、基于关联的因果分析方法和递归分析方法。每种方法具有不同的侧重点,解决不同的知识实现问题。大数据归纳方法主要通过对大数据的存储、分析、挖掘和可视化等,发现大数据中包含的潜在知识。基于关联的因果分析方法主要通过对大数据潜在知识进行因果分析确证其真,实现大数据信念向大数据知识的转换。递归分析方法主要通过向后递归和向前递归彰显大数据知识的实践价值。大数据知识的实现方法彰显大数据技术的重要性、大数据知识实现方法的整体性、大数据知识价值的多元性和知识实现方法的继承性和创新性。

关键词:大数据知识;大数据归纳方法;因果分析方法;递归分析方法

中图分类号:N031,C03

文献标识码:A

文章编号:1008-7699(2019)01-0020-07

“早在古希腊时代,西方学者即已自觉探讨知识概念的内涵。苏格拉底坚持从功能、意义方面定义知识,认为知识即人的智力、自我认识和道德的修炼。”^[1]其后有很多哲学家研究知识的定义和构成条件,“知识构成的三个条件分别是信念、真和证实。”^[2]国内有学者在对知识不同概念的分析基础上认为,“知识是人类正确的认识及合理运用与组合的总和。”^[1]可以说,随着社会的进步,知识的内涵在不断发展。大数据时代,大数据来源于传感器、互联网、社交网络、政府统计部门以及各专业医疗部门、金融部门、科研部门等。大数据成为知识新的来源,大数据技术对大数据的存储、分析、挖掘、可视化等发现知识,并将这些知识应用于实践。大数据知识是相对于传统小数据知识而言的,其不仅追求对客观实在的真实反映,而且还应该保护不同主体的数据安全,最后还能解决现实的经济、环境、交通、医疗等治理问题。这样大数据知识是真、善、效、美的统一。可以说,大数据知识是基于大数据被证实为真并具有善、效、美的命题或信念。“知识论的讨论要能够寻找出存在于一切可能的知识中的共同的本性。”^{[3]46}大数据知识论在于寻找大数据知识中共同的本性。

图灵奖得主、美国计算机专家吉姆·格雷提出最著名的第四科学范式理论。在他看来,科学研究具有经验、理论、计算机模拟和数据密集型科学四种范式。这四种范式也是知识实现的四种方法。对于知识论来讲,我们不仅研究知识的发现问题,还研究知识的确证问题和实践问题。传统知识论主要侧重研究认识论层面知识主体如何发现知识,如重视知觉、经验、想象、理性等在知识发现过程中的作用。知识作为真的信念是需要确证的,知识的的确证问题主要关注知识与经验世界的相符合性等。知识的实践问题关注知识的经济、社会、生态等方面价值。大数据时代,大数据知识的实现过程包括大数据知识的发现、确证和实践,大数据知识的实现方法包括发现、确证和实践三个层面,包括大数据归纳方法、基于关联的因果分析方法和递归分析方法,不同方法解决不同的问题,彰显大数据知识实现方法的新范式。

收稿日期:2018-10-30

基金项目:国家社会科学基金后期资助项目“大数据知识论研究”(18FZX004)

作者简介:(1975—),女,山西稷山人,中共山西省委党校理论研究室研究员,理学博士。

一、大数据归纳方法：发现大数据包含的潜在知识

大数据方法起源于20世纪宇宙学,当时积累了大量观测和计算机模拟数据。面对海量数据,如何发现这些数据中的价值成为急需解决的时代课题,大数据方法因此登上历史舞台。“继演绎法和归纳法之后,大数据方法正在成为人类认识和改造世界的新工具。”^[4]“传统的数据通常是结构化数据,结构化数据(行数据)是指存储在数据库里可以用二维表结构来逻辑表达实现的数据,大数据则是混合形态的数据。在大数据中,结构化数据占15%,更多的则是非结构化的数据,如图片、日志、音频、视频、地理位置等数据,这些数据大都是非结构化的。”^[5]“目前全世界的数据已经有约75%都是非结构化数据。”^[6]大数据方法主要通过归纳结构性和非结构性大数据,发现大数据中包含的潜在知识。这样,大数据方法可以更准确地说是大数据归纳方法。

(一)大数据归纳方法的重要性

不断地发现知识和应用知识是人类最本质的特征之一。知识发现的途径主要有归纳法和演绎法。演绎法主要从少量普遍且可靠的前提出发,推演出众多个别的结论。演绎法具有严谨和理性等特点,但其致命缺点是强调推理,并不能使后验知识增加。归纳法主要通过对实验和经验世界运行状况进行归纳,从中找出事物发展的规律。但是,由于归纳的无穷量特征,如果有反例存在,归纳的结论就可能被证伪,这就需要修正原来的理论或者重新提出新理论。大数据归纳方法不是从某个前提推演出来,也不是对有限样本的归纳。所以,大数据归纳方法不是传统意义上的归纳法,大数据归纳方法主要对海量大数据进行分析,是对经验世界和网络世界的解蔽。与传统的归纳方法相比较,大数据归纳方法的重要性和特性是非常明显的。

1. 扩大知识的来源。小数据时代,知识来源于数据量小、相对不复杂的结构性数据。大数据时代,知识来源于结构性大数据和音频、视频、聊天记录等非结构性大数据。大数据归纳方法是对结构性和非结构性大数据归纳基础上的存储、分析、挖掘和可视化,解蔽这些大数据中隐含的知识。大数据归纳方法扩展了知识来源的渠道,使大量的非结构性大数据成为知识新的来源,同时将结构性和非结构性数据整合起来发现潜在知识。正是大数据归纳方法对复杂大数据的归纳性凸显出其强大的处理能力和归纳能力。

2. 动态的全样本大数据突破传统随机数据的局限性。小数据时代,数据多来源于随机采样,这往往会产生样本的局限性。大数据时代,通过对全样本的海量大数据进行分析,可彰显对经验世界和网络世界图景的全面新展示。目前,大数据归纳方法已应用于政府治理、社会治理和企业治理等领域,如企业通过大数据归纳方法可以筛选出不合格的产品,也可以预测不合格产品的特征和分布等。

3. 非预设能发现更多的潜在知识。小数据时代,我们先预设研究的目标和前提,根据研究目标去采集数据,这个过程中会将不相关的数据过滤掉,目标是很明确的,前提的预设使知识发现建立在预设条件的基础之上。大数据时代,大数据知识的发现并没有提前预设目标和前提,而是“让数据说话”。大数据仓库中包含结构性和非结构性大数据,通过对复杂的大数据的分析,从而实现对经验世界和网络世界的认知。因为没有预设目标和前提,其全样本性分析会发现我们预想不到的结果,如美国沃尔玛超市发现蛋挞与啤酒的强相关性,因而通过对大数据分析会发现更多的潜在知识。

4. 通过归纳彰显客观世界的相关关系。归纳方法的本质在于发现事物发展现象背后的因果性和客观规律。有专家认为大数据知识来源于对现象的解蔽,外在世界的展示过程就是显像的过程,而显像可以通过揭示大数据之间的相关关系展示出来。这种相关性的展示,可能是偶然相关或必然相关,也可能是强相关或弱相关。通过大数据归纳方法可以将经验世界和网络世界中大数据的相关性展示出来。有些强相关性反映事物发展的规律性,某种程度上是一种因果关系,这需要进一步确证。

(二)大数据归纳方法发现大数据潜在知识的路径

大数据知识实现过程包括大数据知识的发现、确证和实践三个层面。大数据知识的发现主要通过大

数据归纳方法来实现。大数据归纳方法通过对大数据的收集、存储、分析、挖掘和可视化,发现大数据知识中包含的潜在知识,其具体路径体现为以下几个方面。

从技术上看,大数据归纳方法是大数据技术的应用。小数据时代,数据的采集、存储、分析等多是依靠人,数据量较小,处理相对容易。大数据时代,海量大数据客观要求一种新的技术能够承担对大数据存储、分析等一系列工作,减轻人类脑力劳动和体力劳动的负担。大数据技术正是在这种需求下诞生的。大数据技术包括大数据存储技术、分析技术、挖掘技术和可视化技术等,正是依托大数据技术,大数据归纳方法才可能实现。目前,大数据技术通过分布式算法实现对大数据进行的一系列处理,在此过程中整合结构性和非结构性大数据,实现对经验世界和网络世界的镜像化,通过可视化技术将这种相关性展示出来。

从实现过程看,大数据归纳方法的运用是多语境整合的过程。大数据归纳方法体现为对结构性和非结构性大数据的归纳,还彰显为对历史语境、伦理语境、技术语境、认知语境、语言语境等的归纳。从经验世界和网络世界到大数据,需要这些客观世界的全样本数据,此过程需要历史语境中社会需要的支撑。由于一些大数据涉及国家安全、个人隐私和企业秘密等,对这些大数据的处理需要相应的伦理语境作支撑,保障大数据的安全。大数据技术从技术层面实现对大数据的存储、分析、挖掘和可视化等,没有大数据技术就没有大数据知识。人类认知决定大数据知识发现可能的边界,而从大数据到大数据潜在知识的展示,需要语言语境,即将潜在的知识用语言表征为显性的知识。所以,大数据归纳方法的运用不仅体现在技术层面,更多的体现在多语境的整合层面。

从实现结果看,大数据归纳方法实现对经验世界和网络世界潜在知识的表征。迈克尔·波拉尼将知识分为隐性知识和显性知识,通常以书面文字、图表和数学公式加以表述的知识,称为显性知识;在行动中所蕴含的未被表述的知识,称为隐性知识。要确证知识和实现知识的价值,必须将隐性知识转化为显性知识。大数据归纳方法通过相关性发现客观世界隐含的潜在知识,需要通过语言、图表、文字等形式将这种知识表征出来,即在可视化表征基础上通过语言的再加工形成大数据潜在知识。当然,对于可视化结果我们需要理性地看,并不是所有的可视化结果都能成为潜在知识。这就涉及到下一个问题即关于知识的确证问题。

二、基于关联的因果分析方法:确证大数据知识

大数据知识发现过程客观要求对全样本共享的大数据进行挖掘。由于这种特性,有些专家特别是大数据经验主义者认为,“用数据说话”“让数据发声”,仅依靠大数据就可以预测了。其实,这是不够的。原因在于事物之间的强相关性并非一定彰显事物运行的客观规律,也就是说强相关性和知识之间并不是一一对应关系。从相关程度看,有些事物之间的相关度是很低的,可以说是低相关性或偶然性相关,这种相关性并不能构成知识,原因在于并没有反映一种可靠的必然的关系。没有因果的相关性,只能是一种建议或者地方性知识。有些相关性较高,但获得的结果也不一定是知识。原因在于即便强相关也不一定存在必然性或因果关系,如病例与疾病的关系,虽然二者存在高相关性但是二者之间并不存在因果性,而基因突变与疾病之间存在因果性。所以,我们需要透过现象看事物的本质,这就需要基于关联的因果分析方法,确证大数据知识的真。

(一)基于关联的因果分析方法的重要性

大数据时代,仅依靠大数据归纳方法获得的相关性分析结果往往具有时效性和地域性特点,即小知识的特点。所以,大数据经验主义者认为大数据知识依靠相关性获得,时效性强,“我们无法追求每个问题的因果性,或者来不及细究因果性就必须即刻给出问题的解决之道,在此情境之下,相关关系分析法就显得更加有效”^[5]。但是,这种急于求成而获得的解决之道并不是事物客观运行的真实反映。基于关联的因果分析方法能够通过因果分析挖掘现象背后真正的原因,以确证大数据潜在知识的真,将潜在知识

确证为知识。

1. 挖掘大数据潜在知识背后的原因。不忘初心,知其然更要知其所以然,这是知识的本质要求。只有被确证为真的潜在知识才能构成真正意义上的知识。大数据归纳方法彰显的是经验世界和网络世界的相关性,这只是知其然的展示,这种相关性的展示结果只是一种潜在知识,其能否转化为知识,需要进一步确证。因果分析是确证知识较常用的一种方法。相关性有些是偶然的,不具有因果性,也就不构成知识,就被剔除。有些强相关性也不一定是知识,其本质需要通过对现象分析挖掘进而形成知识。这样,即使对强相关性也需要分析其原因,因为这种强相关性只是现象的展示,现象本身并不构成事物的本质,也不构成客观知识。

2. 确证大数据潜在知识为真。从知识实现进程看,当一种方法无法发现知识时客观要求一种新的方法来代替或完善。大数据归纳方法主要用于发现大数据中存在的潜在知识,但其具有局限性。“大数据强调相关性而非因果性的研究取向限制了其探究因果关系的能力……大数据缺乏发现因果关系的优势,应该将其与实验设计和观察研究相结合来获取有价值的知识。”^[7]目前,大数据归纳方法主要通过归纳形成相关性关系的展示。这样,就需要因果性分析方法作为弥补,在知其然基础上明白其所以然,以确证潜在知识与经验世界、网络世界运行规律的相符合性,确证潜在知识为真。大数据所获得的有限的总体并不能反映经验世界和网络世界的普遍性。正是由于此,因果分析方法作为补充是很必要的。

(二)基于关联的因果分析方法确证大数据知识的路径

基于关联的因果分析方法是确证大数据潜在知识为真的重要方法,要实现该方法,我们需要做好以下工作。

1. 对强相关性进行因果分析。对于知识的确证目前有三种理论,即基础主义、融贯论和外任主义,基础主义认为“知识的证实建立在某种基础之上,这一基础就是证实的初始前提”^[2];融贯论认为“所有的信念都将由它们与其他信念,将由所有这些信念相互之间的一致关系,而得到证实”^[2];外任主义认为“真正所必需的只是信念和外任事物之间的某种外在关系”^[2]。对于经验世界和网络世界来说,通过大数据归纳方法获得的相关性,有弱的偶然性的相关性,有强的必然性的相关性,也有强的偶然性的相关性等。可见,大数据知识确证并没有明显的初始前提,相关性某种程度上反映大数据潜在知识与其他信念的一致性,而这种一致性的存在,关键在于潜在知识与外在事物之间的因果性关系,仅依靠相关性获得的表征只是现象的描述并不能构成知识。对相关性背后的原因进行挖掘是知识确证很重要的途径和方法。这些原因,有一果一因、一果多因、多果一因等情况,我们需要具体问题具体分析。

2. 以思维创新对相关性进行因果分析。对经验世界来说,传统意义上追求普遍性和客观性是其使命。而通过大数据归纳方法获得的是经验世界和网络世界相关性这个果,这个果是否有原因,需要执果索因。相关关系指两种或两种以上的社会经济现象间存在着相互依存关系,但在数量上没有确定的对应关系。统计学上研究有关社会经济现象之间相互依存关系的密切程度叫做相关系数。通过相关分析,还可以测定和控制预测的误差,掌握预测结果的可靠程度,把误差控制在一个范围内。社会经济现象之间的相互关系是非常复杂的,表现出不同的类型和形态。从变量之间相互关系的方向来看,分为正相关和负相关。从变量之间相互关系的表现形式来看,可分为直线相关与非直线相关。对于大数据知识来讲,大数据归纳方法将客观世界复杂的相关关系展示出来,我们可以通过辩证思维、理性思维创新确证大数据潜在知识是否为真。

三、递归分析方法:实践大数据知识

“对大数据认识论的批判更多地指向数据主义者‘重相关轻因果’‘重事实轻理论’‘重技术轻研究’的倾向。”^[8]大数据知识作为知识的一种形态,特别是作为对经验世界和网络世界等客观世界的镜像反映,追求普遍性和客观性是其重要任务,要形成对大数据知识因果分析、理论分析是很必需的。因果分析方

法和递归分析方法正是通过弥补大数据归纳方法的不足,彰显大数据知识的因果性、理论性、普遍性和客观性。对于大数据知识来说,我们不仅需要发现和确证大数据知识,还要实践大数据知识,即实现大数据知识的效。而对大数据知识的实践需要递归分析方法。

(一)递归分析方法的重要性

“所谓递归就是把未知的归结为已知的,把较复杂情形的计算,递次地归结为较简单情况的计算,并得到计算结果为止。”^[9]也就是说,递归是把一个不能或不好直接求解的“大问题”转化成几个“小问题”来解决,“小问题”再分解,直到每个“小问题”都可以直接解决。某些问题的解决是环环相扣,前一步完成才能到后一步,这样的问题可以用递归方法来解决。递归方法要求有递归程度的设计。“递归设计就是要给出合理的‘较小问题’,然后确定‘大问题’的解与‘较小问题’之间的关系,即确定递归体,最后朝此方向分解,必然有一个单基本问题解,以此作为递归出口。”^[9]大数据知识的产生直接来源于社会实践的需求。目前,大数据知识已被广泛应用于交通、环保、医疗等领域。递归分析方法是实现大数据知识效用的重要方法。

1. 彰显大数据知识实现的递归性。递归分析方法分为向前递归和向后退回。从大数据知识实现过程看,一是向后递归,彰显大数据发现、确证、实践的动态性。首先,要借助大数据技术发现大数据中包含的潜在知识,其次需要确证其为真,构成大数据知识,最后将大数据知识应用于社会实践。可以说,大数据知识实现过程就是从大数据——大数据潜在知识——大数据知识——大数据知识应用不断向后递归的过程。二是向前递归,彰显大数据知识应用对大数据知识确证、发现的反作用。大数据知识作为知识的一种形态,在被应用过程中会发现大数据知识的不完善或缺陷,这就需要向前递归,修正大数据知识的确证方法,完善大数据仓库等,进一步推动大数据知识的发展。

2. 彰显大数据知识实践从“大问题”向“小问题”不断递归。递归分析方法需要将“大问题”递归为“小问题”,并寻找一个单基本问题解,以此作为递归口。大数据知识的实践是从“大问题”向“小问题”不断递归的过程。具体来说,大数据知识实践的主体包括政府、企业、民众、科研院所等,大数据知识实践是个“大问题”,可以递归为不同主体在观念、生产生活方式、制度、文化等方面应用大数据知识的“小问题”,以实现大数据知识的实践价值。如对于医疗大数据应用的“大问题”,我们可以递归为具体的生活方式、观念变革、文化创新等。我们发现很多疾病的产生都与民众不良生活方式有很大的关系,政府需要宣传与健康生活方式相关的观念,使民众在观念层面认识到生活方式对健康的重要性,与此同时塑造和形成良好的生活方式,并在文化层面形成健康文化。对于企业来讲,“企业如果能在这些非结构化数据中挖掘出新的知识并与业务融合,不但其决策的依据将会更加全面和准确,而且有可能形成新的核心竞争力,进而在生产模式、商业模式、管理模式等方面发生深刻变革。”^[10]大数据知识可用于筛选不合格产品,提高企业的管理效率,这就促进企业生产方式的变革。

(二)递归分析方法实践大数据知识的路径

为了更好地利用递归分析方法,将大数据知识的效表征出来,我们需要做好以下工作。

构建大数据知识递归的双向路线图。对于大数据知识来讲,一是向后递归,从对大数据的存储、分析、挖掘和可视化发现大数据中包含的潜在知识,进而通过关联的因果分析确证大数据潜在知识的真,再到大数据知识的实践应用。这个递归过程必须是畅通的。二是向前递归,即从大数据知识实践向前不断递归。大数据知识应用过程不仅彰显大数据知识的价值,而且会发现大数据知识存在的缺陷和不足。这可能与大数据知识的发现、确证等都有直接的关系,进而需要修正发现、确证等层面存在的不足。正是大数据知识递归双向路线的不断修正和完善,才能实现大数据知识的螺旋式向前发展。如对于医疗大数据知识的应用,我们发现由于患者填写个人信息的不真实性,使医疗大数据仓库所获得的大数据存在虚假性,这直接影响大数据知识的发现、确证和实践。我们需要剔除虚假大数据,在此基础上构建大数据知识递归双向路线图。还有些时候,可能是确证或实践方面存在问题,最后影响大数据知识的实现。总之,构

建大数据知识递归双向路线图,就是发现大数据实现存在问题并不断修正的过程。

寻找“小问题”的递归口。从实践层面看,大数据知识的应用是通过改变政府、企业、民众、科研人员的观念、生产生活方式、制度、文化等实现的。对于不同方面又可以进一步递归为单问题基本解,这构成“小问题”的递归口,也成为大数据知识实践最基本的落脚点。大数据知识与传统知识最大区别在于大数据知识直接服务于实践需要。因此,对于大数据知识实践应用必须递归到最基本点,才可能有真正价值。如对于医疗大数据知识,我们通过宣传将相应知识内化于民众观念中,进而影响民众的生活方式,并通过生活作息制度和饮食制度等规范民众生活习惯,在更高层面塑造健康文化。对于企业来讲,可以通过大数据知识对不合格产品的筛选,进一步分析产生不合格产品的原因,如流水线设计不合理、材料质量问题、人为因素等,我们就需要在这些方面进行创新,提高企业产品的合格率。

四、几点启示

大数据时代,大数据知识作为知识的新来源,彰显大数据技术的重要性、大数据知识实现方法的整体性、大数据知识的多元价值性和知识实现方法的继承性和创新性。

(一)彰显大数据技术的重要性。科学技术作为知识的一种形态,在知识实现过程中起着至关重要的作用。古代,由于科学技术比较落后,“古人对自然的认识是从最简单的外部现象开始的,建立在直观基础上,当直观材料不够用时,就用猜测来弥补。”^[11]随着科学技术的发展,人们对客观世界的认识可以借助技术工具来实现。“现代知识论要讨论的问题是作为主体的人通过什么样的方法或途径才能认识或达到作为客观的外界对象。”^{[3]408}大数据知识的实现直接来源于大数据技术。大数据技术实现对零散的结构性与非结构性大数据的存储、分析、挖掘和可视化,才使大数据可能转化为有价值的大数据知识。大数据技术提高了人们对客观世界的解蔽能力。

(二)彰显大数据实现方法的整体性。从发展历程看,知识论从关注知识的本体向知识的认识论和实践论转向。不同时期关注知识实现方法的层面是不同的。古代,注重感性、理性、经验等在知识发现过程中的重要作用,重视归纳方法和演绎方法的应用,以发现客观世界中的知识。近现代,随着实验方法、计算方法和模拟方法的不断应用,知识发现的方法逐步走向多元化。伴随着知识发现方法的不断进步,知识确证的方法也在不断发展,主要有基础主义、融贯论和外任主义。知识经济的兴起,知识的实践方法显得越来越重要。可以说,古代侧重知识发现方法的挖掘;近代在知识发现基础上重视对知识的确证;当代知识经济的兴起,如何挖掘知识的经济价值已成为方法论研究的重点。大数据时代,大数据技术成为发现大数据知识重要的工具,基于关联的因果分析方法用于确证大数据知识的真,而递归分析方法彰显大数据知识在实践中的治理价值。可见,大数据知识实现方法具有整体性,是发现方法、确证方法和实践方法的辩证统一。

(三)彰显大数据知识价值的多元性。人类掌握知识不仅在于认识世界,更重要的在于改造世界,所以知识具有两个很重要的价值维度,即真与效。大数据知识当然也追求真与效,但是由于大数据本身可能涉及个人隐私、企业秘密和国家安全,所以大数据知识在实现过程中还有一个重要的价值维度,即善,保护大数据本身的安全性。大数据被应用于环境环保、医疗等领域,其价值还体现美和健康等社会价值。

(四)彰显知识实现方法的继承性和创新性。从知识实现的方法看,归纳方法、演绎方法、因果分析方法、观察方法、统计方法等一直以来都是知识实现的重要方法。只是随着科学技术进步,这些方法应用的内容和形式在不断创新。大数据知识作为知识的新形态,大数据知识实现方法既具有继承性,又体现为创新性。一方面,大数据归纳方法继承了传统归纳方法的精神实质,即对结构性和非结构性大数据进行归纳分析形成潜在知识,但是又不同于传统归纳方法,即通过利用大数据技术不局限于对结构性大数据的分析,还包括对非结构性大数据的分析,也就是说归纳的对象和方法都具有创新性。另一方面,基于关联的因果分析方法继承了传统因果分析方法的精髓,即通过对强相关性分析挖掘其背后的因果性;但是

它这种强相关性的因果分析是建立在对大数据全样本基础上的因果分析,而不是简单的随机数据基础上的因果分析,体现了创新性。递归分析方法的整体性更多体现为创新性。

总之,对于大数据知识来讲,大数据归纳方法、基于关联的因果分析方法和递归分析方法主要用于解决大数据知识的发现、确证和实践问题。不同方法与不同问题之间并不是简单的一一对应关系,在现实中三种方法往往交织在一起,各自需要解决的问题侧重点不同。另外,除了这三种方法,大数据知识的实现还离不开人的主观能动性的发挥及其他的辩证分析方法、历史分析方法和逻辑分析方法等。

参考文献:

[1]孙恒志.从已有知识定义的缺陷看知识定义的科学整合[J].山东科技大学学报(社会科学版),2002(3):14-17.

[2]胡军.关于知识定义的分析[J].华中科技大学学报(社会科学版),2008(4):13-23.

[3]胡军.知识论[M].北京:北京大学出版社,2006.

[4]汪大白,徐飞.大数据:科学方法的新变革[J].自然辩证法研究,2016(1):109-114.

[5]张启良.大数据的基本知识及其比较[J].统计与咨询,2016(1):37-39.

[6]涂子沛.大数据及其成因[J].科技与社会,2014(1):14-26.

[7]孟天广,郭凤林.大数据政治学:新信息时代的政治现象及其探析路径[J].国外理论动态,2015(1):46-56.

[8]罗小燕,黄欣荣.社会科学研究的的大数据方法[J].系统科学学报,2017(4):9-12+44.

[9]刘辉,等.论递归方法的本质及其应用的领域[J].衡阳师范学院学报,2006(3):103-105.

[10]姜浩端.大数据的本质及其可能的影响[J].中国经济报告,2012(6):18-22.

[11]林德宏.科学思想史[M].南京:江苏科学技术出版社,2004:66-67.

The Realization Methods of Big Data Knowledge

SU Yujuan

(Shanxi Provincial Party School, Taiyuan 030006, China)

Abstract: From the perspective of the development of theory of knowledge (TOK), contemporary TOK not only studies the discovery of knowledge, but also studies the justification and practice of knowledge. As an important achievement in the development of contemporary TOK, big data TOK mainly studies the discovery, justification and practice of big data knowledge. From the perspective of methodology, the realization methods of big data knowledge mainly include big data induction method, correlation-based causal analysis method and recursive analysis method. Each method has a different focus and solves different knowledge realization problems. Big data induction method mainly discovers the potential knowledge contained in big data through its storage, analysis, mining, and visualization. Correlation-based causal analysis method mainly confirms the truth of big data potential knowledge through causal analysis, and realizes the conversion from big data belief to big data knowledge. Recursive analysis method mainly demonstrates the practical value of big data knowledge through forward and backward recursions. The realization methods of big data knowledge highlight the importance of big data technology, its own integrity, the diversity of big data knowledge value, succession and innovation of knowledge realization methods.

Key words: big data knowledge; big data induction method; causal analysis method; recursive analysis method

(责任编辑:黄仕军)