

大数据时代知识生产中的价值负载

丁大尉

(烟台大学文学与新闻传播学院,山东烟台 264005)

摘要:大数据算法并不是价值无涉的,算法开发过程中不可避免地嵌入了社会性目标,带来了所谓的算法偏见甚至算法歧视等新问题。算法偏见是社会偏见非常隐蔽的智能复制,包括训练数据中的偏见、算法设计者的偏见,以及算法本身的偏见等不同形式。必须进一步区分算法偏见和算法歧视,防止算法设计中的政治内嵌和资本介入,避免大数据算法沦为利益集团和权力部门实现经济、社会甚至政治目标的工具,同时辅之以有效的算法监督和审查机制,以算法正义保证并推动未来的社会正义。

关键词:大数据;算法偏见;算法歧视;价值负载

中图分类号:N031

文献标识码:A

大数据、物联网、云计算、区块链、人工智能、虚拟现实等新技术不但深刻重构了当代的社会经济生活,也带来了科学知识生产方式的深刻变革。放弃真实的物理世界,转而分析各种数据集合的知识生产方式被称为数据密集型或数据驱动型科学发现^[1]。这种知识生产的新范式在推动当代科学研究的对象、方法和逻辑发生深刻变化的同时,也给当代科学的组织管理以及科技政策的制定带来了新的挑战。

大数据时代的科学知识生产已成为当前学术界重要的研究问题域。总体来看,已有研究主要集中在三个层面。一是关于大数据知识生产哲学方法论的研究。比如英国学者舍恩伯格在其畅销书《大数据时代》中论述了“世界的本质就是数据”“一切皆可量化”的大数据世界观^[2];伯格曼研究了全数据定量分析方法及其合理性问题^[3]等。二是关于大数据知识生产新范式特征与趋势的研究,以实证研究为主。比如罗格斯关于大数据采集、建模,以及数据挖掘过程中相关性问题的研究^[4];卡布瑞拉关于大数据从精确到模糊、从因果到相关的知识生产黑箱,特别是数据分析有效性的研究^[5]。三是关于大数据知识生产社会、伦理、法律等问题的研究。比如托尼丹关于基于大数据的科研合作新网络所形成的虚拟研究社区的分析^[6];罗杰斯研究了数字产权界

定以及数据鸿沟、数据独裁、数据危机等新的社会问题^[7]等等。与国外相比,国内学者的研究稍显滞后,但也取得了很多很好的研究结论。

然而,已有研究多聚焦于大数据技术对于当代科研活动产生的积极影响的事实判断层面,鲜有基于哲学批判性视角的价值判断层面的深入考察和反思。本研究试图基于哲学方法论视角,深入分析大数据知识生产中价值负载的主要形式、产生根源及其可能带来的社会问题,以期回答大数据技术条件下当代科学、技术与社会(STS)新的互动和互构模式,从而为大数据的算法设计以及大数据规律的应用提供更全面、更深入的理解视角。

1 大数据时代知识生产中的价值负载

大数据时代的科学知识生产中,科学家放弃了小数据时代的因果性追求,将数据视为科学发现新的逻辑起点,通过挖掘分析海量数据之间或隐或现、或明或暗的相关关系寻找科学规律,实现了“让数据发声”。在大数据成为科学研究对象的同时,自然和社会也被同步映射成为一个虚拟的数据世界。从科学方法论上看,这种研究模式突破了小数据时代实验科学和计算科学中数据处理的局限。实际上,科学家排除个人偏见,尽量不要人为“污染”原

始数据,以客观、理性的科学态度保证研究对象的“本真”状态,一直是小科学时代的科学理想。然而,科学家的参与式观察和介入式操作却都不可避免地对实验数据形成了“污染”,实验和计算结果往往因个体科学家不同的认知水平、实验方法和分析能力呈现出明显的主体间性。

但大数据知识生产时代,不管是通过传感器、智能监测设备还是计算机终端所采集的海量数据所得到的“全数据样本”,还是可以代表“客观、理性”的数据模型和计算机算法,似乎都可以有效避免实验数据的污染,保证科学研究的价值无涉。舍恩伯格曾明确指出,“大数据的相关关系分析法更准、更快,而且不易受偏见的影响”^[8],甚至明确地提出了“科学始于数据”^[8]的观点。

那么,大数据时代的知识生产究竟是价值无涉的还是价值负载的?海量数据采集和处理是否可以避免小科学时代的观察渗透理论^①?数据驱动的科学研究是否将建立一种新的科研范式?对于这些问题的深入思考和回答,既是理解大数据时代科学研究的方法论基础,也丰富和拓展了当代科学哲学研究的问题域。

实际上,寻找相关性规律并不意味着否定因果性规律,大数据只不过是建构数据处理的黑箱为寻找数据规律的相关性建立另一种意义的因果性解释。不过这种因果性解释并不具有单个数据之间必然的关联性,而是按概率出现的数理统计意义上的规律性描述。从这个角度看,相关性解释并不是大数据规律的最终目标,而是为因果性解释所做的必要准备。然而,大数据操作的各个环节,从科学数据仓储的设定、数据挖掘工具的选择、数据模型的建构到大数据结果的解读,都离不开科学家敏锐的观察力、判断力和解释力。相关性关系的抽象和建立不但受科学家个人的认知水平、教育背景、专业素养的影响,还不可避免地渗透着科学家个人的主观意志、情感偏好甚至意识形态等不可描绘的默会内容。与小数据时代一样,大数据时代的科学研究规律同样存在明显的主体间性,甚至带来了科学实验的可重复性、科学结果的解释柔性等新问题^[9]。

不过,算法设计中的“黑箱化”使得普通人无法近距离感知大数据算法设计的中间过程,同时数字

计算的专业性和代码的复杂性又强化了人们对于算法的过度信任。这就导致大数据算法设计中看似纯粹的数学逻辑和数字计算实质上是建构了一种新的“技术乌托邦”,它强调以数据和代码支撑的算法的客观公正性,但实际上这种“数据客观主义”所建构的所谓“算法中立”同样也只能是一个科学理想。

算法设计过程中,各种以偏见的形式嵌入算法的价值负载形成了大数据科学研究中的“算法偏见”,算法偏见不可避免地贯穿于数据采集、算法设计与算法验证等各个环节,使得算法成了社会偏见更为隐蔽和智能的复制。同时,对于包含算法偏见的结果不加审查的应用甚至刻意使用带有偏见的数据结果,则将带来更为严重的所谓“算法歧视”。与算法偏见的不可避免不同,算法歧视将带来社会公平以及道德和伦理等新问题。不过,算法偏见并不必然地导致算法歧视,必须有效规避大数据算法设计中可能出现的算法歧视问题。

2 大数据算法生成中的社会偏见分析

大体来看,大数据算法的生成过程包括几个阶段。首先算法开发者针对具体问题构建算法模型(深度学习算法中也可以产生神经元和权值),随后为该模型输入训练数据集,算法会根据输入的数据集进行自我训练,并视训练情况初步建立一套算法系统。之后开发者将输入验证数据检验算法的科学性,如果初步形成的算法系统给出的计算结果没有通过验证,则需要通过算法的自我调整或人工调整权值对算法进行重新调整,如此循环,直至计算结果通过验证。可见,算法设计主要包括选择训练数据、设计算法和验证算法等三个环节,整个过程包括了数据的采集、清洗、分割、标注、选择、建模以及计算等不同操作。然而,这些操作并非完全中立和价值无涉,而是负载了各种类型的社会性偏见。这种社会性偏见嵌入算法、使得算法负载了一定社会价值的现象,统称为“算法偏见”。弗里德曼和尼森伯曾将计算机系统的偏见分为先行存在偏见、技术性偏见和突发性偏差等三种情况^[10]。但不管哪种偏见,都将导致数据算法具有巨大的个体差异。除了个体差异,算法的自身因素以及潜在的漏洞也经常导致

① “观察渗透理论”是美国科学哲学家 N. R. 汉森提出的著名观点,认为观察并非人对观察对象“刺激”的消极的机械反应,而是受观察者背景知识的影响。“观察渗透理论”是对逻辑实证主义所坚持的“中性观察理论”科学主义观的质疑和批判。

算法运行中的各种偏差,使得算法结果往往呈现出明显的主体间性。总体来说,除去数字量化计算本身的技术缺陷,比如量化计算对于有机和情感世界的遮蔽之外,常见的算法偏见包括以下情况:

第一,训练数据中的偏见。计算科学领域中著名的 GIGO(garbage in,garbage out)定律表明,用于机器训练的数据集合将直接影响机器的学习效果,甚至将进一步决定算法的正确性和可信度。此规律自然地带来了大数据算法训练中的 BIBO(bias in,bias out)现象。也就是说,如果用于算法训练的数据集合带有社会偏见,那么由此生成的数据规律也必然带有相应的偏见。尽管各种传感器、计算终端、网页抓取等大数据采集过程看起来是价值无涉的,但性别、民族、城乡、经济地位、教育状况等差别必然会使得大数据的采集做不到绝对的全面、客观和公正。实际情况是,生活在大数据边缘的人总会被非随机的、系统性地遗漏,他们很难成为数据样本的主体。比如,波士顿曾利用智能手机应用程序 Street Bump 来检测路面状况,公民驾车行驶过程中其手机将自动上传所采集的路面状况。但通过这种方式采集到的所谓“全样本”的大数据集本身却已经包含了明显的社会性偏见,一是驾车行驶的区域更集中在城市中心,郊区以及农村地区的原始数据集很少;二是经济条件好、受教育程度高的中青年群体上传的数据集在整个数据集合中的占比更高。这就导致通过这种数据集训练的算法将更有利于市中心高层次社会群体所居住的区域。

可见,大数据采集过程中不可避免地嵌入了社会价值、文化传统,甚至意识形态等各种隐性偏见,这些带有偏见的数据使得大数据算法内部渗透着很多默会的社会结构性偏见。数据选择的准确程度一定程度上决定了计算模型的有效性,用这些带有社会偏见的数据训练出来的算法将不可避免地带有社会偏见的烙印。实际上,算法设计中训练数据所包含的偏见经常导致算法设计的“选择性失明”。比如,“男性比女性更为理性、勇敢”等偏见导致了“女生不容易学习理科、女司机驾车容易出危险”等一般性社会偏见,用这类带有性别偏见的数据进行算法训练将不可避免地生成带有“偏见”的算法,而带有社会偏见的算法的使用必然将降低女性在公司高管、程序员之类工作中的录取比例。

第二,算法设计者的偏见。算法设计者个人不同的教育背景、价值观念、道德标准、情感偏好等主观偏见也将不可避免地嵌入算法的设计过程。科学家在进行数据挖掘之前,需要将算法的目标任务转换成能够被计算机系统识别的计算语言,这种话语转换过程具有很强的个人主观性,科学家经常“无意识地将个人偏见以及社会性偏见代入计算程序”^[11],算法的实现逻辑甚至“取决于算法设计者对该问题所具备的知识、技术水平、占有的资源以及调整参数的能力”^[12]。实际上,不但数据的选择、标注、清洗等工作明显带有开发者的个人认知痕迹,而且在算法结果的验证分析中,开发者对于数据的对比判断、结果与期望值的比较、权值的重新调整等环节也都嵌入了开发者的各种社会偏见。如果这些偏见主要基于算法设计者准确的直觉或丰富的经验,那么这类偏见将使得算法更加高效、准确,进而形成了所谓的“偏见优势”。但实际情况是,不自觉的偏见经常影响算法的公正性,带有偏见的算法的应用进一步放大了新的社会问题,带来了所谓的“算法正义问题”。

所以,科学家应有意识地区分显性和隐性社会偏见,明晰算法设计中的社会责任,对于算法的决策和应用保持审慎的态度,充分预测和评估其可能带来的社会后果,尤其是避免刻意的利益植入所带来的算法的歧视性应用,防止算法设计沦为利益攫取以及社会控制的工具和手段。美国国家标准与技术研究所(National Institute of Standards and Technology,NIST)通过对 Idemia 算法以及 Cognitec 算法的研究发现^②,即使性能最好的人脸识别系统,识别较深肤色人脸的能力也比识别浅肤色人脸的能力差,这就导致在这类软件的使用中,黑人的错误识别率往往比白人高很多。这种存在偏差的算法进一步带来了更为严重的性别、种族歧视等社会问题。尽管有人辩称其原因是黑人(尤其是女性)脸部的辨识度较低,但其根本原因却是很多公司用于测试或训练面部分析系统的数据集根本不具有代表性,这些开发者所收集的人脸数据更多地偏向白人、男性以及西方人,这就导致由这些带有偏见的数据训练出来的结果带有相应的性别、种族等偏见。

第三,算法本身的偏见。算法本身的偏见包括以下几个层面。其一,算法本身的偏见是由算法的

② 法国的 Idemia 公司和德国的 Cognitec 公司都是生产人脸识别软件的著名企业。

工作原理决定的,“算法的按优先级排序、归类处理、关联选择、过滤排除等特点使得算法本身就是一种差别对待系统”^[13]。关于这一观点更为直接的表述是,“在算法系统中,基于算法分类进行的优先化排序、关联性选择和过滤性排除是一种直接歧视,涉及差别性地对待与不公正”^[14]。可见,算法往往给人客观、公正的心理预设,但实际上,算法原理本质上却是一种“贴标签”的行为,该过程同样带有隐性甚至显性的社会偏见,这些偏见也带来了算法中的偏见。其二,算法也将不断制造偏见。机器学习算法在自我学习的过程中经常制造各种偏见,由于这种过程非常隐蔽、复杂,甚至这种“算法黑箱”使得科学家都很难对机器学习可能的结果进行预测,“技术的复杂性使算法偏见的发生十分隐蔽,程序设计、数据挖掘、数据分析的每一步都可能会使偏见悄然嵌入机器代码”^[15]。然而,看起来“黑箱化”的计算机算法本质上仍是对人类认知的学习和模仿。从这个角度看,算法中的偏见不仅是不可避免的,甚至是算法技术得以应用的前提。其三,从认识论上看,算法学习中偏见无法根本消除的原因在于,通过机器学习所得出的相关性规律,本质上是基于对过去数据和经验的复制与强化(历史经验往往被量化为各种数据),但过去已有的经验并不必然地与预测未来的相关性产生关联。然而,当这些偏见隐藏在“计算黑箱”中的时候,人们对于计算机客观性的信任却在不断加深这种“自动化偏见”。

可见,算法偏见是不可避免的,它以多种形式嵌入了从数据采集、数据训练、数据验证到开发者的最终决策的各个环节。然而,如果在非正义的训练数据不断放大各种算法偏见的同时却对其不加审查地使用,甚至算法开发中刻意嵌入利益集团特定的经济、社会目标,则将导致所谓的“算法歧视”问题。

3 大数据“算法歧视”的成因、表现及反思

3.1 大数据“算法歧视”的成因与表现

如果将大数据的相关性关系强化成实验科学中的因果关系,忽视了大数据算法应用的伦理、法律和社会影响(ELSI),则往往导致算法偏见的歧视性应用,即决策中的所谓“算法歧视”。算法歧视的表现是盲目信任机器算法,而忽视了算法设计中必然带有的先验目标,以及这种目标是否符合社会伦理规范、是否将挑战社会的公平正义等道德标准等问题。

算法歧视往往是由数据训练过程中的偏见导致的。利用带有预期目标的偏见数据训练机器算法,则将进一步强化这种偏见。验证数据为了保证这种倾向,需要不断调整算法的框架和权重,以使其适应这种预期。于是,“黑箱化”算法设计中被隐蔽地植入了社会性歧视,“算法晦涩难懂,人们很难理解算法黑箱中包含的各种潜在风险”^[16]。并且,算法歧视往往披着“科学理性”的外衣,更具隐蔽性。2020年12月,斯坦福大学医学中心(Stanford Health Care,SHC)的数十名医护人员对于医院公布的新冠疫苗接种顺序进行了抗议,因为他们无缘第一批疫苗接种。事后,医院将原因推给了计算机算法。但真实原因却是,算法设计者只是简单地按年龄进行了排序,并没有将在病毒环境中的暴露程度作为验证结果的参数和指标。

“算法歧视”现象为算法设计者提出了更高的要求。既然个人的道德标准、价值判断以及意识形态将不可避免地嵌入算法的设计过程,那么算法设计者必须明确算法设计中所肩负的社会责任和伦理规范,并以此严格要求自己。同时,决策者需要对大数据算法的应用进行宏观科学的综合考量,尽可能规避算法设计中的各种显性和隐性歧视性偏见,尤其是那些应用歧视性算法可能会引发社会问题的领域。实际上,“大多数的算法歧视归根结底是人的歧视”^[17],即使是所谓的自主决策算法,其本质也是人的理性设计的结果。从这个意义上讲,必须从道德、社会甚至法律层面对大数据算法的开发者和政府部门、科技企业等算法的使用者进行监督和管理,明确算法歧视可能发生逻辑,强化算法使用者对于歧视性问题的认识,明确不同行动者对于歧视性算法所造成的社会后果所承担的责任。

算法歧视的恶意使用将带来严重的社会后果。比如近些年广受诟病的电商平台利用大数据算法对平台客户进行的带有价格歧视的“算法杀熟”以及“算法欺生”,就是有目的地将商品以较高的价格卖给用户黏性较高或新接触该商品的用户。必须看到的是,在当代社会的政治、经济和文化决策越来越依赖大数据算法的同时,被大数据边缘化的群体却已成为“被遗忘的数据主体”,甚至逐步“失去了社会流动、经济机会甚至政治平等的权力”^[18]。大数据算法不但进一步区隔、分化了既有的社会分层,包含性别、种族、地域、宗教等偏见的算法歧视甚至制造了社会公共生活中新的更多的不平等,带来了“大

数据驱动技术的安全性、自主性、公平性和可持续性等相关问题”^[19]。

3.2 大数据“算法歧视”的反思

必须注意,算法歧视与算法偏见是不一样的。首先,算法偏见既可以是消极的,也可以是积极的。而算法歧视却将产生不良的社会后果,其结果一定是消极的。其次,算法偏见是一种认知状态,可以分为显性偏见和隐性偏见,其中的显性偏见较容易发展成为算法歧视。而算法歧视则指涉一种具体的实践行为,并非所有的偏见都会转化成歧视。再次,算法偏见是不可避免的,从某种角度看甚至是合理的。但算法歧视则应尽可能避免,不仅要在算法设计过程中辅之以必要的社会监督,还要从法律法规层面予以规约,建立相应的算法问责机制。

避免算法歧视的关键在于明确算法设计者的社会责任,算法设计者必须有能力对算法黑箱进行控制和调整,以确保算法偏见不会转化成算法歧视。首先算法设计者必须尽可能避免算法中歧视性观点的植入,尽可能地辨别算法中的隐性偏见和显性偏见,尽量避免带有显性偏见的算法转化成算法歧视。其次算法设计者应该对歧视性算法可能导致的社会后果负责,尤其是明晰算法应用可能涉及的群体以及后果。实际上,“虽然算法可以呈现多种不同的文化内涵,但它们最终仍然与定义和部署它们的人、机构以及它们所嵌入的权力关系紧密联系在一起”^[20]。在此基础上,必须加强对包含偏见和歧视的数据算法的社会性审查,算法偏见一定要在规范的、可控的范围内。在此基础上建立并完善对于算法使用部门的有效监管和有效补偿机制,确保算法设计的可控,保证算法的公平公正,避免算法沦为利益集团实现经济和社会目标的工具。

实际上,利益集团经常通过操控算法设计,有意识地输出各种带有歧视性态度的信息,从而实现预期的经济、社会甚至政治目标。然而,不管是在社会事务中还是在政治、经济等生活中,算法对于恶意植入的各种歧视却毫无抵抗能力,大数据算法在与社会的互动过程中常常表现为社会结构性偏见或歧视的智能复制。比如,谷歌的图像识别训练数据库 ImageNet 中,大部分图像数据(约 3/4)来自欧美国家,中国和印度尽管人数众多,但数据量总和所占的比重却不足 3%。于是也就不难理解为什么用 ImageNet 训练的智能算法在识别第三世界国家时经常出现“选择性失明”的现象了^[21]。

更为严重的是,利益集团或权力部门往往利用人们对于大数据技术客观性的盲目信任,在制造虚假公平的基础上,隐蔽地完成了对于不同圈层群体的经济掠夺或政治控制,大多数人却无法察觉也无力抵抗这种植入了社会歧视性目标的算法。比如在 2011 年 9 月美国民众发起的声势浩大的“占领华尔街”示威事件中,美国媒体选择了集体噤声。以 Twitter 为例,尽管网民已就该事件进行了激烈的讨论,但其“趋势发现”板块并没有根据大数据计算的结果将该事件定义为热搜事件。Twitter 算法引擎的失灵充分说明,被政治和资本所裹挟的大数据算法已成为利益集团实现其特殊的经济、社会甚至政治目标的工具和手段。

总之,不管是算法设计者还是算法使用者,都应该对大数据算法保持敬畏的心态和审慎的态度。尽管“开发者会辩称,他们的算法是中立的,只是容易被置入有偏见的数据和社会不当的错误环境中。而使用者则会声称,算法很难识别,更不用说理解,因此排除了使用者在使用中的道德含义的任何罪责”^[22],但他们必须对带有歧视的大数据算法保持高度警惕,并对算法应用可能导致的社会后果负责。总而言之,大数据算法本质上仍是一种器物层面的技术性工具,技术理性不能代替更不会超越人类理性。

4 小结

大数据算法将成为未来社会规范、经济秩序,甚至政治规则的重要驱动,以算法为中心重建人类社会秩序已成为大数据时代的社会共识。然而,大数据驱动的科学知识生产新范式并非价值无涉,大数据世界也并非自然世界的完备镜像。从本质上看,大数据算法的价值负载问题仍是技术理性和技术决定论等传统技术哲学问题的当代追问。

算法偏见不同于算法歧视,算法偏见是不可避免的,要尽可能防止算法偏见转化为算法歧视。算法正义是大数据算法设计的重要遵循,而数据公正则是保证算法正义的基本条件。必须认真审视大数据算法背后复杂的利益博弈权力重构关系,明晰科学共同体、数据公司、政府部门、社会公众等不同利益相关者的责权利关系,防止算法的滥用、误用甚至垄断。在算法成为大数据以及人工智能时代驱动社会变革和塑造社会新秩序重要力量的同时,保证算

法设计者坚持正确的道德规范、价值判断和社会伦理责任,以算法正义保证并推动未来的社会正义。

参考文献:

- [1] Hey T, et al. The Fourth Paradigm: Data – intensive Scientific Discovery [M]. Redmond: Microsoft Research, 2009.
- [2] Schonberger V M, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work and Think [M]. London: John Murray, 2013.
- [3] Borgman C L. Big Data, Little Data, No Data: Scholarship in the Networked World [M]. Massachusetts: MIT Press, 2015.
- [4] Cabrera F. Correlation isn't good enough: Causal explanation and big data [J]. Metascience, 2021, 30: 335 – 338.
- [5] Cabrera F. The fate of explanatory reasoning in age of big data [J]. Philosophy & Technology, 2021, 34: 645.
- [6] Tonidandel S, Kling E, Cortina J. Big data methods: Leveraging modern data analytic techniques to build organizational science [J]. Organizational Research Methods, 2018, 21(3): 525 – 547.
- [7] Rogers R. Digital Methods [M]. Massachusetts: MIT Press, 2013.
- [8] V · M – 舍恩伯格, K. 库克耶. 大数据时代 [M]. 杭州: 浙江人民出版社, 2013. Schonberger V M, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work and Think [M]. Hangzhou: Zhejiang People's Publishing House, 2013.
- [9] Paul B. de Laat. Algorithmic decision – making based on machine learning from big data: Can transparency restore accountability [J]. Philosophy and Technology, 2018, 31: 525 – 541.
- [10] Batya F, Helen N. Bias in computer Systems [J]. ACM Transactions on Information Systems, 1996, 14(3): 334 – 335.
- [11] Barocas S, Selbst A D. Big data's disparate impact [J]. California Law Review, 2016: 674.
- [12] Tene O, Polonetsky J. Taming the golem: Challenges of ethical algorithmic decision making [J]. North Carolina Journal of Law & Technology, 2017, 9(1): 137 – 142.
- [13] Diakopoulos N. Algorithmic accountability: Journalistic investigation of computational power structures [J]. Digital Journalism, 2015, 3(3): 398 – 415.
- [14] 杨成越, 罗先觉. 算法歧视的综合治理初探 [J]. 科学与社会, 2018, 8(4): 8 – 19. Yang C Y, Luo X J. A preliminary study on comprehensive treatment of algorithm discrimination [J]. Science and Society, 2018, 8(4): 8 – 19.
- [15] 郭小平, 秦艺轩. 解构智能传播的数据神话: 算法偏见的成因与风险治理路径 [J]. 现代传播, 2019(9): 19 – 24. Guo X P, Qin Y X. Deconstructing the data myth of intelligent propagation: The causes of algorithmic bias and the path to risk governance [J]. Modern Communication, 2019(9): 19 – 24.
- [16] Ziewitz M. Governing algorithms: Myth, mess, and method [J]. Science, Technology & Human Values, 2018, 41(1): 3 – 16.
- [17] 孟令宇. 从算法偏见到算法歧视: 算法歧视的责任问题探究 [J]. 东北大学学报(社会科学版), 2022(1): 1 – 9. Meng L Y. From algorithmic bias to algorithmic discrimination: An inquiry into the responsibility of algorithmic discrimination [J]. Journal of Northeastern University (Social science edition), 2022(1): 1 – 9.
- [18] Lerman J. Big data and its exclusions [J]. Social Science Electronic Publishing, 2013(66): 5.
- [19] Dencik L, Hintz A, Cable J. Towards data justice? The ambiguity of anti – surveillance resistance in political activism [J]. Big Data & Society, 2016, 3(2): 1 – 12.
- [20] 阿里尔 · 扎拉奇, 莫里斯 E. 斯图克. 算法的陷阱: 超级平台、算法垄断与场景欺骗 [M]. 北京: 中信出版社, 2018: 116. Ezrachi A, Stucke M E. Algorithmic Traps: Super Platform, Algorithmic Monopoly and Scene Deception [M]. Translated by Yu X. Beijing: China Citic Press, 2018: 116.
- [21] 果壳. 当 AI 学会性别歧视 [EB/OL]. http://www.huxiu.com/article/258405.html? h_S = fl. 2023 - 03 - 16. Husk. When AI learns sex discrimination [EB/OL]. http://www.huxiu.com/article/258405.html? h_S = fl. 2023 - 03 - 16.
- [22] Martin K. Ethical implications and accountability of algorithms [J]. Journal of Business Ethics, 2019, 160(4): 835 – 850.

Value – laden in knowledge production in the age of big data

DING Da – wei

(School of Literature and Journalism, Yantai University, Yantai 264005, China)

Abstract: Big data algorithms are not valueless. Social goals are inevitably embedded in the process of algorithm development. And it brings new problems such as algorithm bias and even algorithm discrimination at the same time. Algorithm bias is a very hidden intelligent copy of social bias, including the bias in training data, the bias of algorithm designers, and the bias of the algorithm itself. It is necessary to further distinguish algorithm bias and algorithm discrimination, prevent political embedding and capital intervention in algorithm design, and avoid big data algorithms becoming tools for interest groups and power departments to achieve economic, social and even political goals. At the same time, effective algorithmic supervision and review mechanisms should be employed to ensure and promote future social justice through algorithmic justice.

Key words: big data; algorithm bias; algorithmic discrimination; value load

(上接第 2080 页)

The research findings are as follows: (1) Technological equity is not a simple issue of technological diffusion; rather, it is a progressive composite proposition encompassing accessibility of access, empowerment of capabilities, and justice of outcomes. Its ideal goal is to achieve equal access to resources, fair distribution of benefits, and reasonable sharing of risks through institutional arrangements and technological design. However, in practice, due to the interweaving of multiple tensions, paradoxical characteristic that the coexistence of empowerment effects and alienation risks is presented. (2) The root cause of the paradox of technological equity lies in the interweaving and superposition of four tensions: the intersection of equity ideal and technological practice, market logic and social goal, social structure and technological system, and governance system and technological evolution. (3) The governance approach needs to complete the paradigm shift from fragmented regulation to systematic resilience – based collaborative governance. Through multiple dimensions such as institutional reconfiguration, technology domestication, entity collaboration, and global cooperation, a dynamic coupling effect can be formed between institutional adaptability supply, technology manageable design, and social initiative activation. This will provide a resilient governance practical approach for systematically resolving the paradox of technological equity.

This study reveals the deep generative mechanism of the technology equity paradox, providing theoretical support and practical approaches for its systematic governance. However, this study also has some limitation. Methodologically, it mainly relies on normative analysis and case induction, and lacks empirical verification using large sample data to test the effectiveness of the governance approach for the technology equity paradox. Future research can deepen in two aspects: Firstly, at the methodological level, cross – national panel data can be adopted to empirically test the synergy effect of the three? dimensional governance elements of institutions, technology and society; Secondly, at the contextual level, attention should be focused on emerging technology iteration scenarios to explore the dynamic adaptability of governance tools.

Key words: paradox of technological equity; empowerment effect; alienation risk; governance pathways