

基于社会实验的循证式智能社会治理研究

苏 竣^{*}

党的二十大报告明确提出,要以中国式现代化全面推进中华民族伟大复兴。当前,以人工智能为引领的新一轮科技革命和产业变革,正快速推动着全球创新版图的重构、文明秩序的重塑和人类社会的智能化转型,人类即将从传统工业社会迈向以科技进步与智能革命为基础的智能社会。在此背景下,社会智能化转型的新态势对国家治理体系和治理能力现代化提出了新的要求,智能社会治理现代化也应是中国式现代化的题中应有之义和必然要求。

随着智能社会的不断发展和演化,人类将在隐私伦理、道德、法律、安全等方面迎来更多现实挑战,人类社会在过去数千年积累的社会治理理论和方法也随着人工智能技术的蓬勃发展而面临失效和重构的困境,智能社会治理的难度和不确定性增加了。探寻行之有效的智能社会治理新路径,不仅需要当下摸着石头过河的渐进式实践探索,更需要具有前瞻性和预测性的循证式科学研究。在这场人工智能掀起的人类社会巨变中,我们需要用科学的方法来研究和应对科学技术带来的种种风险和问题,开展长期、长线的观察循证研究,深度理解人类社会转型的机理,全方位回应来自学术界、实务界的关切,探寻智能社会治理的新方向和新路径。

一、社会实验:循证式智能社会治理研究新路径

实验法为基于观察与经验习得提出的科学假说上升为科学理论搭建了桥梁,在科学发展历程中具有重要意义,是现代科学发展的基石。经过李比希、费希尔等一代又一代科学家的努力,实验研究逐渐走出实验室,进入真实的社会情境,形成了社会实验(social experiment)的研究范式,在政策方案的效果评价、科学技术的社会影响评估等领域得到了广泛应用。与自然科学领域盛行的实验室

^{*} 苏竣,清华大学公共管理学院。

研究相比,社会实验主要在真实的社会情境中进行,能在很大程度上克服个体的能动性反应导致的不确定性,所得到的结论更加贴近社会实际,具有较高的外部效度,因而成了研究人类社会变革、发掘真实社会干预因果效应的重要渠道。在过去的一百余年里,社会实验的研究逻辑被不同学科和专业的学者广泛应用于探索工业化与城市化、政策与制度创新、新兴技术发展与应用等因素引致的社会影响的软科学研究实践中,例如霍桑实验、普遍基本收入实验和邹平实验,为准确把握社会变革的规律与机制、解决社会问题提供循证的知识基础。

基于社会实验的研究方法,高度契合智能社会的治理需求。智能社会是基于数字手段和智能技术建立的、虚拟空间和物理空间高度融合的社会系统。开展智能社会治理,就是要对人工智能技术带来的社会性风险进行系统性预判和防范,更好地解决人类面临的信息不对称问题,提升物质和精神服务供给,并基于智能技术构建充满活力的、富有人文温度的、促进经济发展和社会进步的崭新社会。与传统工业社会迥异的生产方式和运转模式,使人类在过去数百年中建立的工业文明政治与社会秩序不再适用于智能社会的需求,智能社会中的国家治理和社会治理面临价值观的极端对立与撕裂、经济活动中的新型垄断模式、社会管理的超复杂形态、文化科技领域的认知风险等新兴风险和挑战。这些问题的解决,迫切需要以循证式科学研究为支撑。习近平(2018)总书记指出:“要加强人工智能发展的潜在风险研判和防范,维护人民利益和国家安全,确保人工智能安全、可靠、可控。要整合多学科力量,加强人工智能相关法律、伦理、社会问题研究,建立健全保障人工智能健康发展的法律法规、制度体系、伦理道德。”

由于智能社会治理议题的新颖性、超前性,可供观察和分析的案例、数据有限,现有研究多是基于对未来情境预设而进行的演绎与评述,以理论思辨为主,缺乏以科学循证逻辑为基础、以现实数据记录为依据的实证研究,难以形成对技术和产业实践的有效反馈(苏竣、魏钰明、黄萃,2021)。《自然》(*Nature*)2016年刊登的《人工智能研究中的盲点》也提出了类似观点:目前对于人工智能等新兴技术的社会影响的应对,主要依靠技术研发者与企业对于规则和伦理的自觉遵从,或是科学家们的主观预测和哲学思辨,而对于人工智能技术应用落地带来的真实社会影响的测量和分析,仍然是人工智能研究的盲区(Crawford & Calo, 2016)。因此,运用长周期、跨领域、多学科的实证方法,记录、描述和预判这些

正在发生和即将发生的变革,是深入研究人工智能对人类社会的影响、探索智能社会治理新路径的先决条件和基础性工作。这不仅仅具有鲜明的方法论意义,也是“实验主义治理”这一经典公共管理思潮和“干中学”实践精神在智能社会的复归。

在人工智能技术广泛应用推动社会转型的背景下,以及实证科学研究范式的推动下,利用社会实验持续跟踪和分析人工智能等新兴技术的社会影响,探究智能化时代的社会发展规律,成为推动循证式智能社会治理研究的新路径。2019年,本文作者和国内数位知名专家学者率先提出了“开展人工智能社会实验,探索智能社会治理的中国道路”的倡议。这份倡议的主旨就是利用政府和市场的力量,推动特定人工智能技术在社会治理中广泛应用,遴选有代表性的人工智能应用场景,建立实验组和对照组,开展科学抽样和伦理审查,将应用过程中产生的泛意性影响转变为内涵清晰、概念准确的可测度变量,最终形成技术规范、技术标准、政策建议等,并将之反馈给企业和政府相关部门,促进人工智能技术发展路径优化,助力国家治理体系和治理能力现代化建设(苏竣、魏钰明、黄萃,2020)。这份倡议体现了习近平总书记“以人民为中心”的治国理政思想,突出了“实践是理论创新之源”这一自然法则,复归了“实验主义治理”这一传统的公共管理思潮,遵循了科学、规范、量化、循证的研究范式,发挥了我国人工智能技术应用广泛、多样的优势,得到了学术界、产业界和实务部门的广泛关注与积极响应。

二、国内智能社会治理的实践和探索

作为一项紧扣时代需求、契合社会关切的理论方法与实践创新,国家高度重视开展人工智能社会实验,探索智能社会治理中国道路,社会各界统筹规划、顶层设计、重点发展、示范推动,现阶段人工智能社会实验的场景建设和标准制定已经在全国有序开展,实现了“理论研究—政策建议—政治决策—行政执行—组织实施”的有序循环。

2022年4月,中共中央、国务院发布《关于加快建设全国统一大市场的意见》,明确指出要“深入开展人工智能社会实验,推动制定智能社会治理相关标准”。这是中央文件第一次明确提出开展人工智能社会实验,并为开展人工智

能社会实验提供了基本遵循和具体任务。中央网信办、发改委、教育部等八部委在全国 22 个省市建设了 92 个国家智能社会治理实验基地,遴选了十大行业数百个应用场景,组织开展为期五年的智能社会治理实验探索。科技部将“开展人工智能社会实验”作为全国 18 个人工智能创新发展实验区的四项建设任务之一。2022 年 9 月,国家市场监督管理总局批准组建“智能技术社会应用与评估基础标准化工作组”,并在全国范围内征集工作组成员,标志着标准化工作即将进入正式组织建设和标准制定阶段。清华大学也积极推动智能社会治理组织建设、能力建设和专业人才培养,率先成立了智能社会治理研究院,设立了“智能社会治理创新领军工程博士项目”和“智能社会治理 MPA 项目”。

三、基于社会实验的智能社会治理整合研究框架

智能社会转型发生于社会生活的各个领域,但其带来的影响却具有很强的可迁移性。寻找人工智能技术在各领域的共性问题,利用人工智能社会实验方法对其进行解剖分析,需要智能社会治理研究学者的共同努力。2021 年,基于对人工智能应用场景生态的分析,我们提出了“微观—中观—宏观”的人工智能社会影响整合分析框架。2022 年,在多次研讨论证的基础上,结合近年来智能社会治理的研究实践,我们提出基于社会实验的智能社会治理整合研究框架。具体如下。

(一) 智能社会风险识别与形成机理研究

技术引致的社会风险,具有动态演变和不断演化的特性。智能社会中的风险隐匿于政治、经济、文化、医疗等领域,全方位地影响着智能社会下的微观个体、中观组织和宏观社会。智能社会治理研究需要有效识别各个领域和各个维度的潜在风险,并且探寻不同风险产生的影响因素,以及对个人、组织和社会的影响机制。在人工智能社会实验过程中,通过物理感知与社会感知等手段,结合数据融合、知识网络推理与社会计算等方法,对人工智能技术发展过程中的社会秩序、社会经济、生态环境、人类安全、财产隐私、生命健康、组织变迁等维度的收益与风险进行逐一定位分解,实现对相关风险态势的实时监测与把控,探索构建

敏捷智能社会治理的路径方案。

（二）智能社会演化路径与内生动力研究

社会演化常常与技术演化相伴相随。以人工智能为代表的智能技术日新月异,在梅特卡夫定律、摩尔定律、泰思勒定律的加持下,智能社会以远超农业社会和工业社会的速度快速发展。信息资源数字化、网络化、智能化背景下政府组织的扁平化,公共政策决策的透明化和规范化,公共治理的多元参与和智慧化,智能社会中的新问题、新挑战催生出的新社会组织类型、社会组织形态和社会组织参与路径,彰显着智能社会复杂多变、包罗万象的成长路径,进一步增加了预见性智能社会治理的难度。以社会实验为研究路径,能够识别出人机交互、自主系统等新技术与工业互联网、智能制造等新业态下的企事业组织形态变革和组织任务目标调整方向,以及个体心理认知极化和行为方式转变规律,掌握智能社会转型下的内生动力,从而预测智能社会演化的未来趋势,为智能社会治理提供前瞻性研究支撑。

（三）智能社会治理模式与治理工具研究

智能社会治理可以分为“以智能技术为工具基础”和“以智能技术为治理对象”两类。智能技术在社会治理领域的应用已经屡见不鲜,例如城市大脑、数字政府、工业大脑等,以智能技术为核心治理工具,对传统社会治理领域的治理流程和模式进行再造,达到更加高效和精准的治理目标。基于社会实验,可以对智能技术赋能不同领域社会治理的效果进行有效评估,寻找智能技术在提升政府公共服务绩效、优化企业组织管理模式、便捷个人社会生活等方面的最佳途径。与此同时,为应对智能技术产生的社会风险,政府部门、互联网平台、技术开发企业都在寻找智能技术带来的风险与收益平衡点,尤其是在运用政策法规和政策工具规范智能技术运用的同时,又促进智能技术良性发展。开展局部地区可控的社会实验,能够在不影响智能技术宏观发展节奏的同时,寻找出符合公共利益的智能社会治理模式和治理工具。

人工智能技术发展仍处于起步阶段,智能社会是人类从未触及的蓝海。开展人工智能社会实验,探索智能社会治理中国道路,是一项政治性、社会性、科学性兼而有之的事业。我们这代人生逢其时,正在经历着人类从工业社会向智能

社会转变的历史机遇期。开展基于社会实验的循证式智能社会治理研究,是找寻降低人工智能技术发展过程中人与社会的冲突、跨界回应人民群众的关切的必由之路。这一面向人类未来的伟大事业,需要更多具有人民中心人文精神、坚守科技创新初心和定力的研究者参与。让我们共同努力,用我们的智慧和力量,深入探索智能社会治理的中国道路,建设有人文温度的智能社会,引领人类命运共同体走向更加光明的未来。

参考文献

- 苏竣、魏钰明、黄萃,2020,《社会实验:人工智能社会影响研究的新路径》,《中国软科学》第9期。
- 苏竣、魏钰明、黄萃,2021,《基于场景生态的人工智能社会影响整合分析框架》,《科学学与科学技术管理》第5期。
- 习近平,2020,《推动我国新一代人工智能健康发展》, http://www.xinhuanet.com/politics/leaders/2018-10/31/c_1123643321.htm。
- Crawford, K. & R. Calo 2016, "There Is A Blind Spot in AI Research." *Nature* 538(7625).