智能社会研究

（双月刊）

2024 年
第 2 期
2024 年 3 月 10 日出版

2022 年 11 月 10 日创刊

总第 9 期

目　次

论文

走向负责任和可问责的金融大模型伦理治理 …… 段伟文（ 1 ）

数字民族志：数字化社会的意义共享

——社会学的方法论反思 …… 曾晨妤（ 19 ）

智慧照护系统对养老护理员的工作影响研究

——以浙江省建设智慧养老院为实证场景 …… 林　苗　张兴文（ 36 ）

“成为更好的女性”

——小红书与生活方式平台的文化政治 …… 林　欣　苗伟山（ 59 ）

沉默裁员与技能迭代

——制造业数字化升级的影响效应研究 …… 魏　丹　刘曙蕾（ 80 ）

智能时代“微粒社会”的治理议题 …… 王仕军（ 95 ）

研究报告

智慧城市赋能城市治理的发展现状及路径解析 …… 张　博　邓芳芳（117）

译文

反思“数据和人工智能向善”

——当前趋势与未来之路 …… 维勒·奥拉　詹姆斯·鲍尔斯　著

阚天颜　吕　鹏　译(139)

书评

重思深度媒介化时代的社交与关系

——读《重构关系:数字社交的本质》…… 何秋红　陈新毅(167)

平台即工厂:数字泰勒主义、劳动裂变激增与基础设施空间

——评莫里茨·奥腾立德《数字工厂》 …… 蔡振华(182)

访谈

计算技术的历史的当代启迪

——专访历史学家由杰夫 …… 由杰夫　叶瀚璋(196)

CONTENTS

THESIS

Towards Ethical Governance of A Responsible and Accountable Financial Big Model
…… Duan Weiwen(1)

Digital Ethnography: The Meaning Sharing in the Digital Society from A Sociological Methodological Perspective …… Zeng Chenyu(19)

Research on the Impact of AI Caring System on the Work of Elderly Caregivers: Evidence from the Construction of AI Nursing Homes in Zhejiang
…… Lin Miao, Zhang Xingwen(36)

"Becoming A Better Woman": Xiaohongshu (Red) and the Cultural Politics of Lifestyle Platforms …… Lin Xin, Miao Weishan(59)

Silent Layoffs and Skill Iteration: Research on the Impact of Manufacturing Digital Upgrading …… Wei Dan, Liu Shulei(80)

Governance Issues in the "Microparticle Society" of the Intelligent Era … Wang Shijun(95)

RESEARCH REPORTS

Analysis of the Current Status and Path of Smart Cities Empowering Urban Governance
…… Zhang Bo, Deng Fangfang(117)

TRANSLATED TEXT

Stepping Back from Data and AI for Good: Current Trends and Ways Forward

…… written by V. Aula, J. Bowles; trans. by Kan Tianyan, Lv Peng (139)

BOOK REVIEW

Rethinking the Social Contact and Relationship in the Era of Deep Mediatization: Review of *Together with Me: How Digital Media Changes Social Relationships*

…… He Qiuhong, Chen Xinyi (167)

Platform as Factory: Digital Taylorism, Multiplication of Labor and Infrastructure Space: Review of *The Digital Factory* by Moritz Altenried …… Cai Zhenhua (182)

INTERVIEW

Contemporary Enlightenment from the History of Computing Technology: Interview with Historian J. Yost …… J. Yost, Ye Hanzhang (196)

走向负责任和可问责的金融大模型伦理治理*

段伟文**

摘要:作为计算智能和数字思维发展的新阶段,近期兴起的基于大模型的生成式人工智能已初具通用人工智能的特征,且实现了从现实空间、数据空间到生成空间的突破性进展。然而,生成式人工智能一方面在安全和伦理上存在可能危及人类生存的风险,另一方面存在一定的社会伦理风险。生成式人工智能在金融领域的应用,是金融行业在数字化转型大潮推动下的必然趋势。在金融领域,生成式人工智能的创新应用包括优化客户体验和定制个性化推荐、金融欺诈检测和预防、风险评估和信用评分、交易和投资策略优化、借助自然语言处理提升合规效率等。与此同时,金融大模型的应用也进一步加剧了人们的担忧,数据隐私、偏见、可解释性等伦理风险在该领域备受关注,而上述风险的共性问题主要在于生成内容的"幻觉"、数据投毒以及合成数据的使用。面对金融大模型的伦理和法律风险,对其进行伦理治理将更有利于对大模型潜在伦理和法律风险的整体治理。对于政府和行业主管部门而言:首先,应深入理解不同层级的伦理和治理原则,明确其背后的价值观和优先考量,从而根据具体需求确立金融大模型伦理治理的基本理念;其次,通过明确我国科技伦理和人工智能伦理治理的指导思想,探索金融大模型伦理治理的工作思路;最后,积极开展科技伦理审查工作,走向负责任和可问责的金融大模型伦理治理。

关键词:生成式人工智能　金融领域应用　伦理治理　数据隐私　偏见和歧视　负责任和可问责

---

* 本文系国家社会科学基金重大项目"智能革命与人类深度科技化前景的哲学研究"(项目批准号:17ZDA028)、中国社会科学院"登峰战略"新兴学科和交叉学科项目(项目批准号:DF2023XXJC02)的阶段性研究成果。

** 段伟文,中国社会科学院哲学所、中国社会科学院大学哲学院。

近年来,数据驱动的人工智能及其社会应用取得迅猛发展,呈现出诸多具有颠覆性社会影响的创新前景,迫使人们不得不系统研究如何认识其可能带来的社会、伦理、安全风险,进而采取合理的应对之策。特别是继深度合成技术之后,基于大规模预训练语言模型的生成式与对话式人工智能展示出强大能力,使其在包括金融领域在内的诸多商业领域拥有巨大的应用前景,但其在社会、伦理和法律等领域的风险也不容小觑。而从风险的动态感知和敏捷治理的角度来看,实现大模型有序创新和向善发展的关键在于系统认识其潜在社会伦理风险,并使伦理治理等规制伴随其全生命周期。

毋庸置疑,生成式人工智能的巨大变革力量将对社会和经济发展产生深远影响。国际货币基金组织最近发布的《金融中的生成式人工智能:风险考量(2023)》指出:人工智能在塑造经济和金融部门发展方面发挥着越来越重要的作用,并被视为通过提高效率、改进决策流程以及创造新产品和产业来提高生产力和促进经济增长的引擎;与此同时,人工智能还通过重塑金融中介、风险管理、合规性和审慎监管的性质,迅速改变金融业格局。随着基于大模型的生成式人工智能在金融领域的广泛部署,政府和行业主管部门必须充分认识这些部署可能导致和加剧的风险,探寻通过伦理治理等规则措施有效地加以审慎应对的措施。

一、走向通用人工智能的大模型与生成空间的形成

从工程技术上讲,近期兴起的基于大模型的生成式人工智能开启了对类人类智能的探索,其最具革命性的突破在于初步展现出通用人工智能的特征。一般而言,认知科学意义上的通用人工智能是指机器人和智能软件及算法具备像人一样可以普遍泛化的智能,它们能够作为独立的智能体进行学习、认识和决策,甚至可能具有独立的自我意识。虽然当前兴起的

ChatGPT、文心一言等大语言模型还不是认知科学意义上的通用人工智能，但由于它们的智能已经可以在一定程度上实现泛化，这也就意味着“通用人工智能”不再是一个理论概念，而是已经走在工程技术实现的路上，因此在研发活动和产业政策层面已将生成式人工智能归类为通用人工智能。

（一）大模型与生成式人工智能

大模型横空出世之后，人们就其所带来的突破性创新形成了一些基本共识。

首先，大模型生产的类似人类的连贯和流畅的语言内容表明，它已突破了连贯和流畅的人机交互的临界点。大模型不仅可以针对任意话题与用户进行高质量的对话，而且能准确地按照用户意图完成分类、问答、摘要和创作等若干应用场景的自然语言理解与生成任务，进而在流畅的人机对话指引下完成各种工作。

其次，大模型获得成功的关键是通过“模仿学习+强化学习”等新的学习模式，包括预训练和人工标注等工程方法，在工程上对齐了人机目标。通过系统地实施人机价值对齐工程，大模型能够对生成内容中存在的价值观念冲突进行较为有效的调节和矫正，从而对由此带来的风险加以控制。人机对齐工程固然建立在一些数学和算法理论之上，但本质上是一种工程，其地位和重要性就像航天工程和建筑工程中的可靠性工程和质量工程一样。人机对齐工程对于大模型创新政策的重要启示在于，大模型生成内容在价值观念上的偏差原则上是可控的，但考虑到价值的模糊性和过度的价值对齐对成本和效率的影响，也应该看到阈值的设置应该具有一定的包容性。

最后，在生成式人工智能的相关讨论中，主张暂停大模型研究的人工智能研究者的基本理由是大模型相当于人工智能的“奥本海默时刻”，强调大模型所开启的通用人工智能可能会带来威胁人类生存这一长期风险。但反对这一说法的人工智能研究者认为，目前的人工智能并不会带来这种长期

风险或生存风险,真正值得关注的是大模型等人工智能创新应用所导致或加剧的偏见、歧视、虚假信息、幻觉以及过度的能源与资源消耗等现实风险。尽管人们对大模型的风险认知存在分歧,但基本的共识是,这一强大的人工智能应用将非常广泛,应该密切关注其导致的包括伦理风险在内的多重风险,并对由此带来的相关社会、伦理、法律问题加以规制和治理。

值得指出的是,当前建立在大模型基础上的生成式人工智能并不是一种独立存在的智能体,而是一个由人类和机器智能构成的巨型智能生态系统,这个生态系统的运转实质上是通过工程和系统方法在人类智能和机器智能的复杂组合之上实现的。因此,生成式人工智能不能被简单地理解为文字、图像、音视频内容的自动化生产工具,其所开启的通用人工智能将成为未来社会的基础设施:一方面,它是建立在全数据基础大模型之上的知识生产的引擎;另一方面,它正在从根本上改变人类文明的操作系统。

(二) 从现实空间、数据空间到生成空间

为了深化对基于大模型的生成式人工智能的实质的认知,应该进一步看到,生成式人工智能是计算智能和数字思维发展的新阶段,其所带来的突破性在于其在现实空间(物理空间与社会空间)、数据空间的基础上形成了全新的生成空间。

计算机和互联网应用普及之后,逐渐完成了对现实世界的数据化,由此开启了一个新的空间——数据空间。数据空间可以被视为现实空间的镜像,如消费者的数字画像。现实空间的数据化所获得的数据可以分为不同类型:(1)特征数据,如人脸、基因组等生物特征数据;(2)位置数据,如反映现实空间中人和事物的坐标参数;(3)行为数据,如人们的搜索点击数据反映的是人们的行为倾向;(4)内容数据,如语言、图像和音视频等传递了有意义的信息内容。而且,数据的内涵会随着技术的演进被重新定义,随着情感计算等技术的发展,人脸和步态等生物特征数据还会进一步转化为情感

数据。不难看出,人工智能的伦理问题在很大程度上与这些数据的使用相关。

近来,随着深度合成和生成式人工智能的发展,内容数据的合成和语言的合成逐渐达到了连贯、流畅、合理甚至以假乱真的程度,合成的文字、图像、音频和视频内容呈现出海量趋势,这就在数据空间的基础上形成了由自动合成内容构成的生成空间。显然,生成空间来自现实空间和数据空间。认识到这一点,就能认识到网络信息、文献和各种数据的数量和质量等对生成式人工智能发展的要求是基础性的,就能理解现实空间的歧视偏见、数据空间的数据毒性可能导致生成空间对相关问题的放大与强化。反过来,生成空间中的错误内容——从虚假信息、看似有理但违背事实的陈述、随意编造的文献来源到以肯定的方式表达的幻觉——存在对数据空间和现实空间的反向侵蚀,以及进一步导致生成空间内容退化的可能性。

二、生成式人工智能的长期风险和现实的社会伦理风险

这一拨基于大模型的生成式人工智能发展中呈现的一个显著特征是,从研发者到普通用户都意识到了其研究开发和应用部署所导致的潜在风险和现实风险。有的人尤其关注生成式人工智能在安全和伦理上可能危及人类生存的长期风险或生存风险,另一些人则将其正在引发的社会伦理风险作为优先关切的事项。不论持哪种立场,人们普遍认为生成式人工智能的研发应该采取审慎发展和负责任创新的态度。

(一) 生成式人工智能的长期风险

最近,作为研发者代表的开放人工智能研究中心(OpenAI)将其首要核心价值设定为“我们致力于建立安全、有益的通用人工智能(AGI),这将对

人类的未来产生巨大的积极影响”,并强调“任何对此没有帮助的事情都超出了我们关注的范围”。OpenAI 的人工智能安全和价值方面的基本理念是通过多次优化和迭代实现安全和有益于人类的通用人工智能,力图避免强大但风险不为人类所能控制的通用人工智能在人们没有意识到的情况下突然降临。但是提出暂停研究人工智能的学者的主要理由是,在伦理和安全跟不上人工智能大模型的加速创新时,应该对相关研究采取更加审慎的态度,甚至有必要暂停或暂缓相关研究。问题是这一呼吁显然不能改变各国在该领域展开的激烈竞争,而且它与 OpenAI 的理念主要基于长期风险和生存风险的视角,不论该视角是否恰当,对于目前大多数以应用为目标的大模型创新来说并非优先事项。

尽管如此,国际科技界和产业界在应对人工智能的长期风险和生存风险方面还是达成了一些基本共识,如强调大模型和生成式人工智能的研发应该更加公开、透明,不能隐藏其研究、应用和部署。为了共同应对人工智能可能导致的长期风险,在科技界和产业界进一步推动开源创新势在必行。为此,一方面,应构建合作研究和开源创新平台,科学家以个人身份共同参与开发,共同研究生成式人工智能的形成机制与导致伦理安全问题的技术原因;另一方面,应通过合作研究和开源创新,发展一系列测试工具,对其安全和伦理问题进行量化测试。

然而,欧美的科技脱钩和去风险等政策的推行,使得生成式人工智能和通用人工智能的竞争不再是单纯的科技竞争,同时也为中国参与包括应对长期风险在内的国际人工智能伦理安全治理带来了困难。

(二) 生成式人工智能的社会伦理风险

当前,大模型呈现出泛在应用的趋势,人们更关心的是基于大模型的大量生成式人工智能应用所导致的社会伦理风险。

首先,生成式人工智能可能导致知识生产方式的根本性变革,对传统的

知识产权产生颠覆性影响。目前,人工智能在艺术创作领域的应用受到了大量的侵权指控,但作为一种内容生成技术,其不能没有数据作为原料,而其所生成的内容与生产原料存在区别,因此,是否侵权较难界定,并且会随着技术的发展而越来越难以界定。这使按照现有法律对人类创作内容的知识产权保护变得越来越困难;同时,对于人工智能生成内容的版权是否应该得到保护也成为一个新问题。鉴于生成式人工智能是一种全新的内容和知识生产方式,其模型训练必须大量使用现有内容,但目前法律上对内容侵权的界定是在这种技术出现之前提出的,因此,如果要发展生成式人工智能并对其实现有效监管,须重新界定其内容使用边界。

其次,生成式人工智能将对工作方式和就业产生根本性冲击。生成式人工智能与人类高度类似的对话能力表明,它将使人机交互的便捷程度大大提升,各行各业现有的工作流程都可能因此得到简化。值得指出的是,大模型生成语言和知识的能力迫使人们将其拥有的类似知识和能力当作“水分”挤了出来,这必然带来对现有工作内容的改变和工作岗位的挤压。特别值得关注的是,这次被取代的可能是受教育程度比较高的专业技术人员,如文秘、咨询人员、翻译人员、文创人员、设计人员甚至医疗和教育等领域的人员都可能面临失业。这实际上会导致大量专业技术人员教育投入的加速折旧,加大并形成新的数字鸿沟,造成规模更大和更复杂的社会问题。

再次,在社会认知层面,人们有可能因过度依赖生成式人工智能而将其视为知识权威、道德权威乃至精神权威,形成人工智能无所不知、无所不能的认知幻觉。随着生成式人工智能进一步发展,其很可能成为普通人日常生活中的助手,帮助其解答知识、辨别是非乃至区分善恶。鉴于人工智能并不能真正理解其所生成的知识内容以及判断是非善恶,而且会产生错误或随意堆砌和编造的内容,故对人工智能的过度依赖难免放大其生成的不准确内容和知识上的错误,甚至对社会认知产生结构性的负面影响。在金融

理财的相关应用中,基于大模型的人工智能顾问可能会出现这方面的问题。另外,生成式人工智能存在过度拟人化趋势。随着人机对话的场景应用日益普遍,在商业服务等应用场景中可能会出现滥用人工智能的风险,如利用客户的情感偏好设计误导性的人机对话等。

最后,对基于大模型的生成式人工智能的伦理风险的基本认识是,生成式人工智能的广泛运用会强化目前已经显现的各种社会伦理问题。一是偏见和歧视。如果用于开发人工智能算法和模型的训练数据有偏见,那么算法和模型也会有偏见,从而导致生成式人工智能的回应和建议出现歧视性的结果。二是信息误导。人工智能语言模型所生成的对话可能会向用户提供不准确或误导性的回应,进而导致错误信息的传播。三是信息滥用。人工智能大模型及其预训练需要收集和处理大量的用户数据,其中必然涉及技术和商业保密数据以及国家安全数据,隐私数据和敏感个人信息可能会被滥用。四是虚假内容及恶意使用。尽管生成式人工智能目前还没有实现大规模商业化社会应用,但从信息网络媒体、虚拟现实和深度合成等技术的发展经验中不难看到,生成式人工智能可能被用于制造不易识别的虚假内容,甚至被恶意使用,从而影响、干预和操纵社会舆论和政治过程。五是对个人自主性的干预。例如,生成式人工智能在商业上可能被用来影响或操纵用户的行为和决策。

三、生成式人工智能在金融领域的应用及主要风险

金融大模型应用创新与伦理治理之所以备受关注,关键在于人工智能在金融领域的应用是金融行业数字化转型大潮推动下的必然趋势。众所周知,金融领域无疑是最具竞争压力、最需要风险管控和复合监管的行业。在这些压力和要求的驱使下,提高效率、节省成本、重塑客户界面、提高预测准确性以及改善风险管理和合规性成为金融领域自我变革的方向。

（一）生成式人工智能在金融领域的创新应用

当前，拥抱数字技术、实现数字化转型成为金融领域推动创新和保持竞争优势的首要战略选择，因而生成式人工智能在金融领域得到迅速应用。根据相关报道和研究报告，摩根大通、高盛等知名企业已开启生成式人工智能在金融领域的应用，相关用例包括构建自动化文档处理功能、虚拟聊天机器人助理、欺诈检测和预防、后台流程自动化、内部软件开发和信息分析等。

从拥抱生成式人工智能的潜力以推动创新并获得竞争优势的目的来看，金融领域出现了一些值得关注的功能强大的创新应用：

（1）优化客户体验和定制个性化推荐。生成式人工智能不仅可以进行自动化日常交互和提供个性化建议，而且具有类似人类表达能力的聊天机器人可以快速高效地响应与客户的互动，实时回应查询。人工智能算法还可以分析客户数据，包括交易历史和浏览行为，为金融产品和服务生成量身定制的推荐，以提高客户的参与度和忠诚度。

（2）金融欺诈检测和预防。生成式人工智能可以通过分析大量交易数据并发现欺诈活动的模式来检测欺诈，金融机构可以运用这些根据历史数据训练的人工智能模型开发出高度复杂的欺诈检测系统，以此实现对金融欺诈的实时监控、异常检测、提前预警和主动打击。据报道，Capital One 和 JMPC 利用生成式人工智能增强欺诈和可疑活动检测系统，显著减少了误报，提高了检测率，降低了成本并提高了客户满意度。

（3）风险评估和信用评分。生成式人工智能算法可以通过分析包含财务记录、信用记录和其他相关信息的广泛数据集来生成预测模型，使贷方更准确地了解借款人的信用度，从而就贷款审批做出更明智的决策，优化贷款策略，降低用户违约的可能性。

（4）交易和投资策略。基于市场历史数据训练生成模型并整合实时市场信息，人工智能算法可以生成个性化的投资建议，优化投资组合策略，乃

至预测市场趋势。值得注意的是:一方面,生成式人工智能具有强大的洞察力,可识别人类观察者可能忽视的有价值的见解;另一方面,也要看到,人类的专业知识和判断力在投资过程中仍然至关重要。因此,生成式人工智能洞察与人类决策的集成将可能创建一种结合人工智能算法和人类直觉优势的协同方法。

(5) 借助自然语言处理提升合规效率。金融科技等行业受到严格监管,在保护客户数据隐私、防止洗钱、反欺诈等方面有严格的合规要求。面对日益强化的监管要求,金融科技等行业可以将生成式人工智能与自然语言处理(NLP)相结合,通过自动分析法律和监管文件来简化合规流程。在此过程中,生成式人工智能可以更高效地识别相关信息,提取关键见解并标记潜在的合规风险,以此减少工作失误、确保合规性(Shabsigh & Boukherouaa, 2023)。

(二) 金融领域生成式人工智能应用的主要风险

近年来人工智能在金融领域特别是金融科技领域的应用,引发了诸多对大数据分析和深度学习等相伴随的固有风险的普遍担忧,当前生成式人工智能在金融领域的应用及其部署前景则进一步加剧了人们对相关风险的关切。其中,备受关注的风险包括数据隐私、算法与模型中嵌入的偏见和歧视、机器学习过程的不透明和不可解释、系统稳健性、网络安全以及对金融稳定的影响等。

在数据隐私方面,机器学习等人工智能在金融领域的应用引发了一些伦理和法律难题,主要包括训练数据集中的数据泄露、通过推理披露匿名数据的可能性、在使用和丢弃数据后对训练数据集中个人信息的记忆、输出结果直接或通过推理导致敏感信息泄露等。导致这些问题的底层原因是人工智能的应用不能不使用个人数据,而使用个人数据就存在侵犯个人隐私和导致数据泄露的可能。因此,对这些问题的应对与缓解,就成为数据驱动的

人工智能应用中伦理、法律和监管等治理与规制的主要关切点及改进方向。随着大模型和生成式人工智能越来越多地嵌入金融业务,数据隐私方面的挑战变得更加严峻。一方面,为了获得足够的欺诈检测、信用评估等方面的能力,大模型需要从互联网和社交媒体等平台抓取信息,必然涉及个人信息的收集使用以及知情同意等各种可能存在争议的问题;另一方面,生成式人工智能系统需要不断通过用户的输入对其大模型进行训练和微调,可能存在敏感财务数据与个人信息泄露的风险。

在偏见嵌入方面,人工智能应用可能会因为训练数据的不完整和缺乏代表性、算法设计受到人类偏见的影响等导致对某些个人或群体不公正的偏见和歧视,如金融决策可能会因为人工智能系统嵌入的偏见而导致金融排斥。大模型和生成式人工智能可能会加剧这一问题。其中值得关注的问题包括:(1)偏见矫正更加困难,大模型在接受大量的在线文本和其他内容数据训练时,难免受到其中嵌入的人类偏见等数据毒性的影响,而且鉴于数据的广泛性和多样性,通过数据选择等方法减少偏见变得更加困难;(2)新内容的产生可能强化人类偏见,在人类提示下产生的新内容可能带有提示者的偏见,而这些新内容本身会成为后续生成内容的数据原料,有可能使得嵌入其中的偏见进一步固化;(3)对生成内容的过度依赖导致其生成内容造成误导,必须强调的是,人工智能目前并没有意识和人格,大模型不能像人那样理解其生成内容的意义和价值取向,人工智能生成内容需要人类强化判断,以尽可能减少不准确、歧视性和错误的内容。

类似地,在可解释性、稳健性、网络安全和金融稳定等方面,大模型和生成式人工智能也加剧了人工智能金融应用的相关风险。其中,具有共性的重要问题包括生成内容的“幻觉”、数据投毒和投毒以及合成数据的使用。生成内容出现“幻觉”是与大模型生成新内容的能力相伴随的一种输出风险,在用于金融服务对话的场景中更加严重,在金融安全和消费者保护方面的风险更大,可能极大地影响金融业务的稳健性。这就要求企业级大模型

的开发应通过更有针对性、质量更好和更透明的训练数据集,最大限度地抑制这一现象。数据中毒是指在训练数据中出现一些特定输入,其可能对大模型的构建产生破坏性的影响,数据投毒是利用这一缺陷破坏大模型训练的准确性等恶意行为。合成数据是通过算法创建的、可模拟真实数据的统计分布的人工数据,可用于测试模型的稳健性,用合成数据替代真实数据有助于降低数据隐私和数据安全方面的挑战,用它模拟消费者的行为模式可以获得更具可行性的见解。合成数据还可以生成更多样化的可能的数据集,更好地捕捉现实世界事件的复杂性,从而减少真实数据集合的代表性不平衡和偏见等问题,因而有助于构建能力更强、更具可解释性和满足监管要求的大模型。但同时,对合成数据的过度使用也会导致数据退化等问题。

四、负责任和可问责的金融大模型伦理治理初探

近年来,面对人工智能大模型的巨大机遇和风险,大模型的伦理和法律风险及其治理成为全球关注的话题。如上文所述,由于大模型一方面涌现出类似人类的强大智能,另一方面又存在数据来源难追溯、易产生难以分辨的虚假信息和知识幻觉、知识产权难界定、模型难审计以及对技能和就业的巨大冲击等新问题,对建立在其上的生成式人工智能的伦理治理和法律规制带来了全新的挑战。

当前,大模型部署应用的伦理治理主要聚焦于几个关键问题。首先,如何实施负责任且可问责的人工智能审计,这涵盖了大模型的决策机制、构建方式以及使用途径。有研究提出了包含治理审计、模型审计和应用审计在内的三级审计框架,以应对这一挑战。其次,如何以透明的方式向消费者传达人工智能的伦理标准,明确揭示潜在风险,这是政府与行业主管部门的重要任务。这意味着它们需要为大模型的构建方式制定清晰、广泛的标准,并向消费者和员工明确这些标准。最后,政府和企业的人力资源部门须迅速

感知并应对人工智能对技能和就业市场的冲击。人力资源部门应迅速识别这一冲击,并从战略层面应对这一前所未有的压力。他们须为员工提供技能提升的机会,并根据预测所需的新技能和劳动力调整相关策略。

由于生成式人工智能功能强大,且已经呈现出走向通用人工智能的可能,这就使得通过伦理治理和法律规制发展负责任和可问责的人工智能显得更加迫切。由于大模型和生成式人工智能技术的发展日新月异,鉴于法律规制的相对滞后性及其主要功能在于对结果的监管,再考虑到对伦理风险的识别有助于尽早感知法律风险,从全生命周期治理和过程管理的角度来讲,对大模型的伦理治理无疑更有利于对大模型潜在伦理和法律风险的整体治理。因此,一方面,应将大模型的伦理治理作为包括法律规制在内的大模型治理的基础性工作,对数据隐私和安全、偏见和公平、可解释性和透明度、可靠性和稳健性、模型风险评估和管理、人类监督和把关等基本的伦理关切提出必要的要求;另一方面,为了保持大模型治理的整体性,伦理治理与法律规制等应该遵循相同的治理原则,法律规制也应该与相关的伦理原则相互协调。

(一) 确立金融领域大模型伦理治理的基本理念

为了更好地开展大模型的伦理治理,首先必须确立伦理原则和治理原则等基本理念。在具体的大模型研究应用中,相关研发者、部署者和管理者可以对其所进行的研究与应用制定出有针对性和简洁可行的伦理原则、治理原则。

这些伦理原则的制定不应闭门造车,其实质是对目前已经发布的各个层级的规范中的伦理原则和治理原则的具体落实。目前,这些规范大致包括以下三个层级:

一是科技伦理原则和科技伦理治理原则。《关于加强科技伦理治理的意见》明确提出了增进人类福祉、尊重生命权利、坚持公平公正、合理控制

风险、保持公开透明等科技伦理原则,同时提出了伦理先行、依法依规、敏捷治理、立足国情、开放合作等治理要求。

二是人工智能伦理规范和治理原则。《新一代人工智能治理原则:发展负责任的人工智能》明确提出了和谐友好、公平公正、包容共享、尊重隐私、安全可控、共担责任、开放协作、敏捷治理等八项原则;《新一代人工智能伦理规范》明确了增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当、提升伦理素养等六项基本伦理要求。此外,我国参与制定的《人工智能伦理问题建议书》提出了相称性和不损害、安全和安保、公平和非歧视、可持续、隐私和数据保护、人类的监督和决定、透明度和可解释性、责任和问责、认识和素养、多利益攸关方与适应性治理和协作等原则(联合国教科文组织,2021)。

三是金融领域和生成式人工智能相关的伦理指引和管理规定。《金融领域科技伦理指引》提出,在金融领域开展科技活动需要遵循守正创新、数据安全、包容普惠、公开透明、公平竞争、风险防控、绿色低碳等七个方面的价值理念和行为规范;《生成式人工智能服务管理暂行办法》提出,“国家坚持发展和安全并重、促进创新和依法治理相结合的原则,采取有效措施鼓励生成式人工智能创新发展,对生成式人工智能服务实行包容审慎和分类分级监管”,从中可见我国对生成式人工智能的监管和治理的基本原则。

确立金融领域大模型伦理治理的基本理念的第一步就是全面了解这些原则,认真理解和思考它们背后的价值观和优先考量,并由此展开价值观反思和优先事项讨论,进而明确大模型的研发和应用中应遵循的价值观和优先考量。然后,在理解和讨论的基础上,对这些原则进行归纳,并列出其所针对的问题;在此基础上探寻这些原则的基本内涵,反思其现实有效性和可能的含混之处,探讨不同的规范中相同或相近的原则之间的互操作性。最后,根据需要解决的问题和可能面对的问题,拟定出内涵相对明晰、互操作性强、既全面系统又重点突出的伦理原则和治理原则。而确立大模型伦理

治理的基本理念的过程,无疑有助于研发和创新应用者将价值观等理念纳入其心智模式和行为指南。

(二) 探索金融大模型伦理治理的工作思路

金融大模型的伦理治理要求进一步明确我国科技伦理和人工智能伦理治理的指导思想,并据此确立主要的工作思路。

正如《科技伦理审查办法(试行)》所强调的,我国开展科技伦理治理的宗旨是“强化科技伦理风险防控,促进负责任创新”。而《新一代人工智能治理原则:发展负责任的人工智能》这一文件,再次体现了我国开展科技伦理治理与人工智能伦理治理的指导思想是坚持促进创新与防范风险相统一,推动负责任的研究和创新。要使这一宗旨得到贯彻,关键在于以下三个方面的认知。

其一,对大模型和生成式人工智能兼具高度创新性和不确定性风险的技术的伦理治理和法律规制本身是一种综合性的协同创新,目的是使其充分发挥出技术创新潜力的同时又为社会所接受和信任。

其二,伦理治理和法律规制所面对的综合性的协同创新实际上是非常复杂和困难的系统工程,需要多方协同共治。从系统和整体创新的角度来讲,研发创新主体和规制监管主体之间的协同尤为重要:一方面,研发创新主体要开展负责任的创新,认真预见并实时发现各种不能忽视的风险,积极采取应对措施;另一方面,规制监管主体要充分了解生成式人工智能给社会与生产生活带来的颠覆性变化,把握其中技术逻辑的变迁,寻求对促进创新发展更具弹性和更为有效的治理。

其三,金融领域大模型的伦理治理应基于相称性原则,根据其产生的具体风险确定治理力度。《生成式人工智能服务管理暂行办法》提出,“对生成式人工智能服务实行包容审慎和分类分级监管”,强调主管部门将“针对生成式人工智能技术特点及其在有关行业和领域的服务应用,完善与创新

发展相适应的科学监管方式,制定相应的分类分级监管规则或者指引”。这一规定无疑体现了促进负责任的创新与敏捷治理的治理原则相结合的精神。对此,有法律学者指出,生成式人工智能的治理应改变我国原有的“技术支持者—服务提供者—内容生产者”的监管体系,实施“基础模型—专业模型—服务应用”的分层规制,并针对不同的层次适配不同的规制思路与工具。具体而言:基础模型层的治理应以发展为导向;专业模型层的治理应以审慎包容为理念,关注重点领域与场景的分级分类,设置合理的法律责任水平;服务应用层的治理应沿用原有的治理理念与监管工具,保证我国人工智能监管的协调性与一贯性,同时还应建立敏捷治理的监管工具箱,细化合规免责制度,给新兴技术发展留下试错空间(张凌寒,2023)。

最近,英国政府发布的《促进创新的人工智能监管方法》白皮书提出,在监管观念上,应避免可能扼杀创新的高压立法,而应采取更具适应性的方法来监管人工智能。白皮书概述了监管机构应该考虑的五项原则,以更好地促进在其所监控的行业中安全和创新地使用人工智能。这些原则分别是:第一,安全性和稳健性,人工智能的应用应该以安全和稳健的方式运行,并仔细应对管理风险;第二,透明度和可解释性,开发和部署人工智能的组织应该能够沟通何时以及如何使用它,并以适当的细节解释系统的决策过程,以与使用人工智能带来的风险相匹配;第三,公平,人工智能的使用方式应符合现行法律,并且不得歧视个人或造成不公平的商业结果;第四,问责制和治理,需要采取措施对人工智能的使用方式进行适当的监督,并对结果进行明确的问责;第五,可竞争性和补救,人们需要有明确的途径来对人工智能产生的有害结果或决定提出异议。这些观念对于探索大模型的治理具有一定的启发和借鉴价值(Secretary of State for Science, 2023)。

(三)加强科技伦理审查,走向可问责的大模型治理

使金融大模型的研发和应用实现负责任的创新,关键在于可问责的风

险评估和监督审查。对此,联合国教科文组织《人工智能伦理问题建议书》指出,应建立适当的监督、影响评估、审计和尽职调查机制,包括保护举报者,确保在人工智能系统的整个生命周期内对人工智能系统及其影响问责。为此,技术和体制方面的设计都应确保人工智能系统运行的可审计和可追溯。

2023 年 10 月,为规范科学研究、技术开发等科技活动的科技伦理审查工作,科技部会同教育部、工信部等十部门联合印发了《科技伦理审查办法(试行)》。这一办法的颁布无疑是我国加强科技伦理治理的里程碑,大模型的治理可以以此为抓手强化问责机制。根据审查办法,从事生命科学、医学、人工智能等科技活动的单位,研究内容涉及科技伦理敏感领域的,应设立科技伦理(审查)委员会。

值得关注的是,《科技伦理审查办法(试行)》对科技伦理(审查)委员会的职责提出了很高的专业要求和能力要求。除了开展科技伦理审查工作之外,还要按要求跟踪监督相关科技活动全过程,为科技人员提供科技伦理及其风险评估方面的咨询和指导,组织相关的业务和知识培训。为了更好地履行科技伦理审查责任,有效推进可问责的大模型治理,应该在加强相关认识和提升科技伦理素养的基础上推进相应的人才建设。

具体的建议是,从事金融领域大模型的机构和企业要注重培养乃至设置内部的科技伦理分析师和科技伦理架构师或科技伦理专员。其中,科技伦理分析师通过把握相关事实和证据,澄清相关观点、概念、问题与选择,为科技伦理审查等工作提供支撑;科技伦理架构师负责机构和企业的科技伦理治理的框架设计与整体协同,他们要在全面了解和参与内外相关政策和决策过程的基础上,负责协调机构和企业的科技伦理风险防范,在科技伦理审查工作实施中保障审查主体职责的履行。

参考文献

联合国教科文组织,2021,《人工智能伦理问题建议书》,https://unesdoc. unesco. org/ark:/48223/pf0000380455_chi。

张凌寒,2023,《生成式人工智能的法律定位与分层治理》,《现代法学》第 4 期。

Secretary of State for Science 2023, "A Pro-Innovation Approach to AI Regulation." https://www. gov. uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper.

Shabsigh, G. & E. Boukherouaa 2023, "Generative Artificial Intelligence in Finance: Risk Considerations." https://www. imf. org/en/Publications/fintech-notes/Issues/2023/08/18/Generative-Artificial-Intelligence-in-Finance-Risk-Considerations-537570.

数字民族志:数字化社会的意义共享*

——社会学的方法论反思

曾晨妤**

摘要:在人类社会数字化趋势下,数字民族志被视为一个新兴的方法论范式。如果说数字民族志相比于传统民族志更适应数字化时代,那么其中会有哪些新特征?此种讨论又何以能推进社会学的方法论发展及其学术使命?为回答此问题,本文从时空条件、研究主体与数据收集三方面界定数字民族志,认为它既指向一种"切身共在"与"意义共享"的概念内涵,同时也强调一种"主体共存"的概念外延。换言之,田野工作者通过拟象空间与实体空间切身共在的方法,共享局内人生活世界的双重意义,继而向局外人转译局内知识。这种对民族志之数字化的方法论反思,或有助于学术共同体意识到重新审视人类社会主体关系共存的必要性及意义。

关键词:方法论　田野调查　数字社会学　数字民族志　拟象空间

一、引子:如何在数字化社会做田野调查

自20世纪末以来,随着互联网、信息与通信技术、人工智能在全球范围内的普及,人类社会的数字化上升为全球性议题,如"智能城市"(smart city)、"数字孪生"(digital twin)的兴起便是其表现(Batty, 2018; Townsend, 2014)。在此背景下,以人类社会为核心研究对象的社会学,也开始出现数

* 本文撰写过程中受到北京大学社会学系田耕、范新光、张力生三位老师的启发与建议,特此致谢。文责自负。

** 曾晨妤,北京大学社会学系。

字民族志(digital ethnography)的研究方法创新,与之伴随的新概念包括计算社会科学(computational social science)、计算社会学(computational sociology)、数字社会学(digital sociology)等(Antunes & Dhoest, 2019; Lane & Lingel, 2022; Lupton, 2015; Murthy, 2008; Wynn, 2009)。

早在几个世纪以前,民族志便被运用于人类学、社会学研究中。它主张对特定时空之局内人(insider)生活经验的解释(interpretation)(Geertz, 1977: 21),强调不带预先假设的、细致深入的田野工作(field work),即在特定社会情境中通过足够充分的在场观察,搜集图像、文本、实物等材料以及进行深度访谈,来理解研究对象如何赋予其行动和生活以意义(田耕, 2019)。同时,民族志也是关于社会故事的叙说(Murthy, 2008),它指向一种局内外不同群体(如美国黑人贫困群体与白人中产阶级)的生活世界意义共享(Liebow, 2003)。后疫情时代,随着数字技术(digital technology)给社会互动模式带来颠覆性改变,人类社会从工业社会转向数字社会(邱泽奇,2022)。由此,数字民族志新诞为田野工作者的一种资料收集与分析方法,与之类似的概念还有赛博民族志(cyber ethnography)、网络志(netnography)、虚拟民族志(virtual ethnography)、网络民族志(online ethnography)。

在此历史演进与学科发展背景下,笔者产生了困惑:若数字化是人类历史的一个必然趋势,那么田野工作者如何研究日益数字化的社会空间?换言之,田野工作者如何通过切身共在与意义共享,来追求主体关系的共存?首先,随着数字技术的广泛应用,社会关系显然不再局限于实体空间(physical spaces),而是拓展到包括虚拟互联网、混合现实(mixed reality)在内的拟象空间(simulated spaces)。那么,有哪些新的质性方法可被用于研究这两个空间的关系?其次,如果数字民族志相比传统民族志确实更适用于研究拟象空间,其中存在哪些新的机遇与挑战?最后,这种探讨何以推进社会学的方法论发展,延续社会学乃至社会科学的学术使命?为解答这些困惑,笔者将通过文献梳理来加以廓清,在方法论层面为田野工作者提供新的

可能路径。

二、数字社会学与数字民族志

(一) 数字社会学与数字化社会

数字民族志的兴起,与数字社会学紧密相关。“数字社会学”概念据称最早可追溯至韦恩(J. Wynn)的论文,他指出数字技术可用于研究和教学目的的各种方式,比如使用录音机、照相机等数字记录设备(digital recording device)进行民族志研究(Wynn, 2009)。杜涅尔(M. Duneier)在纽约街头研究黑人摊贩时,不仅用录音机收集数据,还邀请黑人摄影师卡特尔(O. Carter)进行了纪录片制作,便是其一实证案例(Duneier, 1999)。

在随后的十多年间,“数字社会学”概念逐渐主流化。根据拉普顿(D. Lupton)的说法,2012 年英国社会学会批准建立了一个新的数字社会学研究分会,2013 年澳大利亚社会学会年会首次以“数字社会学”为主题举行了两次会议(Lupton, 2015)。近两三年,中国也出现了极大的“数字转向”,比如北京大学社会学系推进建设数字治理研究中心,哈尔滨工程大学等高校举办学术论坛,讨论诸如数字治理、数字劳动等数字时代的社会学议题。

与此同时,海内外学者也从各自角度对“数字社会学”概念加以界定,比如拉普顿基于四重类型学的定义(Lupton, 2015: 16)(表 1),或如赵一璋等人的定义:

> 数字社会学是以社会学的分析视角和研究方法探究数字技术的发展和应用、关注数字技术对人类行为和社会运行规律的影响、研究数字技术与社会环境相互建构的作用机制的社会学分支学科。(赵一璋、王明玉,2023)

表 1 拉普顿对“数字社会学”的定义

类型	内容
职业数字实践（professional digital practice）	将使用数字工具作为社会实践的一部分：建立网络、构建在线档案、宣传和分享研究成果、指导学生
数字技术使用分析（analysis of digital technology use）	研究人们使用数字技术构建他们的自我意识、具身化和社会关系，以及数字媒体在社会制度和社会结构的创造或再生产中的功能
数字数据分析（digital data analysis）	利用自然产生的数字数据进行社会研究，无论是定量的还是定性的
批判性数字社会学（critical digital sociology）	在社会文化理论的指导下，对数字技术进行自反性分析

尽管数字社会学的内涵与外延仍有待明晰，但既有的讨论或多或少反映了同一个时代背景，即数字技术的发展正在推动人类社会形态的转变。以城市规划与高频数据为例。在过去，城市规划的数据基础大多以年为单位，而规划编制的计划时长往往也在 5 年甚至 10 年以上。随着摄像头、气象监测仪、通信基站等在城市空间内的布设，这些感应设备在单位时间内所产生的总体数据量急剧增加，产生了诸如“高频城市”（high frequency city）的现象；而对大数据（big data）的技术应用反向提高了城市规划从业者的时间敏感性，并推动了城市规划编制的时间刻度精细化。

此类例子仍有很多，比如：人类本体（ontology）“赛博格”（cyborg）化，即人类从自然人过渡到机器人的倾向（包括近年来的人体脑机接口芯片植入技术试验）；人类社会空间的拟象化，即应用数字技术对真实生活的一种高度仿真。这些人类社会形态的转变，往往绕不开物人关系、互动、组织、制度等社会学核心议题。更为重要的是，对这些新兴现象的观察与解释，都有待新的研究方法与分析框架。

（二）民族志的数字化前景

在“数字社会”“数字社会学”概念兴起的同时，数字民族志作为方法论

意义上的新说法也逐渐显现。与数字社会学类似,数字民族志的定义也处在尚待丰富的状态。根据既有文献,它被视为“一种用来研究数字空间(digital space)的社会与文化的方法”(Góralska, 2020),能够结合“参与式观察与导引技术”(elicitation techniques)来观察数字空间中的“赛博文化方面”(cybercultural aspect)(Antunes & Dhoest, 2019)。在这个意义上,数字民族志与传统民族志具有内在的关联性:除田野地点不同,二者都基本关注研究者对局内人之特定文化意义的解释。

于笔者而言,“数字民族志”概念的定义往往需要先确定“空间”概念的边界。在本文语境下,人类社会的空间可被基本分为离线的实体空间与在线的拟象空间——这两个概念是相对的,二者分别对应日常生活相对真实(real)与相对数字化(digital)的一面。具体而言:实体空间反映现实(reality),即“自然的实体空间”(natural physical spaces,包括自然形成的山川河湖)和“人工的实体空间”(artificial physical spaces,包括城市建筑、道路);拟象空间则更多反映数字化现实(digital reality),包括互联网社交媒体平台(social media platform)的虚拟现实,以及利用新兴数字设备(如 Apple Vision Pro)重叠在线虚拟现实与离线实体空间的混合现实。

在人类社会数字化进程中,经典民族志方法如参与式观察、随意闲聊、半结构访谈等的实践范围将从实体空间(如三维的城市街头)拓展到拟象空间(如二维光滑的手机屏幕),而田野工作者与信息提供者的互动也会从面对面的直接关系转向以数字设备为中介的间接关系。就“数字民族志”的概念定义而言,它作为研究数字化社会的一种新方法论范式,不仅指向一种“切身共在”与“意义共享”的概念内涵,同时还强调一种“主体共存”的概念外延。数字民族志意味着:田野工作者通过切身共在于双重空间的民族志实践,可以尽可能借由自身感官浸润于局内的日常生活(包括人、物品、仪式过程等),敏锐地把握局内人的言外之意,实现生活世界的意义共享;继而,田野工作者得以通过民族志写作、影像记录等表达方式,向局外人转

译(translate)局内知识,最后可能推动人类社会问题的解释与改善,以及对主体关系共存秩序的重新思考。

(三) 数字民族志的基本特征

由于数字民族志新诞于世纪之交的数字化社会——后者比工业社会具有一些更显著的特征(如不确定性、虚拟性、匿名性)——因而数字民族志也体现出与传统民族志有着细微差异的三个主要方面,即时空条件、研究主体与研究数据。

1. 数字民族志的时空条件

通常,传统民族志实践包括四个步骤(Small, 2022):前往一处地点(travel to a site)、观察其社会互动(observe its social interaction)、与人访谈(interviews whoever will talk)、做大量田野笔记(take copious fieldnotes)。第一步往往为后面的三步提供时空前提,无论是前往一个传统社区(费孝通,2001)、一处都市街头(Liebow, 2003),还是观察一个社会组织,不同田野地点往往要求民族志学者采取不同的资料收集方式。

20 世纪末,民族志学者提出“多点民族志”(multi-sited ethnography)、“扩展地点法”(extended place method),逐渐拓展田野地点的空间边界,加深对局内的全貌理解,以适应全球流动的现代社会(Duneier, 1999; Marcus, 1995)。随着互联网的普及,聚焦于网络社群的网络民族志则进一步拓宽了田野调查的时空边界,使田野工作者无须移动地理位置便可进行线上田野(online fieldwork)和跨越时间限制的对话。

如今,数字民族志的时空条件比过去更灵活多变,田野工作者所要面对的时空维度也更加交织重叠。除了实现网络互联的智能手机、手提电脑,进行数字田野(digital ethnography)的田野工作者更能触及体现“万物互联”(Internet of everything, IoE)的在线联网设备,并更有可能通过拟象空间与实体空间的双重在场来收集不同时间线程的数据。随着田野地点的虚拟

化、拟象化,田野调查方式不仅会从“地面田野”(land-field)转到“桌面田野”(desk-field),也会由后者归至前者,实现田野工作者对田野时空的双重感知。

2. 数字民族志的对象选择

相比传统民族志方法,数字民族志学者在研究对象方面或有更多机会来优化样本结构。在现实生活的离线互动(offline interaction)中,研究者的个人特征通常会影响其对研究对象的接触。比如,青年女性可能比青年男性更容易接触到中老年女性群体,或是视障者比起“明眼人”更容易接触到残障群体。因此,田野工作者在撰写研究计划时往往要审慎考虑研究的可及性(accessible),思考诸如“我能不能找到守门人(gatekeeper)顺利进入田野”“我是否可以找到足够多的受访者”“我如何获得对方的信任”等问题。

不过,由于互联网的匿名性和公共性,田野工作者可以更容易检索到研究对象的在线互动区域,并与之建立研究关系。安图内斯(D. Antunes)等人的研究显示,Web 2.0 技术为那些呈现孤独症状的人(autistic people)提供了一个相互交流的包容性环境,研究者可在“孤独症相关社区”(autistic communities)中观察孤独症状者及其家属与其他照料者(caregivers)的社会互动,而这往往在实体空间中难以捕捉(Antunes, Dhoest, 2019)。不过,如果想要知道目标社群的真实生活,仅仅在网络上做田野调查是远远不够的。据米勒团队(Miller, Rabho & Awondo et al., 2021)的数字人类学研究,他们不仅观察东亚、南美洲等地区居民使用 Line、微信、WhatsApp 等应用软件的习惯,还结合居民的线下生活与社会背景来理解智能手机对全球人类社会文化的总体性影响。

3. 数字民族志的数据来源

在数据方面,仰赖于与数字技术有关的事物(如数字设备、数字平台),数字民族志学者可接触到海量数据,如文本(text)、图像(image)、视频(video)等。这些海量数据不仅包括对实体空间或拟象空间的单一数字化记录,更

涉及两个空间的双重数据交叉。

通常,数字民族志研究有两类倾向:一是以数字技术为研究对象来探究其社会影响;二是以数字技术为研究工具来推进研究。研究者将数字技术应用(digital technology use)融入民族志的传统实践中,能获得丰富、高度情境化的“厚数据”(thick data)。例如,贝尔(G. Bell)对印度、马来西亚、新加坡和印度尼西亚的 50 多个家庭进行调查,观察这些家庭成员使用数字技术的方式与具体活动,并进行采访、拍照和制作技术清单(Bell, 2006)。当然,这两类倾向也可以在一项研究中结合起来,如米勒(D. Miller)等人不仅将数字设备作为研究对象,以此来探讨智能手机(smartphone)与老龄化(aging)的关系,还将这一数字设备作为研究工具(Miller, Rabho & Awondo et al., 2021)。总之,无论是哪一类倾向,数字民族志学者都能够从中获取来自实体空间或拟象空间的双重数据,这不仅拓宽了资料收集的渠道,也增大了数据的规模量,提高了数据的颗粒度。

三、数字民族志的机遇与挑战

上文提到数字民族志的三种新特征(或机遇),分别对应三种新挑战(表 2)。在方法论层面,我们可以通过理论上的廓清,来一窥社会学田野调查的数字化前景及其应对的可能路径。

表 2　　数字民族志的特征/机遇及其带来的挑战

维度	数字民族志的特征/机遇	数字民族志带来的挑战
时空条件	跨越时空限制的数字田野	时空边界的不清晰
研究主体	加大研究对象的接触机会	研究关系的处理
研究数据	增加研究数据的收集来源	研究数据的收集及储存

(一) 时空边界的不清

数字民族志带来的挑战之一,是如何应对田野工作时空边界不清晰的问题。在线下民族志的经典实践中,田野调查的时空延伸是田野工作者进入田野后获得局内人信任、深度感知局内情境的重要条件之一。有的民族志学者会以纪录片形式来记录少数民族祭祀仪式的全过程,包括前期人力物资的准备、不同人物的仪式实践、祭祀仪式后的人物对谈,以此描绘当地传统风俗在当代中国的流变。与单纯参与式观察或结构式访谈不同,民族志传统不止步于白描局内人干了什么或说了自己干了什么,更希望通过比较二者之偏差来进一步问题化(problematize)。

一旦田野工作者将视野从实体空间拓宽到拟象空间,就要理解研究对象在双重空间中的话语与实践。这种民族志实践的转变有助于整体性地理解研究对象,借助拟象空间里的田野观察来加深对实体空间的思考。比如,里特尔(C. Ritter)在挪威一家游戏公司进行民族志研究时,不满足于仅仅在现实的办公室内观察软件开发者们,还在 Slack 协同办公平台上观察他们是如何工作的——这被他称为针对数字平台界面(interfaces)的"技术漫游"(technical walkthrough),从而分析人与技术之间的动态关系(Ritter, 2022)。里特尔指出,数字平台劳动的实践社区之双重意义结构,能被这种漫游法和参与式观察的结合所捕捉。这种融合法也有助于拓展"数字-实体连续统"(digital-physical continuum)中的民族志知识范畴。

与此同时,往返于拟象空间与实体空间的数字民族志特征,对田野工作者来说是不小的挑战。一方面,比起传统的线下民族志,数字民族志意味着田野工作者要更长时间蹲守在数字设备前;他们或许会无意识地拉长"桌面田野"的工作时间,甚至 24 小时都在线,以免错过任何可能的重要线索。另一方面,相比全然驻扎线上的网络民族志,数字民族志又意味着实体空间—拟象空间的反复移动,以及个人生活与学术工作的边界模糊化;田野工

作者可能过于沉浸在双重时空的田野调查中,继而因为对个人身心健康的疏漏而间接影响研究质量。

针对数字民族志带来的这一挑战,田野工作者须自行划定清晰的田野工作边界(clear fieldwork boundaries)。戈拉斯卡(M. Góralska)给出了一些建议:通过规律性休息、限定在线时长、设定离线时间点来提高研究的质量与可持续性(Góralska, 2020)。值得注意的是,数字化时代背景下的时空边界并不清晰,甚至进一步模糊了日常生活的真假边界。随着虚拟人物形象技术的广泛应用,“数字主播”“数字客服”等新兴现象会干扰田野工作者识别人工智能生成产品的准确度,继而重构民族志学者对于时空之真实(感)的理解与把握。

(二)研究关系的处理

将数字技术引入民族志研究,可能会影响研究者与研究对象的关系。数字民族志带来的挑战之二,是研究者如何(或在何种程度上)公开自己的学术身份,以及如何(或在何种程度上)考虑研究对象的匿名性(anonymity)与保密性(confidentiality)——这些是社会研究中非常重要的伦理议题(Babbie, 2009;陈向明,2000)。

公开研究(overt research)还是隐蔽研究(cover research),一直是困扰民族志学者的问题。一方面,如果研究者公开自己的学术身份,也许符合了尊重研究对象、确保其知情的研究伦理规范,那么这可能不利于研究资料的收集;另一方面,如果研究者选择不公开学术身份,这虽然有违伦理规范,但或许有助于观察研究对象在自然状态下的言行举止,避免研究者对他们施加的可能干扰。总之,“公开还是不公开,这是一个问题”,并且“公开或不公开,都会成为问题”。

对于这一问题,至少有两派主流说法:一派坚决主张公开,另一派则主张根据情境与后果来选择是否公开。借由《欲望交易:亚洲的崛起、西方的

衰落和全球性工作的隐性货币》(*Dealing in Desire: Asian Ascendancy, Western Decline, and the Hidden Currencies of Global Sex Work*)一书,我们可以理解如何有选择性地公开。作者黄(K. Hoang,音)在胡志明市的四个酒吧工作以便接触性工作者的过程中,会刻意闭口不提自己的真实身份,但也会在需要访谈男性客户时透露自己的学术身份或部分研究问题(Hoang,2015)。作者认为,一旦学术身份公开,则与自己接触的人可能会被认为是与性产业有关的人,有所保留的公开有助于保护研究参与者的隐私权利。

在现实生活中,研究者往往要对访谈对象的个人信息进行保密和匿名化处理,从而保证研究伦理的核心之一——尽量避免研究参与者受到伤害(no harm)。当田野地点从实体空间转移至拟象空间,处理研究关系的问题并不会骤然消失,反而可能更为紧张。在高度匿名化的拟象空间里,田野工作者更有可能成为"看不见的观察者"(invisible observer),即只要他们选择潜伏(lurk),几乎没有一个网络社区成员会意识到自己正在被窥探(spy on)。尽管这在某种程度上会减轻研究者的研究伦理压力,但其中的张力在于:一旦研究者向某网络社群透露了自己的学者身份,其行为所留下的痕迹数据将迅速被转送至其他网络社群,从而可能影响田野工作的实质进展。

进一步的问题与"双重匿名"有关。在拟象空间中,田野工作者可能需要对本就已经匿名化的对象进行第二次匿名。假如有学者想通过某款热门健身类应用软件及其网络社群进行观察来分析日常生活空间使用的性别差异,可能难以通过用户自设的标签(如性别、年龄、地区等)来直接判定其真实情况。尽管通信运营商所掌握的手机信令数据可以提供较准确的用户信息,但这些数据并不能够轻易获取,其数据规模之大与更新频率之高也会给民族志学者带来极大挑战。与此相关的新问题仍有很多,比如:数字民族志学者如何获取研究对象的人口学特征?在多大程度上能确证研究对象提供的个人信息契合实际?此外,对读者来说,如何判断数字民族志学者在文章中呈现出的研究对象是真实存在的?如何判断数字民族志学者是否为了其

他目的而虚构出所谓的“研究对象”？这一系列疑惑不仅涉及匿名性，还关乎研究本身的真实性与可靠性。

（三）数据的收集及储存

信息时代的主要特征是数据太多、太杂且更新频率太高，而数据的时间刻度单位也从（十）年缩小到周、日，甚至分、秒——这无疑增加了资料收集的难度。尽管海量数据意味着数字民族志学者更有机会接触到“不方便样本”（inconvenient sample），即尽量扩展可能会影响原初结论的样本结构，实现更令人信服的民族志研究（Duneier，2011），但面对这么多的数据，田野工作者难免会更加困惑：如何知道自己收集到了足够多的研究数据？在田野调查进行中或结束后，如何妥当储存所收集的研究数据？

通常，对质性研究者或民族志研究者而言，“数据是否饱和”是判定资料收集程度的一个主要标准。而在往返于实体空间与拟象空间的过程中，田野工作者可能会担忧数据不够充足，因而想要尽可能多地收集数据。这里存在两个挑战：一是数据收集未必那么顺利，二是所收集的大量数据未必都能用得上。虽然数字民族志学者有机会接触到海量数据，但这并不意味着他们在拟象空间展开访谈是容易的。戈拉斯卡在研究数字平台的健康知识生产时，就意识到在网上做深度访谈并没有参与式观察那样行之有效，因为前者的拒访率较高（Góralska，2020）。与此同时，田野工作者也需要花费更多时间、精力来整理这些数据，包括识别哪些数据更加契合研究，剔除那些与研究议题无关的数据，或是对数据进行清洗使之便于分析。

可以说，伴随着人类社会数字化而来的挑战，也许不是数据收集不够饱和，也可能是过于饱和或者假性饱和（看似收集了很多数据，但无助于对研究问题的回答）。或许，数字民族志学者在反复切换于双重空间的过程中，可以通过拟象空间与实体空间的相互限定（比如借由面对面访谈的线索来

限定线上观察的范围),于数据浪潮中抓到一条主线,继而在数据收集方面更游刃有余。

除研究数据的收集之外,研究数据的安全也十分重要,这也往往涉及研究关系的处理。即使研究者能够恰当处理身份公开问题,收集到经研究对象知情同意的数据,但接下来的问题是:如何规避因个人数字设备里的数据泄露而曝光研究对象的意外情况?在过去信息与通信技术尚不发达的年代,学者们通常采取手写、手绘等朴素方式记录田野中的所见所闻,比如曾对美国唐人街做过城市民族志的吴景超先生便善于使用手写卡片。如今,具有录音、摄像等功能的在线联网设备极大提高了研究数据收集与储存的效率,但诸如黑客窃取用户隐私、Cookie 技术应用等也带来了数据安全风险。万巴伦(S. van Baalen)为此提出了十条基本原则,包括做数字风险评估(digital risk assessment)、在所有电子设备中设置复杂密码、在与研究参与者沟通时采用加密技术等(van Baalen, 2018)。总体而言,研究数据的收集及其初步阶段的储存、整理,仍是数字民族志学者需要高度重视的方面之一。

四、数字民族志的方法论反思

(一) 数字民族志的余论

基于对既有文献的梳理与评述,笔者从时空条件、研究主体与数据收集三个方面讨论了数字民族志的新特征,也探讨了这三个特征可能带给田野工作者的机遇与挑战。此外,还有许多可以展开的重要议题。一是在民族志实践方面,鉴于有的文献存在以“数字民族志”之名而冠以“内容分析”之实的情况,如何进一步明确数字民族志的核心特征,使之展现时代新貌又可复归民族志的传统(包括局内外视角与意义共享)。二是在研究伦理方面,

如何斟酌田野过程中研究关系的处理(包括匿名性、保密性等),使田野工作者更恰当有效地理解、转译数字化的日常生活。三是在民族志写作方面,尽管已有不少文献声称采用数字民族志,但其中关于研究方法的讨论是相对缺失的,如何在实质性作品中呈现数字民族志的运用过程,从而提高自身研究论点的说服力。四是在田野工作者一端,如何应对资料收集及初步整理阶段可能出现的心理压力。因文章篇幅有限,笔者不便就此展开,而留待后续探讨。

值得注意的是,尽管数字民族志可作为研究数字化社会的一个新方法,但它并不是一个必选项。毕竟不是所有的民族志学者都具备相应条件,如充足的资金、先进的设备或相关的敏锐性等。正如在过去工业化社会里的快速城市化趋势下,民族志学者愈加难以找到一处时空条件相对稳定的社区进行长期参与式观察,如今伴随着数字技术应用而来的实体空间拟象化与时空不稳固化,在数字化社会里做田野工作的难度也无疑大大增加了。

可以说,数字民族志看似轻松,实则挑战与机遇伴生。仅仅在网络上收集文本或是做做在线访谈,或许难以称得上真正意义上的数字民族志。因为它不仅涉及研究技术的实际操作,更触及方法论的反思与认识论的深究。

(二)方法论反思:共在、共享与共存

最后,笔者希望借由新兴的数字民族志来完成方法论反思的下一步任务。前文谈到的三个特征其实并不新鲜,相反,这些议题早已置于田野工作者的"论坛圆桌"上了。尽管人类社会的数字化趋势为这些议题附上了新貌,但它们的内在机理往往与社会科学史相关联。

如果把民族志比喻为一棵洋葱,则它最表层的机理是具体操作方法,比如怎样通过参与式观察收集到有针对性的线索、在街头偶遇他人时如何展开闲聊。第二层机理是方法论探讨,比如怎样看待研究主体间的深厚情感

而干扰研究者价值中立的具身性代价(embodied cost),以及如何应对因过于深描细节而忽略结构性因素的民族志谬误(ethnographic fallacy)。第三层机理则从方法论层面转向认识论层面,包括如何界定民族志与其他质性研究方法(如内容分析)之异同的民族志性(ethnographic)问题(福伯格、希尔特,2023)。最内核的机理则抽象到社会学乃至社会科学的历史脉络上,回应更为长期的根本性问题:应秉持一种怎样的研究立场及使用什么实践方法来完成解释人类社会的学术使命?在既有民族志作品中,关于前三层机理的讨论通常体现在方法论附录里(Duneier, 1999; Hoang, 2015; Vaughan, 1996),同时也会体现于正文的重点段落中(Desmond, 2016; Klinenberg, 2015);对最内核机理的阐述,往往折射出学者对自我与他者之主体关系的深层次反思,而非仅仅将田野调查视为研究局内的一种科学方法。

换言之,对(数字)民族志的方法论反思是十分重要的。一套科学系统、逻辑自洽的研究方法往往有助于夯实一部作品之研究论点的论证基础,使之令读者不仅感到无比信服,还能对局内生活感同身受。不过,与其把民族志视为一种方法,倒不如说它是一种观察力(a way of seeing),是"我们活在世上的本质状态"(Desmond, 2016; Wolcott, 1999)。因而对于社会科学工作者而言,更为关键的是思考方法论之上的学术使命:作为和常人一样的人类个体,我们可以用何种方式来知道别人在想些什么或做些什么,以及要以何种姿态来加以看待?并且,通过这样一系列的实践与理论,我们如何描绘、解释自身以外的生活世界之变化,使之不仅有助于人类社会问题的解决与改善,还能推动人类与其他主体(如动物、植物等)关系的共存?

总之,从方法论的角度出发,无论用数字设备来完成田野工作,还是分析数字技术的社会性影响,也无论未来涌现出何种新兴的研究方法,更为关键的是学术共同体意识到重新审视人类社会主体关系共存的必要及意义——而笔者对上述议题的思考便以此为期待。

参考文献

陈向明,2000,《质的研究方法与社会科学研究》,北京:教育科学出版社。

费孝通,2001,《江村经济》,北京:商务印书馆。

福伯格、希尔特,2023,《何为数字民族志中的民族志性——一个社会学视角》,倪燕萍、丁旖译,《智能社会研究》第 4 期。

邱泽奇,2022,《数字社会与计算社会学的演进》,《江苏社会科学》第 1 期。

田耕,2019,《社会学民族志的力量:重返早期社会研究的田野工作》,《社会》第 1 期。

赵一璋、王明玉,2023,《数字社会学:国际视野下的源起、发展与展望》,《社会学研究》第 2 期。

Antunes, D. & A. Dhoest 2019, "We Are People and So Are They: Shared Intimacies and the Ethics of Digital Ethnography in Autism Communities." *Research Ethics* 15(2).

Babbie, E. 2009, *The Practice of Social Research*, New York: Wadsworth.

Batty, M. 2018, "Digital Twins." *Environment and Planning B: Urban Analytics and City Science* 45(5).

Bell, G. 2006, "The Age of the Thumb: A Cultural Reading of Mobile Technologies from Asia." *Knowledge, Technology & Policy* 19(2).

Desmond, M. 2016. *Evicted: Poverty and Profit in the American City*, New York: Crown.

Duneier, M. 1999, *Sidewalk*, New York: Farrar, Straus & Giroux.

Duneier, M. 2011, "How Not to Lie with Ethnography." *Sociological Methodology* 41(1).

Geertz, G. 1977, *The Interpretation of Cultures: Selected Essays by Clifford Geertz*, New York: Basic Books.

Góralska, M. 2020, "Anthropology from Home: Advice on Digital Ethnography for the Pandemic Times." *Anthropology in Action* 27.

Hoang, K. 2015, *Dealing in Desire: Asian Ascendancy, Western Decline, and the Hidden Currencies of Global Sex Work*, Berkeley: University of California Press.

Klinenberg, E. 2015, *Heat Wave: A Social Autopsy of Disaster in Chicago*, Chicago: The University of Chicago Press.

Lane, J. & J. Lingel 2022, "Digital Ethnography for Sociology: Craft, Rigor, and Creativity." *Qualitative Sociology* 45(3).

Liebow, E. 2003, *Tally's Corner: A Study of Negro Streetcorner Men*, Washington: Rowman & Littlefield Publishers.

Lupton, D. 2015, *Digital Sociology*, New York: Routledge.

Marcus, G. 1995. "Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography." *Annual Review of Anthropology* 24.

Miller, D., L. Rabho & P. Awondo et al. 2021, *The Global Smartphone: Beyond A Youth Technology*, London: University College London Press.

Murthy, D. 2008, "Digital Ethnography." *Sociology* 42(5).

Ritter, C. 2022, "Rethinking Digital Ethnography: A Qualitative Approach to Understanding Interfaces." *Qualitative Research* 22(6).

Small, M. 2022, "Ethnography Upgraded." *Qualitative Sociology* 45(3).

Townsend, A. 2014, *Smart Cities: Big Data, Civic Hackers, and the Quest for A New Utopia*, New York: W. W. Norton & Company.

van Baalen, S. 2018. "Google Wants to Know Your Location: The Ethical Challenges of Fieldwork in the Digital Age." *Research Ethics* 14(4).

Vaughan, D. 1996, *The Challenger Launch Decision: Risky, Technology, Culture, and Deviance at NASA*, Chicago: The University of Chicago Press.

Wolcott, H. 1999, *Ethnography: A Way of Seeing*, Lanham: Alta Mira Press.

Wynn, J. 2009, "Digital Sociology: Emergent Technologies in the Field and the Classroom." *Sociological Forum* 24(2).

智慧照护系统对养老护理员的工作影响研究*

——以浙江省建设智慧养老院为实证场景

林　苗　张兴文**

摘要：本研究在浙江省建设智慧养老院的大场景下，基于杭州市两所养老机构的社会实验和实证调研数据，对智慧照护系统如何影响养老护理员的工作进行研究。结果显示，在控制变量的情况下，使用智慧照护系统会显著提升养老护理员的工作满意度，但前提是养老护理员对智慧照护系统是认可和满意的。此外，养老护理员的信息化水平与其使用智慧照护系统之间的关系，还须进一步探索其内在作用机制。本研究认为，未来人工智能技术与养老服务产业的融合，需要全面考虑供需各方的真实需求，包括老年人的服务需求和养老护理员的使用需求，设计和开发便捷、易用的养老服务系统和产品，同时也要有针对性地提升养老护理员的信息化水平，让人和技术更好地兼容耦合，有效推动养老服务行业发展提质增速。

关键词：智慧照护系统　养老护理员　工作满意度　信息化水平

随着人工智能技术的兴起，物联网、云计算、大数据等新一代信息技术和智能硬件产品在养老服务领域得到了深度应用，正在促进形成新的养老

* 本文系浙江省软科学研究计划项目“人工智能养老发展趋势、路径及浙江应对研究”（项目批准号：2021C35104）的阶段性研究成果。

** 林苗，之江实验室智能社会治理研究中心、浙江省哲学社会科学试点实验室之江实验室智能社会治理实验室。张兴文，之江实验室智能社会治理研究中心、浙江省哲学社会科学试点实验室之江实验室智能社会治理实验室。

服务业态(何妮、霍聪聪、徐功铖等,2021)。近年来,许多学者开始致力于智慧养老、“互联网+”养老等相关研究,旨在进一步探索各类人工智能技术如何促进养老服务领域的发展,推动养老服务产业更加健康良性发展(睢党臣、曹献雨,2019)。

我国的养老模式以“9073”为主要配置,即 90%左右的老年人居家养老,7%左右的老年人依托社区养老,3%左右的老年人入住机构养老。受传统观念影响,大部分老年人只有在生活不能自理或半自理的情况下,才会选择入住养老机构(牟羡,2010),再加上养老机构建设缺少标准规范,因而我国的机构养老模式长期存在供需不平衡、资源利用率不高、服务形式单一、管理低效且不规范等问题(周学馨、刘美华,2020)。

为有效破解机构养老的管理和服务困局,2021 年 10 月,工信部联合民政部、卫生健康委发布《智慧健康养老产业发展行动计划(2021—2025 年)》,明确指出要重点面向智慧养老院建设,打造智慧化解决方案。2022 年 3 月,浙江省率先出台《浙江省智慧养老院建设方案(V1.0)》,提出了全国首个省域层面的智慧养老院建设框架,集智慧服务、智慧照护、智慧交流、智慧管理、智慧安防于一体。其中,智慧照护系统是智慧养老院的核心业务模块,以老年人照护服务为中心,通过精准能力评估、精准照护计划、精准床位使用、精准照护实施、精准服务响应、精准紧急救助六个子场景,规范照护服务,提高服务质量,提升专业水平。

一、研究背景及问题

据国家统计局数据,截至 2023 年底,我国 60 周岁及以上的人口共 2.97 亿人,占总人口的 21.1%。按照每万名老年人口拥有 20 名持证养老护理员的比例计算,我国目前存在超百万的养老护理员缺口,这也是世界各国都面临的重大挑战。人才队伍的培养是短期内无法马上解决的痛点,但是飞速

发展的人工智能技术却为日益严峻的养老问题提供了新的解决思路,被认为是降本增效的关键武器(Unbehaun, Aal & Vaziri et al., 2020),大量面向养老照护的智能化产品和服务被设计、应用到实际场景中,以期为养老服务行业带来根本性的变革。

(一) 人工智能技术对养老服务的影响

理论上,这些人工智能技术的应用可以有效提升老年人的生活质量,降低老年人发生意外的风险,同时也能降低照护者的劳动强度,提高养老服务的精细化水平与质量(何妮、霍聪聪、徐功铖等,2021)。例如,远程家庭诊疗设备、人工照护产品可以让老年人足不出户就能享受到专业的医疗照护服务,有利于提高老年人对养老服务的满意度(蒲新微,2018);生命体征监测器可以实时监测老年人的呼吸率、心率等,让养老护理员随时随地掌握老年人的身体状态,及时发现异常情况,减少夜间巡床次数;智能呼叫器可以突破传统服务的时空限制,快速直接地将老年人的服务需求和紧急救助需求精准地传递给服务方,提升服务质量(朱海龙,2016)。关于人工智能技术对传统养老服务领域发展的促进,研究者认为主要体现在以下四个方面(睢党臣、曹献雨,2019)。

第一,有助于提升老年人的生活质量。随着年龄的增长,老年人开始出现视力、听力下降,记忆力、注意力衰退,行动能力受限等问题,不仅行动不便,还会因此导致忘关火、易跌倒等意外风险。以物联网为主要技术手段的智能家居产品,如烟雾报警器、电动窗帘等,通过加强安防监测、提升设备可控性,可以有效降低老年人的行动风险。智能卫生护理机器人通过感应功能,可以帮助失能半失能老年人进行日常大小便清理和局部按摩康复训练,既方便又专业,进一步提升老年人的生活质量。

第二,有助于满足老年人的精神需求。根据马斯洛需求层次理论,个体在满足了生理和安全需求以后,还有社交、尊重、自我实现等需求。尤其是

老年人,更容易随着工作退休、子女离家、社会关系网逐渐缩小而产生孤独感(于晓琳、陈有国、曲孝原等,2016)。研究发现,孤独感等负面情绪与老年人的幸福感呈显著负相关,对其生活质量有反向预测作用(王希华、周华发,2010),孤独感越高,老年人的幸福感越低。以人工智能技术为主的情感陪护类产品,可以通过语音交流、表情交互等功能,在一定程度上缓解老年人空虚寂寞的生活状况,更能提供各种互联网多媒体资源,满足老年人的兴趣爱好,提升老年人的精神生活水平。

第三,有助于提升养老服务的精细化水平。传统养老服务模式只能通过提供公共服务来解决共性问题,无法满足老年人的个性化需求。比如,一些患有高血压、糖尿病等慢性病的老年人,需要长期控制饮食,就只能对着社区老年食堂的"大锅饭"感慨。多部门数据协同则可以实时掌握老年人的身体状态,为社区老年食堂采购食材、制定菜谱等提供辅助决策功能,满足老年人的实际需求;也可以通过互联网实现个性点单、送饭上门,提供更加精准的养老服务。

第四,有助于缓解行业人才短缺的现状。养老护理是人口老龄化程度进一步加深背景下的一个重要职业。然而,由于要长期护理失能半失能老年人,工作琐碎而繁重,工作时间长、报酬低,无法吸引年轻人的目光,养老护理员队伍普遍存在从业人员年龄偏大、文化程度低、工作压力大、流动性高等问题(谭玉婷、徐依、谭雅琼等,2017),服务能力亟待提升。随着智能产品的不断迭代升级,诸如洗澡、翻身、康复等工作都将可以由机器人来完成,既能降低养老护理员的工作强度,又能提升其专业水平,有助于加强养老服务队伍的专业性和稳定性。

人工智能技术对养老服务行业的推动和促进作用毋庸置疑,但其优势是否能得到有效发挥,推进是否顺利,仍有待验证。有研究表明,当前智能产品的科技化和专业化水平还有待提升,不能完全满足养老服务领域持续增长的各类需求(张博、韩俊江,2019),且无法完全替代养老护理员开展工

作。例如,语音识别技术还不够完善,当老年人的普通话不标准或者讲方言的时候,情感陪护机器人就无法发挥理想作用;全自动洗澡机等设备虽然能一条龙式完成所有助浴工作,但也须养老护理员在旁操作、监督;剪指甲、喂食等对精细度要求比较高的工作,目前只能由养老护理员完成。另外,生命体征监测设备也普遍存在数据不准确的问题,容易误报漏报,反而增加了照护工作的负担,甚至还会出现意外事故责任划分不清等问题。

(二)人工智能技术对养老护理员的影响

有关现有人工智能技术如何影响养老服务的研究,多集中于某单一类型的人工智能产品或由产品延伸出来的服务(何妮、霍聪聪、徐功铖等,2021)。并且,这些产品和服务都是面向老年人或管理者设计的,而忽视了养老服务行业中非常关键的一个角色,即养老护理员。作为养老服务的直接提供者,养老护理员的角色、责任和行为,都随着智慧养老院建设的不断深化而发生了相应的改变(徐青雨,2022)。

第一,从体力劳动向脑力劳动转变。为了解决养老护理员工作压力大、生理负荷重(谭玉婷、徐依、谭雅琼等,2017)、缺少情感和社会支持(Hashizume, Suzuki & Takazawa et al., 2021)等问题,研究者们充分发挥技术优势,研发能辅助养老护理员工作的产品,在产品中增加沟通平台,支持养老护理员表达需求(马婷、张银华、陈燕,2016);开发面对养老护理员的自主学习和培训平台,加强技能培训;利用技术提升养老护理员的自我管理等能力(Hensely-Schinkinger, Carvalho & Glanznig et al., 2015),希望能为养老护理员的工作减负增效。然而,这些新技术的应用却意味着,养老护理员不仅要承担照护老年人的主要责任,同时还要承接让技术发挥作用的任务,不断学习各种新兴技术,通过技术来提高养老服务、提升专业水平,由传统的体力技术人员演变成现代意义上的数字劳工(tech labor)(Graham, Hjorth & Lehdonvirta, 2017),将信息技术与数字技术作为生产资料

和生产力。

第二,对信息素养的要求不断提高。人工智能技术日新月异,各种新技术如雨后春笋般层出不穷。然而,养老护理员却是一批年龄较大、文化程度较低,即将步入老年的“准老年人”,其认知能力、信息加工水平、信息素养都无法与年轻人相提并论,存在明显的数字鸿沟。他们需要长时间学习操作智能产品及其服务,才能达到理想的应用状态。再加上新技术普遍存在限制和障碍,短期内不仅无法提升养老护理员的工作效率,反而会给他们带来额外的学习和使用负担(Marasinghe, 2016),让本就烦琐、沉重的照护工作更加难以完成,不仅影响养老护理员的工作效率,也影响了养老服务的可持续发展。

第三,对新技术的认知态度影响其整体状态。有研究表明,人工智能技术环境下的养老护理员并没有表现出更加积极的工作状态,反而在承担了更大压力的同时,面临着更低的生活质量和幸福感,并且存在受技术影响越深健康状况越差的趋势。其原因在于,人工智能技术的应用会对养老机构的组织和管理流程进行重塑,进一步影响养老护理员的公平感、身份认同感和主观能动性(徐青雨,2022)。当养老护理员认为其工作被人工智能技术所取代,或者增加了更多的工作负担时,其整体健康状况就会更差;反之,当养老护理员对人工智能技术表现出更多的信任感和积极态度时,他们主观上就会认为自己更健康。

(三)研究问题

综上所述,人工智能技术对养老护理员的影响受养老护理员所应用的技术类型影响较大。养老护理员对技术的接受程度和信任度,直接影响了技术应用的效果,以至于大部分研究都要求在应用人工智能技术时,更多地考虑养老护理员的需求和态度,以最大限度地发挥技术应有的作用。与此同时,养老护理员自身的素质,如技能水平、信息化水平等,也会影响其对人

工智能技术的信任和使用。

基于此,本研究在浙江统一建设省智慧养老院的场景下,以智慧照护系统为小切口,聚焦于影响养老护理员工作满意度的因素,进一步讨论智慧养老院建设的理念和实践,对当前的养老护理员培训、智慧照护系统建设等问题提出意见和建议,寻找提升养老护理员工作满意度的新路径。

二、研究方法与数据

本研究以浙江省建设智慧养老院为实证场景,采用社会实验设计思路,将实验情境引入现实情境,定量与定性相结合,综合评估、分析养老护理员在使用智慧照护系统过程中的工作满意度,以获取经验知识,发现因果关系,进而指导和解决现实问题,达到实践应用和理论建构的双重目的(苏竣、魏钰明、黄萃,2020)。

(一) 数据选取

本研究根据养老服务行业专家意见,在浙江省杭州市选取了两家机构性质、等级定位、规模体量等基础属性几乎一致,但在信息化、智能化应用水平方面有着天然差异的养老机构。其中,A 机构为实验组,其智慧养老院建设水平属全省前列,是一家有活力的、创新型的养老机构;B 机构为对照组,其对人工智能技术的使用局限在智能感知设备方面,且使用率较低,是一家以传统经营模式为主的养老机构,暂未启动智慧养老院系统建设(表 1)。由于 A 机构位于风景秀丽的郊区,B 机构位于热闹繁华的市区,因此两家机构的占地面积虽相差较大,但老年人的入住情况、机构的服务水平等属性相差并不大,可作为实验对照组。

表 1　　A、B 两个机构基本属性对比

机构	机构性质	占地面积（亩）	床位数（张）	老年人数（人）	智慧化水平
A 机构（实验组）	公办	169	2000	1244	浙江省智慧养老院首批试点单位
B 机构（对照组）	公办	60	1450	1197	无智慧养老院系统

在 A、B 两个机构中随机抽取一定比例的养老护理员，通过问卷调研、个别访谈等方式，对其技能水平、信息化水平、工作情况、工作满意度等内容进行度量和分析。

（二）数据采集

本研究采用随机抽样法，在 A、B 两个机构随机发放问卷 135 份，回收问卷 135 份，问卷回收率为 100%。回收问卷中，A 机构的问卷有 76 份，B 机构的问卷有 59 份。为了保证问卷的回收率，本研究采用匿名形式在“问卷星”上发放，以选择题为主要填写形式，降低填写难度，且提交时会提示漏填部分，因而问卷回收率和有效率均达到 100%。

调研问卷分为三部分：第一部分为人口统计学特征，包括性别、年龄、文化程度、从业年限、月收入、所拥有的相关技能证书或资格证书、日常使用的智能设备种类等；第二部分为护理工作内容，包括所照护的老年人护理等级、年龄、数量，平均每天照护时长、夜间巡视次数、翻身频率，以及对目前工作的满意度等情况；第三部分为对智慧照护系统的使用情况，包括是否使用智慧照护系统、床位码、线上点单服务、紧急呼叫设备、数字化管理等内容。其中，第三部分仅由 A 机构的养老护理员填写，B 机构无须填写。

（三）数据处理

本研究的自变量为是否使用智能设备、是否使用智慧照护系统和对智

能设备、对智慧照护系统的满意度,因变量为养老护理员的工作满意度,控制变量则是人口统计学特征和养老护理员的工作情况等。

对于养老护理员的各项满意度情况,本研究采用李克特五级评分法进行衡量,分别为“很不满意”“不满意”“一般”“满意”“很满意”;对于控制变量技能水平和信息化水平,本研究参考前人方法进行处理(何妮、霍聪聪、徐功铖等,2021)。技能水平指养老护理员是否拥有相关的技能证书和资格证书,包括职业技能等级认定证书、健康证、社会工作者资格证及其他证书,有则定义为 1,没有则定义为 0,将养老护理员在各类证书选项下的得分相加,以衡量其技能水平。信息化水平指养老护理员对智能设备的使用情况,包括智能手机、智能手环、电脑、智能家电及其他,有则定义为 1,没有则定义为 0,将养老护理员在各类智能设备选项下的得分相加,以衡量其信息化水平(表 2)。

表 2　　变量的描述性统计结果

变量	变量名称	变量赋值	样本量	均值	标准差	最小值	最大值
控制变量	性别	男=1;女=2	135	1.92	0.275	1	2
	年龄		135	50.14	5.753	32	61
	户籍	本地户籍=1;非本地户籍=2	135	1.75	0.436	1	2
	文化程度	不识字=1;小学=2;初中=3;中专/高中=4;大专=5;本科及以上=6	135	2.90	0.558	2	5
	技能水平	拥有的证书数量	135	1.42	0.496	1	2
	工作年限		135	6.31	5.217	0.125	25
	月收入	1000 元以下=1;1000—1500 元=2;1500—2000 元=3;2000—2500 元=4;2500—3000 元=5;3000—3500 元=6;3500—4000 元=7;4000—5000 元=8;5000—6000 元=9;6000 元以上=10	135	7.81	1.121	2	10

（续表）

变量	变量名称	变量赋值	样本量	均值	标准差	最小值	最大值
控制变量	信息化水平	日常生活中使用的智能设备数量	135	1.61	0.828	1	4
	夜间巡视次数	0次=1;1次=2;2次=3;3次=4;4次=5;5次=6;6次=7;6次以上=8	135	5.03	2.334	1	7
	护理老年人总数	0人=1;1人=2;2人=3;3人=4;4人=5;5人及以上=6	135	3.84	2.080	1	6
	护理工作时长	6小时以内=1;6—8小时=2;8—10小时=3;10—12小时=4;12小时以上=5	135	3.91	1.504	1	5
	为老年人翻身频率	无卧床病人=1;1小时1次=2;2小时1次=3;3小时1次=4;4小时1次=5;4小时以上1次=6	135	2.27	0.950	1	4
因变量	工作时间满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	135	4.12	0.763	1	5
	排班灵活满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	135	4.04	0.867	1	5
	工作满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	135	4.20	0.644	3	5
自变量	是否使用智慧照护系统	是=1;否=0	76	0.56	0.498	0	1
	智慧照护系统满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	76	4.38	0.588	2	5
	床位码满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	76	1.93	2.201	0	5
	线上点单服务满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	76	2.44	2.228	0	5
	紧急呼叫服务满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	76	2.40	2.200	0	5
	数字化管理满意度	很不满意=1;不满意=2;一般=3;满意=4;很满意=5	76	2.48	2.232	0	5

三、研究结果与分析

本研究使用 SPSS 28.0 软件,按照 A、B 两个机构的实验分组,对养老护理员的基础信息、工作情况和工作满意度等内容进行统计分析,并进一步探讨各因素之间的关系。

(一) 组别间异质性分析

A 机构为实验组,其养老护理员均使用智慧照护系统开展工作;B 机构为对照组,其养老护理员未使用智慧照护系统开展工作。本研究采用独立样本 *t* 检验方法,对两个组别的养老护理员基础信息、工作情况和工作满意度进行比较,结果表明:

第一,在基础信息方面,A、B 两个机构存在显著差异。相对于实验组 A 机构而言,对照组 B 机构的女性占比更多、年龄更大、文化程度偏低、工作年限更长、月收入更高,且差异显著(表 3),符合一家传统大型养老机构的基本特征。

表 3　　A、B 两个机构基础信息比较分析结果

变量	A 机构	B 机构	*t* 值
性别	1.87±0.34	1.98±0.13	2.450**
年龄	48.00±5.87	52.90±4.27	5.398***
文化程度	3.08±0.51	2.68±0.54	−4.417***
工作年限	4.39±2.68	8.78±6.52	5.322***
月收入	7.63±1.14	8.05±1.06	2.186**
技能水平	1.39±0.49	1.46±0.50	0.730
信息化水平	1.80±0.88	1.37±0.69	−3.083***

注:变量数据为平均值±标准差;***p<1%,**p<5%,*p<10%。下同。

第二,在技能水平方面,A、B 两个机构不存在显著差异,可以认为两个机构的养老护理员专业技能差异不大,不会影响其工作表现和工作满意度。

第三,在信息化水平方面,A 机构的养老护理员要显著高于 B 机构,这可能是由于 A 机构的养老护理员更年轻,文化程度更高,更有开放性和包容性,乐于接受新事物;也有可能是因为 A 机构的养老护理员长期使用智慧照护系统,其信息化水平得到了提升,需要进一步分析原因。

第四,在工作情况方面,A、B 两个机构存在较大显著差异。在护理老年人总数、护理工作时长两个方面,B 机构要显著高于 A 机构。A 机构每位养老护理员平均需要护理老年人 3.17 名,护理工作时长为 10 个小时左右;B 机构每位养老护理员平均需要护理老年人 4.72 名,护理工作时长为 12 个小时左右。在夜间巡视次数方面,A、B 两个机构不存在显著差异,都要巡视 4 次左右。而在为老年人翻身频率方面,A 机构要显著高于 B 机构(表 4),几乎每个小时都要为老年人翻身一次,而 B 机构是每 2 个小时翻身一次。

表 4　　A、B 两个机构工作情况比较分析结果

变量	A 机构	B 机构	*t* 值
护理老年人总数	3.17±2.15	4.72±1.63	4.576***
护理工作时长	3.11±1.57	4.93±0.37	8.663***
夜间巡视次数	4.92±2.76	5.17±1.65	0.612
为老年人翻身频率	1.99±0.97	2.64±0.78	4.233***

护理老年人总数、护理工作时长以及为老年人翻身频率都是养老护理员工作负担的指标之一。虽然 A 机构养老护理员护理的老年人数和工作时长都较 B 机构要少,但为老年人翻身的频率却更高,可进一步探索使用智慧照护系统是否会对工作内容产生影响。

第五,在工作满意度方面,A、B 两个机构存在显著差异。不管是工作时间、排班灵活性,还是工作整体满意度,A 机构得分都要显著高于 B 机构

(表 5),表明 A 机构养老护理员更满意自己的工作状态,即使他们的月收入要显著低于 B 机构。

表 5 A、B 两个机构工作满意度比较分析结果

变量	A 机构	B 机构	*t* 值
工作时间满意度	4. 42±0. 62	3. 73±0. 76	−5. 843***
排班灵活满意度	4. 45±0. 62	3. 51±0. 86	−7. 378***
工作总体满意度	4. 47±0. 55	3. 85±0. 58	−6. 379***

(二) 工作满意度与基础信息相关性分析

为进一步观察工作总体满意度与养老护理员基础信息各变量之间的关系,本研究采用斯皮尔曼相关系数对各变量进行分析,结果表明(表 6):

第一,养老护理员的年龄与月收入、工作总体满意度相关显著。养老护理员的年龄与月收入之间呈正相关,表明年龄越大,月收入越高;年龄与工作总体满意度之间呈负相关,表明年龄越大,其工作总体满意度越低。

第二,养老护理员的月收入与文化程度、技能水平、信息化水平、工作总体满意度等各变量之间的相关性均不显著。

第三,养老护理员的文化程度与技能水平之间呈显著负相关,表明文化程度越高,技能水平越低。这可能是因为养老护理员对护理工作的技能水平有一个总体平衡的认知,当自身的文化水平较低时,就需要各类资格证书来证明自己的能力;当文化程度较高时,则认为不需要资格证书来证明自己的能力。

第四,养老护理员的技能水平与信息化水平、工作总体满意度等变量之间的相关性均不显著。

第五,养老护理员的信息化水平与工作总体满意度之间存在显著正相关,表明信息化水平越高,养老护理员的工作总体满意度就越高。

综上所述,养老护理员的年龄会影响其月收入,文化程度会影响其技能

水平,但月收入和技能水平均不影响工作总体满意度。养老护理员的工作满意度与其年龄和信息化水平有关,年龄越大则满意度越低,信息化水平越高则满意度越高。

表 6　　工作满意度与基础信息各变量之间相关系数矩阵

变量	年龄	月收入	文化程度	技能水平	信息化水平	工作总体满意度
年龄	1.00	0.16*	−0.12	0.05	−0.10	−0.32***
月收入		1.00	−0.10	−0.11	−0.10	0.03
文化程度			1.00	−0.23***	0.12	0.10
技能水平				1.00	0.11	−0.01
信息化水平					1.00	0.29***
工作总体满意度						1.00

值得注意的是,根据数据分析结果,养老护理员的技能水平与工作满意度之间并不存在相关性。将两个组别的养老护理员按照技能水平进行分组,做进一步独立样本 t 检验,结果并未发现二者的工作满意度存在显著差异($t=0.108$,$p>0.05$)。

(三) 工作满意度与工作情况相关性分析

为进一步观察工作满意度与养老护理员工作情况各变量之间的关系,本研究采用斯皮尔曼相关系数对各变量进行分析,结果表明(表 7):

第一,工作满意度与护理老年人总数、护理工作时长之间呈显著负相关,即护理老年人总数越多、护理工作时长越长,养老护理员的工作满意度越低。

第二,护理老年人总数、护理工作时长、夜间巡视次数各变量之间均呈显著两两正相关。

第三,由于为变量“为老年人翻身频率”赋值时,数值越小表明为老年

人翻身频率越高,因此工作满意度与为老年人翻身频率之间呈显著正相关,即翻身频率越高,工作满意度越高;护理老年人总数、护理工作时长、夜间巡视次数与为老年人翻身频率之间均呈显著负相关,即翻身频率越高,护理老年人总数、护理工作时长、夜间巡视次数均越少。这可能是由于不同的养老机构分组造成的。笔者在进行异质性分析时发现,实验组 A 机构养老护理员护理老年人总数、护理工作时长均比对照组 B 机构要少,但是为老年人翻身频率比对照组 B 机构要高,需要进一步控制变量,以讨论工作情况与工作满意度之间的相关性。

表 7　　工作满意度与工作情况各变量之间相关系数

变量	工作满意度	护理老年人总数	护理工作时长	夜间巡视次数	为老年人翻身频率
工作满意度	1.00	−0.31***	−0.51***	−0.01	−0.24***
护理老年人总数		1.00	0.55***	0.45***	0.51***
护理工作时长			1.00	0.30***	0.50***
夜间巡视次数				1.00	0.34***
为老年人翻身频率					1.00

（四）线性回归模型分析

为进一步探索是否使用智慧照护系统、智慧照护系统满意度是否会影响养老护理员的工作满意度,本研究构建了两类线性回归模型:一类是采用全样本数据,探索是否使用智慧照护系统对护理员工作满意度的影响(表 8 模型 1—3);另一类是仅采用实验组 A 机构的数据,在智慧养老院场景下探索智慧照护系统满意度如何影响养老护理员的工作满意度(表 8 模型 4—6)。采用线性回归模型对各变量进行实证研究,具体形式为:

$$y_i=\alpha+\beta_1 * \textit{是否使用智慧照护系统}_i+\beta_i * X_i+\varepsilon_i$$

$$y_i=\alpha+\beta_1 * \textit{智慧照护系统满意度}_i+\beta_i * X_i+\varepsilon_i$$

其中，y_i 为因变量，表示第 i 个养老护理员的工作满意度；X_i 是所有影响工作满意度的其他控制变量，包括个人基础特征和工作情况；ε_i 为随机扰动项。自变量 1 为是否使用智慧照护系统，自变量 2 为智慧照护系统满意度。

为提高模型回归结果的稳健性，本研究逐步调整了控制变量：模型 1、模型 4 分别为两类模型的基准回归结果，只包含自变量和因变量；模型 2、模型 5 分别在基准回归结果的基础上，引入与因变量相关性较高的控制变量，包括信息化水平、年龄、工作年限、文化程度、月收入、护理老年人总数和护理工作时长；模型 3、模型 6 进一步引入性别、技能水平、夜间巡视次数、为老年人翻身频率等相关性不显著的变量。

表 8　　**线性回归分析结果**

变量	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
是否使用智慧照护系统	0.626*** (6.34)	0.291** (1.99)	0.249 (1.63)			
智慧照护系统满意度				0.781*** (12.58)	0.731*** (11.49)	0.742*** (10.06)
性别			0.245 (1.49)			−0.059 (−0.49)
年龄		−0.029*** (−3.21)	−0.028*** (−2.97)		−0.011 (−1.52)	−0.013 (−1.63)
工作年限		0.010 (0.96)	0.012 (1.13)		−0.001 (−0.07)	0.006 (0.30)
文化程度		−0.140* (−1.90)	−0.153* (−1.89)		0.025 (0.42)	0.039 (0.67)
月收入		0.149** (2.12)	0.146** (2.01)		0.060 (0.93)	0.061 (0.87)
技能水平			−0.044 (−0.43)			−0.014 (−0.19)
信息化水平		0.146*** (2.74)	0.156*** (2.77)		0.051 (0.94)	0.035 (0.67)

（续表）

变量	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
护理老年人总数		−0.068** (−2.51)	−0.077*** (−2.84)		0.020 (0.81)	0.017 (0.78)
护理工作时长		−0.104* (−1.82)	−0.151** (−2.49)		−0.101** (−2.11)	−0.108** (−2.19)
夜间巡视次数			0.030 (1.13)			−0.009 (−0.36)
为老年人翻身频率			0.016 (0.24)			0.063 (1.20)
_cons	3.847*** (50.86)	5.103*** (7.53)	4.714*** (5.63)	1.050*** (3.49)	1.453** (2.55)	1.522** (2.03)
N	135	135	135	76	76	76
R^2	0.234	0.387	0.404	0.690	0.740	0.747

表 8 的线性回归模型结果显示，使用智慧照护系统对养老护理员的工作满意度有显著正向影响，即使在增加控制变量之后，其影响程度虽有所下降但并不显著。在控制变量不变的情况下，使用智慧照护系统可以让养老护理员的工作满意度提升 24.9%。

同时，年龄、护理老年人总数、文化程度、护理工作时长会对工作满意度产生显著负面影响。值得一提的是，在控制变量不变的情况下，为老年人翻身频率并不会对工作满意度产生影响，进一步验证了工作情况的差异部分是由于两个机构的分组造成的。

智慧照护系统满意度对养老护理员的工作满意度有显著正向影响，且在引入控制变量后没有出现下降情况。在控制其他变量不变的情况下，对智慧照护系统满意度的评价提升一个等级，能够使养老护理员对其工作满意度提升 74.2%。

本研究采用相关分析，进一步探索在智慧照护系统使用场景下，对床位码、线上点单服务、紧急呼叫服务、数字化管理等场景的使用，是否会影响养老护理员对智慧照护系统的满意度评价。结果（表 9）表明，智能设备满意

度与床位码满意度、线上点单服务满意度、紧急呼叫服务满意度、数字化管理满意度均呈正相关,且非常显著。

表 9　　满意度相关系数矩阵

	智慧照护系统满意度	床位码满意度	线上点单服务满意度	紧急呼叫服务满意度	数字化管理满意度
智慧照护系统满意度	1.00	0.62***	0.73***	0.72***	0.76***
床位码满意度		1.00	0.71***	0.62***	0.66***
线上点单服务满意度			1.00	0.86***	0.83***
紧急呼叫服务满意度				1.00	0.85***
数字化管理满意度					1.00

四、研究结论与展望

本研究发现,在当前浙江省建设智慧养老院的大场景下,使用智慧照护系统确实有助于提升养老护理员的工作满意度,但前提是智慧照护系统能够得到养老护理员的认可和信赖,只有养老护理员满意了、愿意用了,其才能发挥出人工智能技术应有的作用,有效提升服务效率,提高服务质量。

(一) 产品设计考虑供需双方

智慧照护系统的目标是以老年人照护服务为中心,规范照护服务,提升专业水平,提高服务质量,其主要的使用对象为照顾失能(失智)、须长期卧床老年人的养老护理员。然而,调研和访谈结果显示,由于智慧照护系统的建设主体是民政部门,因此系统建设时仅考虑民政部门的监管职责和机构的管理功能,以数字化台账管理、全过程监督规范为主要目标,与养老护理员的工作内容耦合度并不高。这也是当前养老护理研究存在的主要问题,虽然一直强调以供需矛盾为切入点,但始终都未从养老护理员的角度提出

切实有效的供给改善策略(丁雪萌、孙健,2019)。

调研结果显示,本研究中的女性养老护理员有 124 人(91.85%),总体年龄在 32—61 岁,平均年龄为 50.14±5.75 岁,其中 50 岁及以上的养老护理员占比达 57.04%;文化程度较低,以初中及以下学历为主(90.4%)。由于起步晚,职业认同、社会认同、自我认同较低(袁群、易霞、张银华等,2015),我国养老护理员的待遇普遍较低,主要存在两方面的问题:一是整体工资水平不高,本研究中两个机构的养老护理员工资均在 2500—6000 元,其中 4000—5000 元最多(45.19%);二是养老护理员的工资水平与其工作时长、工作强度不成正比,76.3%的养老护理员每天要工作 10 小时以上,负责多位老年人的生活起居,更不用说每隔 1—2 小时就要给老年人翻身,以防止其长褥疮。另外,由于缺少明显的上升通道和职业规划体系,在岗的养老护理员没有主动学习提升的意识和动力,也无法吸引更高层次、更年轻的专业技术人才进入养老服务行业,导致养老服务人才队伍建设停滞不动、毫无活力。

虽然实验组 A 机构的智慧照护系统建设情况较好,但也存在个别没有实现业务闭环的功能,比如:养老护理员为老年人进行线上个性化点餐后,还须手抄一份用于分餐;部分紧急救助设备存在误报情况,需要养老护理员仔细辨认处理,甚至多跑几趟。事实上,对于养老护理员而言,收入并不是最重要的。实验组 A 机构的养老护理员工资比对照组 B 机构的要少,但并不妨碍其工作满意度更高,反而是智慧照护系统满意度、工作内容等对工作满意度的影响更大。在此情况下,若是继续忽视养老护理员的实际需求,不考虑实际业务闭环,没有实实在在地帮助养老护理员减负增效,将会导致本就流动率高的养老护理员队伍更加动荡,人才缺口继续扩大。

因此,在设计智慧养老服务或产品时,要充分考虑人与机器、技术之间的交互关系,使其既简单易用,让养老护理员快速上手,又能提供准确可靠的数据,切实减轻养老护理员的工作负担。同时,也可以提供服务支持类应

用项目(王杰、董少龙,2021),让养老护理员真实受用。

(二)全面加强养老护理员的信息化水平

本研究发现,养老护理员的技能水平并不会影响其工作满意度,文化水平低的养老护理员会通过考取资格证书来提升自己的竞争力,总体趋于平衡状态。经访谈确认,除非是康复训练等需要专业资质的服务内容,大部分养老护理员承担的都是照顾老年人起居的生活类服务,技能水平相当,并不存在特别好或特别差的情况。并且,大部分失能(失智)老年人意识不清或有认知衰退问题,对养老护理员的服务态度感知较少,不影响对养老护理员的考核和评价。

调研结果显示,实验组 A 机构有 25%的养老护理员会使用智能手环、电脑,而对照组 B 机构仅有 8%的养老护理员会使用智能手环、电脑,A 机构养老护理员的信息化水平要显著高于 B 机构,这与其工作满意度密切相关。虽然暂无证据表明是信息化水平提升了养老护理员使用智慧照护系统的能力,还是智慧照护系统使用提升了养老护理员的信息化水平,但二者之间确实存在一定的相关性。信息化水平并不等同于信息素养。信息素养是指信息的获取与检索、理解与评估、创建与分享等能力(曲绍萍,2022),而本研究中的"信息化水平"指的是个体运用信息化设备进行工作和生活的能力,更多地强调对信息化设备的熟练使用和掌握。

随着技术的飞速发展,各类设备功能层出不穷且日益多元化,随之而来的是不断增加的操作难度,大大提升了使用门槛,这就对使用设备的人提出了更高要求。信息化水平高的人,可以从一类设备的操作迁移到另一类设备,使用上更加流畅便捷;而信息化水平低的人,日常使用智能手机也仅限于打电话、发信息、看视频,在面对需要安装 APP 或者用小程序进行工作时,就会因为操作困难而出现畏难情绪,甚至会下意识地抵制和反对,不愿花时间学习。另外,也有部分养老护理员使用的手机版本过低、配置不够,

导致无法使用相关养老系统或功能。

因此,相关部门或机构可以根据年龄、基础水平等基础信息,对养老护理员进行分层分组,从手机操作、手机摄影、平板电脑使用等一些常规、趣味性较强的培训入手,让年龄大的养老护理员能够消解抵触情绪,更好地拥抱信息化社会。同时,在进行技术开发时也要关注系统配置问题。养老护理员的信息化水平提高了,对人工智能技术更加开放和信任了,将更加有利于养老产品的使用、推广和优化迭代。

(三) 不足与展望

考虑到社会实验设计的需要,本研究仅在杭州市内选取了两个工作环境、薪酬待遇、规模体量、人员资质相差不大的养老机构作为研究对象,并随机抽样了部分养老护理员进行调研,没有做到全样本考察,样本量略显不足,需要进一步完善。

由于样本量较少,本研究在变量控制方面稍显薄弱,两组养老护理员的年龄、文化程度等基本信息存在显著差异。尽管线性回归模型 2 结果显示,即使对变量进行控制,使用智慧照护系统仍然能显著提升养老护理员的工作满意度,但是当引入相关不显著的控制变量时,这种影响效果会变差,以至于不显著。另外,信息化水平与使用智慧照护系统之间的关系,以及工作时长、工作强度等对工作满意度的影响,都还有待进一步探索和验证,后续可以在扩大样本量的基础上进行组内分析比较,以确定不同因素对工作满意度的影响。

对于智慧养老,“智慧”只是手段和支撑,“养老”才是根本和核心,养老服务的供给双方都应该是参与的主体。尤其是在养老机构的信息化建设过程中,更要关注养老护理员的真实需求,通过设计和开发便捷易用的系统和产品,有针对性地提升养老护理员的信息化水平,让人和技术更好地兼容耦合,有效推动养老服务行业发展提质增速。

参考文献

丁雪萌、孙健,2019,《近二十年我国养老护理人员研究的现状与趋势》,《江汉学术》第6期。

何妮、霍聪聪、徐功铖等,2021,《人工智能应用对养老服务质量的影响——基于杭州、合肥、武汉三地调查的实证分析》,《社会保障研究》第5期。

马婷、张银华、陈燕,2016,《养老护理员离职意愿与社会支持的相关性研究》,《护理研究》第7期。

牟羡,2010,《浅析机构养老在养老保障体系中的作用》,《社会工作》(下半月)第6期。

蒲新微,2018,《失能老人智能化养老何以可能:一个系统性施行框架的阐释》,《兰州学刊》第10期。

曲绍萍,2022,《农村老年人信息素养研究》,济南:山东大学硕士学位论文。

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。

睢党臣、曹献雨,2019,《人工智能养老的内涵、现状与实现路径》,《新疆师范大学学报》(哲学社会科学版)第2期。

谭玉婷、徐依、谭雅琼等,2017,《养老护理员工作压力及其影响因素分析》,《中国护理管理》第7期。

王杰、董少龙,2021,《智慧养老技术及落地应用指南》,北京:电子工业出版社。

王希华、周华发,2010,《老年人生活质量、孤独感与主观幸福感现状及相互关系》,《中国老年学杂志》第5期。

徐青雨,2022,《数字养老环境下护理员的协同照护行为分析与建模》,上海:华东师范大学硕士学位论文。

于晓琳、陈有国、曲孝原等,2016,《影响老年人主观幸福感的相关因素》,《中国心理卫生杂志》第6期。

袁群、易霞、张银华等,2015,《养老护理员工作压力研究现状与进展》,《中国护理管理》第1期。

张博、韩俊江,2019,《积极老龄化政策下智能养老产业发展研究》,《广西社会科学》第7期。

周学馨、刘美华,2020,《我国失独家庭养老体系中机构养老兜底保障作用研究——基于对

全国 709 个失独者调研数据的分析》,《重庆社会科学》第 1 期。

朱海龙,2016,《智慧养老:中国老年照护模式的革新与思考》,《湖南师范大学社会科学学报》第 3 期。

Graham, M., I. Hjorth & V. Lehdonvirta 2017, "Digital Labour and Development: Impacts of Global Digital Labour Platforms and the Gig Economy on Worker Livelihoods." *Transfer: European Review of Labour and Research* 23(2).

Hashizume, S., I. Suzuki & K. Takazawa et al. 2021, "Discussion of Intelligent Electric Wheelchairs for Caregivers and Care Recipients." International Conference on Human-Computer Interaction.

Hensely-Schinkinger, S., A. Carvalho & M. Glanznig et al., 2015, "The Definition and Use of Personas in the Design of Technologies for Informal Caregivers." International Conference on Human-Computer Interaction.

Marasinghe, K. 2016. "Assistive Technologies in Reducing Caregiver Burden among Informal Caregivers of Older Adults: A Systematic Review." *Disability and Rehabilitation: Assistive Technology* 11(5).

Unbehaun, D., K. Aal & D. Vaziri et al. 2020, "Social Technology Appropriation in Dementia: Investigating the Role of Caregivers in Engaging People with Dementia with A Videogame-Based Training System." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

“成为更好的女性”*

——小红书与生活方式平台的文化政治

林　欣　苗伟山**

摘要：在中国最为流行的生活方式平台小红书上，成长类内容的文化表达对于女性用户的身份与主体性建构具有深刻影响。本文以文化政治为核心理论，借助文本分析和线上参与式观察的研究方法，详细考察小红书的发展状况、女性成长内容的文本特征，以及相关评论界面的内容互动，从生产前提、话语特征和生产机制三个层面剖析小红书平台上女性成长内容背后的权力关系。研究发现：首先，小红书的社区自治模式体现了新自由主义和中国礼俗社会的结合；其次，以平台塑造的中产阶级身份想象为基础，女性成长博主生产出一套对抗现代性焦虑的成长叙事；最后，在平台希望实践的调控下，女性成长和分享被紧密地联系在一起。此外，本文还从“平台基础设施—平台和用户共同塑造的身份想象—实现机制”三个方面提出了分析平台文化政治的概念框架。

关键词：生活方式平台　文化政治　社交媒体　女性成长　现代性

一、引言

20 世纪 90 年代，生活方式电视节目在全球范围内兴起，它们关注的是指导人们如何管理他们的日常生活，包括“改造”他们的家庭和衣柜、提供有关健康与人际关系的建议。刘易斯（T. Lewis）认为，生活方式电视节目，

* 本文系中国人民大学“双一流”建设马克思主义新闻观创新研究“新时代中国传媒的国际传播能力研究”（项目批准号：MXG202309）的阶段性研究成果。

** 林欣，中国人民大学新闻学院。苗伟山，中国人民大学新闻学院。

尤其是改造节目,集中体现了这种自愿自我改造的精神,因为它旨在改变个人从家庭到自我的方方面面(Lewis, 2011)。如今,在有着“生活方式平台”之称的中国社交媒体平台小红书(Red)上,这类改造类内容也被大规模地生产出来。自 2020 年起,小红书便已从成立最初的境外海淘社区、中期的女性美妆穿搭社区,转型为泛生活方式平台,平台内容因此愈发纷繁复杂,包括穿搭指南、美妆教程、旅游攻略、健身经验、拍照妙招、学习方法、City Walk 路线推荐等,与日常生活有关的一切经验与感受都被收录其中。小红书上的生活方式内容大致可以分为以下两类:改造自我,聚焦女性的自我提升与成长;改造生活,关注日常生活的美学体验。在女性成长领域,小红书博主往往以第一视角记录或传授女性如何变得更好的经验,对女性情感、家庭、职场、审美、人际交往、教育、消费等问题提供私人化的详尽建议,如“听劝”改造挑战、如何与人沟通、如何教育孩子、人生建议、别提前消费……总的来说,小红书作为女性用户群体主导的社交媒体平台,其女性成长类内容不仅体现了生活方式内容在社交媒体时代的新发展,而且对女性用户的身份与性别建构产生了重要影响,值得进一步关注和探讨。

既有研究探讨了小红书上的女性身体(Liu & Li, 2023)、网红治理(Ju, 2022),以及数字劳动与平台基础设施发展(刘亭亭、许德娅,2023)、性别创业主义的关系(Guo, 2022)等议题,然而这一社交媒体平台在传播研究领域仍未被充分探索,尤其是女性成长议题方面。自五四新文化运动开始,从谋求自由的女性解放,到追逐平等和个性的女性发展,再到如今的女性成长,中国现代女性解放史已百年有余。女性成长同样是媒介建构的关键主题,这一主题不仅贯穿于中国文学发展史(郭力,2020),在过去十年,女性消费力量的崛起促成文化消费对于女性群体的迎合,有关独立女性的“大女主”成长内容作为占据主导地位的媒介女性叙事,成为新近影视剧的重要主题,且充盈着女性虚拟社区的讨论。线上社区既帮助女性用户建构了

多元身份认同,在网络时尚论坛的自由氛围下,消费压力也被掩盖并逐渐内化于个体(同心,2014)。一方面,小红书作为女性用户主导的社交媒体平台,制造了女性共居的社区生态,引导用户生产出大量的女性可以怎样成长的经验;另一方面,起步于“种草”社区的小红书同样是为女性提供消费决策的重要平台,小红书上的女性成长无法不受到消费主义的浸染。在这一情境下,小红书上的女性究竟诉说了怎样的成长?本研究以文化政治(cultural politics)为核心理论,以文本分析为主、线上参与式观察为辅的研究方法,将小红书上以“女性成长”为标签热度的排行前 100 的博文内容,以及相关新闻报道作为分析对象,探讨小红书上的女性成长内容体现出何种权力关系和平台文化政治。

二、文献综述

(一) 文化政治理论的溯源

文化政治理论是根植于马克思主义思想传统的一种文化批判研究范式,承续意大利结构主义马克思主义学者葛兰西(A. Gramsci)的文化霸权理论、英国新左派思想,并吸收了法国后结构主义学者福柯(M. Foucault)的话语权力学说等后现代理论,在 20 世纪下半叶勃兴的文化研究中应运而生。文化政治作为一种社会政治分析框架,承认文化是政治斗争的重要场所,在这一领域,不同的社会群体就身份、归属、包容和排斥等问题进行协商(Korf, 2020)。

文化政治是一种将文化和政治相结合的理论立场与文化批评方法(汪民安,2020:939)。厘清文化政治理论,需要在这一概念范畴中重新理解文化和政治。此处的“文化”并非狭隘的艺术生产(Angus & Jhally, 2022: 2)或高雅文化(Korf, 2020),而是社会群体生产的生活意义和方式,其通过一

系列物质实践与符号表征反映和塑造社会身份与认知。二战后,立足于马克思主义与英国经验主义传统,英国文化评论家威廉姆斯(R. Williams)将文化阐释为一种整体的生活方式,开启了文化唯物主义思想;同时期的英国文化研究学者霍加特(R. Hoggart)、汤普森(E. Thompson)关注工人阶级的日常生活,赋予文化以阶级维度。随后的几十年里,性别、性、种族等其他维度不断被提出以丰富文化的内涵。

文化政治理论的关键还在于对政治的新理解。政治通常被理解为参与政权的形式或政党权力对社会的控制。进入 20 世纪 60 年代后,葛兰西的文化霸权理论、阿尔都塞(L. Althusser)的意识形态国家机器、福柯的权力观纷纷揭示了权力在社会生活中的弥散,革新了政治社会学的传统理论视角,宏大的、整体性的政党政治、阶级政治和革命政治开始让位于市民社会领地的微观政治。作为文化转向的结果,所有的社会生活都被视为潜在的政治,而政治即权力关系的争夺(Nash, 2001)。文化政治中的“政治”指的便是权力关系,“权力是文化政治学的核心……所有的指意实践,即所有承载意义的实践,都涉及权力关系”(Jordan & Weedon, 1995: 11),不仅表现为常规政治(formal politics),还更多地体现为非常规政治(informal politics),即权力通过文化产品体现出来的无意识的、隐秘的调节和控制(Painter, 1995: 27-57)。

文化政治理论在伯明翰大学当代文化研究中心(CCCS)得到了学者的应用,他们关注:战后作为反文化和亚文化的流行文化的风格政治,这些具有象征意义的风格是对霸权的挪用和抵抗,最终成为一种新的霸权惯例(Rycroft, 2009)。正如霸权是一个过程而非既定的系统或结构,文化政治也是一个斗争的领域,围绕核心问题——权力在制度和意识形态两种形式下的行使,以及权力的行使对文化实践的影响方式——文化政治探讨的是文化和政治之间的互动过程。在这个过程中,人们探寻日常生活的文化实践如何受到政治制度和意识形态结构的限制(Angus & Jhally, 2022: 2),并

寻找抵抗(resistance)与僭越(transgression)的能动策略进行反权力实践(鲍尔德温,2004:263—274)。总的来说,文化政治作为一种生活政治,关注的是日常生活领域的意义实践和权力运作,覆盖性别、种族、身体、阶级、生活方式等议题,并探查其中的权力斗争关系(Angus & Jhally, 2022: 2;汪民安,2020:940;李艳丰,2023)。

文化政治的理论视角可以帮助我们剖析文化现象中异化和不平等的权力关系,寻找反权力的抵抗途径,推动文化变革。文化政治学者往往试图从一切文化现象中发掘权力关系(鲍尔德温,2004:221):一方面,权力的弥散使得文化政治理论几乎涵盖所有社会关系,泛政治化正是其招致批评的原因之一;另一方面,文化政治的斗争都是在生活方式领域进行的,这意味着异化和压迫的结构并没有得到阐明(Rycroft, 2009)。一些马克思主义学者提醒到,文化政治分析应该与经济、阶级以及现实社会制度结合起来(汪民安,2020:945)。当代文化的商品化对文化政治产生了进一步的影响。当代文化政治不仅要承认文化产品所有权的权力,还要承认文化表征中的权力,把对新的表征形式的分析转向对接受表征的外部条件的分析(Angus & Jhally, 2022: 2),在文化的生产和生产的文化的辩证关系中剖析新媒体文化的结构性因素(孙皖宁、苗伟山,2016)。那么,聚焦小红书平台上文化形式的权力关系,便需要和经济环境、企业愿景、社会文化制度、意识形态结构联系起来。

借助文化政治这一分析形式,学者探讨了各种文化现象中的权力关系。文化政治既包括物质实践,也包括符号实践(Korf, 2020):一方面,有学者关注实体建筑、场所(鲍尔德温,2004:258)或文学文本中的权力关系,探讨文本的生产、语境和使用中所涉及的权力关系(Apple, 2014),在此基础上,后来的学者探讨了政府主导的电视媒体中的文化政治问题(Abu-Lughod, 2003;吴靖、云国强,2016),以及商业公司的文化生产,如房地产广告的身份政治(吴靖、王颖曜,2007);另一方面,有学者关注社会底层的阶级(Sun,

2009；Sun，2012）、种族（Louisa，2020）、性别与性（Hooks，2014）、越轨（Sun，2022）、不平等（Sun，2013；Sun，2018）等议题中的身份政治。社交媒体平台上的内容生产包含了用户、平台等更加多元的主体互动，这使得文化政治又面临哪些新的变化？这个问题亟待探讨。

（二）文化政治与数字平台

近年来，随着数字平台的发展，文化内容的生产、发行和流通逐渐由平台所中介，平台的经济、政治和基础设施向网络和应用生态系统的渗透，从根本上影响文化产业的运作。尼伯格（D. Nieborg）等人总结到，平台的政治经济学研究者特别关注平台权力和政治，以及文化商品转型所产生的劳工问题，但忽视了这如何促进文化商品的转型，以及平台治理与基础设施存在的问题（Nieborg & Poell，2018）。本文引入文化政治理论，深入剖析文化形式背后的权力关系，一定程度上弥补了政治经济视角的不足。

在过去，女性人生价值的评判标准往往受限于贤妻良母等传统的家庭和母职角色。随着经济和社会的发展，性别平等观念崛起，女性对个体价值的追求受到鼓励。女性不再被视为男性的依附和客体，而是同样具有发展权利的主体，也是需要成长和承担责任的社会公民。女性察觉到了当下的困境，积极进行新身份的探索，谋求事业上的成功，使女性成长逐渐成为现代女性的普遍需求。但是女性成长仍然深受父权制度的桎梏，婚姻、家庭和生育等问题仍是女性在个人发展上面临的阻碍。小红书作为女性用户主导的社交媒体平台，搭建了女性用户共居的数字社区，女性在这里分享个体成长经历，交流个人境遇，为他人提供积极建议，为女性如何成长给出了经验参考和多样可能。这些内容大至人生选择、性格命运，小至职场生存、情感处理，在这一系列内容组成的符号环境中，女性用户的身份认同、主体性问题得以形成。因此，考察小红书平台上的相关内容，可以理解女性成长这一

套话语叙事背后的结构性因素。

已有研究除了关注文本中既定的意识形态、经济力量、阶级等问题，也聚焦不同群体基于特定文本的互动过程中所体现的权力关系。例如，在对进城务工人员系列婚纱照的研究中，孙皖宁指出，爱情/浪漫是由以下三个领域互动形成的一套文化叙事：包括相关政治经济的社会实践、分配不公正、阶级政治的组成部分（Sun，2018）。也有学者强调，任何对中国平台上的文化政治的考察都必须考虑到国家、市场、技术创新和社会之间的关系（Cunningham，Craig & Lv，2019）。既有的文化政治分析忽视了社交媒体平台上的文化形式和实践，而小红书作为用户生成内容（UGC）模式的生活方式分享平台，用户自发生成的内容既体现了用户的主体性实践，又顺从了平台的意识形态，反映了用户和平台之间复杂的权力关系和互动。小红书上的女性成长内容不仅反映了平台的价值偏向，也是一套文本叙事和话语，还暗示了平台的希望实践，是这三个领域相互作用的产物，这三个领域最终构成分析平台文化政治的概念框架：平台基础设施—平台和用户共同塑造的身份想象—实现机制。基于这一分析框架，本文从生产前提、话语特征和生产机制三个层面探讨小红书上女性成长内容的文化政治。

综上所述，本文探讨以下问题：

1. 小红书搭建了何种基础设施，以支持女性成长内容的生产？

2. 小红书提倡何种女性成长的平台话语？小红书上的女性成长内容体现出怎样的叙事特征？

3. 小红书制造出女性成长的何种生产机制？

三、研究方法

为了回答以上问题，本文综合运用了文本分析和线上参与式观察两种研究方法。

首先,本文对小红书相关的新闻报道进行文本分析。搜狗搜索是中国颇具影响力的搜索引擎之一,根据《2023 年中国搜索引擎行业研究报告》,搜狗搜索位列传统搜索引擎第三名(CTR 洞察,2023)。研究者以搜狗搜索为引擎,在搜狗资讯中输入关键词“小红书”,系统性地检索了 2013—2023 年的相关报道。由于搜狗资讯的技术屏蔽,研究者只能查看部分新闻报道,本文对可检索到的 300 余条新闻报道逐一进行阅读。

其次,研究者对小红书上女性成长领域热度排行前 100 的博文内容进行文本分析。截至 2024 年 2 月 17 日,在小红书上,以“女性成长”为标签的话题浏览量为 191 亿次。在这一话题下,研究者以“最热”为排序方式,选择内容热度排行前 100 的博文内容作为分析对象,这些内容覆盖了女性情感、家庭、职场、审美、人际交往、教育、消费、心理、人生选择等领域。以内容热度而非粉丝数量为排行依据,可以尽可能地覆盖素人博主,也体现了平台对内容质量的评判标准。自 2017 年以来,小红书采用核心指标“用户参与评分”(customer engagement score, CES)评价指向粉丝和创作者的互动情况。CES=点赞数+收藏数+评论数+转发数+关注数(真格基金,2023),每一项指标都被赋予不同的权重,最直接地反映了一篇小红书博文质量的高低。本文在研究对象的选取过程中遵循以下筛选标准:(1)博主性别为女性;(2)博主运营主体为个人,剔除明显由机构运营的内容;(3)博文为原创内容,即排除搬运他人发布的内容;(4)博文为图文形式,小红书是由图文形式主导的社交媒体平台,为保证文本形式的统一性,剔除视频内容;(5)剔除含有商业广告的博文。

最后,辅以线上参与式观察。在 2024 年 1—2 月,研究者在以“女性成长”为标签的内容的评论区,观察网友的留言和互动,并重点观察上述博文的评论区。

总的来说:首先,本文对小红书相关新闻报道进行文本分析,尤其是有关平台发展模式的访谈文本,提炼出平台的基础设施模式;其次,结合对新

闻报道和100条女性成长博文的文本分析,总结平台话语和用户内容生产的文本特征,研究两者塑造出何种成长话语;最后,基于对小红书女性成长博文的评论区的线上参与式观察,辅以对新闻报道的文本分析,分析平台制造了怎样的女性成长机制。

四、社区自治的基础设施:礼俗社会的新自由主义

20世纪90年代,随着改革开放的逐步深化,市场化改革和私有制经济迅速发展,单位制解体,人们的社会交往形态和生活状态发生了剧烈变革。90年代末期,中国互联网中形成以论坛为代表的虚拟社区(virtual community),乡村社会逐渐消失的共同体形式延续至线上。“社区”(community)这一概念最早由德国社会学家滕尼斯(F. Tönnies)提出,与现代商业社会不同,社区主要存在于传统的乡村社会,指一定地域范围内关系密切、守望相助的生活共同体(滕尼斯,1999:52—57)。虚拟社区类似于滕尼斯所归纳的心灵社区,只不过是由电子通信媒介形成的(Rothaermel & Sugiyima, 2001)。传统互联网时期的综合性虚拟社区往往由特定兴趣的爱好者所组建或因公共事件而集聚,如天涯社区、猫扑社区,其内容覆盖了时事、娱乐、生活等领域,用户通过发帖和回帖的形式表达想法和分享信息。进入21世纪后,中国深度融入全球市场,文化趋向多元化发展,得益于Web 2.0时代互联网的普及和宽带技术的发展,网民数量激增,他们不再是信息的被动接收者,还参与到内容生产和交互过程中。为满足网民创作内容和自主管理的需求,天涯等网络社区引入了版主选举、用户投票系统、社区规则共同制定等自治机制。2010年以来,移动互联网技术兴起,为了承载更大规模的用户量级并匹配差异性的用户喜好,虚拟社区向多元化、专业化发展,网民自治的社区运营逻辑也得以延续。网民自治的模式既保证了用户的参与度,又隐秘地将海量网络用户的内容创作活动转化为平台的

免费劳动。

在众多社交媒体平台中,小红书是采纳社区自治模式的典型案例。在相关报道中,其高层管理人员多次用“城市”隐喻来类比小红书。随着城市化进程加速,现代社会完成了从熟人社会向陌生人社会的转型,城市中的水泥森林无疑是冰冷的,但小红书勾勒的“城市”意象并非如此,而是对应了上海这座城市的一系列符号表征,是充满烟火气和人情味的弄堂小巷,以及具有海派特色的市民文化。社区里的居民彼此相识,真诚地关怀和分享,小红书便是通过制造出这样一种拟社区形态下的有机团结,实现了原子化社会中集体主义的复归。社区自治是一种基层民主政治形式,在小红书上的治理模式中,用户被放在十分重要的位置。受《美国大城市的生与死》中邻里自发监督行为的启发,小红书联合创始人毛文超带领团队探索社区内的自治机制和行为准则,“良好的社区氛围,需要热心的‘朝阳群众’一起维护。我们花了很多时间,找出并激励平台上的热心‘群众’,邀请他们一起维护社区的真实美好多元……并设置举报、陪审团等机制,去建立一个互助和维护的体系”(GGV 纪源资本,2019)。小红书上的社区自治还体现为用户和平台的关系重构。对于 2021 年推出的《社区公约》,小红书首席运营官丁玲解释道:“之所以称为‘公约’而非‘价值规范’,是因为‘价值规范’更像是自上而下的要求,是从社区的运营者视角向社区的居民提出的要求,而‘公约’则是小红书与不同类型的用户反复讨论、修改,最终成型。”(张洋,2022)

小红书上的社区自治模式,体现了新自由主义在当下情境的嬗变。从成立伊始提供境外购物资讯,到中期提供美妆购物参考,再到如今提供泛生活方式决策,小红书的内容生产一直围绕着生活方式的方方面面。生活方式内容兴起于新自由主义社会背景下,新自由主义强调福利国家权力下放后自我治理的霸权文化,以及以消费者为导向的公民身份的转变。在这一背景下,现代晚期的自我被构建为进取的、具有可塑性的个体(Lewis,

2011),而亚洲地区的压缩现代性带来了集体矛盾的个体化问题,大众媒体便成为解决这些现代生活矛盾的一剂良方(Martin & Lewis, 2016)。生活方式电视节目作为新自由主义的重要话语表征,为人们的房屋装潢、服饰打扮、健康决策、人际关系等日常生活提供了建议。在小规模共建共治的虚拟社区中,用户彼此亲密互助、真诚交流,提供在线社会支持,以拟亲属关系的“姐妹”相称,陌生的网友便变成熟识的居民。这种弱集体治理模式体现出新自由主义和中国传统乡土社会的结合,是一种礼俗社会的新自由主义。在小红书上,不再是自己一个人进行自我管理,而是一群女性用户共同进行自我管理,如她们会自发团结起来进行性别对抗的社区自治实践。在由熟人组成的礼俗社会中,人们习惯并热衷于介入、评价和参与他人的事项,社会交往常常运行为一种包裹着个人竞争的集体主义。小红书的社区自治便嵌入于这种集体主义文化,与西方语境下的自我治理模式不同,小红书的女性用户往往会吸收大量来自其他女性用户的经验和建议,并反复参考比较,直到最终做出决定;而评论区中同性之间的欣赏和赞美总是不吝流露,这似乎暗示着集体主义和竞争性的个人主义并不决然对立,社群主义俨然成为女性实现个体管理和自我改变的必要途径。例如,毛文超在小红书创办 10 周年的内部分享会上提到,希望用户可以在小红书上“感受到一个‘温暖的小社区’,让陌生变得不冰冷,让活着不再孤单”(Morketing, 2023)。社区自治作为小红书最为关键的基础设施模式,是女性成长内容生产的前提。

五、话语制造:中产阶级抱负与现代性焦虑

进入 21 世纪第一个十年,中国经济快速发展,居民消费升级。在这样的时代机遇下,聚焦于境外购物资讯的小红书诞生了。毛文超说他相信大

的趋势,并愿意为此做长期投入:“在 2010 年,中国只有 5000 万中产阶级,麦肯锡报告预测 2030 年中国会有 8 亿中产阶级,也是全球最大的中产阶级群体。很多时候答案都写在书里了,只是你愿不愿意尝试,而且它要 20 年的时间。”(潘乱,2021)小红书自成立之初便瞄准了中国崛起的中产阶级群体,这群早期用户具有如下特征:居住在一线城市,受过良好教育,热衷时尚,拥有较强的购买力。和常被诟病为“土味”的抖音、快手截然相反,小红书常常被媒体类比为中国的 Instagram,都专注于“年轻的中产阶级生活方式的方向”(潘乱,2021),精致的平台调性由此定下。小红书首席运营官丁玲曾将小红书比作“线上安福路”(财经无忌,2023)。正如商铺林立、装潢优雅的上海安福路,在小红书上,女性用户分享精致生活,排列的内容体现了中产阶级的消费能力和品位,进一步唤起女性用户对匮乏的焦虑,以及购买和占有商品的欲望。尽管“中产阶级”的概念范畴在亚洲语境下需要谨慎使用,但对中产阶级身份的想象确实是亚洲地区一个关键的抱负概念(Martin & Lewis, 2016)。中产阶级代表着一种开放的优绩主义(meritocracy),暗示了人的主体性是实现社会流动的关键,个体既可能因为努力而上升,同样也存在滑落的风险(Weiss, 2019: 23-24)。在现代社会的发展过程中,中产文化逐渐转变成一种主流生活方式和主导文化,对中产阶级身份的追求具象为对美好生活的向往。小红书便基于对中产阶级身份的想象,鼓励用户分享日常生活中美好的部分,在庸常的日子中制造“小确幸”。

在小红书上“女性成长”的热门内容中,女性成长博主往往拥有高等学历、互联网大厂就业经历或动辄年薪几十万元的高收入。光鲜的履历不仅是女性成长博主传授经验的社会资本,同样使得女性成长具象化,宣告着女性成长的终点:经济独立、自由、拥有丰富的人生阅历和稳定的亲密关系的中产阶级女性。例如:“一晃就毕业了,论文也发表了,房子也买好了,对象

也找好了。”“28 岁,旅居并不上班,但体验了 15 种不同的生活。”那么,当遥不可及的“更好的自己”得以具象化,在现状和缥缈的许诺之间,它提倡了何种切近方式呢?

首先,小红书社区中的女性成长被量化为精细的事项清单。计量作为现代社会的重要特征,不仅是衡量客观事物的方式,还形成了一套反思与提升自我的实践装置。自我量化作为一种现代自我治理技术和认识自身的经验模式,帮助女性记录与监测自己的生活状态,以及进行全方位的、精细化的身体意象管理。现代人的成长本是模糊的、多义的和不确定的意象,但在小红书上,成长不仅被拆解为一个个可被执行的清单和每日细微变化,而且成为一个可以被推广和应用的普遍方案,中产阶级也就变成了一个通过努力可以实现的身份。无论想要成长为何种女性,女性成长叙事都遵循同一套逻辑——打造健康的身体、学习技能、提升学历、锻炼口才、观看纪录片、通过执行清单——女性用户获得掌控自我和维持秩序的快感。例如,在“布局自己的 2024 年”中,这名博主给出了 18 条详细的操作方案,如:“每周敷面膜 4 次,每天早上 6 点起床,晚上 11 点半准时睡觉,体重保持 95 斤,每天抽出 1 小时用来阅读。”似乎实践了这些方法,女性用户便能够和博主发生同样的人生蜕变。在不确定的时代,当日常生活成为清晰可见的量化指南,自律和勇气便被作为一种内在道德品质建构出来,成为应对现代性焦虑的手段。

其次,克服物和情感的冗余成为女性成长的第二种方式。对物的占有似乎构成女性成长的首要障碍。现代社会中的物和心灵被决然地对立起来,物被认为是虚妄的。女性成长的第一步,便是宣告与物欲割席。女性成长博主纷纷提倡断舍离和极简生活,将时间和金钱投资于考研考证等学历深造、技能培训上,用于充实头脑、丰盈内心,而非消耗于外在穿搭。同时,感性和理性被决然地对立起来,如“学会‘不响’”“你富养自己最好的方式:

藏”。社会往往要求女性承担责任、表达和分享情感,但在女性成长内容中,这类感性思维是需要舍弃的,女性被要求压制情绪,敏感脆弱甚至被视为低劣的、有害的品质,只有理性才是有用的特质。同时,原生家庭、恋爱关系、人际关系都被视为负担,只有自身才是值得去经营和培育的一份事业。从以上可以看出,“大女主”的成长被绘制为一个内隐的不动声色地努力和取得成功的过程,这一套叙事策略不仅顺从了效率至上的现代社会对现代人的话语规训,还彰显出中产阶级对自身的纪律要求和防止坠落的策略。“中产阶级把自身定义为具有领袖气质的阶层,因为它有诸多美德:极高的道德标准,自我规训和节制,勤勉与理性,对科学和进步的坚定信念”(洛夫格伦、弗雷克曼,2011:215),对这些美德的追求体现了小红书女性用户的中产阶级抱负。

对比小红书提倡的平台话语和女性成长内容的叙事逻辑,可以发现,女性成长博主的内容生产极大地顺从了平台价值观,最终被简化为缓解现代性焦虑的同质化的叙事模板。在图文叙述中,女性成长博主多数采用一套阶层跃迁的叙事逻辑,如“改命”“疯狂地重构自己”“走出人生低谷”“往死里培养自己”,一场场弱者的胜利看似在宣告女性主义意识的觉醒,却暴露了现代女性受到的约束。这些成长叙事与成功学具有极其类似的话语逻辑,和资本主导的女性文化场域共享同一套“打怪升级”的快感机制。同时,底层身份被挪用为建构中产阶级身份的政治元素,成为促进社区团结、激发力量和获得共鸣的手段,而真正的底层群体自始至终都被排除在外。

六、生产机制:分享作为一种希望实践

作为依托于用户生成内容的社区,社区内容生态的优劣直接决定了小

红书的发展。因此,小红书采用去中心化的算法分发机制、低门槛的用户激励、图文并茂的内容形式,确保普通用户生产内容的便捷和可见度。根据《2022 新媒体内容生态数据报告》,小红书腰部账号、尾部账号几乎占了平台总数的 99. 97%,头部关键意见领袖(KOL)占比不到 0. 03%,而视频号、抖音、快手、哔哩哔哩则拥有 1% 以上的头部 KOL 占比(财经故事荟,2024)。在这样的流量分发机制下,即使是粉丝寥寥的新用户,只要有优质内容产出,同样可以获得大量曝光。同时,根据小红书蒲公英平台规则,粉丝数量过千的用户,即可以通过平台开展商业广告推广,这为普通用户的平台创业实践提供了更多的机会。其中,准入门槛更低的便是女性成长博主。女性成长博主在展露自身变化的同时,还纷纷号召更多的女性用户加入这项平台创业实践中。“建议所有普通女生都来做成长博主”,这是小红书女性成长内容中出现的高频标题,理由往往是“适合普通人”“受众范围广,变现空间大”“粉丝增长速度快,黏性强”“倒逼自己快速成长”。做成长博主并不需要艰深的专业知识,只需要有敢于改变自我的决心,以及能够持续不断地记录自己的改变。对于普通女性用户而言,其优势在于“不用你多厉害,而是有一批人能见证你一点点变厉害,这样才能给更多的普通人提供更具参考的成长意义”。女性成长内容热度排行前 100 的博主,其粉丝数量在 1000—12 万不等,也就意味着她们都不同程度地获利。小红书女性博主的创业实践被阐释为一种性别创业主义,在网红经济的推动下,这些“野生”女性生活方式博主展现出一种后女权主义创业精神(Guo, 2022)。浏览女性成长博主的简介和动态可以发现,她们要么是在本职工作外兼职自媒体月入过万元,要么是逃离体制内工作启动全世界旅行的数字游民。无一例外,她们通过运营自媒体账号,在经济和思想上成为更加“自由”的人。评论区的女性用户不仅流露出对这种生活方式的欣羡,还表达出对这种创业精神的向往。

对于小红书这一生活分享类平台(彭兰,2022)而言,内容分享是社区运行的关键。通过向普通用户倾斜的算法分发机制,小红书营造出了良好的 UGC 社区内容生态:人人都可以做博主,并有机会成为网红。平台的底层技术逻辑实现了广泛的用户入场动员,开启了希望实践的数字空间(谌知翼、李璟,2022)。在女性成长领域,成绩斐然的博主还会受到小红书平台的合作邀请。例如,女性成长博主"野蛮生长段老师"便被邀请参与小红书组织的话题讨论:"我被小红书邀请给大家一起放'烟花'啦!……我送给大家一朵'生命力烟花'。年龄和生命力相比,完全不值一提。"通过参与该话题,该博主获得了更多的流量支持。"有关进步的许诺"(promise of progress)被认为是希望生产的中介(Mattingly, 2011)。通过邀请出色的女性成长博主参与官方活动这一展演行为,小红书平台制造出有关出名、成功与财富梦的承诺。基于平台的希望实践,小红书试图唤起女性用户的创业主义精神,激发和调控女性的分享欲望,使其成为持续的内容生产者,最终把数字创业主义的内涵化约为分享这一动作,即分享与被看到即一种改变。在平台的网红经济逻辑和希望实践下,女性成长最终演变为一系列命题:女性不仅要实现个体的成长,还要分享自己的变化,具有创业主义精神并成为女性成长博主,保持持续的内容输出,最终通过平台创业,实现经济独立,这才算是改变了现状,完成了最终的成长。而女性博主进行平台创业的手段——接广告,则是更彻底地对物的占有的宣誓,再次背离了断绝物欲的初衷。在这个过程中,小红书正是利用持续分享的女性,隐秘地进行了商业运营。

七、结论与讨论

基于文化政治的理论视角,本文以社交媒体平台小红书上的女性成长

内容为分析对象,详细剖析了在平台和用户的互动中,女性成长内容中的阶级、身份和现代性等问题。首先,小红书搭建了社区自治的平台基础设施,试图在高度原子化的现代社会中营造一个有机团结的女性共居社区,这体现出新自由主义的当下嬗变,以及与传统乡土社会的结合,即礼俗社会的新自由主义;其次,小红书创设了中产阶级的身份想象,用户在内容创作中不仅迎合了平台上中产阶级的抱负,还为此生产出一套高度同质化的对抗现代性焦虑的成长叙事;最后,女性成长既是平台塑造中产阶级身份的由头,也是构成女性分享和创业精神的前提。平台内容分享作为一种希望实践,成为女性成长内容的重要生产机制,拒绝成长的女性不仅不被社会接纳,也不被平台所包容。

总的来说,在女性用户群体主导的小红书上,女性成长俨然成为一个必然的命题。在平台和用户的共同作用下,女性成长成为一种雷同的无限繁殖的媒介叙事,女性成长的终点被同化为对权力和金钱的追逐,单一的评判标准取代了多元的成长,所谓的“女性成长”实际上只是偷换了性别,本质上仍然是一场对父权制的成功逻辑的复制。此外,小红书上的女性成长被运作为一种性别化新自由主义,女性似乎都具有先在的自驱力和自由,其面临的生育、家庭照料等结构性困境被掩盖了。小红书的平台文化政治就在于,采用本真和多元的话语表征来漠视平台的阶级特征,用社群主义和技术人文主义的温情来掩盖平台经济和父权制残酷运转的本质。

需要注意的是,在基于数字平台的用户内容生产中,分析这类新媒体文化的权力关系,不仅需要纳入对平台基础设施、平台话语、生产机制的考察,还需要关注用户的主体性实践,并在此基础上考察多元主体的权力互动。本文考察小红书上的女性成长内容生产的文化政治,不仅探讨了新媒体文化形式和实践的权力关系,还观照了数字平台本身的文化政治。这既是对既有数字平台研究忽视文化商品的转型,对平台治理和基础设施问题

(Nieborg & Poell, 2018)的补充,也将对表征形式和接受表征的外部条件的分析结合起来(Angus & Jhally, 2022: 2),是文化政治理论在数字社会和消费社会情境中的新发展。

参考文献

鲍尔德温,2004,《文化研究导论》,陶东风译,北京:高等教育出版社。

CTR 洞察, 2023,《2023 年中国搜索引擎行业研究报告》, https://www. 36kr. com/p/2298513609122179。

财经故事荟,2024,《小红书"长尾"变现:超头缺位,KOC 雄起》, https://www. woshipm. com/it/5969341. html。

财经无忌,2023,《微信逼了一下小红书》,https://www. jiemian. com/article/8948985. html。

谌知翼、李璟,2022,《"三无直播间"何以维系:抖音平台素人主播的希望实践》,《新闻记者》第 11 期。

GGV 纪源资本,2019,《小红书 CEO 毛文超:好的内容社区,需要"朝阳群众"》, https://www. sohu. com/a/351830126_323203。

郭力,2020,《文学史的对话——以"互文性"探寻现代女性成长主题》,《中国文学批评》第 2 期。

李艳丰,2023,《文化·政治·权力:西方文化政治理论关键词辨析》,《暨南学报》(哲学社会科学版)第 1 期。

刘亭亭、许德娅,2023,《作为时尚文化基础设施的小红书与时尚观念的重塑——基于三类时尚博主的田野研究》,《国际新闻界》第 6 期。

洛夫格伦、弗雷克曼,2011,《美好生活:中产阶级的生活史》,赵丙祥、罗杨译,北京:北京大学出版社。

Morketing,2023,《小红书 CEO 星矢十周年分享:科技为笔,人文精神为墨,创造广阔而温暖的社区》,https://new. qq. com/rain/a/20230804A06U2Q00。

潘乱, 2021,《小红书的晚熟和非理性执着》, https://mp. weixin. qq. com/s/l36P51hMJIbxC7a5u406bw。

彭兰,2022,《媒介化、群体化、审美化:生活分享类社交媒体改写的"消费"》,《现代传播(中

国传媒大学学报)》第 9 期。

孙皖宁、苗伟山,2016,《底层中国:不平等、媒体和文化政治》,《开放时代》第 2 期。

滕尼斯,1999,《共同体与社会:纯粹社会学的基本概念》,林荣远译,北京:商务印书馆。

同心,2014,《“成为更好的自己”:时尚论坛与女性身体消费——以天涯论坛时尚版为例》,《新闻大学》第 1 期。

汪民安,2020,《文化研究关键词》,南京:江苏人民出版社。

吴靖、王颖曜,2007,《可见的地产和不可见的家——解读房地产广告的文化政治》,《广告大观(理论版)》第 3 期。

吴靖、云国强,2016,《新自由主义现代性阴影下的家、审美权威与阶级认同——〈交换空间〉与生活方式电视节目的文化政治》,《开放时代》第 4 期。

张洋,2022,《小红书 COO 对谈学者刘擎:小红书发展不算快,需要慢慢融入更广大用户》,https://new.qq.com/rain/a/20220119A05Z6A00。

真格基金,2023,《1 亿人的小红书:时机、摇摆和决心|Z Circle》,https://www.sohu.com/a/727696487_355029。

Abu-Lughod, L. 2003, “The Objects of Soap Opera: Egyptian Television and the Cultural Politics of Modernity.” in D. Miller ed., *Worlds Apart: Modernity Through the Prism of the Local*, London: Routledge.

Angus, I. & S. Jhally 2022, *Cultural Politics in Contemporary America*, New York: Taylor & Francis.

Apple, M. 2014, “Cultural Politics and the Text.” in S. Ball ed., *The Routledge Falmer Reader in Sociology of Education*, London: Routledge.

Cunningham, S., D. Craig & J. Lv 2019, “China’s Livestreaming Industry: Platforms, Politics, and Precarity.” *International Journal of Cultural Studies* 22(6).

Guo, J. 2022, “The Postfeminist Entrepreneurial Self and the Platformisation of Labour: A Case Study of Yesheng Female Lifestyle Bloggers on Xiaohongshu.” *Global Media and China* 7(3).

Hooks, B. 2014, *Yearning: Race, Gender, and Cultural Politics*, New York: Routledge.

Jordan, G. & C. Weedon 1995, *Cultural Politics: Class, Gender, Race, and the Postmodern*

World, Oxford: Blackwell.

Ju, R. 2022, "Producing Entrepreneurial Citizens: Governmentality Over and Through Hong Kong Influencers on Xiaohongshu (Red)." *Policy & Internet* 14(3).

Korf, B. 2020, "Cultural Politics." in A. Kobayashi ed., *International Encyclopedia of Human Geography*, Amsterdam: Elsevier.

Lewis, T. 2011, "Making Over Culture? Lifestyle Television and Contemporary Pedagogies of Selfhood in Singapore." *Communication, Politics & Culture* 44(1).

Liu, Y. & X. Li 2023, "Pale, Young, and Slim Girls on Red: A Study of Young Femininities on Social Media in Post-Socialist China." *Feminist Media Studies* X(X).

Louisa, S. 2020, *Minority Rules: The Miao and the Feminine in China's Cultural Politics*, Durham: Duke University Press.

Martin, F. & T. Lewis 2016, "Lifestyle Media in Asia: Consumption, Aspiration, and Identity." in F. Martin & T. Lewis eds., *Lifestyle Media in Asia: Consumption, Aspiration, and Identity*, London: Routledge.

Mattingly, C. 2011, *The Paradox of Hope: Journeys Through A Clinical Borderland*, Berkeley: University of California Press.

Nash, K. 2001, "The 'Cultural Turn' in Social Theory: Towards A Theory of Cultural Politics." *Sociology* 35(1).

Nieborg, D. & T. Poell 2018, "The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity." *New Media & Society* 20(11).

Painter, J. 1995, *Politics, Geography and "Political Geography": A Critical Perspective*, London: Edward Arnold.

Rothaermel, F. & S. Sugiyama 2001, "Virtual Internet Communities and Commercial Success: Individual and Community-Level Theory Grounded in the Atypical Case of TimeZone.com." *Journal of Management* 27(3).

Rycroft, S. 2009, "Cultural Politics." in R. Kitchin & N. Thrift eds., *International Encyclopedia of Human Geography*, Amsterdam: Elsevier.

Sun, W. 2009, *Maid in China: Media, Morality, and the Cultural Politics of*

Boundaries, New York: Routledge.

Sun, W. 2012, “Poetry of Labour and (Dis)Articulation of Class: China's Worker-Poets and the Cultural Politics of Boundaries.” *Journal of Contemporary China* 21(78).

Sun, W. 2013, “The Cultural Politics of Recognition: Rural Migrants and Documentary Films in China.” *Journal of Chinese Cinemas* 7(1).

Sun, W. 2018, “Romancing the Vulnerable in Contemporary China: Love on the Assembly Line and the Cultural Politics of Inequality.” *China Information* 32(1).

Sun, W. 2022, “Dark Intimacy and the Moral Economy of Sex: Rural Migrants and the Cultural Politics of Transgression.” *China Information* 36(1).

Weiss, H. 2019. *We Have Never Been Middle Class: How Social Mobility Misleads Us*, London: Verso Books.

沉默裁员与技能迭代*

——制造业数字化升级的影响效应研究

魏　丹　刘曙蕾**

摘要：制造业日益成为国际经济的竞技场，数字化则是经济发展的重要驱动力。自2017年始，中国将数字化提升至国家战略高度，但国内制造业数字化程度仍面临地域、行业及企业间的发展不均衡状态。典型案例研究显示，纸品包装印刷行业目前正处于数字化装备普及阶段，呈现出工序简化、柔性化的生产范式转变。相比于规模性裁员，数字化转型过程中的失业更多体现为，在维持既有员工规模的基础上，产业规模增大与订单数增多的沉默裁员。与此同时，数字化转型呈现出技能迭代效应，体现为核心操作技能被计算机程序替代，对非常规性任务、认知性任务与软技能的要求提升。这一过程也意味着技能的控制权从工人向工程师与管理人员转移。

关键词：数字制造　沉默裁员　技能迭代

一、引言

人类制造业经历了“手工制造—链式机器制造—大机器制造—标准化规模制造—规模定制”的范式转型（杨青峰，2018：1—2），目前正向数字革

* 本文系国家社会科学基金项目“智能制造时代农民工技能养成研究”（项目批准号：18BSH142）、江西省高校人文社会科学项目“数字社会背景下数字技能对农民工创业决策的影响与对策研究”（项目批准号：SH22108）的阶段性研究成果。

** 魏丹，南昌大学公共政策与管理学院。刘曙蕾，南昌大学公共政策与管理学院。

命和智能革命范式演进。数字经济成为经济发展的重要驱动力,中国更是将数字经济上升为国家战略高度。2017 年,“数字中国”首次被写入党和国家政策纲领文件,此后更是强调加快建设制造强国、数字中国,加快发展数字经济。《数字经济及其核心产业统计分类(2021)》指出,产业数字化旨在“应用数字技术和数据资源为传统产业带来的产出增加和效率提升,是数字技术与实体经济的融合”。产业数字化转型是借助数字技术进行传统产业升级,提升生产数量和效率的过程(肖旭、戚聿东,2019)。推进产业数字化转型,要对传统产业实行全方位、全角度、全链条改造,提高全要素生产率(李洋,2022:53)。

数字技术已在制造业领域应用了 30 多年(Chen, Heyer & Ibbotson et al. , 2015)。20 世纪 90 年代,数字信息被应用于提高生产效率;21 世纪初,数字信息应用被拓展至供应链系统;2010 年以后,数字信息被应用于制造流程、供应链、产品和服务的全过程。“数字制造”概念起源于数控机床和计算机数控技术。随着 CAD 制图、计算机网络、虚拟现实、快速成型等技术的发展,制造业借助信息的分析、规划、重组、协调、共享功能,实现了产品设计、仿真与原型制造的快速实现。数字制造是数字技术、网络技术和制造技术的深度融合,是制造企业、制造系统、生产系统的数字化过程。数字制造以个人、车间、设备、市场以及企业为节点,构成了整个系统网络,包括设计中心系统、控制中心系统、制造中心系统和管理中心系统(Oztemel & Gursev, 2020: 127-182)。数字制造是利用数字技术、网络技术进行从设计到制造的数字化处理过程,是数字化技术与制造技术融合的新产业、新业态、新模式。在传统制造领域,企业的竞争对手及竞争关系相对清晰。数字制造则打破了既有的行业界限,改变了员工、客户、供应商和合作伙伴的关系,是对整个生态系统的重塑(梅菲特,2018:7—8)。在这个过程中,数据成为新的生产要素和最具价值的资源,将引发生产组织结构的变化,以及劳动者工作内容和工作方式的深刻变化(王天夫,2021)。

人类制造业生产范式的每一次变革,都相应影响着工人就业和技能发展。譬如马克思(K. Marx)认为,手工工场向机器工厂过渡的过程,即发生了上述变化(马克思,2004:469—487)。布雷弗曼(H. Braverman)指出,资本主义生产既延长了空洞的"教育",又将劳动变成简单的、无须知识的工作。管理部门或参谋机构,借助科学和工程技术知识,分解了传统工艺,留给工人不完整的技能、有限和重复的动作,工人的技能发生了退化(布雷弗曼,1973:397—399)。关于工业革命对技术工人的影响,形成了"去技能"(deskilling)和"技能提升"(upskilling)的不同观点(Grugulis, Keep & Warhurst, 2004: 111)。"去技能"观点认为,到 19 世纪末 20 世纪初,布雷弗曼观点的解释力逐渐减弱。随着工厂的日益复杂和规模的日益扩大,工厂的管理、控制、协调和计划发生了系统性变化,工厂的生产过程和劳动的组织基础不同以往。技术的进步催生了新工作,通常要求工人有较高的技能水平和责任感(Heisig, 2009: 1639-1651)。在早期研究中,技能偏好往往等同于受教育水平,即技术变革导致对受教育水平较高的劳动力的需求增加(Tinbergen, 1974),同时也导致了生产中人际关系的变革(谢新水,2021)。

阿西莫格鲁(D. Acemoglu)等人研究了计算机技术对工人技能的影响,并将工作划分为常规性任务(routine)、非常规性任务、认知性任务和操作性任务四种类型。其中,常规性任务的特点是对许多中等技能的认知和操作,如记账、文书、重复性生产和监控。计算机等技术的发展导致欧美企业中的常规性任务量大大减少,工作呈现两极化现象(Acemoglu & Autor, 2011: 1043-1171)。与美国相比,在大多数中东欧国家,常规性任务、认知性任务却增加了(Frank, Autor & Bessen et al., 2019: 6531-6539)。

既有研究揭示了工业革命与技术变迁对就业替代、技能方向的影响,以及基于任务模型呈现的技能替代率的差异影响。但这些分析与论述中的技能是高度抽象的,并未考量具体产业情境、组织生态等因素。此外,人类前几次工业革命是物理介质层面的变化,而数字变革突破了物理介质层面,向

软体化、虚拟仿真、智能思维方向发展,预示着全新的竞争模式与生产理念的形成。这一变化趋势将如何影响工人就业与技能发展,即在数字化变革、技术升级过程中,对工人就业与技能发展的影响仍然是两极化吗?既有研究从横向截面角度,对不同工作种类的工人就业与技能发展进行了较多研究,但从行业纵深层面看,数字化技术究竟会对工人就业与技能发展产生何种影响,这种影响又是如何发生的?这些问题需要深入产业及企业组织具体形态和场域中去解答。本文基于纸品包装印刷行业的跟踪研究与相似案例比较发现,制造业数字化升级呈现出就业,沉默裁员效应及技能迭代效应。

二、田野空间及案例说明

早在2019年底,浙江、福建和广东便被国家发改委和中央网信办列为首批国家数字经济创新发展试验区。《中国中小企业数字化转型发展研究报告(2022版)》指出,江浙沪、广东、福建等东南沿海地区贸易及制造业发达,区域产业集中,中小企业数字化转型优势突出,且总体数字化水平较高(腾讯研究院,2022)。因此,笔者选择浙江、福建和广东作为案例观测点。研究对象选择纸品包装印刷行业的缘由如下:第一,笔者曾入职纸品包装印刷企业,两年的工作经验及相应的社会关系为笔者了解和探索相关问题创造了现实条件;第二,纸品包装印刷行业是典型的外部技术推动的技术应用型产业,设备、材料领域的创新,特别是数字化技术、智能化技术的发展,对此行业产生了革命性影响。换言之,纸品包装印刷行业是外部技术变化敏感型行业,更容易借此捕捉及管窥外部技术变化引起的行业及企业变化。研究对象最终选取了浙江台州D企业、福建泉州T企业、广东东莞Z企业与广东中山X企业。尽管近年来中国数字经济取得了令人瞩目的成就,中小企业数字化升级快速发展,但绝大部分企业仍然处于初级转型阶段。

《中小企业数字化转型分析报告(2021)》显示,2021 年处于数字化转型初步探索阶段的企业达 79%(中国电子技术标准化研究院,2022)。纸品包装印刷行业发展多处于数字化的普及期、网络化的推广期、智能化的探索期,因此本研究选择已购置数字印刷设备的工厂进行研究。

纸品包装印刷行业通常将工厂生产链划分为一级厂、二级厂和三级厂:一级厂的产业链最全,包含原纸制造、瓦楞纸板、纸箱加工;二级厂指购买原纸以生产瓦楞纸板、进行纸箱加工的工厂;三级厂指购买瓦楞纸板进行纸箱加工的工厂。

D 企业成立于 1995 年,是一家生产瓦楞纸板、进行纸箱加工的二级厂,其中瓦楞纸板生产占比 70%,纸箱产品生产占比 30%。2020 年,D 企业生产总值超过 1.4 亿元。2019 年以前该厂设备以传统链式印刷机械为主,2020 年购置了数字印刷设备,开始生产环节的数字化转型。当前,D 企业印刷配套设备未完全更新,经营管理理念正在转型中。

Z 企业成立于 1995 年,坐落于东莞工业区。2023 年春节前夕,Z 企业购置了数字印刷设备,主营瓦楞纸品包装印刷业务。Z 企业有相对固定的销售客户群,当前正处于数字化转型磨合初期。

X 企业成立于 2009 年,由数字印刷厂、彩印厂、数字设计部三个部门构成,不仅提供纸品包装加工产品,且为客户提供产品服务解决方案。X 企业于 2021 年购置了三台数字印刷设备,2022—2023 年又迅速更新了印刷配套前后端的自动化设备,且引进金软软件,形成了设备结合完善的 ERP 系统,成为行业内数字化转型迅速且完整的后起印刷企业。

T 企业成立于 2010 年,主要经营纸品包装印刷加工业务。T 企业的数字印刷设备更新较早,2018 年购置了万德数字印刷设备一台,2020 年购置了汉华数字印刷设备一台。经过近 5 年的摸索,其生产环节、经营理念、企业管理等数字化转型相对成熟,数字化订单数占总订单数的一半,成为当地的纸品包装印刷数字化领跑型小微企业(表 1)。

表 1　　数字化转型企业案例

企业	成立时间	企业类型	数字印刷设备购置时间	当前转型阶段
D 企业	1995 年	二级厂	2020 年	配套转型摸索
Z 企业	1995 年	三级厂	2022 年	转型磨合初期
X 企业	2009 年	三级厂	2021 年	转型迅速完整
T 企业	2010 年	三级厂	2018 年	转型成熟领先

X 企业实现了印前、印中及印后的自动化与智能化，且在产品设计、生产与物流管理层面数字化水平较高；D 企业、X 企业拥有高速扫描数字印刷机，设备水平相当，但 D 企业印前、印后的自动化水平相对较低，全自动机械的更新相对滞后；Z 企业的自动化水平较高，数字印刷环节仍在磨合期；T 企业的数字印刷环节使用效率较高，印前自动化水平较高，但印后自动化水平中等。

三、数字化升级生产范式的简化与柔性化转向

数字印刷（digital printing）又称“无版印刷”，是通过计算机制图处理后将图文信息直接成像于介质表面，无须提前制作印版、输出菲林的新型印刷方式。因其拥有保护商品、便于储运等多种优势，瓦楞纸包装产业迅速发展。随着电商、互联网经济的发展，中国物流纸品包装需求迅猛发展（王志星，2009）。中国数字印刷设备制造兴起于 2010 年前后，其应用在 2020 年以后开始逐渐普及，但总体而言投产比例仍然较低，不超过 30%。

生产工序及流程简化是数字化印刷生产范式的最大特征。传统纸品印刷采用柔版印制方式，需要十多道复杂的工艺流程；数字印刷应用工业打印机原理，实现了“图稿制作—工业设备打印—印刷输出”的简单过程。在瓦楞纸板印刷方面，数字印刷实现了从计算机到瓦楞纸板的直接印制，节约了大量工序和材料，制作周期更短，印刷准确率更高。例如，X 企业将传统印

刷两三天的时间缩短为 30 分钟,最快甚至可以实现 5 分钟出产品。

数字印刷设备的使用不仅缩短了生产周期,提高了成品效率,而且生产全过程更加柔性化。在成品批量生产前,传统印刷需要业务员与客户进行大量的沟通、制版、打样、确样等工作;需要开版、制版,制版费为 500 元/m^2—1000 元/m^2,颜色多样、图文复杂的印刷品费用更高。为控制成本,印刷文件需要保持相对稳定,印刷字体、大小、颜色等任何内容的变化,都将导致原有印版的失效,需要重新调版、制版。在数字印刷中,图文信息经由计算机处理可直接印制,实现了计算机成像与印刷品的“所见即所得”,省去了制版公司的大量印前工作。数字印刷一张起印,即时纠错,数据随时可变,按需印刷,工厂、客户可即时沟通,省去了大量的沟通协调工作,印刷生产更加柔性高效,但也使得企业沉默裁员成为可能。数字化技术产生了对数字思维的要求(郭琳,2022),也将对工人就业与技能发展产生影响。

四、制造业数字化升级的沉默裁员效应

工业革命时期,资本家通过提高资本有机构成、投资规模越来越大的生产资料、雇佣越来越少的劳动力以提高生产效率,获取更多的利润(马克思,2004:262—263)。20 世纪 50 年代初开始,美国制造业雇佣工人的比例急剧下降。1995—2002 年,全球有 2200 万个工厂工作消失。与此同时,制造业产出增长 30%(福特,2015:59)。自动化及新式制造等技术使得生产就业岗位一直在减少,受冲击最大的是机械工。与此同时,对工程师及技术人员的需求则在成倍增长。但是新技术带来的沉默裁员效应却被忽视了,这体现为公司停止招募新的员工。这一现象隐秘发生,但影响深远(诺布尔,2013:76)。所谓“沉默裁员”,是指随着新技术的发展与应用,企业扩大规模、增加产值的实现不再以增加工人数量为前提,而是用更少劳动力生产更多效益。

与中国传统印刷业集聚于长江三角洲、珠江三角洲、环渤海的区域性分布不同，目前数字化印刷企业呈现区域离散化、碎片化的特点。印刷行业的数字化贯穿于设计、生产、管理、服务等各个环节，但目前中小企业的数字化主要表现为印刷工艺的数字化、数字印刷设备的升级与运用。例如，佛山集聚了一批自动化、数字化与智能化程度较高的企业，如美盈森、山鹰国际等。国内众多中小企业同本文中的 4 个企业一样，处在印刷环节数字化设备升级阶段，但生产全流程的自动化、数字化水平差异较大。在数字化设备升级后，4 个企业呈现出同样的沉默裁员效应，比起规模性裁员，更多体现为企业在既有保持工人数量的基础上，订单结构多元化，订单规模增加，生产效率、产值和利润率提升（表 2）。

表 2　　4 个企业数字化升级前后变化对比

企业	单日完成订单数（个）		订单增加量	产值变化（万元）		员工数	利润率增长
	升级前	升级后		升级前	升级后		
D 企业	50	80—100	20%	300	500	不变	30%
Z 企业	20（磨合期）		30%	30	50	不变	20%
X 企业	50—60	120—150	60%	400	900	印刷环节减少	45%
T 企业	60	80—120	40%	50	80—120	总数减少	60%

在数字印刷设备更新后，4 个企业的生产工人数量总体变化相差不大，但印刷岗位工人数量减少，且生产效率、产值、利润率都出现了增长。

D 企业：纸箱部生产人员总数保持 32 人不变。升级前传统印刷机单日完成订单量不多于 50 个，加班频率高。引入数字印刷设备后，单日完成订单量增长至 80—100 个，客户总体订单量增加了 20%，月产值从原来的 300 万元增加至 500 万元，利润率增长约 30%。

Z 企业：数字印刷主要用于生产单一外贸客户订单，生产员工总数保持 41 人不变。2023 年 3 月为新机器磨合期第一个月，完成订单量 20 个（受订

单市场影响,且该客户单个订单量较大),订单产值从 30 万元增加至 50 万元,未来目标产值在 80 万元左右,利润率增长约 20%。

X 企业:引入数字印刷设备前,员工数量为 50—60 人,月产值 400 万元。引入数字印刷设备后,员工总体数量不变,月产值提升到 900 万元,单日最高可完成订单量 150 个,完成纸板数量 1.2 万—1.5 万张,利润率增长约 45%。

T 企业:引入数字印刷设备前,生产车间的工人有 12—15 人,月产值最高 50 万元。引入数字印刷设备后,印刷工人从 4 名减少至 2 名,另有 1 名开槽兼印刷工人,固定工人减少至 9 人,订单量增加了 40%,月产值提升至 80 万—120 万元,利润率增长约 60%。

数字化设备升级并未带来大规模的裁员,但却在没有增加工人数量的前提下大大提高了工厂的劳动生产率和产值,员工单位产值和单日处理订单量大大增加,表现为更隐秘的沉默裁员效应。数字化生产除了对工人就业产生影响外,同时将引发工人技能的重大变革。

五、制造业数字化升级的技能迭代效应

数字化设备的更新与使用简化了传统的工序,体现为印前环节和印中环节的相关技能被大范围替代。

第一,印前环节的制版技能被制图技能替代。数字印刷实现了从图稿到机器工业打印的转变,开版、制版、打样、调版确样、挂版等一系列印前环节均被计算机制图程序替代了。4 个企业的印前环节都体现了制版技能为制图技能所替代,但由于企业间数字化程度不同,因此制版技能被替代的程度也不同。

第二,印中环节的机械网纹辊压力调试、印版位置调整以及分色印刷机操作技能,被数字印刷机通过程序设计替代为参数控制面板操作,体现为传

统印刷机、分色印刷机操作技能被兼容替代。传统印刷行业通常因机器设岗,单色印刷机、双色印刷机、四色印刷机的操作原理及技能差别很大,机械压力、印版位置、套印技能及印刷效果各不相同,岗位之间往往不可调配、不可兼容。

从4个企业设备升级的岗位对比发现,数字印刷升级打破了分色印刷的岗位刚性,呈现出更为灵活的趋势。D企业引入数字化设备后,30%的传统订单向数字订单转移,其制版方由外包制版公司转向企业内部跟单员;2名传统双色印刷岗员工转成数字印刷岗员工。Z企业数字印刷业务只针对外贸客户,招聘了1名数字印刷技术人员。X企业因水墨印刷部数字化设备更新比较彻底,传统制版及分色印刷岗位不愿意转岗的员工流向了其他传统企业,制版、印刷岗位都换成了数字印刷适配岗位人员。T企业50%的传统印刷业务向数字印刷业务转型,既有的分色印刷通过技能重塑转岗成为灵活岗位,其中1名由开槽工转岗而成(表3)。

表3　　4个企业数字设备升级后的技能迭代

企业	印前环节		印刷岗位	
	升级前	升级后	升级前	升级后
D企业	30%制版外包	跟单兼职制图	5个分色印刷	3个分色印刷、2个数字印刷
Z企业	3个制版	1个制图	6个分色印刷	增设1个数字印刷
X企业	12个设计制版	8个设计制图	8个分色印刷	3个灵活数字印刷
T企业	50%制版外包	数字印刷兼职制图	4个分色印刷	3个灵活数字印刷(含1名开槽转岗)

奥特尔(D. Autor)等人主张,有关技术升级影响工人技能的分析,应从技能偏好模型(受教育程度)转向任务偏好模型(Autor, Levy & Murnane, 2003: 1279-1333)。从任务偏好模型视角出发,数字设备升级后技能迭代呈现以下特征:

第一,从常规性任务向非常规性任务转型。传统印刷机械因制版成本

等因素,印刷品颜色越多、印刷版面越复杂,开版成本、制版成本、印刷成本越高,除高档化妆品、礼品包装袋等附加值较高的商品印刷外,多采用颜色单一、版面固定、单个订单数量较大的大批量、标准化生产模式以节约成本。里茨尔(G. Ritzer)将这种有限产品、标准化时间作业的生产范式称为“麦当劳化”(里茨尔,1999:29)。传统生产技能呈现出重复的、高度程序化的机械操作特征,更接近于常规性任务。数字印刷无须印版,通过计算机喷墨直接印制的方式,从图稿制作到成品印刷“所见即所得”,能够满足客户个性化的印刷需求。生产订单重复率低,单款数量少,且在图片设计上体现个性化、渐变、不规则、多色等特点,需要工人在操作上进行个别化、具体化、情境化处理,在技能上进行抽象理解、复杂思维等转向。因此,所需要的技能更多呈现为非常规性特征。

第二,从操作性任务向认知性任务转型。工业 4.0 背景下,工业生产越来越重视技术使用过程或系统知识,这源于生产过程日益一体化、网络化,以及数字技术使过程变得更加看不见和难以掌握。工人需要过渡技能,如编码、数据分析、人机交互和对信息技术的理解(Antonazzo, Stroud & Weinel et al., 2023: 25-46),以适应数字化生产。传统印刷生产依赖于对单一机械的简单操作,如分纸机操作、单色印刷机操作、双色印刷机操作、开槽机操作等;数字印刷兼容多色,且整合开槽等功能,并通过电脑集成设备将印刷技能程序化、参数化。一些新兴的智能工厂、数字化程度较高的车间引入联动生产线,将纸箱印刷生产从分纸到成箱进行了工序整合、联动生产,对工人的计算机原理、设备原理等认知性任务的要求大大提高。原有单一机械的操作技能被一体机的复合工序技能所取代,还需要计算机操作、集成设备等技术,提高了对工人计算机认知性任务的要求。

第三,对沟通、协作等软技能要求的提升。数字印刷实现了生产从“客户—制版公司—工厂”向“客户—工厂”的转型。传统印刷生产中的“客户—工厂”关系更多体现为企业被动执行生产任务,而数字印刷生产中工

厂则提供包装工艺设计、产品全周期一体化等服务。例如,X 企业提供了设计、包装、生产的一体化服务,Z 企业提供了部分设计、包装服务,D 企业、T 企业则主要依托配套公司提供服务。但无一例外,4 个企业都实现了从既有的外包组织关系转向组织内部关系,也即实现了从市场关系向组织内部科层关系的转化(周雪光,2003:28),以降低不稳定性、无法按时交付等因素的影响。数字化生产范式的柔性化、服务化趋向,将催生更多组织提供制图甚至设计等增值服务。这对员工通过技术加强与客户的沟通,与其他岗位员工协作以实现设计意图和达到服务水平,提出了更高要求。

数字化升级同时提升了对技术人员产品设计、工艺创新等技能素养的要求。例如,多数电脑、家电、手机等设备包装为保护内装物,提升防震、抗压、耐破性能,需要在进行纸箱外部包装之前加入保利龙泡沫固定。为节约材料,降低工艺成本,提高包装效率,提升环保指数,数字印刷使用瓦楞纸板新工艺包装设计,不仅能使防震抗压性能达到同等效果,且易装易取,用户使用友好,深受市场青睐,这一工艺创新是研发人员与技术工人的协作成果(韦康博,2016:245)。

六、结语

在本研究中,4 个企业是中国沿海千千万万个中小纸品包装印刷企业在数字化转型前期的缩影。在数字化设备升级的沉默裁员效应下,工人的生产节奏更快了,他们每天需要处理更多的订单,这对工作的复杂度、衔接度提出了更高要求,需要工人更加专注。虽然部分企业实现了扫码换单的简单操作,但印刷环节仍然是最重要、最关键的生产环节,需要工人认真审查、反复核对,更多呈现出经验性、重复性的操作特征。相关行业的招聘信息中,无一例外强调了对工人细心品质的要求。数字化设备升级的技能迭代效应,部分替代甚至完全淘汰了传统的制版及分色印刷技能。数字制造

通过计算机程序设计、联动一体化设计,将生产操作技能从印刷工人的隐性知识领域推向了计算机控制面板的显性知识领域,技能进一步从工人手中转向工程师、管理人员手中。如果说 20 世纪的机械化和劳动分工使工人的技能被切割成一道道工序,工人成了“局部人”,那么在数字化设备升级背景下,自动一体化、数字化机械通过整合工序,用计算机、集成设备进一步替代了工人的核心技能。数字化设备的更新与迭代,降低了核心技能岗位工人在工厂车间的地位及议价能力,对工人技能产生了毁灭性打击(诺布尔,2013:76)。

但技术变革不只有破坏性影响,同时具有生产效应或创造效应,熊比特(J. Schumpeter)称之为“创造性破坏”(熊彼特,2011:8)。从工人就业与技能发展来看,数字化增加了印前图稿处理和制图岗位需求,催生了产品设计、包装定制、需求解决、方案设计等服务需求,提高了对工人软技能的要求。随着人工智能技术的进一步发展,工人就业与技能发展势必受到更深刻的影响。因此,了解数字化、智能化对工人就业与技能发展的具体影响,一方面有助于制造业提前应对这一趋势,为人才培养提供重要的方向,另一方面有助于行业甚至整个社会从宏观层面做好相应准备,为我国制造业转型提供人才储备,为走向制造强国提供重要参考。

参考文献

布雷弗曼,1973,《劳动与垄断资本:20 世纪中劳动的退化》,方生译,北京:商务印书馆。

福特,2015,《机器人时代:技术、工作与经济的未来》,王吉美、牛筱萌译,北京:中信出版社。

郭琳,2022,《数字思维、数据知识生产与技术隐喻研究》,《南昌大学学报》(人文社会科学版)第 5 期。

李洋,2022,《产业数字化转型精要:方法与实践》,北京:人民邮电出版社。

里茨尔,1999,《社会的麦当劳化:对变化中的当代社会生活特征的研究》,顾建光译,上海:上海译文出版社。

马克思,2004,《资本论》,北京:人民出版社。

梅菲特,2018,《从 1 到 N:企业数字化生存指南》,沙莎译,上海:上海交通大学出版社。

诺布尔,2013,《生产力:工业自动化的社会史》,李风华译,北京:中国人民大学出版社。

腾讯研究院,2022,《2022 中小企业数字化转型发展报告》,https://www.sgpjbg.com/baogao/100014.html。

王天夫,2021,《数字时代的社会变迁与社会研究》,《中国社会科学》第 12 期。

王志星,2009,《中国纸包装行业发展历程》,《上海包装》第 1 期。

韦康博,2016,《国家大战略:从德国工业 4.0 到中国制造 2025》,北京:现代出版社。

肖旭、戚聿东,2019,《产业数字化转型的价值维度与理论逻辑》,《改革》第 8 期。

谢新水,2021,《数字技术对人的解析:社会功能与社会失范》,《浙江学刊》第 3 期。

熊彼特,2011,《经济发展理论:对于利润、资本、信贷、利息和经济周期的考察》,何畏、易家祥等译,北京:商务印书馆。

杨青峰,2018,《未来制造:人工智能与工业互联网驱动的制造范式革命》,北京:电子工业出版社。

中国电子技术标准化研究院,2022,《中小企业数字化转型分析报告(2021)》,https://www.xdyanbao.com/doc/j5uhnpa94q?bd_vid=9835979659845864268。

周雪光,2003,《组织社会学十讲》,北京:社会科学文献出版社。

Acemoglu, D. &D. Autor, 2011, "Skills, Tasks, and Technologies: Implications for Employment and Earnings", in D. Card & O. Ashenfelter eds., *Handbook of Labor Economics*, DOI: 10.1016/S0169-7218(11)02410-5.

Antonazzo, L, D. Stroud & M. Weinel, 2023, "Institutional Complementarities and Technological Transformation: IVET Responsiveness to Industry 4.0, Meeting Emerging Skill Needs in the European Steel industry." *Economic and Industrial Democracy* 44(1).

Autor, D., F. Levy & R. Murnane 2003, "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118(4).

Chen, D., S. Heyer & S. Ibbotson et al. 2015, "Direct Digital Manufacturing: Definition, Evolution, and Sustainability Implications." *Journal of Cleaner Production* 107(16).

Frank, M., D. Autor & J. Bessen et al. 2019 "Toward Understanding the Impact of Artificial Intelligence on Labo." *Proceedings of the National Academy of Sciences* 116(14).

Grugulis, I., E. Keep & C. Warhurst 2004, "The Skills that Matter: Teaching Interpersonal

and Intrapersonal Competencies in Any Classroom.” Oxford: Red Globe Press.

Heisig, U. 2009, “The Deskilling and Upskilling Debate.” in *International Handbook of Education for the Changing World of Work*, https://doi.org/10.1007/978-1-4020-5281-1_110.

Oztemel, E. & S. Gursev 2020, “Literature Review of Industry 4.0 and Related Technologies.” *Journal of Intelligent Manufacturing* 31(4).

Tinbergen, J. 1974, “Substitution of Graduate by Other Labour.” *International Review for Social Sciences* 27.

智能时代“微粒社会”的治理议题*

王仕军**

摘要：本文采用质性研究方法：首先，分析了智能时代的社会模式——“微粒社会”——的基本特点，包括数字化、监控化、解析化、评价化、单体化等；其次，探讨了“微粒社会”生成的技术逻辑、社会逻辑和文化逻辑；再次，剖析了“微粒社会”的三大革命（差异革命、智能革命和控制革命）及其引发的重大治理议题，如预防国家滑向“数字利维坦”、遏制平台巨头成为“大他者”、防止社会的两极分化、阻止智能机器风险失控、化解“微粒人”面临的难题等；最后，简要阐述了应对“微粒社会”治理议题的宏观思路。

关键词：智能时代　微粒社会　逻辑　治理

一、引言

人类演进的历史充分表明，重大科技创新每每担当推动社会变迁强大动力的角色，往往引致社会形态和社会模式的剧变，不以人的意志为转移地形塑人们的生产方式、生活方式、思维方式乃至国家社会的治理方式。概言之，会“导致社会和经济的范式转换”（扎卡达基斯，2017：295）。进入21世纪以来，在信息化基础上发展起来的智能化科技创新浪潮奔涌浩荡，互联网、物联网、大数据、云计算、人工智能等新兴技术蓬勃发展，以“计算无所

* 本文系江苏省社会科学基金项目“大学生创新创业胜任力测度评价的话语互动模型研究”（项目批准号：23JYD005）的阶段性研究成果。

** 王仕军，南京工程学院马克思主义学院。

不在、软件定义一切、数据驱动发展”(吴军,2016:VII-VIII)为主要特征的智能时代曙光初现。未来已来!面对智能化革命浪潮的冲击,我们所处的社会将发生怎样的改变?这样的社会有哪些突出的治理议题?这些治理议题该如何应对?德国著名社会学家库克里克(C. Kucklick)在其著作《微粒社会》中对上述问题进行了系统思考,提出了“微粒社会”的理论。他指出,随着数字化、智能化技术的发展,所有人都将迈入被高度解析的新型社会——“微粒社会”。我们正在经历的是一场新型的“解析—解体”,即技术的解析和一些社会制度的解体,并由此创造一个全新的世界,社会治理的“异步”困境、“治理赤字”等成为重大挑战,社会治理体系的涅槃重生蓄势待发。“微粒社会”理论为我们认识已经或即将到来的社会及其治理提供了崭新视角,成为我们在新的时代条件下推动国家和社会治理现代化进程中值得借鉴的理论资源。

本文从“微粒社会”理论视角切入,主要回答如下三个问题:其一,智能时代的“微粒社会”究竟是一种怎样的社会模式,其特性何在?其二,“微粒社会”是如何形成的,其内在逻辑是什么?其三,“微粒社会”的治理将面临哪些重要议题,其影响如何?最后,本文简要地就这些治理议题谈谈宏观应对思路。

二、“微粒社会”:智能时代的社会模式

“粒度”是计算机科学领域用来反映数据解析程度的一个概念,数据的解析程度越高,粒度就越低。随着社会的进步,网络、大数据、人工智能等技术的发展将推动我们所有人一步步进入精细解析的社会,库克里克将之称为“微粒社会”。在他看来,“这种新型的社会是高度解析的,而且不再关注平均值,因为它有了更好的东西:高密度的、更详细的认知。这会从根本上改变我们的生活方式、我们的世界观以及我们对自身的理解”(库克里克,

2018:VI)。在这样的社会中,我们熟悉的很多事物变得相当不同甚至面目全非。在此之前,我们所熟悉的社会可被称为“粗粒社会”。与“微粒社会”相较,“粗粒社会”好像“由无数的台球组成,随着时间的推移,我们慢慢学会将其建成一个可以负重的结构”;而在“微粒社会”中,“这些台球逐渐由微小的铅丸替代,这将急剧地改变社会的聚集状态和静力学结构,同时迫使我们寻找新的路径,以便在细小的微粒中建构一种稳定的秩序”(库克里克,2018:XIII)。从社会模式的角度看,“微粒社会”的特性至少包括如下几点:

1. 数字化社会。人类社会的演进伴随着数字化程度的不断提高。数字化社会是人的生存方式数字化的社会。1995 年,尼葛洛庞帝(N. Negroponte)在其影响巨大的名著《数字化生存》的前言中鲜明指出:“计算不再只和计算机有关,它决定我们的生存。”(尼葛洛庞帝,2016:61)他粗线条地描绘了数字化生存的社会图景,包括比特时代、个性化界面、数字化生活等,并分析了数字化生存的四大特征——分散权力、全球化、追求和谐和赋予权力。在这样的社会中,比特与原子共存;人们的生存空间由“物理—社会”两维空间扩展为“物理—社会—数字”三维空间;人们的生活有了线上与线下之分,“永远在线”成为必然,线上线下生活日益融合;众多数字化设备把人与人、人与物、物与物连接在一起,数据逐渐成为重要的甚至最为重要的生产要素;人的生活的方方面面,包括生产、消费、社交、娱乐等,都拖着或长或短的“信息尾巴”,都可在很大程度上以数字化的形式呈现出来。值得关注的是,“人的实体会越来越多地被数据化,数据从不同维度映射着人的‘虚拟实体’,数据也可以用来对个体的某个身体‘元件’进行描摹与复制,人的虚拟实体与现实个体越来越多地捆绑在一起”(彭兰,2019)。

2. 监控化社会。社会监控的历史源远流长。古代监控手段落后,总体属于感官型监控,监控的对象、监控的时间、监控的空间具有针对性、有限性等特点。随着科技的进步,现代社会的监控更多采用大数据、人工智能等先

进技术,通过层级的视频监控、数据监控和赛博格监控等方式,全面深入人们的日常生活,所有人都将遭受“深度”观察,监控呈现出全面性、隐蔽性、分散性、自动化、数字化、智能化等特点,甚至还能吸引被监控者主动配合。对此,中国社会科学院哲学所研究员段伟文一针见血地指出:“从人脸识别到智能手环,从行车记录仪到布满天际的电子眼,人们对世界的认识终于结出了世界时刻不知疲惫地关注我们的奇异果。”(段伟文,2020)现代监控之下,个体抑或群体都被简化成机器可读的数据,监控与惩戒的难度大大降低。需要特别指出的是,在现代社会,监控不再只是单纯的治安或规训的手段,更是一种商业模式或“致富”途径,在挖掘商业情报、预测用户行为和监控社会舆情方面发挥着重要作用。

3. 解析化社会。对人的数据解析可以追溯至被称为“政治算术”的早期统计学研究。在数字化和监控化的社会中,人们的身份、位置、行为甚至思想等信息都可以在很大程度上被一些机构甚至个人凭借数字化、智能化设备观察、搜索、跟踪、记录,并以数字化的形式被存储、挖掘和分析,从而“像社会之镜一般揭示和构建它们所表征的社会实在”。一方面,数据所扮演的角色如同 13 世纪出现的透镜,使人的行为得到量化的记录与透视;另一方面,数据分析与智能算法的应用正在使人的行为规律得到洞察和解析(段伟文,2019)。比如,脸谱网仅凭公开的点赞就能以 88%的准确率判断某个用户的性取向,以 75%的准确率推测用户是否购买过毒品。这种点击行为会像心理学中的标准测试一样,不多的点赞数量就足以推测出一个人性格的基本特点(库克里克,2018:91)。

4. 评价化社会。社会评价古已有之。以往社会对一个人的评价,在评价主体、评价客体、评价手段、评价节奏、评价结果等方面相对有限。在数字化、监控化、解析化的社会中,对一个人的评价呈现全面性、实时性、隐匿性、精准性和智能化等特点,所有人都将面对新的评价标准,这些标准比以往更加精确和无情。那些“从社交网站、银行汇款信息、信用卡消费记录、视频

摄像头等当中得来的数据,将会允许人们对某人的信誉等级、购买行为、职业道德以及更多的信息做出评价和分级”,“那些是数字化的评价,有时被称作分数,有时被称作画像”(库克里克,2018:109)。在这样的社会里,人们被笼罩在一张重叠而又矛盾的评价大网之中,评价渗透到人们生活的各方面和全过程,每一个人除非与各种网络隔离,否则就可能被各种机构、智能设备等通过多种算法进行评价打分甚至精准画像而无处可逃。

5. 单体化社会。在“粗粒社会”中,尽管人们可以感知某个人行为的独特,但并不能准确地命名这种“独特”,并将其测量归类;对人们行为的研究主要通过各种社会调查、统计学分析等手段进行;每个人被当作“个体”看待。“个体”是“问卷和民意调查中的人,是统计中间值和平均值代表的人”(库克里克,2018:7),每个人的独特性被性别、阶层、学历等粗略的分类遮蔽起来,因此对人的行为分析是相对模糊的。“微粒社会”凭借数字化、智能化技术对人类生活方方面面大数据痕迹的记录、采集、评价和分析,将每个人的独特性都清晰地显露出来,从而使每个人都成为不能混淆的“单体”,并且每个人还可以通过量化过程进行自证。

三、“微粒社会”形成的内在逻辑

根据技术社会学的基本原理,某一社会模式的形成往往是技术、社会、文化等因素互动耦合的产物,具有内在的逻辑链条。“微粒社会”的形成亦不例外,其形成的内在逻辑包括但不限于如下三个层面。

(一) 技术逻辑

1. 从“图灵革命”到计算机发展。弗洛里迪(L. Floridi)在《第四次革命:人工智能如何重塑人类现实》中指出,科学的发展推动人类自我认知发生了四次重大革命,即“哥白尼、达尔文、弗洛伊德和图灵”革命:“哥白尼革

命”使人类居住的地球失去了宇宙中心的荣耀;“达尔文革命”证明了人类并不是万物之灵;“弗洛伊德革命”让人们承认人类的意识不是纯粹和清晰的,不能做到对意识的绝对控制;“图灵革命”使人们认识到人类在逻辑推理、信息处理和智能行为领域的主导地位已经动摇(弗洛里迪,2016:85—87)。1936 年,图灵提出了“图灵机”的计算模型,用纸带式机器模拟数学运算的过程;1946 年,世界上第一台现代意义上的电子计算机诞生;1950 年 10 月,图灵发表了一篇题为“机器能思考吗?”的论文,提出了划时代的“图灵测试”设想和通用计算机的概念。此后的半个多世纪时间里,计算机技术沿着“电子管计算机—晶体管计算机—集成电路计算机—大规模集成电路计算机—智能计算机—生物计算机”的发展脉络,在硬件和软件两个方面迅猛进步、加速迭代,历经“主机型阶段—个人电脑阶段—普适计算阶段”等三大发展阶段,最终计算机与环境融为一体,其应用已渗透到社会的各个领域,正在改变着传统的工作、学习和生活方式,成为当今世界支撑性的通用技术和极为关键的基础设施性技术。

2. 从计算机连接到泛在网络。计算机如果不相互连接,其作用就会大大受限。伴随着计算机技术的发展,计算机网络技术亦快速向前推进。它将人类进行深度捆绑,并将人类生活的物理空间、社会空间、数字空间紧密连接,从而在实体社会之外建构了一个被称为“虚拟社会”的网络社会,其影响已远超信息技术的范畴而被称为“附魅的网络”(enchanted network)(欧阳英,2020)。在互联网技术不断进步的基础上,信息通信领域(ICT)的一系列技术进展如传感器网络、射频识别技术(RFID)、二维码技术、短距离无线通信技术、移动通信模块技术、地理信息系统(GIS)、全球定位系统(GPS)和遥感测绘技术(RS)等也推动着物联网(IOT)的蓬勃发展,从而实现广域或大范围的人与物、物与物之间的信息交换。“互联网+物联网”的叠加进步,使得“泛在网络”(ubiquitous network)由概念变成社会现实。“泛在网络”通过“嵌入式动态元件匹配融合既有的计算机网络、大众媒体

网络、社会服务网络，打造出一个全球的、泛在的、无缝的信息科技平台，从而实时地连接万亿数量级的具有通信地址的实体（包括人、物），并感知周边环境的变化”（刘永谋、吴林海、叶美兰，2012），整个世界由此呈现出“超级互联”状态。

3. 从小数据到大数据。数据是人类建造文明的基石，但其作用在很长的一段历史时期内并未引起人们的足够重视。根本原因在于记录、存储、分析数据的技术或工具不够多、不够好，由此导致：一方面，数据量不足，积累数据需要的时间太长，短期内数据的作用不明显；另一方面，数据和所想获得信息之间的联系通常是间接的，两者的相关性不能被很好地揭示出来（吴军，2016：19）。于是，在人的“数字化生存”模式到来之前，我们的制度、处理过程和激励机制中要尽可能减少数据的使用，大多采用统计学上随机抽样的方法，用最少的数据获得最多的信息，属于小数据模式。大数据是与小数据相对而言的，指的是大量的、多种类的、复杂的、分散的数据集，这些数据集来源于各种传感器、电子交易、电子邮件、视频、点击流等。随着信息通信技术的巨大进步，特别是以云计算为基础的信息存储、分享和挖掘手段的发展，我们可以廉价高效地将这些大量、高速、多变的终端数据存储下来并随时进行分析和计算，人类社会的大数据时代随时开启，数据的“透镜”作用进一步凸显。“人类第一次有机会和条件，在非常多的领域和非常深入的层次获得和使用全面数据、完整数据和系统数据，深入探索现实世界的规律，获取过去不可能获取的知识，得到过去无法企及的商机。”（迈尔-舍恩伯格、库克耶，2013：V）

4. 从人工智能 1.0 到人工智能 N.0。人工智能历经 60 多年的艰辛探索，终于在 21 世纪的第一个十年与第二个十年的交替之际迎来了真正爆发式增长的黄金时期。当前和今后一段时期内的人工智能被称为“新一代人工智能”或“人工智能 2.0”，从而区别于传统的人工智能或人工智能 1.0。人工智能 1.0 是符号人工智能，计算机按人设定的程序运行；人工智能 2.0

最突出的特点就是能够在一定程度上自主学习。人工智能 2.0 仍属于弱人工智能,依靠暴力计算和海量数据,往往是“有多少人工,就有多少智能”。未来,人工智能将沿着“弱人工智能—强人工智能—超人工智能”的路径发展演化,由人工智能 2.0 向人工智能 N.0 迈进,多样的人工智能体将走进我们的生活,人类将在相当程度上实现脑力的外包,并由此进入“第二次机器革命”时代(王仕军、赵利群、胡志彬,2020)。

(二)社会逻辑

1. 消费社会到来。20 世纪 30 年代的生产过剩危机重创了资本主义世界,引发了全球剧烈动荡。二战后在科技进步、科学管理、资本运营革新等因素的推动下,西方各国的社会生产力迅猛发展,资本主义迎来了物资丰盛的时代,并由生产社会进入消费社会。20 世纪 70 年代,西方发达国家再一次遭受经济危机的侵袭。鲍德里亚(J. Baudrillard)在《消费社会》一书中最先提出了“消费社会”的概念。在他看来,消费社会是以消费为本位的社会,消费不仅构成整个经济生活的主导性逻辑,而且成为当代世界的根本特征。在消费社会中,由于“消费作为一种建立关系的主动模式、一种符号的系统化操控活动,主要是由媒介通过对符号进行编码来影响大众而实现其消费企图的”(徐琴,2009),因此作为经济主体的厂商受内在盈利动机的驱使和外在竞争的压力,对信息通信技术特别是媒介技术表现出热切的期盼。一方面,了解并顺应消费者的需求;另一方面,刺激和塑造消费者的各种欲望,强化或诱导消费者的需求。还有一点,就是通过信息通信技术,优化内部管理,降低生产成本。同时,政府部门和军队出于提高效率的目的,也对信息通信技术寄予厚望。总体来看,在这样的社会,扩大消费与信息通信技术进步之间形成了互为表里、双向强化的关系。这促进了信息通信技术特别是数字技术的迅猛发展。

2. 金融资本扩张。消费社会的到来使得金融资本相对于生产资本的

地位大大提升,金融资本有扩张的强烈冲动。金融资本的扩张一方面对计算机、互联网、大数据、人工智能等技术的发展提出强大需求,另一方面又能够为这些技术的迭代提供强大助力。以美国为例,自 20 世纪 80 年代始,金融资本扩张三大趋势交织的样态日渐明朗,即经济金融化、金融科技化、金融全球化。这三大趋势都对信息通信技术的发展提出了更新、更快、更安全、更智能的要求,也为计算机、大数据、人工智能等技术的应用准备了更广阔的现实情境。进入 21 世纪,以新一代人工智能为标志的新一轮科技和产业革命将对金融行业产生颠覆性冲击,大数据金融、人工智能金融、区块链金融和量化金融等金融科技正在传统金融行业的各个领域积极布局,已然成为新的风口(尹志超,2018)。

3. 数据资本凸显。随着人的数字化生存的拓展,特别是在大数据、人工智能等技术进步的加持之下,数据洪流快速席卷我们生活的整个世界,也融入了资本的运营之中,数据资本日渐凸显出来。数据在资本运营链条中可以划分为四个类别:原初数据、采集数据、资本数据与数据资本(田锋、吕金伟,2020)。原初数据是每个个体数字化生存产生的各种痕迹的数字表现;采集数据是被互联网公司等机构运用多种数字化、智能化设备收集的对资本运作有价值的原初数据的集合;资本数据是互联网公司等机构通过数据工程师的劳动对采集数据进行挖掘、萃取而成的数据集;数据资本是互联网公司等机构将资本数据作为资产用于生产、销售与转让的一种新型资本,从而提高公司的生产效率和盈利率。数据资本正成为产业资本、金融资本之后的第三种起支配性作用的资本样态,并在今天占据着产业资本、金融资本和数字资本构成的金字塔结构的塔尖位置(蓝江,2018),是互联网公司巨头强大盈利能力的关键支撑。正如陈本皓所指出的:“数据的资本价值在体验上遇到了无法跨越的鸿沟,一般用户只能将其当成是需要防御和保护的对象,只有在看到互联网巨头惊人的盈利数字时,人们才会意识到当数据作为资本时能够产生何等惊人的价值。”(陈本皓,2020)

4. 平台公司崛起。加拿大学者斯尔尼塞克(N. Srnicek)在《平台资本主义》一书中深入剖析了平台公司的崛起。2008 年金融危机以来,在以美国为代表的发达经济体中,一种强大的企业新模式——平台——崛起,主要包括以谷歌为代表的广告平台、以亚马逊为代表的云平台、以通用电气为代表的工业平台和以优步为代表的精益平台等几大类型。平台起初由处理数据的内部需求产生,随着其能独占、提取、分析和使用数据量的日益增长以及数据挖掘能力的增长,其在整体经济体系中的统摄作用日益彰显,大量公司被纳入平台之中。作为一种新型公司,其特征在于"提供基础设施,以调解不同用户组,显示垄断倾向,使用交叉补贴来吸引不同的用户组以及利用设计好的核心架构,控制交互的可能性"(斯尔尼塞克,2018:49—55)。这些特征使得平台成为提取和控制数据的关键业务模型,数据资本成为平台公司的核心资本。由此,我们不难看出数据资本和平台公司是互为表里、相互支撑的。

5. 新型社会物理学滥觞。在库克里克看来,"利用新的技术手段,我们不仅能够解释自己的身体,还能解释我们的语言、社会和历史"(库克里克,2018:35)。早在 19 世纪,孔德就曾提出过"社会物理学"的概念,他认为可以用物理学的方法来研究人类社会。当今时代,以大数据、人工智能为代表的新一轮科技革命将推动社会科学研究的范式转换,从而更深入地解析我们的社会。尽管"社会物理学"一词目前仍在沿用,但是在新的技术手段加持之下,研究范式已发生重大转变。比如,以彭特兰(A. Pentland)为代表的"现实挖掘"研究,提出了可穿戴、实时、全体、社会测量、流动性、想法流等新概念,采用了社会计量标牌等新型研究工具,因此被称为"新型社会物理学"。凌昀等学者认为,当今时代和技术背景下,新型社会物理学和计算社会科学在含义上是基本相同的,都是通过大数据技术和智能设备等进行社会模拟、建模、网络分析和媒体分析,实现研究方法的革新和研究对象的拓展等(凌昀、李伦,2020)。

（三）文化逻辑

1. 数据主义。在数字化生存的大背景下，人类的活动与数据技术紧密勾连，处在数据技术的无缝覆盖之中。一方面，它对人类社会产生革命性、颠覆性影响；另一方面，它也催生了一种新型社会文化，即“对数据技术无限崇拜，以为数据技术可以解决一切社会人生问题”，这又被称为“数据主义”（高兆明，2018）。从本体论的角度看，数据主义将数据视作世界的新基质，将世界等同于数据。从价值论的角度看，数据主义把最多的数据和最透明的数据认作最大的善。在这样的社会文化氛围中，人们“对人类知识和智能有所怀疑，而倾向于信任大数据和计算机算法”（赫拉利，2017）。人类的数据化过程无疑将大为加速，而这反过来又会进一步扩大数据主义的影响。

2. 安保主义。当今社会，监控之网越织越密。一方面，这与传感器、物联网、人工智能等监控技术不断进步有关；另一方面，又与被称为“安保主义”的安全文化的影响扩展密不可分。杨子飞认为，监控塑造了监控不断扩张的安全文化理念，即安保主义：一方面，它将不安感无限夸大，刺激人们对绝对安全的狂热追求；另一方面，它将透明看作安全的关键要件，推动监控的不断扩张。尽管这一理念在一定程度上回应和缓解了当代人的安全关切及安全焦虑，但也将安全建立在了监控之上，无疑会推动监控在时间和空间的各个维度上持续开疆拓土甚至滥用，“西西弗斯式”的悲剧难以避免（杨子飞，2017）。

3. 赌场隐喻。隐喻是人类的一种认知方式，是用一种事物暗喻另一种事物。在库克里克看来，对我们生活的现代社会而言，赌场是最好的象征。现代赌场的目的是依靠先进的数字化、智能化设备提升赌客的“生产效率”，赌场里的一切都围绕这个目的而进行详细的计划和控制。赌客看似自由不羁，但无不经受“被设计的上瘾”。在那里，“人的可预测性、对人的

持续监视、对人的行为的预测以及人与机器的紧密合作”得到充分展示,“我们将不再受强迫,而是被引诱”,“我们的身体、行为、感觉及资产的流动均处于持续不断的观察之下,但是将不再受指令的支配,而是被调整”(库克里克,2018:94—102)。这种情形将在数据技术发挥作用的地方不断上演,赌场几乎成了社会的镜像。

四、“微粒社会”的三大革命及其引发的治理议题

库克里克认为,“微粒社会”与“粗粒社会”相较将经历三大革命,即差异革命、智能革命和控制革命,这三大革命的表现形式多样复杂且相互交织,并由此引发一系列特别值得高度关注和深入探讨的重大治理议题。对这些议题的认知和应对直接关涉人类的整体福祉和前途命运。

(一) 三大革命及其表现形式

1. 差异革命。在“微粒社会”中,泛在网络把人们紧密地连接起来,众多数字化和智能化仪器或设备如“社会测尺”“会议调解器”等将人们之间的“不同”清晰地揭示出来,每个人都只能把自己作为极端差异化的单体加以理解,结果是极端的差异与极端的连接并存,此之谓“差异革命”。主要表现在:其一,微粒化的工作。在工作世界中,参与者的各种观察始终存在,但大多带有某种主观性,“社会测尺”等仪器带着“上帝视角”闯入之后,团队之间、公司之间、雇员之间的差异都会被无情地、坚定地显示出来。它们以微粒化的方式让每一个雇员都意识到自己的“独一无二”。由于我们往往认同仪器观察数据的客观性,我们的职场行为不得不因此而改变。其二,微粒化的民主。2012 年,奥巴马再次参选美国总统,采用了一种极具创造性的制胜手段赢得了选举,并让全世界领略了微粒化民主的大致图景。奥巴马组建了自己的数据团队,包括 300 名统计学家、编程人员和数据分析

员。在选举中,数据团队建立了庞大的数据库,囊括了 1.66 亿选民的相关信息,每个选民的数据都有 10 000—20 000 条,涉及政治倾向、收入水平、社会交往等细节数据。在此基础上,团队还每天跟大约 30 000 名选民进行电话沟通,并将沟通结果及时传回数据库。团队的数据分析员所关注的不再是某个选举群体中的个体,而是每一个选民的独特性,选民被单体化了。竞选团队与数据团队紧密配合,不断从单体化的选民那里得到反馈,并把选民朋友圈中的私人交流信息用作增强竞选优势的武器。这次选举相当程度上改变了民主的运作方式,奥巴马也因此被称为“第一任数字化总统”。其三,微粒化的身体。随着智能手机、可穿戴设备的普及,“量化自我”运动兴起,相关 APP 产品也层出不穷,我们的身体也被微粒化了。人们每天可以凭借上述技术和设备对自己的身体进行测量、模仿、刺激、检测,并以比以往更深入、更精准、更温和的方式进行管理,还可以将有关数据通过网络进行共享。

2. 智能革命。进入 21 世纪,得益于并行计算、更大量的数据、更深层次的算法这场完美风暴,新一代人工智能应运而生并奠定了持续进步的坚实基础,大量智能机器涌入我们的生活,此之谓“智能革命”。主要表现在:其一,机器像人。智能手机、无人驾驶、智能家居、语音搜索等人性化产品和服务深入人心,翻译、聊天、阅读甚至文学创作等原来只有人自身才能完成的工作,人工智能于此皆有惊艳表现,以至于不少人会不自觉地把智能机器看成类人的实体,赋予智能机器以某种拟主体性。更有甚者,2017 年沙特赋予机器人“索菲娅”以公民身份,让其成为沙特居民。其二,人像机器。随着智能机器的大量涌现,人依靠机器测量自己、解释自己、拓展自己、提高自己,人的灵活性、创造性等反而可能降低,从而呈现出以往机器的一些特性,如程序化、机械化等,人正不断地受到并适应机器的“规训”而成为它的“附件”(余聪、张亮,2021)。其三,人机共生。与功能日益强大的智能机器相比,人类的剩余优势空间日趋逼仄,主要集中在社会知觉、谈判艺术、论证

能力以及创造性、变通性等方面。智能机器的涌现已经并将使就业市场发生结构性的改变,计算机利用在算法上可分解的任务承担那些定义明确的任务,而人类将承担那些模糊、难以定义且复杂的任务(库克里克,2018:64)。

3. 控制革命。德勒兹(G. Deleuze)在“规训社会”概念的基础上根据西方社会的发展,最早提出了“控制社会”的概念。福柯认为,从 18 世纪到 20 世纪末,具有惩戒理性形式的社会可被称为“规训社会”。在那里,人们生活并辗转在一个个相对封闭的规训机构中,从家庭到学校再到医院、军队、监狱等,时刻受到监视,行为被严格规训。与“规训社会”不同的是,“控制社会”不再强调规训机构对民众的纪律控制,而是更重视对民众持续的、广域覆盖的监视,“影响我们生活的不再是守则规定,而是由观察、监视、预测、评价、引诱和劝告所组成的一个多面的复合体”(库克里克,2018:93),此之谓“控制革命”。主要表现在:其一,“在机器世界欲罢不能”。大量智能监控设备成为人们生活的重要背景,甚至直接进入人们的身体之中。这些设备界面友好,甚至看不见界面,机器发挥其计算上的优势,让人们的生活自由便捷,幻想着操控一切的力量,同时在不知不觉中被机器观察、评价和诱导。其二,预测机器。控制革命的目标不仅在于控制现在,更在于控制未来。智能化的机器在大数据输入的基础上,不断优化程序算法,充分发挥其相关性挖掘的优势,对人们的行为进行预测。尽管并不是所有的预测都会精确,但其为有关的机构如政府部门、平台公司等提供了神奇的简化效果,大大降低了决策的复杂性。

(二) 三大革命引发的重大治理议题

1. 预防国家滑向“数字利维坦”。一方面,人们依靠国家提供安全和保护;另一方面,人们又要防止国家变成“吃人的怪兽”,“国家利维坦”的说法由此流传。“数字利维坦”由“国家利维坦”引申而来。在“微粒社会”中,在

三大革命的推动下,国家可能具有超强的解析和控制社会的能力,更深入地渗入社会的各方面,而单体化的大众难以进行有效抵抗。原因至少有四:其一,技术鸿沟。政府自身或与科技巨头合作,在拥有数据、程序算法、数据挖掘、社会行为解析等方面往往具有强大的技术优势。政府可以对公众进行深度解析、评价和预测,而民众不能同等地回应政府。其二,财力鸿沟。政府能够动用的社会财富甚巨,民众由于“搭便车”等问题的存在,难以聚集起足以同政府抗衡的财力。其三,暴力鸿沟。政府可以以公共安全为名义,合法地运用国家财力和先进技术不断升级自己的暴力水平,民众在合法性、财力、技术等方面皆与政府存在巨大差距,其暴力水平提升缓慢。其四,公众难以紧密团结。一方面,差异革命将民众的细微差距清晰揭示,打碎了使公众团结起来的“无知之幕”,人们因缺乏同类和联合者而陷入斗争的泥潭之中;另一方面,政府在对民众深度解析的基础上差别对待不同的群体或个体,可以有效降低公众团结抵抗的能力。“数字利维坦”的形成可能蕴含巨大风险,包括:威胁公民的隐私信息和合法权益;禁锢人类的自由意志和道德选择;导致数字威权和数据独裁;成为极端主义的温床;加剧社会碎片化的风险;等等(唐皇凤,2018)。

2. 遏制平台巨头成为“大他者”。数字经济是“微粒社会”时代的经济模式,平台公司是数字经济的基本组织形式,也是引领、控制行业的手段。一方面,平台提供了其他行业运行的基本格局;另一方面,平台引导或迫使新行业从产品生产转向服务提供。在发展过程中,平台公司具有强烈的垄断趋势,原因包括:其一,网络效应,平台用户越多,平台价值越高,而平台价值越高,平台用户就越多;其二,平台的早期优势会成为巩固平台行业领袖地位的利器;其三,平台具有连接并整合多个网络的特殊能力;其四,平台获取大量数据,使得预测更加准确,更加刺激了平台内数据的集中;其五,平台大量促进互补产品开发;其六,平台设法建立起与竞争对手相隔离的产品和服务生态系统。上述因素耦合在一起,使得一些平台企业成为垄断组织或

平台巨头,并对越来越多的用户及其产生的数据进行集中控制(斯尔尼塞克,2018:106—107)。平台巨头的不断发展,在缺乏有效监管的情况下,极有可能形成一个威胁社会整体的“大他者”(the Big Other)。“大他者”是美国学者祖博夫(S. Zuboff)提出的概念,在她看来,“大他者”与人的数字化生存、数据的资本属性是紧密相关的。随着互联网的发展、人类数字化生存的深化,人类活动会产生大量行为剩余——数据,最初这些数据被一些网络科技公司用作改进服务的资源,在新一代人工智能技术的推动下,这些数据第二次、第三次利用的价值被厂商发现,数据的资本属性凸显出来。美国一些数字平台巨头在“网络上监控并挖掘个人所留下的信息,并将其用于独家分析及算法产品,最终通过宣扬广告精准发放用买家竞价的模式将这些数据售出,整个过程构建立了一种监视型产业的模板”(陈本皓,2020),这一模板扩散开来,形成了一种通过占有人们的行为剩余而不断扩张的新型资本增值模式——监控资本主义。这些数字平台巨头的迅猛发展,逐步形成了一种全面管理个人和社会的极端能力——“大他者”。祖博夫指认:“这种力量不仅使美国的自由民主之梦化为泡影,而且使整个社会都被绑架在资本逐利及其政治等相关领域的混乱无序中。”(高斯扬、程恩富:2020)

3. 防止社会的两极分化。近 30 年来,经济不平等的幽灵在世界飘荡,“富者愈富,穷者愈穷”的两极分化趋势日渐清晰。《21 世纪资本论》的热销在相当程度上反映了世人对两极分化的不满和忧虑,因为“经济不平等的社会更容易出现政治分化、制度失灵、垄断行为、腐败以及政府由富人主导而不是由人民治理的情况”,民主国家“转向财阀统治,甚至盗贼统治”(阿拉赛维奇、索奇,2018:29)的风险剧增。在“微粒社会”背景下,三大革命相互交织,可能使得两极分化的问题更为突出和严重,创造出历史上最为不平等的社会。其一,微粒化工作的背景下不仅会不平等地分配工作,而且会单体化地发放报酬,职场中劳动者的收入差距迅速拉大。其二,智能机器

将大量普通劳动者驱逐出工作领域,拉大了工作者与失业者的收入差距。其三,“赢者通吃”逻辑让各行各业的“超级明星”赢得了巨额财富。数字化市场往往遵循“赢者通吃”逻辑,“最好的或者最幸运的某个市场细分的供应商几乎将所有顾客都吸纳走了,而稍次一点的供应商获得只是残羹剩饭”,那些具有绝佳创意、出色产品以及特别才艺的人和企业,那些只多做了一点努力(或只是因为幸运)的人,将从数字化网络和全球市场获得呈指数级增长的收益而迅速暴富(库克里克,2018:55)。其四,平台巨头依靠垄断地位分割了大量社会财富。对于未来社会的两极分化问题,赫拉利(Y. Harari)在《今日简史》中不无忧虑地指出:“所有财富和权力集中在一小群精英手中。大多数的人类的痛苦将不再是受到剥削,而是更糟的局面:再也无足轻重。”(赫拉利,2018:67)

4. 阻止智能机器风险失控。随着人工智能技术的发展,大量的智能机器进入人类生产、生活以及国家和社会治理等各个领域。一方面,智能机器带来了生产效率、治理效率的提高以及生活的便捷,具有“天使”的面相;另一方面,不断蔓延的智能机器风险,凸显“魔鬼”的面相。其一,技术基础设施风险。在“微粒社会”中,技术基础设施变得前所未有的复杂且紧密耦合,很多设备由智能化的附属系统控制,彼此之间还存在相互反馈。一方面,小的疏忽有可能通过连锁反应造成大的事故或系统性事故;另一方面,容易成为一些别有用心的极端分子觊觎的目标。其二,智能机器决策风险。当我们越来越多地把一些重要的决定交由智能机器或在智能机器的辅助下做出的时候,智能机器决策的风险便潜入我们的社会生活之中。智能机器决策依靠数据和算法:从数据看,可能存在数据偏见、数据污染、数据残缺、数据造假等问题;从算法看,可能存在有意或无意的算法偏见、算法垄断、算法出错等问题。这两个方面诸多问题的叠加,推动了智能机器决策风险的生成与扩散。其三,智能机器犯罪风险。伴随人工智能技术的广泛应用,恶意利用人工智能技术的优势或缺陷实施犯罪的问题也日益突出,主要表现

为:人机区分难题所引发的新型网络分离主义;能力延展技术所催生的机器支援型犯罪,如利用深度伪造技术实施诈骗犯罪;基于机器学习模型的“黑箱”操作;等等(商瀑,2020)。其四,智能武器风险。智能武器具有智能、自动、精准、杀伤性强等特点,美国等国家不断加大智能武器的研发力度,新的军备竞赛引发世人忧虑(雷鸿竹、曾志敏、熊帅,2019)。其五,智能机器替人风险。人类大量的工作将由智能机器承担,蓝领工人的工作可能被取代,白领工人的工作也可能被取代,甚至大师级人物的工作也可能被取代。赫拉利指出,智能机器凭借连接性和可更新性等两类非人类能力,在体力和智力两个方面远超人类。智能机器的飞速迭代,很快就会让数十亿人失业,并创造出一个人数众多的新“无用”阶级,带来现有意识形态无法应对的问题(赫拉利,2018:17—19)。

5. 化解“微粒人”面临的难题。在“微粒社会”,人类的领地越来越受制于数字化、智能化等技术,整个社会呈现出深度科技化状态,“微粒人”面临一系列的难题。其一,矛盾的人。人的独特性需要在网络中才能显现,可谓“不连接无个性”;人们无法看穿机器,机器却可以看穿人们;人们渴望确定,但世界处处充满不确定;人们看似自由,却时时被困在各种牢笼之中。其二,透明的人。隐私的幕布被各种机构甚至个人凭借数字化、智能化监控、解析、评价技术撕碎,越是紧密地连接到各种网络,人的透明度就越高。其三,受压的人。每个人“都会处于极端的压力之下,将自己预设为灵活的、可变的并且可以自我优化的”(库克里克,2018:9)。压力来自职场的激烈竞争(人与人竞争、人与机器竞争)、社会无所不在的监控、智能机器的不断迭代等,人们要忍受智能机器的精确观察,要忍受各种各样的评分,要忍受智能机器对人们行为的矫正,甚至要忍受智能机器的不公正对待,如“大数据杀熟”等。其四,孤独的人。差异革命将每个人都变得独一无二,人们将失去在群体中的依靠;同时,人们将更多的时间花在“面对屏幕”而不是“面对面”的世界,科技让人们与身体、感官和真实环境越来越疏离,结果就

是人们感到越发孤独。其五,退化的人。智能革命使得人机共生成为生活常态,智能机器功能日益强大,人们将身体能力和认知能力越来越多地外包给智能机器。根据用进废退的原理,人们的一些能力如创造力和感知力等将会退化,增强人类的“自然愚蠢”,“最坏的局面就会是退化的人类滥用进化的计算机,伤害自己,也伤害世界”(赫拉利,2018:67)。

五、结论与讨论

在智能时代,作为社会模式的“微粒社会”,具有数字化、监控化、解析化、评价化、单体化等特点。“微粒社会”的形成,不是人为设计的结果,而是技术、社会、文化等因素互动耦合的产物,有其技术逻辑、社会逻辑和文化逻辑等复杂生成机理。这一社会将经历差异革命、智能革命和控制革命这三大革命,并由此引发一系列重大治理议题。其中,有全新的治理议题,如防止智能机器风险失控;也有旧治理议题的新表现,如社会两极分化的新形态。应对这些治理议题,把“微粒社会”变成“美丽新世界”是全人类的共同期盼。

“技术不是命运,命运由我们自己塑造。”(布莱恩约弗森、麦卡菲,2014:291)智能时代,“微粒社会”大幕开启,上述治理议题已经并将深度展开。如何应对这些治理议题,笔者认为,在如下三个层面下功夫,至为关键:

一是推动哲学革命。“接下来的几十年可能是哲学的繁荣时期,因为我们需要做出许多充满挑战和困扰的伦理选择。”(沃尔什,2020:181)哲学是时代精神的精华,既涉及“怎么看”的问题,也关系“怎么干”的问题。“粗粒社会”的哲学可能不再适应“微粒社会”,需要进行创造性转化,甚至需要全新的哲学以指导人们处理“微粒人”与自然、技术(机器)、社会、国家的关系,从而为应对上述治理议题提供世界观和方法论方面的基本遵循,而这将考验人类的智慧。

二是把人确立为立足点和归宿。人是最高目的,“微粒社会”不应成为只剩冷冰冰的数字和机器的世界。尽管智能技术依然需要大力发展,但应对上述治理议题必须把活生生的、有温度的人作为立足点和归宿,人性温度必须统摄技术精度和技术力度,价值理性必须统摄工具理性和技术理性,应当始终“怀有谦恭之心,铭记人性之本”(迈尔-舍恩伯格、库克耶,2013:247)。

三是重塑社会治理体系。在“微粒社会”,我们既面临技术的全面深度解析,也面临旧有社会(国家)治理体系的调整。实现治理体系的重塑乃大势所趋,重点在于打造三个牢固的“铁笼”,而且每个“铁笼”都有民众团结、法律制度、道德伦理、科学技术四大支柱:首先,打造制约公共权力滥用的“铁笼”,划清公共权力和民众权力的边界;其次,打造制约平台巨头无限扩张的“铁笼”,采取把私人平台变成公共平台、反垄断、数字税收等措施,防止其变成“大他者”;最后,打造制约科技失控的“铁笼”,如为机器嵌入道德、情感等,使其服务人,而不是压制人、奴役人、伤害人。在此基础上,还需要注重保护广大民众特别是普通劳动者的权益,如建立基本收入制度、完善技能教育等,防止社会撕裂,保证社会团结。

参考文献

阿拉塞维奇、索奇,2018,《不平等简史》,罗海蓉、智艳译,上海:上海社会科学院出版社。

布莱恩约弗森、麦卡菲,2014,《第二次机器革命》,蒋永军译,北京:中信出版社。

陈本皓,2020,《大数据与监视型资本主义》,《开放时代》第 1 期。

段伟文,2019,《人工智能与解析社会的来临》,《科学与社会》第 1 期。

段伟文,2020,《面向智能解析社会的伦理校准》,《上海交通大学学报》(哲学社会科学版)第 4 期。

弗洛里迪,2016,《第四次革命:人工智能如何重塑人类现实》,王文革译,杭州:浙江人民出版社。

高斯扬、程恩富,2020,《监控资本主义视阈下的技术权力探析》,《内蒙古社会科学》第 4 期。

高兆明,2018,《“数据主义”的人文批判》,《江苏社会科学》第 4 期。

赫拉利,《今日简史:人类命运大议题》,林俊宏译,北京:中信出版社。

赫拉利,2017,《未来简史:从智人到神人》,林俊宏译,北京:中信出版社。

库克里克,2018,《微粒社会》,黄昆、夏柯译,北京:中信出版社。

蓝江,2018,《数字资本、一般数据与数字异化——数字资本的政治经济学批判导引》,《华中科技大学学报》(社会科学版)第 4 期。

雷鸿竹、曾志敏、熊帅,2019,《人工智能武器的全球发展、治理风险及对中国的启示》,《电子政务》第 11 期。

凌昀、李伦,2020,《计算社会科学研究:范式转换与伦理问题》,《江汉论坛》第 9 期。

刘永谋、吴林海、叶美兰,2012,《物联网、泛在网与泛在社会》,《中国特色社会主义研究》第 6 期。

迈尔-舍恩伯格、库克耶,2013,《大数据时代:生活、工作与思维的大变革》,盛杨燕、周涛译,杭州:浙江人民出版社。

尼葛洛庞帝,2016,《数字化生存》,胡泳、范海燕译,北京:电子工业出版社。

欧阳英,2020,《唯物史观视阈中的数字资本主义》,《国外理论动态》第 3 期。

彭兰,2019,《智能时代人的数字化生存——可分离的“虚拟实体”“数字化元件”与不会消失的“具身性”》,《新闻记者》第 12 期。

商瀑,2020,《从“智人”到“恶人”:机器风险与应对策略——来自阿西洛马人工智能原则的启示》,《电子政务》第 12 期。

斯尔尼塞克,2018,《平台资本主义》,程水英译,广州:广东人民出版社。

唐皇凤,2018,《数字利维坦的内在风险与数据治理》,《探索与争鸣》第 5 期。

田锋、吕金伟,2020,《西方资本主义的时弊:从数据资本到数据剥削》,《自然辩证法研究》第 8 期。

王仕军、赵利群、胡志彬,2020,《新一代人工智能发展与中国应对:基于“第二次机器革命”的角度》,《电子政务》第 4 期。

沃尔什,2020,《2062 终结:人类智能未来史》,罗静译,长沙:湖南科学技术出版社。

吴军,2016,《智能时代:大数据与智能革命重新定义未来》,北京:中信出版社。

徐琴,2009,《鲍德里亚消费社会理论的意义与局限》,《哲学研究》第 5 期。

杨子飞,2017,《作为媒介的监控与“安保主义”——论“监控社会”的安全文化》,《华中科技

大学学报》(社会科学版)第 2 期。

尹志超,2018,《重视金融科技在金融发展中的作用》,《人民周刊》第 23 期。

余聪、张亮,2021,《人工智能时代意识形态批判的哲学反思——“人机关系”视野中的批判向度》,《浙江学刊》第 2 期。

扎卡达基斯,2017,《人类的终极命运:从旧石器时代到人工智能的未来》,陈朝译,北京:中信出版社。

智慧城市赋能城市治理的发展现状及路径解析

张　博　邓芳芳*

摘要：作为数字化发展的关键抓手，智慧城市在应对日趋复杂的城市问题时优势越发明显。本文深入探讨了城市治理数字化转型的内在动因，指出智慧城市可有效破除传统城市治理手段的局限性，强化政府内部赋能，推动外部公共赋权，深度赋能城市治理现代化发展。同时，本文从城市管理、交通管理、治安维护、应急管理和资源管理五大智慧城市主要实践领域出发，分析智慧城市赋能城市治理的典型应用场景，梳理智慧城市治理方式及其对城市治理现代化产生的积极影响，进一步推广智慧城市治理，并提出更多的发展趋势和方向。

关键词：智慧城市　城市治理　数字化转型

一、引言

在城市空间中，经济生产与社会生活不断聚集并演化革新，社会治安、公共服务、自然资源、应急防灾等领域的城市问题也接踵而至。为应对新时期的种种挑战，各地正积极探索城市治理的新方式。智慧城市作为运用物联网、云计算、大数据、空间地理信息集成等新一代信息技术，促进城市规划、建设、管理和服务智慧化的新理念和新模式，被寄予越来越多的期待，并逐渐成为发展战略和工作重点。据不完全统计，我国目前已开展的智慧城

* 张博，湖南湘江新区管理委员会自然资源和规划局。邓芳芳，长沙市自然资源和规划局天心区分局。

市、信息惠民、信息消费等相关试点城市超过 500 个,所有副省级以上城市和 95%以上的地级及以上城市均提出或正在建设智慧城市。《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》明确提出,要“分级分类推进新型智慧城市建设,将物联网感知设施、通信系统等纳入公共基础设施统一规划建设,推进市政公用设施、建筑等物联网应用和智能化改造”(新华社,2021),这也标志着智慧城市将成为未来我国数字化发展的重要举措。

二、城市治理数字化转型的内在动因

城市治理中的资源依赖原则指出,现代城市系统高度开放并日趋复杂,政府亟须向外寻求非政府行为者提供资源,以提高治理能力。随着我国城镇化率突破 65%的临界线,交通出行、公共服务、市政设施、风险防控等各个领域开始出现新的风险,城市治理生态愈发复杂多变,仅凭增加人力资源投入等传统城市治理手段难以进行有效应对,政府需要寻找更高效、更可靠的外部资源。随着我国数字化进程不断推进,城市治理数字化转型成为现代城市应对复杂治理挑战的必然选择。数字化转型通过整合互联网、大数据、人工智能等现代信息技术,为城市治理提供了新的解决方案。

一方面,城市治理的数字化转型可有效实现政府内部赋能。数字时代与信息化社会中充满了不确定性和不可预见性,传统的以政府部门为主体,通过国家力量进行城市治理的模式难以适应时代发展的需求。随着城市人口快速增长和新型经济体的迅猛发展,社会生活中面临的各类风险与不确定性已内化至城市的方方面面,诸如传染性疾病、极端气候灾害等不可预知事件的发生也更为频繁。在危机处置和应急管理的过程中,政府大规模应用大数据和智能技术可有效提升其反应速度与应对能力。

另一方面,数字化转型也让城市治理更为有效地推动对外公众赋权。

随着生活水平的不断提升,人民的需求也越来越多元化,期待享受更高品质、更加便捷、更个性化的服务,也更渴望参与城市治理。数字治理通过运用信息技术和网络平台,推动了传统城市治理中政府间(G2G)、政府与企业(G2B)、政府与公众(G2C)以及政府内部(IEE)互动方式的创新。这种治理模式的变革,为政府、企业与市民之间构建了一种全新的互动机制。在这个机制下,公众(包括企业和市民)不仅可以通过数据平台直接参与到政府的决策过程中,还能通过他们产生的数据间接为政府决策提供参考依据。为此,数字技术为城市治理的精细化转型提供了有力支撑。通过深入分析和刻画数据,政府部门能够更精确地把握民众的需求,并有效进行组织内部及跨组织的协调工作。这不仅有助于满足人们对美好生活的需求,也符合城市高效治理的目标。

三、智慧城市赋能城市治理的发展路径及典型应用场景

(一) 智慧城市可深度赋能城市治理现代化发展

随着移动互联网、云计算、区块链和人工智能等现代信息技术的快速迭代,数据、算法、算力等辅助决策机制全面升级,智慧城市建设也迎来了新一轮发展。《国务院关于印发新一代人工智能发展规划的通知》提出,到 2025 年,新一代人工智能将在智能制造、智能医疗、智慧城市等领域得到广泛应用(国务院,2017)。自 2019 年起,国家发改委连续 4 年强调新型智慧城市建设工作,提升城市智慧水平。各省(直辖市、自治区)、地级市和园区等不同层次的区域,也纷纷开展了相关智慧城市实践活动。例如,湖南省 2023 年开展的“数字湖南十大应用场景”评选中,涵盖了智能交通、智慧能源、智慧文旅、智慧社区等多个城市治理领域。《2023 年四川省新型城镇化和城

乡融合发展重点任务》明确提出,将开展第三批省级新型智慧城市试点,新建智慧社区 36 个,并开展智慧教育、智慧旅游、智慧医疗、智慧体育等试点示范(夏婉琳,2023)。已有的实践表明,智慧城市可深度赋能城市治理现代化发展。

一是可赋能城市治理的可视化。根据现有的发展趋势,无论是智慧城市为城市治理带来模式上的创新,还是与现有的政府管理相融合,图形化、可视化的城市治理轨迹都会成为必然。目前,借助“互联网+物联网+云计算”等信息技术,智慧城市已基本实现治理思维、治理工具、治理行为和治理结果的可视化,这能在一定程度上提升城市治理的公开性、透明度和决策效率。

二是可提升城市治理的精细化水平。随着现代信息技术的高速迭代,由智能感知、数据分析、智慧决策等技术高度集成的智慧城市平台,使得传统城市治理模式逐渐由治理主体(政府)与治理客体(事务)之间的单向信息联系转变为人机共融的多回路治理体系(图 1)。具体表现为,智慧城市建设可增加城市治理网络中的信息传递回路,各回路之间相互协调,以高质量的智能分析和信息反馈能力,快速分析城市内生发展、外生发展下的演化进程,从而为城市治理提供更为精准的公共服务与政策指引。此外,对于突发性的重大城市事件,智慧城市系统亦能通过信息网络传递建立更为迅速的响应反馈机制。

三是可助力城市治理能力的高效能转型。随着我国城镇化的快速发展,社会环境、公共需求与重大事件的发生也更为瞬息万变与错综复杂,传统以“政府—事务”为主体的单链路城市治理方式显得捉襟见肘。随着新型智慧城市建设的推进,大数据、云计算、移动互联网、空间信息、物联网和人工智能等六大关键技术的兴起,使得城市治理各个环节的智慧化成为可能。信息技术的集成应用进一步推动了城市治理模式的创新,使之向更为高效的现代化城市治理体系迈进。例如,城市大脑充分发挥其智慧中心的

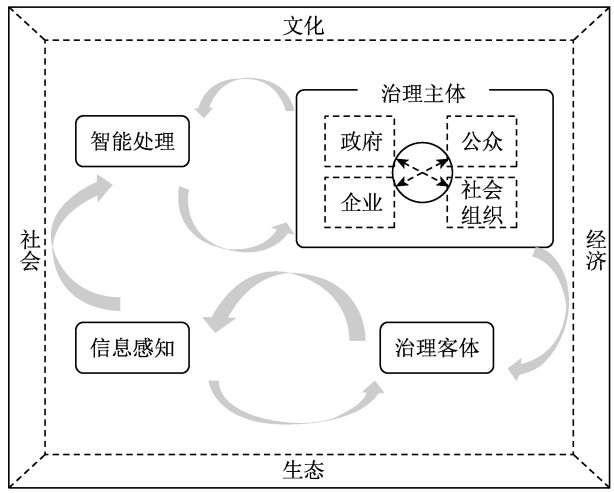


图 1　智慧城市精细化治理模式

作用,利用强大的数据分析、环境感知、智能计算和风险管控等能力,使人与人、人与物、物与物之间的联系更为便捷,极大提升了城市治理效能。

(二) 智慧城市赋能城市治理的实践路径

城市治理在迈向现代化的进程中面临着提升效能、转型升级等方面的需求和契机,而智慧城市通过智慧化、智能化手段,可在城市精细化治理、提升公共服务水平等方面发挥重要作用,进一步提升城市治理现代化水平。

1. 智慧城市赋能城市治理的实践渠道

智慧城市通过整合城市信息资源,将 P2P 的无限沟通时代带入 M2M 的物联时代,为市民提供无处不在的公共服务,为政府公共管理提供高效而有竞争力的手段。目前的智慧城市实践大多基于“智慧、互联、协同”的原则,在城市管理、交通管理、治安维护、应急管理和资源管理五大领域发力(图 2)。本文将从这五大领域出发,通过分析典型案例,进一步梳理智慧城市建设的核心内容和对城市治理的提升作用,为未来智慧城市治理提供参考。

城市管理方面,智慧城市能有效推动城市管理向治理转型和服务转变,提高城市精细化管理水平,提升城市治理效能和公众服务能力。

应急管理方面,智慧城市通过建设感知和通信网络,绘制综合风险防控地图,构建应急指挥救援体系,打造应急管理信息化新模式,推动防灾减灾救灾能力跨越式发展。

交通管理方面,智慧城市通过建设协同业务应用系统,通过城市综合交通一体监测、一点即视、一键掌控、全域预警等方式,实现城市交通整体运行态势感知及交通运行智慧化管理。

治安维护方面,智慧城市以网格化管理为基础,推进大数据分析应用,建设“人工智能+创新社会治理平台”,进一步提升城市治安管理部门精准预防、精准管控的能力。

资源管理方面,智慧城市以自然资源“一张图”为基础和核心,构建全面的业务审批支撑系统、行政事务管理系统及综合监管系统。

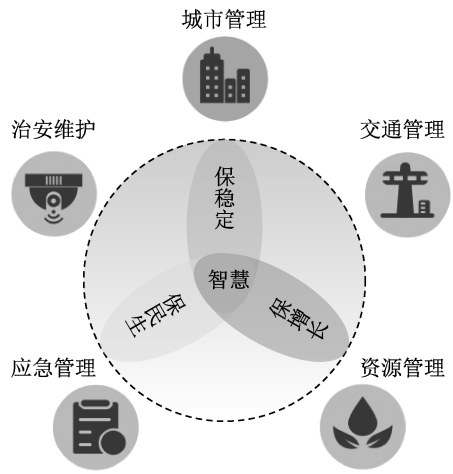


图 2　智慧城市治理五大领域

2. 智慧城市赋能城市治理的技术架构

通常而言,城市治理导向的智慧城市,其技术架构可分为三层:信息采

集层、运作操作层、领导决策支持层(图 3)。其中,信息采集层以互联网、通信网和物联网为传输通道,利用各类传感技术、无线射频识别技术(RFID)、视频实时监控技术等手段进行各类城市数据和事件信息的识别、抓取和采集;运作操作层主要对采集后的数据和事件信息进行加工,并建立相应的工作流程,进行后续的建模与信息处置,提出应对建议,同时可实现历史信息归档查询、网络设备实时监控等功能;领导决策支持层直接面向城市管理者使用,可进行仿真演习、信息查询与监控、工作流程可视化监控、历史数据分析、专家协同分析等操作,为城市的智能化管理提供数据支撑与决策辅助。

上述技术架构通过信息采集层和运作操作层进行信息的集成和处理,并通过领导决策支持层的各类应用场景面向城市管理者使用,有效实现城市治理的内部赋能;同时,广泛的公众行为信息可被各类感知技术识别、抓取和采集,进行信息处理后反馈至领导决策支持层的各类应用场景,为城市治理决策提供支撑。此外,系统还能打通政府与企业、政府与公众之间的信息渠道,在领导决策支持层建立畅通的政务网络平台,让公众、企业和城市管理者的沟通更为直接。

(三) 典型应用场景浅析

随着智慧城市与城市治理各大领域的进一步融合,城市治理过程中遇到的矛盾更容易得到解决,政府公信力与执行力也得到了较大提升。下面,本文将基于城市管理、交通管理、治安维护、应急管理和资源管理五大智慧城市建设典型应用场景,分析智慧城市治理的方式以及对城市治理现代化产生的积极影响。

1. 城市管理应用案例:海淀城市大脑

(1) 海淀城市大脑概述

城市大脑是互联网大脑与城市建设和运行结合的产物,通过建构诸如城市神经元、城市智能管理中枢、视觉神经系统、听觉神经系统等城市级类

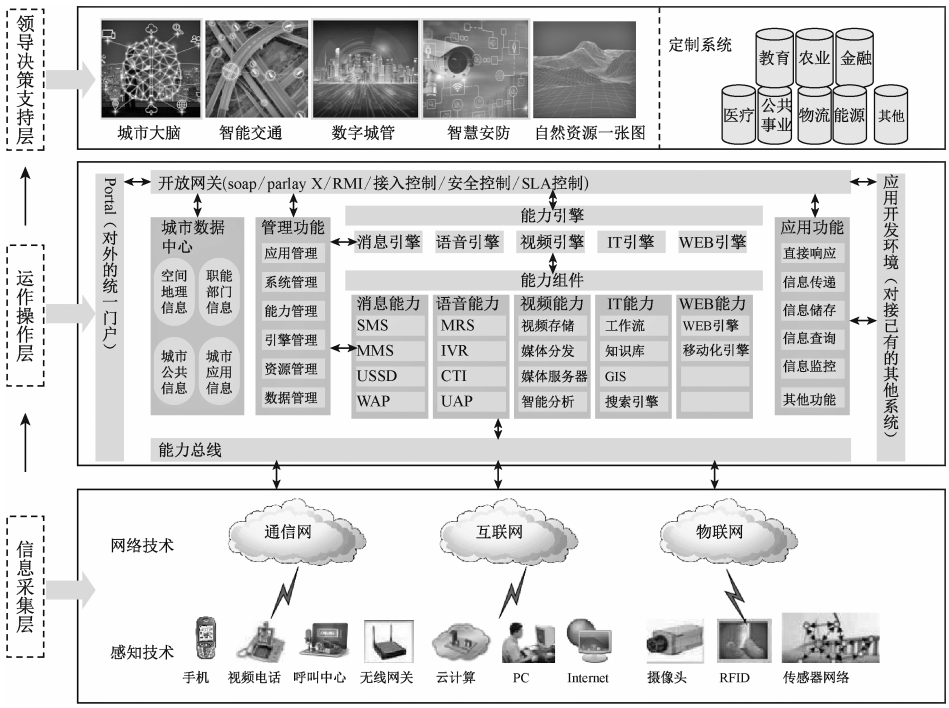


图 3　城市治理导向的智慧城市技术架构

脑功能而形成的复杂智能巨系统。城市大脑的作用是提高城市的运行效率,解决城市运行中面临的复杂问题,更好地满足民众的不同需求。

为贯彻北京市海淀区“两新两高”重要战略部署,海淀区政府委托中科大脑公司开展海淀城市大脑的建构和运营工作。海淀城市大脑以二中心(城市大数据中心、AI 计算处理中心)、二平台(城市智能云平台和共性平台)和一网络(城市感知神经网络)为核心底座,以城市大脑智能运营指挥中心(IOCC)为智慧中枢(图 4),充分发挥态势研判、事件分析、综合决策、统筹调度、仿真推演等能力,让其成为海淀区领导指挥调度的驾驶舱、部门协同会商的参谋部、街镇吹哨报到的调度室、城市平稳运行的中控台。

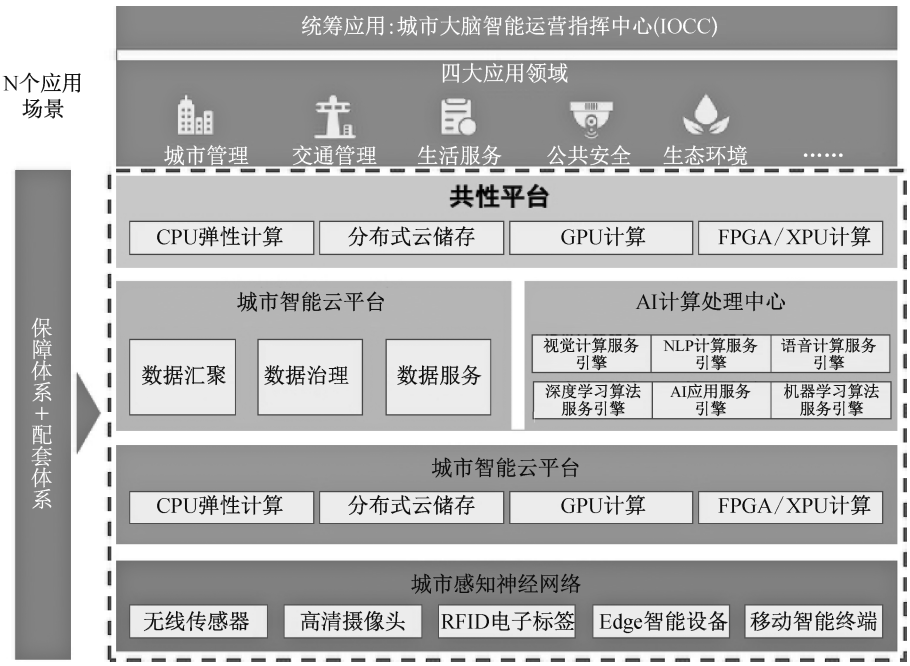


图 4　海淀城市大脑的组织架构

（2）海淀城市大脑高效推动城市治理现代化转型

海淀的 IOCC 城市智能运行管理中心，可融合城市运行、政务、物联网和互联网等多维数据，在时空“一张图”的支撑下，通过可视化的角度总览海淀区城市运行态势，厘清城市发展运行中存在的问题，统一调度城市治理中的各类事件，并实现多个端口的整合，以提高城市管理部门在不同场景下的指挥调度能力。具体而言，海淀城市大脑可在以下方面高效推动城市治理现代化转型。

第一，有助于精减城市管理流程，实现智能化运行。针对城市管理中出现的融合低、视角窄、流程长、效率低、考核难等问题，海淀城市大脑可利用 AI 计算处理中心进行城市管理案件的自动发现、智能推送、预案推荐和自动考核，并将传统城市管理的七步流程简化为三步流程，以达到发现即处

置、处置即考核的目的。

第二,有助于从“被动响应”转变为“主动治理”,从源头减少市民热线投诉问题。针对市民热线投诉数量大、类型多,基层人手不足、处置经验不够等问题,海淀城市大脑融入各行业数据、物联网数据及网络舆情数据,利用区块链、智能分析和智能对比技术,提前发现需要预警的苗头性诉求,实现未诉先办;结合历史数据对未解决的疑难问题进行深层次的原因分析,形成经典案例库,为后续办理相关议题提供依据;在投诉接受、转办、办理、回访、核查等环节,运用智能化技术减少办理步骤,进行智能机器人回访,提高办理效率,增强民众的体验感。

第三,有助于优化交通、环境、居住等方面的城市综合治理能力,提高居民幸福指数。海淀城市大脑通过建设中关村西区综合交通治理科技系统,实现交通视频实时监控,以及执法数据、过车数据、交通流数据智能化分析,从而快速提升对交通事故的执法效率,同时实现信号灯联动控制等功能;通过建设垃圾分类智慧管理体系,实现垃圾分类智慧宣传、垃圾投放智慧监管、垃圾源头身份识别、垃圾车辆监管调度和垃圾站场精准称重五大功能,提升垃圾分投运处的全生命周期社会治理能力水平;围绕智慧平安小区建设 18 个具体场景,实现多维度数据汇集,有效提升小区治安水平,全区刑事、治安警情分别下降 52.7%和 85.3%。①

第四,有助于提升重大风险预警能力。针对传统城市治理手段的监测感知体系不完善、数据获取渠道不通畅、创新业务场景支撑不足等问题,海淀城市大脑通过建设防汛防涝治理系统,提升城市重大风险预警能力。根据有关新闻报道,2021 年 8 月中旬北京海淀区暴雨期间,系统分别于 8 月 16 日 21 时、21 时 15 分、21 时 25 分连续三次发布暴雨预警,为城市管理者与市民提供了应对时间。

① 数据来源:“中科大脑”城市大脑重点领域建设成果汇报。

第五,有助于创新城市治理场景,应对未来城市发展需要。海淀城市大脑设置了低碳能源监管系统,充分发挥电力数据特点,从区域、行业等维度,对辖区内 11 大类行业、25 672 个用户、370 所不同类型的学校开展监测分析,展示疫情期间企业复工复产、学校复学情况,评估疫情对不同行业的影响,为下一步防控工作提供数据支撑。①

2. 交通管理应用案例:东莞市智慧交通平台

(1) 东莞市智慧交通平台概述

智慧交通是人工智能、物联网、大数据等新一代信息技术与交通运输深度融合的新业态,通过集成信息技术、通信技术、传感技术、控制技术和计算机技术等内容,将人、路、车有机结合,从而建立起的一种实时、准确、高效的综合运输管理系统。智慧交通系统的建设,能有效提升既有交通设施的运行效率,减少交通拥堵与交通事故发生概率,对城市治理现代化发展具有积极作用。

东莞市智慧交通平台依托东莞市城市大脑的整体架构,通过构建前端感知系统、交通大数据中心、交通协调指挥应用系统和综合交通管控系统,实现人、车、路、环境信息的全面监测,数据信息的安全传输,数据快速处理、存储、管理,多源交通大数据融合分析,与智慧城市其他系统之间的数据共享交互等功能,全面提升了城市交通管理与服务水平(图 5)。

(2) 东莞市智慧交通平台让城市交通治理更高效、精准、及时

东莞市智慧交通平台立足于东莞市整体交通基础数据,借助大数据手段,建立日常监测、协调联动、决策支持、信息服务、应急指挥等系统功能,改善东莞市的道路交通环境,让城市交通治理更为高效、精准、及时。

第一,可提高综合交通管理决策辅助支持能力。通过深入挖掘交通基础数据、交通历史数据、交通动态数据,通过指标算法模型、机器学习、交通

① 数据来源:“中科大脑”城市大脑重点领域建设成果汇报。

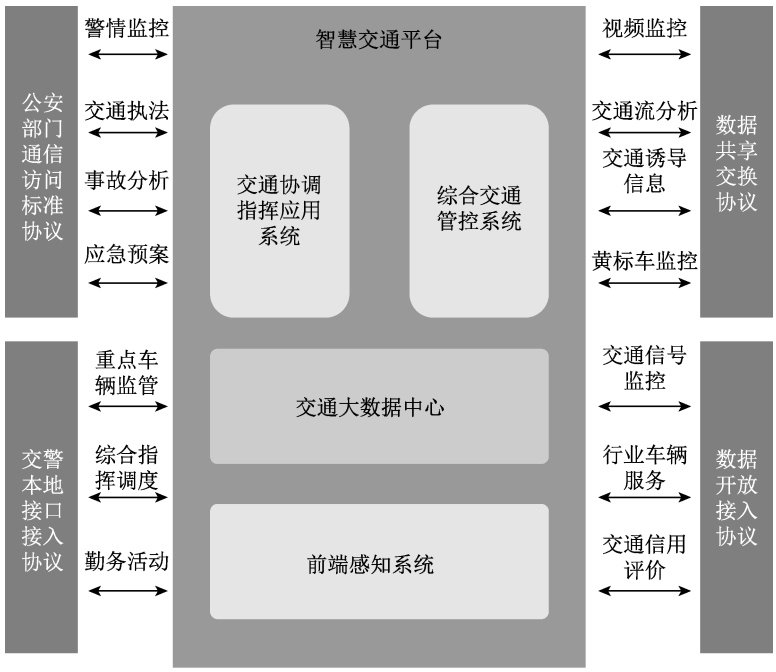


图 5　东莞市智慧交通平台的组织架构

预测、仿真分析等技术，实现综合交通运行情况的统计查询和交通运行演变规律的对比分析研判等功能，为管理调控、交通问题诊断、应急协调联动等提供科学的数据支撑。在系统支持下，东莞市滨海湾新区的道路绿波带上车辆平均行驶时间降低 28.9%，车速提高 63.6%，平均停车次数也大幅下降（约 86.7%），道路通行能力得到明显提升（高鹏飞，2023）。

第二，可全面提升综合交通运行监测与预警能力。东莞市智慧交通平台通过构建覆盖“交通路网—交通设施—出行方式”的综合交通运行监测与预警平台，基于自动化感知、道路视频实时监控、大数据动态融合、智能预测预警等技术，全面提升行业主管部门对行业运行实时状态与发展态势的精准掌控与研判。据统计，在人工智能视频分析技术的辅助下，东莞市交椅湾地区部署的 2500 路监控摄像头（高鹏飞，2023），能够准确统计进出的车

辆和行人数量,还可以自动识别路面异常情况并生成事件报告传递给交警部门。

第三,可强化多运输方式协调联动与突发事件应急处置能力。东莞市智慧交通平台建立覆盖省、市、县三级多部门协调联动通信保障体系,实现平台总中心与各分中心和其他相关单位之间信息的互通与业务的联动。同时,依托信息联动应急处置平台,面向重大交通突发事件、重要时期、重要地点,实现对各类突发事件和计划性事件的及时响应、科学决策、可靠处治。

3. 治安维护应用案例:成都市智慧消防系统

(1) 成都市智慧消防系统概述

智慧城市可通过智能感知、大数据等技术手段,实现对城市安全状况的实时监控和预警,提高警务能力智能化水平,从而切实提高维护社会稳定能力,保障城市居民安全。智慧城市技术在治安维护方面的应用通常聚焦社区、城市两大方面,本文以城市方面的成都市智慧消防系统为例,阐述其在城市治理中发挥的作用。

成都市智慧消防系统建设是《成都市“十四五”消防事业发展规划》的重点项目,旨在加快智慧消防与智慧城市的深度融合,充分运用移动互联网、物联网、大数据、云计算、区块链等技术,全面提升消防科技化、信息化、智能化水平,实现消防综合治理水平和消防综合救援能力的转型升级。成都市智慧消防系统采用“1+4+N”建设模式,搭建 1 个智慧消防大脑中枢管控平台,完善智能化火灾风险预警、综合调度指挥、消防一网通办、消防业务助理 4 个支撑系统,N 个基于智慧消防平台拓展的各类细分领域应用模块,实现接警调度、应急指挥、重大安保和智能决策等多种功能的集成(图 6)。

(2) 成都市智慧消防系统有效支撑城市消防治理高质量发展

成都市智慧消防系统聚焦防火减灾建设“火眼工程”,面向救援指挥建设“火线工程”,针对消防服务建设“火速工程”,立足业务支撑建设“火梯工程”,充分发挥智慧消防系统的优势,精准定位,全程可视化,更有助于提升

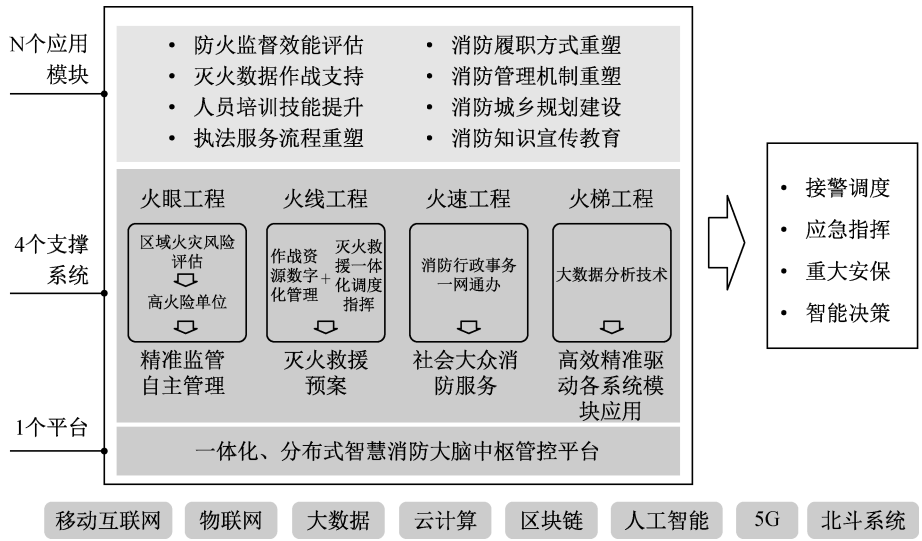


图 6　成都市智慧消防系统的组织架构

消防重大决策统筹和跨区域协同管理水平。

第一,可显著提升社会单位消防安全自我管理效能,提高全社会防火减灾能力。通过综合运用大数据、云计算、区块链等技术搭建智慧消防监管数字化平台,可全面提升数据驱动的火灾风险预警与预防能力,靶向提升高层、地下空间、大型商业综合体等高危场所的火灾预警监测体系建设,从而有效提升社会单位消防安全自我管理效能。有关数据显示,通过 1—2 年的靶向治理,可实现成都市内 70% 消防重点单位的精准监管,促进全市 10 万人口火灾数指标持续有效降低(成都市应急委员会,2022)。

第二,可提升火灾发生时的快速响应与科学施救能力,提高战训准备水平。通过设置智能智慧系统和指挥中心,推动灭火救援作战资源的数字化管理,可有效增强灾害信息采集、分析、展示能力,丰富消防指令传输手段,扩大传输范围,实现成都市灭火救援智能指挥调度与 119 报警电话实时定位,快速提升消防救援指挥能力。

第三,可为面向群众的消防服务和面向管理的消防业务提供支撑,提高

行政效能。一方面,系统可打通“一网通办”消防服务窗口,让消防类行政事务在网上实现申报、审批和结果反馈,全力满足群众“零往返”的办理期待;另一方面,基于大数据中心开发相关服务型应用,为一线防火人员监督执法及消防救援人员熟悉业务提供信息支撑,让数据研判在防火监督效能评估、灭火救援作战支撑、队站评估与规划建设、人员培训与技能提升等方面起到实质性支撑作用。

4. 应急管理应用案例:银川市应急综合管理系统

(1) 银川市应急综合管理系统概述

随着城镇化的快速发展,各种事故灾害、公共卫生和社会安全等问题日益突出,洪涝、地震、山体滑坡、燃气泄漏等突发紧急事件对社会影响之大、持续时间之长、覆盖范围之广,已经不是单一政府职能部门可以解决的问题。为此,银川市政府向科技借力,贯彻智慧银川城市建设总体安排,建设银川市应急综合管理系统。应急综合管理系统可有效调动和协调政府各职能部门,为城市构建一张全面的应急预警和处理安全网,为城市快速响应并处理突发公共紧急事件和向公众提供社会紧急救助服务提供强有力的支撑。

从结构上看,银川市应急综合管理系统可划分为感知交互层、网络云层、云数据中心层,依托银川智慧城市数据平台实现物理上的系统融合,具备监测预警、信息汇总、信息发布、图像接入、决策支持和指挥协调等功能,为各类突发公共事件的预防、监测和预案管理等常态事件处理及重大、特重大事件的协调、决策指挥和监督等非常态事件处理提供系统支撑(图 7)。

(2) 银川市应急综合管理系统让城市应急事件事前、事中、事后全过程管理更高效

通常而言,应急事件处置分为事前、事中和事后三个阶段。银川市应急综合管理系统可实现应急事件的事前电子化预案和预测预警,事中的预案可视化、大数据辅助决策和自动化执行,以及事后的复盘分析、经验总结和



图 7　银川市应急综合管理系统组织架构示意图

预案修订等功能。这一系列功能,可帮助城市管理者在应对城市突发事件时,响应更及时、处置更高效、总结更精准。

第一,可让预案体系从“有”向“有效运转”转变,让城市突发事件的事前应对更快速、更精准。首先,系统通过预案编制集中电子化,将预案名称、编号、发布单位、有效时间、适用范围等内容进行集中整合,建立包含预防预警信息、预防预警行为、预警级别和预警发布等内容的预防预警机制,按灾害程度和范围将应急响应级别分为 I—IV 级,并提供每个响应级别的标准和响应措施,为应急领导小组快速决策提供参考。其次,系统可借助视频人脸布防、视频公交人数统计、广场人群点面结合监控、视频车辆占道分析、视频巡逻等手段,实现视频预防功能。最后,系统可打通公安、交通、城管、消防、卫生等部门的数据信息库,通过大数据算法及时识别、掌握隐患苗头,及时做出预警。当前,该系统已经整合了包括 16 个危险化学品重大危险源、

13 个煤矿、84 处城市道路积水点、53 个易发生内涝的居民区以及 25 个重点防火区域在内的视频监控、资源信息和预警数据，共计 1200 路（中国网中国应急，2023），极大地提高了应急管理效率，为监测预警、指挥救援、决策支持以及政务管理等环节提供了强有力的支撑。

第二，可加快应急事件的事中处理，最大限度降低损失，保障社会公共利益。系统可通过全景地图、指挥标绘、全方位视频监控、定位跟踪、交通实景分析、水文信息分析等方式实现预案执行流程的可视化，以更直观更快速的方式辅助应急指挥，提高决策效率。同时，在预案执行过程中，依托大数据系统进行全城海量数据分析，按照“启动紧急预案—查看波及单位—调动附近视频—告警重大危险源—提示天气情况—划定生命线—展示周边路况—告知人群分布情况—显示周边应急资源”的流程实时输出相关信息，让应急领导小组可快速识别有效信息，从而及时响应事中的各类突发状况，最大限度地降低损失。

第三，可更系统、科学地进行事后复盘，为应对后续突发应急事件提供参考。系统可支撑城市应急事件的多维度复盘分析，做到事件复盘、有据可查、过程分析和时空复现。一方面，可对重特大突发事件的响应效率、战术应用、组织调度、信息传递、资源保障、现场应急预案和任务完成情况等多个方面进行系统性评估分析，精准识别存在的问题并提出意见与建议；另一方面，通过系统可进一步规范事后评估工作的时间、对象、形式和指标等内容，建立事后评估专家库，以提高事后评估工作的权威性和科学性，更好地为后续应对类似突发事件提供参考。

5. 自然资源管理应用案例：湖南湘江新区“多规合一”系统

（1）湖南湘江新区“多规合一”系统概述

广义来看，智慧城市应尽可能整合各种资源，让城市建设更好地适应人的需求。因此，高效的资源管理能力是智慧城市建设的必备功能。《自然资源部办公厅关于加强国土空间规划监督管理的通知》明确指出，要“加快

建立完善国土空间基础信息平台,形成国土空间规划‘一张图’”,并“加强规划实施监测评估预警”(自然资源部办公厅,2020),以更好地进行自然资源管理。

在此背景下,湖南湘江新区基于自身特色,建设了“多规合一”系统,进一步提升了自然资源管理水平。纵向上,该系统以“多规合一”信息平台为基础,对“山水林田湖草”等自然本底资源和“人车房地路”等经济社会发展资源进行规范化表达,同时整合现状数据、规划成果、规划实施和规划监督等数据,共同集成至“多规合一一张图”中,实现自然资源从编制、审批、实施、监测、评估到预警的全过程管控。横向上,湘江新区“多规合一”系统打通湖南省工改系统平台和长沙市政务服务网等省、市级关键政务平台,实现自然资源类业务的无缝对接;强化湘江新区部门内协同,在办理诸如项目策划生成等涉及多部门合作的自然资源类业务时,可实现一键推送、并联审批,极大提升了管理效能。此外,湘江新区作为一个跨市域、跨行政区的国家级新区,“多规合一”系统面向湘潭九华经开区、湘阴县政府开放,进一步加强市级规划协调和自然资源的统筹管理(图 8)。

(2) 湖南湘江新区“多规合一”系统让区域自然资源和规划领域的信息传导更系统、管理更全面、监督更有力

湖南湘江新区“多规合一”系统紧密围绕“一张蓝图干到底”“部门联动,高效审批”“机制保障,监管有力”的核心目标进行平台搭建,破除原有规划自成体系、自然资源相关审批流程烦琐、实施监督力量不足等问题,实现自然资源和规划领域编制、审批、实施、监测、评估、预警全周期管理。

第一,可进一步规范规划编制审批,建立健全规划编制及传导体系。湖南湘江新区“多规合一”系统按照统一标准规范并整合湘江新区 1236 平方千米范围内“山水林田湖草”“人车房地路”等各类空间数据成果质量与形式,形成全域全要素的基础数据库。同时,基于统一的数据库开展各类已编规划的标准化工作,对在编/新增规划统一基础底数和用地分类,统一规划

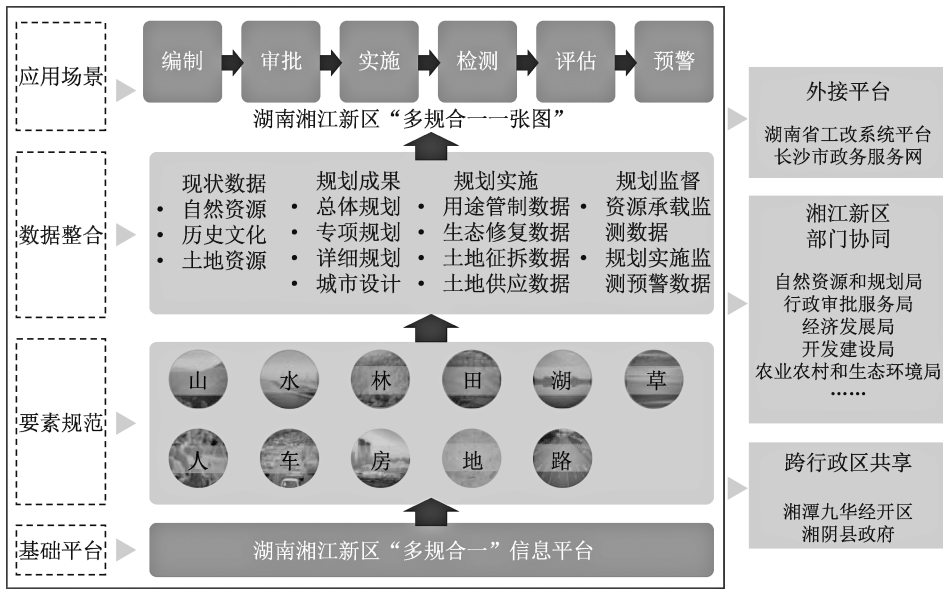


图 8　湖南湘江新区“多规合一”系统的组织架构

基期和规划期限。规划编制与审查全阶段材料可查、可溯、可控,为后续规划编制与审查提供参考。此外,系统可直观体现“基本农田控制线”“生态保护红线”等约束性指标和刚性管控要求,更有利于规划的层级传导与实施管理。

第二,可实现自然资源和规划领域编制、审批、实施、监测、评估、预警全周期管理。湖南湘江新区“多规合一”系统融合汇总规划、管理与专项图层共 8 大类、23 项数据①,实现“规划—土地—审批”空间管控边界精准落地,形成国土空间规划“一张图”;通过健全“一张图”数据采集、管理与更新机制,提供数据共享服务接口,推进与其他相关部门及岳阳、湘潭等地区信息平台的横向联通和数据共享。基于“一张图”,实现从土地调查监控到空间规划、选址预审、用地审批、土地供应、规划许可、工程建设、确权登记的业务

① 数据来源:湖南湘江新区“多规合一”系统平台。

闭环,掌握土地权属、用途、开发利用状态等信息的变化情况,辅助发现批而未供、供而未用、闲置用地等问题,实现对自然资源和规划的全生命周期监控管理。

第三,可进一步加强自然资源领域实施监督的能力。在湖南湘江新区“多规合一”系统中建立国土空间规划动态监测评估预警机制,基于实施条件、实施过程、实施结果、实施效果四个环节建立“13+N”的监测指标体系,依托湘江新区“多规合一”综合管理平台统一开展监测,接入各片区、各部门协同管理。对于长沙范围内的湘江科学城、湘潭范围内的九华兴隆湖创新中心等重点区域,以及生态保护红线面积、耕地保有量等重点监测内容,系统预警后会直接上报至湖南湘江新区管委会巡视整改工作小组,由小组直接领导和推进整改工作、督查问责,推动压力传导到位、整改落实到位,从而较大程度地提升了自然资源领域的监管力度。

四、未来展望

当前,以智慧城市为代表的数字经济发展,已成为我国推动城市治理体系和治理能力现代化的必然选择。智慧城市可运用大数据、区块链、云计算、人工智能等前沿技术推动城市管理理念、管理手段和管理模式的创新,从而推动城市治理模式的现代化变革。从综合管控、精细化治理的“城市大脑”到智能感知、实时监控的“智慧防灾”,再到全流程管理的“自然资源一张图”,智慧城市的治理触角已经扩展到交通、公安、应急、资规、市政、建设等关键部门,可有效促进信息的互联互通,让城市治理更为精准高效。

然而,不可否认的是,目前的智慧城市治理还存在一定问题。(1)顶层设计不足。智慧城市建设需要统一的规划和设计,但目前国家层面智慧城市相关的政策规范多见于各部委的发展规划之中。虽然 2018 年先后颁布了《智慧城市顶层设计指南》《智慧城市信息技术运营指南》等文件,但对智

慧城市的执行标准、资源配置和跨区域协调机制等方面的规定仍有不足。(2)数据治理面临挑战,依然存在“数据孤岛”和个人信息泄露风险。虽然智慧城市建设能有效促进部门协同,但目前“数据孤岛”现象依然存在,部门间数据共享意愿低,缺乏有效的数据整合和共享机制。同时,当前的智慧城市建设更注重技术实现,对数据治理的关注不够,数据安全和隐私保护机制不健全,依然存在数据滥用和数据泄露问题。(3)部门协同缺乏长效机制,融合联动不足。当前智慧城市治理依然存在行政逻辑不明确、部门协调不足等问题,主要表现为治理过程中反馈、受理、派遣、处置各阶段的责任主体不清晰,扯皮现象时有发生。同时,部门间整合大多处于“物理整合”阶段,缺少深层次的“化学整合”,影响了“一网通办”和“一网通管”等目标的实现。(4)区域发展不平衡,难以形成区域网络化发展格局。当前,我国智慧城市的发展重点主要集中在特大型、大型城市和沿海发达地区,鲜有目光投向中西部地区以及中小型城市。此外,当前的智慧城市治理往往以某个城市的“点”为对象,难以在区域层面围绕城镇群建设“成链成网”,滞后于区域一体化发展水平。

为了克服这些障碍,未来的智慧城市治理可在以下四个方面进行突破。(1)优化顶层设计。完善国家级的智慧城市建设指导文件,更为清晰地确立系统建设框架和执行标准,强化跨区域协调机制,确保资源配置更为合理。目前,国家层面正积极开展相关顶层设计工作,相信未来智慧城市建设的顶层设计会更加系统、规范。(2)强化数据治理。建立统一的数据管理体系,推动跨部门的数据流通与共享。同时,加强数据安全和个人隐私保护,防止数据滥用和泄露。(3)创新部门协同机制。技术创新与体制机制创新共同发力,明确智慧治理流程中各部门的职责,构建持久有效的协调与合作机制,推动政务服务的无缝对接和城市运行的统一管理。(4)促进区域均衡发展。通过政策引导和资源配置优化,加速推动中西部及中小城市的智慧治理进程,促进区域内智慧城市集群的发展,实现区域间智慧城市建

设的协同和均衡,构建全国性的智慧治理网络。相信在未来的实践中,智慧城市建设区域发展不均衡的问题能得到有效改善。

参考文献

成都市应急委员会,2022,《成都市应急委员会关于印发〈成都市“十四五”消防事业发展规划〉的通知》,https://yjglj. chengdu. gov. cn//yjglj/cdyj17/2022-03/25/content_b208ad3a311640ee8ea732b9c59da739. shtml。

高鹏飞, 2023,《智慧路灯让道路更“聪明”》, https://pub. timedg. com/a/2023-10/29/AP653e4439e4b00282f5ddbf99. html。

国务院,2017,《国务院关于印发新一代人工智能发展规划的通知》,https://www. gov. cn/zhengce/content/2017-07/20/content_5211996. htm。

夏婉琳,2023,《四川新型城镇化和城乡融合发展,2023 年重点任务来了》,https://www. planning. org. cn/news/view? id=13868。

新华社,2021,《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》,https://www. gov. cn/xinwen/2021-03/13/content_5592681. htm。

中国网中国应急,2023,《宁夏银川全力打造“智慧应急”新模式》,http://www. china. com. cn/txt/2023-03/16/content_85172033. htm。

自然资源部办公厅,2020,《自然资源部办公厅关于加强国土空间规划监督管理的通知》,https://gi. mnr. gov. cn/202005/t20200526_2521189. html。

反思“数据和人工智能向善”*

——当前趋势与未来之路

维勒·奥拉　詹姆斯·鲍尔斯　著**

阚天颜　吕　鹏　译***

摘要:近年来,各种各样的“数据向善”和“人工智能向善”倡议开始兴起,旨在推动和组织各种力量去使用新的计算技术解决社会问题。这些倡议对于计算技术的能力如何被理解为社会和政治变革的工具产生了持续影响。本文分析了以上倡议从一句口号成为一个研究项目的发展过程,这一研究项目把自己定义成各种场景的“场域”。本文讨论了最近的相关学术研究文献,展示了传播这些倡议和提出“善应该为何”的解决方案及二者之间不确定的纠葛。与之不同的是,我们呼吁研究者用一种实用和分析的态度跳出这种纠葛。本文呼吁对这些倡议的构成开展描述性研究,并从社会科学关于计算技术这一更广泛的辩论中汲取营养,开展批判性研究,从而为未来的研究提供框架。本文的经验部分向这个方向迈出了第一步:我们将“数据和人工智能向善”的各种倡议作为一个连续统的一部分看待,并且将其放到一个以“为了发展的信息与通信技术”的各种倡议为直接先驱的历史轨迹中加以理解。

关键词:数据向善　人工智能向善　社会向善的数据科学　为了发展的信息与通信技术　批判性数据研究　技术倡议

---

* 本文原标题为“Stepping Back from Data and AI for Good: Current Trends and Ways Forward”,英文原文见 https://journals.sagepub.com/doi/full/10.1177/20539517231173901,经授权翻译发表。原文由英国经济与社会研究委员会(UK Economic and Social Research Council)通过 ES/P000622/1 和 ES/P000711/1 资助。本文的补充材料可在线查阅。

** 维勒·奥拉(V. Aula),英国伦敦政治经济学院传播学院。詹姆斯·鲍尔斯(J. Bowles),英国伯明翰大学社会政策学院第三部门组织研究中心。

*** 阚天颜,中国社会科学院大学社会与民族学院。吕鹏,中国社会科学院大学数字中国研究中心。

一、导言

过去几年,尝试使用新的计算机工具去解决重大挑战的举措明显增多了。本文考察了两个致力于推广计算技术作为社会问题的解决方案的倡议——“数据向善”(Data for Good)和“人工智能向善”(AI for Good)——的整合过程。

我们认为,“数据和人工智能向善”[①]不只是精辟的口号或者凑巧的拼接;相反,与“数据和人工智能向善”有关的目标、推广、网络和资源,会塑造研究者和从业人员对于潜在的计算机科技的理解。2014—2020 年,“数据和人工智能向善”被创造、普及,并成长为一个虽然松散但是兼具科学性、集体性和实用性的项目。

本文汇总了“数据和人工智能向善”的新旧文献,为读者提供了对这一现象的批判性分析。本文回顾了各种关键的倡议内容,强调了“数据和人工智能向善”倡议与“为了发展的信息与通信技术”(ICT for development)领域的连续性和概念关联。我们为针对倡议的不同分析取向提供了一个框架,对比了规范性研究和描述性研究,比较了关注倡议本身和倡议背后的理论主题这两种不同取向。因此,本文的目标并不是要做一个详尽的实证分析,而是综合先前各种原创性经验分析的研究,勾画一个有助于生发进一步关键研究的实证和理论前景。

具体来说,本文扩展了马蒂亚诺(Madianou, 2021)、麦哲伦和库尔德利(Magalhães & Couldry, 2021)以及霍斯迈尔(Holzmeyer, 2021)所做的初步关键分析,还可以关联上那些针对遵循“科技向善”(Tech for Good)原则的

① 本文并未缩写“数据和人工智能向善”来和倡议保持一个必要的距离(critical distance),并且避免使用倡议自身倾向使用的易记的、带有宣传性的缩写,如 AI4SG 或者 D4G。但是,为了简洁,我们以“向善”(for good)代替了部分倡议名称中的“为社会之善”(for social good)。

科技初创企业的伦理思考进行的批判性评判(Powell, Ustek-Spilda & Lehuedé et al., 2022)。以上成果对单个计划或项目进行了重要解读,但是并未将“数据和人工智能向善”等一揽子倡议汇成一个更广阔的图景。这正是本文的目标,本文的介入正当其时。虽然起初的计算机技术“向善”很可能只是宣传策略,但是后来逐渐成功地将自己确定为一个特定的研究场域。一系列期刊论文(Tomašev, Cornebise & Hutter et al., 2020)强调其重要性,并且在对这一新兴场域进行系统评述的文章中,其技术成果得到了全面的评述(Shi, Wang & Fang, 2020)。因此,本文的分析是非常及时的,既因为“数据和人工智能向善”倡议刚刚“成年”,也因为已经有足够数量的学术文献出现,可以让我们来理解它。

我们提出,研究者应该以既实用又分析的姿态跳出倡议本身:以实用的姿态跳出,意味着计算机研究者应该对不加批判的推广倡议采取审慎态度;以分析的姿态跳出,意味着倾向于认同社会科学和批判性分析的研究者应该聚焦于相关倡议的更广阔图景,从而审视过往和当今各种运动(movements)中的异同,以及运动背后更宏大的理论主题。总而言之,我们反对“数据和人工智能向善”需要专门叙事(specialist literature)的观点——这在“为了发展的信息与通信技术”中已然如此。因为,当我们需要建立一个关于计算机技术的批判性叙事的统一话术时,采取专门叙事只会加重相关学术领域的隔阂。

本文的结构如下。首先,我们概括了不同“数据和人工智能向善”倡议的兴起过程及其内在关联。其次,我们厘清了当前各种倡议和“为了发展的信息与通信技术”的关联。在阐述这些经验内容后,我们讨论了学术文献与倡议的互动,既包括前者通过系统推广使后者凸显其存在,也包括前者批判性评估后者的内容。最后,我们讨论了从更广范围理解计算机技术、政治和社会三者关系的时候,应该采取什么样的方法将“数据和人工智能向善”倡议融入进来。

二、理解“数据和人工智能向善”倡议的结构

我们将追溯“数据和人工智能向善”倡议兴起的关键步骤,从而展现这些倡议紧密的内在关联,并试图建构一个为社会之善的场域。本节不局限于对倡议起源的简单讨论,而将倡议连接起来并且将其视作一个单一连续体的一部分。这种网络搭建在科学领域的社会研究中较为常用(Latour,1987)。我们在此先指出“数据和人工智能向善”倡议的关键角色、事件和发展阶段(表 1)。

表 1 “数据和人工智能向善”倡议的定义、提议和里程碑汇总

倡议或品牌名称	定义和主张	关键节点
数据向善	一种考虑伦理问题并且能带来积极行动的数据科学及其使用,通常以为人类服务为目的,可以与联合国可持续发展目标配合。 “数据向善”对话是彭博一个长久传统的一部分,这一传统是利用数据科学和人力资本去解决社会的核心问题。 我们将支持拥有隐私保护(privacy-preserving)数据授权的伙伴以助力赋能社区和解决社会问题。	2011 年,启动“数据无边界”(Data Without Borders)(DataKind 的前身)活动。 2014 年,举行首次彭博“数据向善对话”(Bloomberg Data for Good Exchange)。 2018 年,举办首届英国“数据向善”大会(UK Data4Good Conference)。
数据科学向善	我们的使命是鼓励数据科学带来积极的社会影响。我们培训和支持新一代的数据科学家,为非营利组织提供直接支持,并且开发工具以保障数据科学和人工智能的使用能够积极地和公平地惠及世界各地的人。	2013 年,在芝加哥大学举行首次“数据科学为社会之善”夏季讲座。 2014 年,“数据科学为社会之善”成为美国计算机协会年度“知识发现与数据挖掘会议”(ACM's Annual KDD Conference)的主题。

（续表）

倡议或品牌名称	定义和主张	关键节点
人工智能向善	“人工智能向善”的目标是发现可以推进联合国可持续发展目标并将其解决方案成规模扩展至全球的人工智能的实际应用。 通过对人工智能生态的研究、设计和发起倡议，我们正在努力使用人工智能去应对社会挑战。 “人工智能为社会之善”研讨会将聚焦那些人工智能有潜力提供有价值的解决方案的社会问题。我们选择聚焦问题的灵感来源于联合国可持续发展目标。	2016 年，“人工智能向善”基金会（AI for Good Foundation）启动。 2017 年，国际电信联盟（ITU）举办首次“人工智能向善”全球峰会（AI for Good Global Summit）。 2018 年，谷歌启动“人工智能向善”项目以统筹工作。 2018 年，神经信息处理系统大会开始举办“人工智能为社会之善”研讨会。

表 1 概括了三个著名的“数据和人工智能向善”倡议的关键定义、主张和时间节点，用以展示它们在使命上的绵延和趋同。从时间顺序看，“数据向善”是第一个兴起的倡议。此倡议以 2011 年启动“数据无边界”为起点，起初只是一个单独的黑客马拉松式（hackthon-style）[①]事件，但是很快风靡起来。此倡议采用“数据向善”的口号，并且在“数据同人”（DataKind）的组织下举行活动。正因为有这种专业和公民编程式（civic-hacking type）的志愿活动，才会出现将借助计算机技术实现社会目的清晰地表述为“向善”的这一初步想法。这一表述策略很快被大学和研究者采纳，他们开始以“数据科学为社会之善”来表述相似的工作。这一表述首次出现在芝加哥大学 2013 年的一次讲座活动上，更著名的是随后在 2014 年作为美国计算机协会年度“知识发现与数据挖掘会议”的指定主题。在某些“数据和人工智能向善”的历史回顾中，2014 年的会议标志着这一表述（moniker）开始引起广泛关注。因此，这也经常被视为此运动的起点。“人工智能向善”是这一运

① 黑客马拉松是一种汇集专家力量合作解决一个特定问题的活动，多聚焦于计算机编程和信息技术领域。——译者

动的最终版本,不仅为 2016 年发起的同名倡议奠定了基础,同时也是联合国 2017 年的首选概念,以代指用计算机方法服务可持续发展目标(sustainable development goals)。

一步步地,“数据和人工智能向善”概念开始传播,最终在公共和学术领域成长为一个应用计算机方法的特定概念(Berendt, 2019; Moore, 2019)。后来发表的一系列关于“人工智能向善”倡议场域的系统性文献评述(Cowls, Tsamados & Taddeo et al., 2021; Shi, Wang & Fang, 2020)具有里程碑的意义,标志着这一表述从一个吸引人的口号,转变为一个代表一个应用领域的战略性概念。在神经信息处理系统大会(NeurIPS)将这一表述采纳为一个特定的研究领域时,这一转变表现得同样明显。同时,“数据向善”由联合国定义的“一个由可持续发展目标勾勒框架的领域”,发展为学术期刊中计划性的推广对象(Floridi, Cowls & King et al., 2020; Tomašev, Cornebise & Hutter et al., 2020)。

为了达到批判性剖析倡议内容的目的,我们将每一个“向善”表述从谓语(它在追求哪一种善)和主语(采取什么样的干预方式)部分分解,并且分析不同倡议践行者的不同理解。我们首先讨论“向善”代表什么。

专业人士、科学从业者和评论人士都曾指出,“向善”的定义在倡议发展早期被刻意地宽泛化,且如今依旧缺乏准确定义。几位学术评论人士曾对概念的模糊性进行过明确探讨,他们指出倡议中的“善”的统一点,也许只能是非纯粹利益驱动(Berendt, 2019; Green, 2019; Moore, 2019)。对“善”所代表的事物进行定义的困难,加剧了将“数据和人工智能向善”表述为既是一个创新性的新事物,又是一个已确立的议程的模糊倾向。倡议的支持者坚持“数据和人工智能向善”是一个应用领域,从而进一步宣称任何由支持者所认定的与“社会善”(social good)某方面相关的计划都是该共享项目的一部分。在研究文献中,一些学者(Shi, Wang & Fang, 2020)将人工智能在所有特定领域的应用都归为“人工智能向善”。这一类例子展现了

倡议内容边界的持续变化,尤其表现为这些相互冲突的倡议会根据自身需要,纳入和排除“社会善”中某些不明确的部分。因此,我们应该主要通过人工智能工具及其应用来理解“数据和人工智能向善”代表什么。这将我们引向对于倡议要求的干预措施的探讨。

“××向善”(X for good)的主语(××)告诉我们造福社会的工具是什么。在数据、数据科学和“人工智能向善”中,我们可以看到每一次计算机方法的微小进步都会用具有辨识度的名称变化以保持新鲜感。如表 1 所示,随着公众关注主题的变化,机构会提出新的倡议变体。因此,在未来,研究者在准备好面对技术变革社会中的新变体的同时,应注意不要将变体误认为全新的发展。

兴起之初,“数据向善”利用大数据现象的热度,将数据作为其创新见解和干预措施的来源。一开始我们可以看到的例子是:博为(J. Porway)在 2011 年发表的博客文章中呼吁“社区驱动的变革”(community-driven change),最终发起了“数据同人”计划,希望“聚集一群想要利用自己空闲时间去做一些有意义的事情的同人,去支持那些需要数据服务的非营利组织或者小企业”(Porway, 2011)。这一模糊的概念在后来的彭博“数据向善对话”(Bloomberg Data for Good Exchange)活动中同样有所体现。这一活动探索性地将焦点放在慈善企业家和计算机研究者的合作上。因此,“数据向善”的第一轮在“善”的表现形式和干预的工具上都表现出最大限度的模糊性。可以说,它除了是一个赋予技术发展正向引导的叙述策略,其他什么也不是。

因面向学术受众和推广特定的计算机工具,“数据科学向善”(data science for good)与其前身“数据向善”有所不同(Mann & Sahuguet, 2018)。“数据科学向善”由机器学习和计算机数据科学方法应用领域的专家,而不是数据的一般使用者主导(Paolotti & Tizzoni, 2018)。尽管如此,数据科学和数据之间的界限有时较为模糊,正如优时(C. Juech)在编著《数据科学为

社会之善:复杂世界中的慈善和社会影响力》(*Data Science for Social Good: Philanthropy and Social Impact in A Complex World*)时所做的那样,通过非营利组织使用大数据的行为来给“数据向善”下定义(Juech, 2021)。我们可以借助“数据科学向善”的高校项目来观察“数据科学向善”的特点。在这些项目中,计算机学科的学生参与由公共部门提议和企业资助的项目,如流浪汉救助、组织社会和医疗服务、进行城市规划以及促进种族公平。综合来看,“数据科学为社会之善”强调数据科学家作为变革行动者的角色,即运用具体的计算机工具为社会公益做贡献。

最后,“人工智能向善”试图运用在机器学习和人工智能研究中的进步以预防、缓和或者解决人类和自然世界的问题,同时不带来伤害和加剧不平等(Cowls, Tsamados & Taddeo et al., 2021; Tomašev, Cornebise & Hutter et al., 2020)。正如公众对话中经常发生的——对于人工智能的定义目前是含混不清的。史(Shi,音)等人就“人工智能向善”的技术研究进行的调查显示,人工智能可以指代多种计算机技术,但是所有技术中,机器学习在所有应用领域中都是最受欢迎的(Shi, Wang & Fang, 2020)。此外,医疗保健是所有分类中最热门的,同时也是机器学习应用率最高的垂类。

以上概述展现了这是一个正在扩张的领域,这个领域包含了诸多将数据、数据科学或人工智能作为造福社会的手段的倡议和机构。将这些倡议作为一个单一连续体的部分进行分析,使得我们可以展现出将这些倡议连接起来的更广阔的轨迹。接下来,我们将回顾更加久远的历史,从而展现“数据和人工智能向善”倡议是沿着“为了发展的信息与通信技术”运动勾画的轨迹发展的。

三、与“为了发展的信息与通信技术”的连续性

以分析的姿态跳出“数据和人工智能向善”倡议的推广逻辑后可以发

现,这些倡议应该被视为致力于用技术解决社会问题的过往和当今运动链条中的一个新联结。我们认为,对“数据和人工智能向善”倡议进行实证研究,是一条发现该倡议与历史上的和同时代类似倡议的相似性和连续性的重要途径。我们通过考察“为了发展的信息与通信技术”(ICT4D)和“数据和人工智能向善”倡议之间的联系,展现了这两种倡议之间的连续性。“为了发展的信息与通信技术”开创了上文提及的通过企业、非营利组织和学界的合作来搭建网络的战略,并且此战略在 ICT4D 有关文献中也得到了广泛探讨。我们并不试图说明 ICT4D 是唯一重要的相似体(parallel),并鼓励更多的研究以发现其他重叠的地方;我们想说明 ICT4D 与“数据和人工智能向善”倡议的共通之处尤其明显,而我们的研究提供了一个重要起点。

ICT4D 出现于 20 世纪 80 年代,并在随后的几十年内发展为一个广受认可的专业性和学术性运动。随着运动的进一步发展,ICT4D 最终成为会议、联合国项目、企业计划、慈善倡议、研究活动和学术期刊使用的名称(关于 ICT4D 研究领域的兴起,参见 Walsham, 2017)。ICT4D 也逐渐从起初重点关注现代化和经济发展,转变为关注联合国千年发展目标中体现的多元发展理念。此倡议的演变显示出,如联合国这样的发展组织和大型非营利组织非常乐于接受西方私营技术公司的资金和技术支持;在这一过程中,被支持者也采纳了技术推动社会进步这一正面叙事。正如上文所述,在联合国、大型技术企业和慈善捐赠者推广的“数据和人工智能向善”计划中,该策略得到了延续。

关于 ICT4D 的学术研究表明,此运动起初发轫于对发展中国家信息系统的研究,并且侧重于技术创新及其应用。不过,人们关于信息与通信技术(ICT)是否以及多大程度上促进了发展通常难以达成共识(Zheng, Hatakka & Sahay et al., 2018)。的确,对于发展是什么难以达成共识(Avgerou, 2010; Crush, 1995)以及定义信息与通信技术的不易(Sein & Harindranath, 2004)几乎一开始就成了困扰这一领域的难题。与之相似的是,对于计算机

技术是什么、“数据和人工智能向善”中的“善”是什么,也是难以统一和定义的。鉴于存在这些模糊之处,ICT4D 同样无法建构能够回答整个组织架构中核心问题——信息与通信技术的进步如何促进发展——的有力理论(Best, 2010; Heeks, 2010; Walsham, 2017)。但是,这些模糊之处并没有阻止 ICT4D 成为一个重要的专业学术性现象。这证明了专业学术性运动即使面临批评也具有推动目标达成的能力。

人们对 ICT4D 和“数据和人工智能向善”的批评也是相似的。ICT4D 因为具有技术决定论和新自由主义思想驱动的市场扩张的倾向而受到批评,这两种倾向将给达成缓解经济和社会不平等或者数据鸿沟的目标带来不利影响(Pieterse, 2010)。类似的批评也被用于指责“数据和人工智能向善”倡议的发展指向(Madianou, 2021; Magalhães & Couldry, 2021)。因此,政治经济学和技术决定论中的相关理论挑战,也同样适用于这两项运动。在关于 ICT4D 的学术研究中就存在针对上述问题的强烈批评声(De, Pal & Sethi et al., 2018; Schelenz & Pawelec, 2021),但是专业实践经常对此类批评充耳不闻。类似的挑战在下面一种“差距”中同样存在——尽管批评声越来越大,实践中缺乏进展,但是对于发展社会场景运用的人工智能的炒作和投资依然不绝。“数据和人工智能向善”的发展再次验证了在 ICT4D 中发现的经验,即在技术研究中展开的关于解决非技术问题的研究,即使存在薄弱的核心理论假设和强烈的社会科学批评,也有着自我维持动力的能力。

四、从 ICT4D 到“人工智能向善”的转变

上述两个运动之间的关联有时会以从前者到后者直接转变的形式表现出来,上文提及的国际电信联盟(ITU)开展的“人工智能向善”的工作是这种转变的一个实例。ITU 隶属于联合国,后者对 20 世纪 90 年代 ICT4D 的

兴起起到了关键作用,联合国采取的政策将技术打造为发展工具,同时与信息与通信技术行业领先的私营公司开展合作。ITU 在协调组织联合国 2005 年信息社会世界峰会论坛以及发起“衡量信息通信技术促发展的伙伴关系”(Partnership on Measuring ICT For Development)倡议后,成为联合国在 ICT4D 工作中的主要伙伴。自 2017 年起,ITU 通过定期召开以“人工智能向善”为主题的全球峰会,成为联合国人工智能工作的领头机构。此项倡议由国际商用机器公司(IBM)、科技慈善机构 X 奖基金会(XPRIZE)和美国计算机协会(ACM)协办。这一由 ITU 主办的新的“人工智能向善”项目,通过联合国可持续发展目标来定义“善”,这一定义方式也被联合国 ICT4D 的工作所采纳。因此,ITU 的“人工智能向善”工作是联合国早前 ICT4D 工作的直接延续。将联合国可持续发展目标作为“人工智能向善”的主题基石的方法也被学术界(Cowls, Tsamados & Taddeo et al. , 2021)和私营企业部门(Tomašev, Cornebise & Hutter et al. , 2020)所采纳。

为了对 ITU 开展的 ICT4D 和“人工智能向善”工作之间的连续性进行评估,我们对该联盟 AI for Good 网站上发布的官方文章进行了主题内容分析。我们收集了所有在 2021 年 1 月 1 日至 2022 年 12 月 31 日之间发表的文章(总计 88 篇),并使用两阶段归纳编码(two-stage inductive coding)进行数据处理。依据西尔弗曼(D. Silverman)提出的方法,我们按照文章的主要兴趣点、讨论的技术和应用、提及的合作者和伙伴的名称与类型以及想要贡献的特定联合国可持续发展目标进行主题编码(Silverman, 2020)。第一轮编码得出 28 类主题。在对这 28 类主题进行同类合并来保障既具有代表性又更方便操作后,我们对文章进行第二轮编码。图 1 显示了在文章中最频繁出现的主题以及它们出现的频率。

有两类主题在这个时间段最突出:一是使用人工智能工具和方法减缓气候变化、解决环境问题(12 篇);二是致力于解决联合国可持续发展目标(SDGs)这类“大挑战”(12 篇)。其他重要的主题包括可持续和发展语境

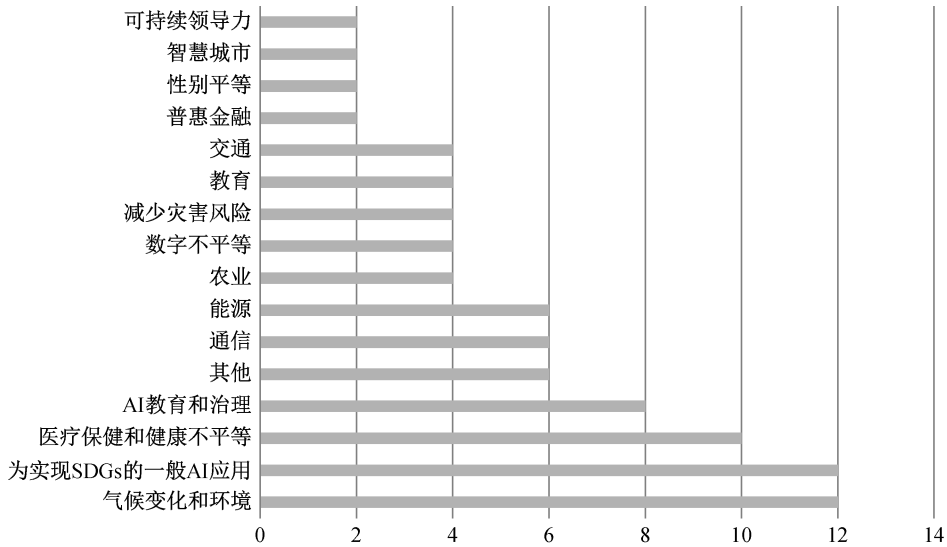


图 1　2021 年 1 月 1 日至 2022 年 12 月 31 日 ITU 发布的“人工智能向善”文章主题

下的人工智能教育和治理(8 篇),以及使用人工智能工具和方法减少健康不平等和提供医疗保健干预措施,尤其是向全球南方(Global South)①(10 篇)。图 1 展示的许多主题直接反映了 ICT4D 学者和 ITU 的 ICT4D 工作者关心的主题。ITU 召开的 2003 年信息社会世界峰会论坛是 ICT4D 体系化的关键节点(Heeks, 2014; Unwin & Unwind, 2009)。当年,ITU 愿景的核心是确保“信息与通信技术的运用应该以造福人类生活的方方面面为追求目标”(ITU, 2003)。据此,ITU 提出了涵盖政府服务、就业、健康、商业、疾病预防和教育的各类明确目标。沃尔沙姆(G. Walsham)指出,这与后来的 ICT4D 研究中的核心目标除了排列上有些不同,内容几乎完全一致(Walsham, 2017)。因此,ICT4D 学者和工作者同“人工智能向善”的支持者所关注的挑战是重叠的。

① “全球南方”是新兴市场国家和发展中国家的集合体,多位于南半球;与此相对,“全球北方”代指高收入国家,多位于北半球。——译者

除了联合国 ICT4D 工作和“人工智能向善”倡议关注主题上的连续性，上述文章还展现了这些工作长期依赖于与外部机构的合作。这些外部机构在开发新计算机技术的同时，侧重于技术的发展潜力而非使用现有技术去解决大挑战。公—私部门（public-private sector）的战略伙伴关系被认为是确保信息与通信技术在发展语境中得到有效利用的关键。同时，我们也发现，“人工智能向善”领域中的私营部门组织受到高度重视。事实上，联合国可持续发展“目标九”（SDG 9）（产业、创新和基础设施）是所有联合国可持续发展目标中被 ITU“人工智能向善”相关文章最频繁提及的目标；同时，55%的文章都将私营部门组织列为合作伙伴。[①] IBM 在 14 篇文章中都被列为重要伙伴，成为最高频次出现的私营合作者，同时是 ITU“人工智能向善”倡议的“领头赞助者”（IBM，2021）。许多文章将 IBM 的人工智能技术和倡议视作实现可持续发展目标的核心。这些文章将首要关注点放在人工智能技术解决大挑战的潜力上，而非放在人工智能技术实现诺言、成功解决问题的案例上。

第二个展现 ICT4D 和“人工智能向善”之间直接连续性的相关事件与微软公司有关。如果说 ITU“人工智能向善”展示了一个重要发展的行动者从一个倡议转向另一个倡议的过程，那么微软则展现了一个私营公司的相同经历。2006—2017 年，微软是信息通信技术和发展（ICT and Development）年度学术会议中提交论文最多的供稿者之一（Mukerji & Chauhan，2020）；同时，微软在联合国发起的 ICT4D 活动中的参与有据可查（Leye，2007）。2018 年，微软启动了“人工智能向善”实验室（AI for Good Lab）。实验室发表研究成果以及与外部伙伴就解决世界性的，尤其是在全球南方的包容、可持续和人权等问题上展开合作。微软所选择的路径与其他深入“人工智能向善”的重要技术公司的选择相类似（Magalhães & Couldry，

① https://journals.sagepub.com/doi/full/10.1177/20539517231173901#supplementary-materials.——译者

2021),这条路径与 ICT4D 采取的战略高度一致。此外,对微软“人工智能向善”计划关于技术还原论、新自由主义和市场扩张的批评(Magalhães & Couldry, 2021),与 20 年前对微软与联合国合作的批评类似(Leye, 2007)。

以上的分析展现了 ICT4D 和“数据和人工智能向善”倡议的连续性和相似性。本文的主要论点是:“人工智能向善”倡议正试图复制当年 ICT4D 运动采取的扩展和自我维持战略。ICT4D 的成功使其拥有了对整个发展部门如何理解技术施加影响的能力。“数据和人工智能向善”倡议使用了相同的战略和治理安排以推广计算机技术。ITU 的“人工智能向善”倡议和微软的“人工智能向善”实验室展示了机构转型的直接证据。

五、文献研究“数据和人工智能向善”的四条路径

由于“数据和人工智能向善”倡议不再只是一个口号,一批文献开始涌现,它们不再仅将倡议看作五花八门的技术应用的集合。此处我们将讨论文献中的关键论点。我们不会讨论技术性论文中的内容或者与倡议相关的计划报告书。因为包含与排除某一个计划是“数据和人工智能向善”倡议自我归类的结果,所以对这个与日俱增的技术性文献进行细节性的讨论并无益于本文的目的。因此,本文采取部分近期“数据和人工智能向善”文献评述所使用的策略,即只考虑自我认定为“数据和人工智能向善”这一部分的论文,而不考虑只是提及发展中国家信息和通信技术的论文(Walsham, 2017)。史等和卡沃斯(J. Cowls)等所做的追溯分类调查不仅对展现倡议成果有价值,也有利于将“数据和人工智能向善”提升为一个独特的应用领域(Shi, Wang & Fang, 2020; Cowls, Tsamados & Taddeo et al., 2021)。

在表 2 中,我们区分了目前兴起的针对“数据和人工智能向善”倡议的四种不同类型的文献。这些路径展现了研究者对倡议采取的多样视角,而关于此主题的讨论正在展开。这些文献的搜集方法是:在重要科学出版物

数据库中检索“数据和人工智能向善”倡议的各种变体和缩写,以及从重要期刊中手动检索。我们基于文献本质上是规范性的还是描述性的,以及是基于“数据和人工智能向善”本身的分析还是将之与社会科学和人文科学中更广泛的理论讨论联系起来,对文献进行分类。我们不想去探求这些文献的实质(substantive)贡献,而是试图通过上述分类策略,聚焦元层面的问题,即这些文献是如何与兴起的“数据和人工智能向善”倡议联系在一起的。

表 2　现有研究文献对“数据和人工智能向善”倡议所采取的分析路径

	基于对“数据和人工智能向善”倡议和实践的分析	基于对社会科学和人文科学中更广泛的理论讨论
规范性研究	1. 倡导性研究文献	3. 关于“善”应该是什么的讨论
描述性研究	2. 将“数据和人工智能向善”作为倡议进行分析	4. 对于“数据和人工智能向善”计划的批判性评估

(一) 倡导性(promotional)研究文献

目前为止,这部分的研究文献聚焦于传播“数据和人工智能向善”领域的存在和重要性。在美国计算机协会 2014 年“知识发现与数据挖掘会议”将“数据科学为社会之善”大众化后,研究者开始使用这一名称作为技术性论文的关键词,并将论文归于这一正在兴起的领域。这些论文不一定都是探讨“数据和人工智能向善”倡议的,也有些聚焦于具体的应用和计划。“数据和人工智能向善”的支持者利用这一倡议,将原先分割独立的技术计划连成网络。接下来,众多技术性和经验总结性(lessons-learned)论文开始给自己贴上“数据和人工智能向善”的标签,来达到分享作为倡议“成员”的好处的目的。

但是,在 2014 年会议召开后发表的那些指向明确的系统性的论文对本研究更有意义,这些论文勾勒和赞扬(celebrate)了一个新研究项目的诞生。最早的学术文献(Catlett & Ghani, 2015; Chandy, Hassan & Mukherji, 2017)

明确聚焦于勾勒和宣传这个研究项目,并且扩展和巩固了初期会议中搭建的框架。其中,一些文章与美国计算机协会“知识发现与数据挖掘会议”和其后的彭博“数据向善对话”,以及其他以“数据向善”命名的会议有直接关联(Mann & Sahuguet, 2018; Paolotti & Tizzoni, 2018)。其中,许多文章是由在私人研究实验室工作的技术专家为发展和推广计算机技术而撰写的。这些论文将早期受追捧的框架弃之不理,而试图建构自己的目标与话语。托马舍夫(N. Tomašev)等人、弗洛里迪等人的文章都较好地例证了一个现象,即越来越多的不属于初期会议的研究者选择加入和推广倡议(Tomašev, Cornebise & Hutter et al., 2020; Floridi, Cowls & King et al., 2020; Cowls, Tsamados & Taddeo et al., 2021)。我们还发现,这些推广活动并不是一蹴而就的,在较长一段时间里不同的论文都在传播这些相同的理念。

倡导性的学术文献经常与智库和咨询公司发表的专业文献观点一致,且前者经常得到后者支持。举例来说,总部位于英国的内斯塔(NESTA)采用了“数据同人”在《数据向善》报告中对“数据向善”的理解(NESTA, 2015);此外,“数据向善”在硅谷引起关注,麦肯锡(McKinsey)也采用了这一愿景(Chui, Harryson & Manyika et al., 2018)。这种内部的变革动力突出了学术与实践显著且紧密的联系,尤其是在倡导性文献中。比如,关于呼吁“人工智能为社会之善”计划中更大程度的跨部门合作的文章的作者,既有来自学术机构的,也有来自支持人工智能应用的头部商业组织的(Tomašev, Cornebise & Hutter et al., 2020)。

因为这些倡导性文献的目的是稳固“数据和人工智能向善”作为一个研究领域存在的地位,所以与 ICT4D 的早期文献具有可比性。正如许多 ICT4D 学术文献是在初期私人部门和非营利组织参与后而非参与前发表的一样,这些倡导性文献大部分是在私营公司和非营利组织开始推广倡议之后发表的。因此,这些文献是基于其推广的领域已经存在这样一个事实

基础而展开的。但是,就这些文献自身来说,其起到了稳固倡议存在的作用。

(二) 关于“善”应该是什么的讨论

“数据和人工智能向善”倡议的成功,很快引发了关于倡议所追求的“社会之善”的定义问题的反思。这些文献与 ICT4D 对于定义“发展”存在的困难具有相似之处,但是由于前者更加抽象而又有所不同。从一开始,这些文献就是关于“善”应该是什么的规范性探讨。尽管许多作者对于“善”所代表内容的模糊性感到惋惜,但是他们依然愿意支持“数据和人工智能向善”,并且将重点放在可以为倡议提供坚定基石的特征阐释上。这一立场出人意料地使作者经常与倡导性文献发生交际,却使他们远离了批判性的社会科学文献。

摩尔(Moore, 2019)和贝伦特(Berendt, 2019)作为第一批撰写对“数据和人工智能向善”进行批判性解读文章的人,提出“数据和人工智能向善”倡议应该将设定“善”是什么作为其应用领域。他们提出,“善”的概念并不是建立在共同原则或者政治目标之上的,而是由计算机技术的应用领域(如医疗保健、环境和发展援助)所决定的。作者认为,这种基于领域(domain-based)的定义使倡议暴露在模糊性和伦理挑战中。因为这种定义方式反映了“向善”在初级阶段定义的不精确性(正如前文所讨论的),所以提出归纳性定义的重要性再怎么强调都不为过。但是,摩尔和贝伦特的文章只是将以上评论作为各类替代性定义的跳板,他们并没有将这一论述发展成关于倡议本身或者其更广泛影响的详细探讨。不过,他们的论述有着更广泛的分析价值并且成为本文的一个重要起点。

一些作者提出,“善”的定义应该由开发和部署计划时遵循的伦理标准组成。这一定义方式来源于关于信息伦理和道德的数据科学(ethical data science)的更广泛的讨论,这使得数据和人工智能成为总体讨论中的一小部

分。例如,卡沃斯就将这一主题视作一个可以进行哲学讨论的松散体(Cowls, 2021)。在对于追求“社会之善”更直接的贡献中,研究者提出了预先设定的伦理准则(Floridi, Cowls & King et al., 2020)以及指导伦理反思的切实问题(Berendt, 2019; Berendt, 2020)。

有学者也试图通过计划希望达成的目标来定义“善”:格林(B. Green)提出应该通过实现社会正义的目标和作为政治目标的反抗压迫二者来定义“善”(Green, 2019; Green, 2020);摩尔提出另一个同样由政治驱动的定义,即定义应该注重提高人的能力(Moore, 2019);卡沃斯等人提出“人工智能向善”倡议应该遵循联合国可持续发展目标这个对于“善”的公认准则,同时也意味着他们认可联合国下属机构 ITU 在 2017 年提出的定义(Cowls, Tsamados & Taddeo et al., 2021)。所有这些作者都将“善”与一个超越学术定义的更广泛的公众讨论相联系。

(三) 将“数据和人工智能向善”作为倡议进行分析

许多探讨“数据和人工智能向善”倡议的论文,都包括关于倡议产生的过程、倡议的赞助者和倡议推广的应用对象的讨论。不过,大部分描述性论述都未详细分析这些倡议的社会、组织、沟通和经济结构。例如,贝伦特和摩尔在关于以领域为基础的“善”的定义的重要论述中讨论了倡议的运作方式,但是文章的中心点还是对于“善”应该是什么的规范性探讨。此外,本文下一小节探讨的很多批判性的社会科学论文都包括对“数据和人工智能向善”倡议运作方式的探讨,但是这类探讨的最终目的还是就当代的数字化特征做更广泛的探讨(Magalhães & Couldry, 2021),或者对使用人工智能解决非技术性问题进行总体批评(Holzmeyer, 2021)。在少数完全聚焦于“数据和人工智能向善”倡议的组织方式的研究案例中,埃斯皮诺萨(M. Espinoza)等人是其中的一个代表(Espinoza & Aronczyk, 2021)。他们的研究展现了数据和气候行动计划如何成为更广泛的“数据和人工智能向善”

讨论的一部分,以及这一行动是如何被推广使用计算机工具的利益相关者所组成的相互依赖的网络所塑造的。

以上所有论述在它们自身的领域范围内都是有价值的,但是当将倡议置于一个由相似尝试组成的长链条中时,这些论述就不尽如人意了。本文的实证贡献就是开始填补这个空缺并为未来研究提供方向。

(四) 对于“数据和人工智能向善”计划的批判性评估

越来越多的实证研究和理论研究文献从更广泛的社会科学文献中汲取力量,挑战“数据和人工智能向善”倡议会带来更好的社会结果这一观念。我们在此讨论的文献与 21 世纪头十年随着大数据讨论兴起的批判性研究潮流紧密相关,并且站在倡导性文献的对立面。此外,这些文献因为并不提出“社会之善”的替代性定义,所以也与关于“善”的讨论不同。

一些学者指出,“数据和人工智能向善”所追求的“善”经常局限于新计算机技术的属性的狭窄定义,并且这些定义最终为技术的商业开发者利益服务(Magalhães & Couldry, 2021)。这些作者提出,当从更宽泛的角度探讨时,“数据和人工智能向善”倡议不仅无法解决社会问题,而且会使问题长期存在,同时产生新的问题。霍尔茨迈耶(Holzmeyer, 2021)和格林(Green, 2020)对此进行了全面的批评,并且将这一问题与迅速积累的批判性数据和算法研究联系起来。此外,马蒂亚诺(M. Madianou)等人认为,“数据和人工智能向善”倡议使发展援助变成了一个私人化的、商业性的且具有强烈殖民主义暗示的行动,并且最终会带来负面影响(Madianou, 2021; Magalhães & Couldry, 2021)。此处只提及了明确指向“数据和人工智能向善”倡议的研究,但针对 ICT4D 的批判性研究也显现出类似的趋势(De, Pal & Sethi et al., 2018; Schelenz & Pawelec, 2021)。鉴于以上研究的广泛基础,这些研究只是针对当今计算机技术更大规模的批判性研究的冰山一角。当然,针对 ICT4D 的批判性研究为“数据和人工智能向善”倡议

研究提供了一个重要起点,尤其是当这些研究与更大范围的批判性研究——社会学、科学与技术研究以及媒体和传播学——联系在一起的时候。

六、未来的研究道路

正如上文所述:一方面,已有研究文献充斥着要么推广“数据和人工智能向善”的概念,要么讨论其标准基础的正确型构的规范性研究;另一方面,几乎没有描述性研究试图理解“数据和人工智能向善”作为倡议是如何持续影响大众对于计算机技术的理解和使用的。即使影响范围只限于学术界,这些倡议仍然影响研究者对计算机技术与社会的关系、个体研究者与政治的关系的理解。在此一节,我们重点讨论与未来研究相关的理论主题。我们认为,未来研究的方向在于分析性描述和批判性调查,而非规范性研究。换言之,我们呼吁研究者跳出来(take a step back),针对倡议的实践、相似体(parallels)和结果开展理论严谨的分析。今后的研究应该探索新的如“数据和人工智能向善”等私营和非营利组织驱动的倡议与科学社会学(sociology of science)中的多种网络构建策略有何异同(Latour, 1987)。这同时需要与当代的技术政治运动,尤其是那些运动驱动或者如用“数据改善黑人生活”(Data for Black Lives)的聚焦解放的运动进行比较研究。这一路数的潜力在于,将分析学术文献和项目当作倡议本身的一部分,而不是当作一种内部的实质性对话。因此,这就要求对学术和倡议之间的关系有一个元层面的理解。在这篇文章中,我们朝这一方向迈出了第一步。研究者对倡议的经济和话语力量应该给予格外关注。由于对 ICT4D 的批判性研究未能改变技术解决方案主义者(techno-solutionist)计划的进程,研究者需要思考的是:面对不断出现的失败案例,这些倡议是如何为计算技术持续性的正向吸引力做出贡献的?此外,研究者需要分析商业和慈善赞助是如何取代社会科学研究和批判性研究所勾画的战略的。最后,假设“数据和人

工智能向善”倡议无法再获得牵引力,或者被其他类似倡议所取代,研究者应该追问为什么一些倡议失败而另一些倡议成功了。“数据和人工智能向善”倡议除了可以作为现象进行分析,还可以对倡议过程和实践与范围更广的理论主题的互动进行剖析。此类研究的大部分是在“批判数据研究”(critical data studies)(Dalton, Taylor & Thatcher, 2016; Kitchin & Lauriault, 2014)这一涵盖范围广泛的名称下进行的,还有一部分是与其类似的算法、人工智能和数字平台的社会研究。在某些情况下,正如马蒂亚诺所采取的措施(Madianou, 2021),研究者可以直接将既有的 ICT4D 批判研究范围加以扩展,以涵盖“数据和人工智能向善”倡议。我们呼吁将“数据和人工智能向善”主题与现有研究文献联系起来,而非发起一个有独属关键词的新研究项目,主要原因就是现有讨论所具有的力量。我们认为,下一步的研究应该是将与“数据和人工智能向善”倡议相关的飞速发展的研究相结合。

采用结构性视角去研究个案与更大范围的社会技术(sociotechnical)、经济和文化结构的互动会生发许多机遇。大型技术企业和发展援助行业对细致地理解渗透在“数据和人工智能向善”倡议中的文化价值和叙述来说只是起点。第一批对倡议进行批判性评估的部分学者,像麦哲伦、库尔德利和马蒂亚诺,已经朝此方向前进了。这些研究者明确地将“数据和人工智能向善”倡议与权力不对等(power asymmetries)和共同的技术想象力(shared technological imaginaries)联系在一起。因为这一策略是将其研究的实证对象(empirical objects)作为更大的结构性主题,如资本主义、新自由主义、殖民主义或者技术支配的潜在表现。这类研究只是未来就此主题可能进行的大范围研究的一部分。虽然地点选择可能会使理论主题不同,但是不论研究的实证地点选取是在全球北方(Global North)还是全球南方,以 ICT4D 为对象和作为组成部分的研究都是我们所指的这类研究的重要关联体。因此,政治经济学视角和批判话语视角为“数据和人工智能向善”的未来研究与批判数据研究的更广泛工作提供了坚实的理论基础(Dalton,

Taylor & Thatcher, 2016)。数据正义(data justice)(Taylor, 2017)提供了一系列超越结构主义观点的研究策略。

关于“数据和人工智能向善”的新实证研究,也具有修改现有理论或者提出计算机技术和政治新动能的潜力。倡导性研究文献和关于规则伦理的规范性分析都试图在技术和社会的交汇处创造新事物。例如,格林和卡沃斯等学者敏锐地认识到了政治悖论,并且试图找到应对当今挑战的方法。这类工作与关于人工智能伦理的更大范围的研究联系紧密。我们认为,忽视此类工作创造新实践和动向的能力是错误的。因此,有一个问题还有待回答,即“数据和人工智能向善”倡议是对现有的结构性动能(structural dynamics)的复制、调整、超越,还是创造了一个全新概念化的新结构。

另一个需要更多关注的更大的主题是数据科学作为专业的兴起,以及“数据和人工智能向善”倡议是如何表达关于“技术与政治”以及“计算机研究者与政治”的具体想法的。贯穿“数据和人工智能向善”倡议的一个现象是,没有计算机背景的从业人员自称数据科学家——一个直到最近才被明确的职业身份。有学者提出,正在形成的数据科学行业需要进一步对其形成过程进行实证关注(Ribes, 2019)。因为很多“数据和人工智能向善”倡议将“善”定义为应用领域,却将实现“善”的工具排除在外,所以对于数据科学行业本身的研究是非常重要的。树立该领域独立地位的策略已经成为更大范围的数据科学的关键策略(Ribes, Hoffman & Slota et al., 2019)。数据科学从业者声称,计算机技术与任何场景都是相关的,并且“数据和人工智能向善”倡议宣传有些应用领域本质上是“善”的,从而达到使此领域的从业人员受到重视和尊重的目的。从这一角度看,“数据和人工智能向善”计划成了展现一个新行业探索、扩展和合法化自身潜力和局限的示例。

给予数据科学家“社会变革的行动者”这一新身份定位的直接结果是,数据科学家自身无法回避数据科学和人工智能伦理讨论中涉及的价值和政治问题。日益严格的道德审查意味着数据科学家被逐渐塑造成伦理学家。

因此,“数据和人工智能向善”倡议为道德和政治如何在以提升社会福祉为目标的数据科学实践中被协商讨论提供了关键研究领域。实际上,谭维尔(A. Tanweer)展示了数据科学家是如何在工作中面临政治问题时塑造自身的道德立场的(Tanweer, 2018),奈夫(G. Neff)等人展示了在实际项目中他们是如何缓和紧张关系的(Neff, Tanweer & Fiore-Gartland et al., 2017)。随着对道德的日益关注,“数据和人工智能向善”倡议自然成了面对道德挑战的输出口。倡议可以向外表现出数据科学愿意直面批评,并且即使数据科学存在短处,其依然可以是道德的和为善的事业服务的。不过,正如鲍威尔(A. Powell)等人在“技术为善”企业网络中所展示的那样(Powell, Ustek-Spilda & Lehuedé et al., 2022),对于道德实践和规范理念的明确追求可能会包含自相矛盾的情况与设定目标矛盾的做法。对于数据科学行业的更深理解也许会逼迫我们不再将对“善”的定义的规范性研究视为外部评论,而是内生于数据科学行业本身的讨论。这会使数据科学关于道德和规范性本身的讨论成为实证研究对象。同时,因为从业人员和学者的交际重叠日益增长,所以这一转变与实施对象是从业人员还是学者无关。分析规范性理念和正当理由的实用主义转向(Boltanski & Thevenot, 2006)提供了一条以实证方法探索此类发展的研究途径。

我们并不能确定上述方法是相互兼容的。严谨的实证和理论讨论机会由此产生。借助有关计算与社会的更广泛的研究讨论,我们可以进一步推动对“数据和人工智能向善”的研究,而对这些倡议的研究同样可以进一步加深我们对更广泛主题的理解。

七、结论

本文跳出了“数据和人工智能向善”倡议本身,并提出如何对这些倡议在学术内外的发展历程进行理解。这些倡议并不是一时的口号,而是对如

何理解计算机技术与重大政治现象的互动方式有着持续影响。我们剖析了重要案例,凸显了它们与 ICT4D 的关联,并对近期文献进行了评述,最终指出了未来发展之路。

本文的主要贡献是指出了各种各样的“数据和人工智能向善”倡议之间的内在关联,以及它们与其他技术政治倡议的相互联系。这引导我们在研究这些倡议时认识到,它们具有塑造一个技术与政治交汇的未来的能力,并且能够展现出更广阔的理论主题。对于“‘善’是什么”的各种个人定义展开批评已没有太大意义。同样,重复“我们需要针对个人在复杂的真实世界中如何解决问题开展更多研究”的口号也并无益处。我们认为,关于以上二者的回答已经有了充分的文献积累,而我们未来的发展方向应该是对我们的已有理解进行综合和理论概况。此外,我们不认为针对“数据和人工智能向善”进行的研究应该形成专门叙事,就像与 ICT4D 相关联的文献一样。相反,我们应该学习 ICT4D 文献的关键部分,并将其融入更广阔的实证和理论主题。创造一个只关注最新发展的倡导性或者批判性的研究项目,只会僵化规范性和技术性的主张。这就要求我们不再围着这些倡议转圈,而正是出于这一原因,我们需要以实用和分析的姿态跳出来。

参考文献

Avgerou, C. 2010, “Discourses on ICT and Development: Information.” *Technologies and International Development* 6(3).

Berendt, B. 2019, “AI for the Common Good?! Pitfalls, Challenges, and Ethics Pen-Testing.” *Paladyn (Warsaw)* 10(1).

Berendt, B. 2020, “(De)Constructing Ethics for Autonomous Cars: A Case Study of Ethics Pen-Testing Towards ‘AI for the Common Good’.” *The International Review of Information Ethics* 28(6).

Best, M. 2010, “Understanding Our Knowledge Gaps: Or, Do We Have An ICT4D Field? And Do We Want One?” *Information Technologies & International Development* 6.

Boltanski, L. & L. Thevenot 2006, *On Justification: Economies of Worth*, Princeton: Princeton University Press.

Catlett, C. & R. Ghani 2015, “Big Data for Social Good.” *Big Data* 3(1).

Chandy, R., M. Hassan & P. Mukherji 2017, “Big Data for Good: Insights from Emerging Markets.” *Journal of Product Innovation Management* 34(5).

Chui, M., M. Harryson & J. Manyika et al. 2018, “Notes from the AI frontier: Applying AI for Social Good.” https://www.mckinsey.com/~/media/mckinsey/featured%20insights/artificial%20intelligence/applying%20artificial%20intelligence%20for%20social%20good/mgi-applying-ai-for-social-good-discussion-paper-dec-2018.pdf.

Cowls, J. 2021, “‘AI for Social Good’: Whose Good and Who’s Good? Introduction to the Special Issue on Artificial Intelligence for Social Good.” *Philosophy and Technology* 34.

Cowls, J., A. Tsamados & M. Taddeo et al. 2021, “A Definition, Benchmark and Database of AI for Social Good Initiatives.” *Nature Machine Intelligence* 3(2).

Crush, J. 1995, “Introduction: Imagining Development.” in J. Crush ed., *Power of Development*, Florence: Taylor & Francis Group.

Dalton, C., L. Taylor & J. Thatcher 2016, “Critical Data Studies: A Dialog on Data and Space.” *Big Data and Society* 3(1).

De, R., A. Pal & R. Sethi et al. 2018, “ICT4D Research: A Call for A Strong Critical Approach.” *Information Technology for Development* 24(1).

Espinoza, M. & M. Aronczyk 2021, “Big Data for Climate Action or Climate Action for Big Data?” *Big Data and Society* 8(1).

Floridi, L., J. Cowls & T. King et al. 2020, “How to Design AI for Social Good: Seven Essential Factors.” *Science and Engineering Ethics* 26(3).

Green, B. 2019, “‘Good’ Isn’t Good Enough.” *Proceedings of the AI for Social Good Workshop at NeurIPS* 17.

Green, B. 2020, “Data Science as Political Action: Grounding Data Science in A Politics of Justice.” https://ssrn.com/abstract=3658431 or http://doi.org/10.2139/ssrn.3658431.

Heeks, R. 2010, “Do Information and Communication Technologies (ICTs) Contribute to Development?” *Journal of International Development* 22(5).

Heeks, R. 2014, "ICT4D 2016: New Priorities for ICT4D Policy, Practice and WSIS in A Post-2015 World." *Development Informatics Working Paper* 59.

Holzmeyer, C. 2021, "Beyond 'AI for Social Good' (AI4SG): Social Transformations—Not Tech-Fixes—for Health Equity." *Interdisciplinary Science Reviews* 46(1-2).

IBM 2021, "IBM AI for Good Sponsorship Focuses on AI in Service of the Planet." https://www.ibm.com/blogs/journeyto-ai/2021/11/ibm-ai-for-good-sponsorship-focuses-on-ai-inservice-of-the-planet.

ITU 2003, "Declaration of Principles. Building the Information Society: A Global Challenge in the New Millenium." https://www.itu.int/net/wsis/docs/geneva/official/dop.html.

Juech, C. 2021, "Building the Field of Data for Good." in M. Lapucci & C. Cattuto eds., *Data Science for Social Good*, Cham: Springer.

Kitchin, R. & T. Lauriault 2014, "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112.

Latour, B. 1987, *Science in Action: How to Follow Scientists and Engineers Through Society*, Milton Keynes: Open University Press.

Leye, V. 2007, "UNESCO, ICT Corporations and the Passion of ICT for Development: Modernization Resurrected." *Media, Culture & Society* 29(6).

Madianou, M. 2021, "Nonhuman Humanitarianism: When 'AI for Good' Can Be Harmful." *Information, Communication and Society* 24(6).

Magalhães, J. & N. Couldry 2021, "Giving by Taking Away: Big Tech, Data Colonialism, and the Reconfiguration of Social Good." *International Journal of Communication* 15.

Mann, G. & A. Sahuguet 2018, "With Great Data Comes Great Responsibility: How Data Science Can Be Applied to Solve Public Interest Problems Without Losing Its Soul." *Journal of Technology in Human Services* 36(1).

Moore, J. 2019, "AI For Not Bad." *Frontiers in Big Data* 2(32).

Mukerji, M. & U. Chauhan 2020, "A Social Network Analysis of ICTD Conferences (2006-2017)." *Information Technology for Development* 26(4).

Neff, G., A. Tanweer & B. Fiore-Gartland et al. 2017, "Critique and Contribute: A Practice-

Based Framework for Improving Critical Data Studies and Data Science.” *Big Data* 5(2).

NESTA 2015, “Data for Good: How Big and Open Data Can Be Used for the Common Good.” https://media.nesta.org.uk/documents/dataforgood.pdf.

Paolotti, D. & M. Tizzoni 2018, “DSAA 2018 Special Session: Data Science for Social Good.” 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA).

Pieterse, J. 2010, “Digital Capitalism and Development: The Unbearable Lightness of ICT4D.” in P. Kalantzis-Cope & K. Herab-Martin eds., *Emerging Digital Spaces in Contemporary Society*, London: Palgrave Macmillan.

Porway, J. 2011, “Doing Good with Data-Data Without Borders.” https://www.datakind.org/blog/doing-goodwith-data-data-without-borders.

Powell, A., F. Ustek-Spilda & S. Lehuedé et al. 2022, “Addressing Ethical Gaps in ‘Technology for Good’.” *Foregrounding Care and Capabilities: Big Data and Society* 9(2).

Ribes, D. 2019, “STS, Meet Data Science, Once Again.” *Science Technology and Human Values* 44(3).

Ribes, D., A. Hoffman & S. Slota et al. 2019, “The Logic of Domains.” *Social Studies of Science* 49(3).

Schelenz, L. & M. Pawelec 2021, “Information and Communication Technologies for Development (ICT4D) Critique.” *Information Technology for Development* 28(1).

Sein, M. & G. Harindranath 2004, “Conceptualizing the ICT Artifact: Toward Understanding the Role of ICT in National Development.” *The Information Society* 20(1).

Shi, Z., C. Wang & F. Fang 2020, “Artificial Intelligence for Social Good: A Survey.” ArXiv Preprint ArXiv: 2001.01818.

Silverman, D. 2020, *Qualitative Research*, London: Sage.

Tanweer, A. 2018, “Data Science of the Social: How the Practice Is Responding to Ethical Crisis and Spreading Across Sectors.” Doctoral Dissertation, University of Washington, ResearchWorks Archive.

Taylor, L. 2017, “What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally.” *Big Data & Society* 4(2).

Tomašev, N., J. Cornebise & F. Hutter et al. 2020, "AI For Social Good: Unlocking the Opportunity for Positive Impact." *Nature Communications* 11(2468).

Unwin, P. & T. Unwin 2009, *ICT4D: Information and Communication Technology for Development*, Cambridge: Cambridge University Press.

Walsham, G. 2017, "ICT4D Research: Reflections on History and Future Agenda." *Information Technology for Development* 23(1).

Zheng, Y., M. Hatakka & S. Sahay et al. 2018, "Conceptualizing Development in Information and Communication Technology for Development (ICT4D)." *Information Technology for Development* 24(1).

重思深度媒介化时代的社交与关系*

——读《重构关系:数字社交的本质》

何秋红　陈新毅**

摘要:深度媒介化时代数字技术的发展引发了从“个体社会”到“自我社会”的研究视角转向,为在“连接泛在”的时代重新思考人类社会的社交与关系打开了新的视野。邱泽奇教授《重构关系:数字社交的本质》一书,以个体与他人和社会的关系为切入点探索数字社交的本质,为审思数字社交如何影响和重构人类社会的基本关系这一问题提供了新的乐观主义态度。该书立足“连接泛在”的时代特征,系统性地阐释了“自我社会”的理论观点和基本主张,从数字家庭关系、数字朋友关系、数字工作关系和数字生活关系四重视角洞悉复杂的人类社会,以丰富生动的具体案例和鞭辟入里的理论分析对数字技术带来的关系重构以及建基于此的本质变革进行了整体探讨,为重新理解数字化社交和媒介化关系注入学术动力,并为个体如何面对数字时代的关系变革从而抵达心灵空间的幸福、避免走向群体孤独提供现实的指导,无论是在学术研究中还是在现实生活中都具有一定的意义。

关键词:深度媒介化　数字社交　关系　自我社会

自人类跃入现代社会以来,媒介技术的发展与应用就始终和人类社会紧密相连,并在很大程度上影响着人们的日常生活实践和社会文化建构。就像媒介化研究领域的代表学者库尔德利(N. Couldry)所说的,数字化技

* 本文系河海大学中央高校项目“新媒体对少年儿童的影响及应对策略研究”(项目批准号:B230207009)、河海大学中央高校项目“区域传播视域下长三角媒体融合发展的动能与路径”(项目批准号:B220201067)的阶段性研究成果。

** 何秋红,河海大学公共管理学院。陈新毅,河海大学公共管理学院。

术在社会世界的建构中持续发挥的作用改变了社会世界的动态和结构(库尔德利、赫普,2023:19)。作为机械化浪潮、电气化浪潮之后的第三次媒介化浪潮,数字化浪潮掀起了一个深度媒介化的高潮(库尔德利、赫普,2023:64)。媒介化是一个历史的演变过程,其作为与全球化、个人化和商业化等并举的元过程之一,代表着新传媒技术以及与之相伴的不同形态的"中介式交往"的兴起,也是对一个媒介作为基础设施深深嵌入社会现实的时代的准确描述(Krotz, 2007)。可以说,深度媒介化已经成为当今时代的一个基本特征,人类社会的各个方面乃至社会现实的建构本身已经与媒介基础设施密不可分,这意味着学界对当代人类社会的深入分析已经无法绕开媒介这一要素。

在一系列因数字技术的渗透和嵌入而发生的革命性改变中,个体与社会和他人的关系也不可避免地面临着重构。当我们将目光投诸数字技术的社会化过程就会发现:一方面,自互联网诞生之日起不过 40 年,伴随着智能手机的普及和社交媒体的兴起,人类提升物质媒介触达范围和效率的努力让"交往在云端"成为可能;另一方面,在感受到新技术带来的曙光的同时,人们似乎逐渐察觉到数字交往的喧嚣背后个体所面临的孤独与淡漠的隐忧。雪莉·特克尔(S. Turkle)曾这样评价我们拥抱的科技时代:"当我们和机器人谈情说爱,与智能手机难舍难分时,我们通过机器重新定义了自己,也重新定义了我们与他人的关系。"(特克尔,2014:4)当人们面临群体性孤独的警世危言而忧心忡忡时,当人们在数字技术带来的无远弗届的交流中渐渐无所适从时,一切都似乎在印证着一个观点,那就是数字化浪潮之下人类社会的交往与关系也不可避免地媒介化,这种本质性的变革给人们的心灵体验带来了极大的冲击,迫使人们重新思考数字社交带来的社会关系革命。数字社交带来的后果一定是群体性孤独吗?数字社交是实现群体性的可选路径吗(邱泽奇,2024:20)?新的时代境况呼唤新的理论体系,新

的理论体系解释新的时代境况。邱泽奇教授的著作《重构关系:数字社交的本质》(以下简称《重构关系》)适时推出,为观察和理解深度媒介化时代被重构的社会关系提供了新的思考。

长期以来,学界面对层出不穷的新兴媒体往往有两种观点:一种观点认为数字技术带来的中介化互动逐渐解构人际关系的神圣性,并将由此产生的所谓"群体性孤独"视为亟须矫治的社会病态;另一种观点则对数字技术为人类关系的赋能心向往之。两种截然不同的观点共同印证着一个事实,那就是数字技术正在重构人类社会关系的本质。然而,人们如何基于数字社交重构个体与社会的关系?这种关系与传统基于身体同场的个体与社会的关系有何不同(邱泽奇,2024:20—21)?人们该如何认识,又该如何面对这种社交关系的重构?一系列问题摆在深度媒介化时代人们的面前。邱泽奇教授师从费孝通先生,深谙"学术的目的不是用老百姓的实践说老百姓听不懂的话,而是用老百姓听得懂的话说老百姓不曾明白的道理"(邱泽奇,2024:21)。在该书中邱泽奇教授删繁就简,透过四种关系的四种典型场景探讨数字社交带来的社会关系革命,生动地分析了亲密又疏离的家庭关系重构、遥远又邻近的朋友关系重构、熟悉又陌生的工作关系重构以及无形又有影的生活关系重构,案例与理论相互支撑、相得益彰(邱泽奇,2024:1—4)。在笔者看来,该书的学术价值与现实意义主要体现在以下几个方面。

一、"连接泛在"与"自我社会":洞悉复杂社会的关系结构

从媒介化浪潮的历史来看,自机械化浪潮到电气化浪潮再到如今的数字化浪潮,一个贯穿始终的主题就是连接的演进。库尔德利等在《现实的中介化建构》中指出,媒介化浪潮的一个主要特征就是媒介之间日益紧密

的相互关联性。机械化浪潮下古登堡印刷机的出现使得传播媒介通过标准化内容的复制将更广泛的人群连接起来,大规模的传播得以“脱域”;在电气化浪潮之下,电网、有线网络等技术使得各种媒介被连接在一起,让高效率、同步性的人际沟通成为可能(库尔德利、赫普,2023:48—59);这种演进趋势在互联网兴起后的数字化浪潮中尤其明显,甚至有学者认为互联网发展的内在逻辑就是连接的演进。在数字化浪潮之下,不只是机器与机器之间,连接广泛地存在于内容与内容、人与人、人与物甚至物与物之间,整个世界因为连接的广泛存在而成为一张巨大的网络,可以说“连接泛在”成为数字化浪潮的底色。在此情境下,一种被“连接泛在”的社会现实所启发,但又没有在“连接泛在”这一社会事实前停止思考的“自我社会”的观点应运而生。邱泽奇教授认为,“连接泛在是前提,自我社会是主题”(邱泽奇,2024:7)。人类通过广泛的连接进行社会互动,而互动恰恰是基于关系进行的。就这样,该书在“连接泛在”的社会前提下找到了一条洞悉深度媒介化时代复杂社会的学术进路,那就是通过对“自我社会”这一主题的思考对数字技术带来的社会关系变革进行探究,从而理解深度媒介化时代的社会现实和个体境遇。正如邱泽奇教授在该书中所说的:

> 在社会学家看来,社会的结构和动态,影响结构和动态的因素,以及结构和动态的变化演化等,也都是复杂的……不过,一旦我们换个思路,将复杂的社会落实为人与他人和社会的关系,从围绕每个人的关系出发,尤其是从结构关系出发,复杂的社会就变得简单了。(邱泽奇,2024:4)

“连接泛在”为以个体为中心建构“自我社会”提供了前提。自贝克夫妇(Becks)在其著作《个体化》中系统阐释个体化社会以来,“为自己而活”

就成为人们经久讨论的话题,个体作为从社会制度中脱离出来的力量逐渐被重视。贝克夫妇的“个体化”试图分析第二现代性下个体从家庭中脱嵌出来,力图寻求自己的活法这一现象,为理解后工业时代个体与家庭的关系提供了思考。在后工业时代,人们不再依靠家庭而渐渐成为自己生活的主宰,社会也越来越关注和尊重个人的需求和权利。然而跃入现代性后的个体就像“钢丝上的舞者”,这一隐喻形象地再现了人们所面临的生活方式的个体化以及这种生活方式的转变带来的“不确定的自由”(贝克、贝克-格恩斯海姆,2011:2)。值得肯定的是,个体化社会中个体作为能动的力量而被看见,个体与社会之间的冲突被讨论,他们有从传统的制度结构中脱离出来追求自己的活法的诉求。但从本质上而言,个体化社会作为一种理论仍没有突破“个体围绕社会”这一社会学研究传统。而这种研究传统在如今的深度媒介化时代已经失去了部分解释力,因为“连接泛在”的深度媒介化时代中,是个体主动建构社会而不是个体被动围绕社会。虽然社会对个体仍具有外在性和强制性,但是在数字化浪潮之下,社会资源的组织开始以个体为中心。个体已不仅仅是个体化社会意义上的从传统社会制度中脱离,而是能动地运用数字化技术组织资源,积极建构属于自己的网络圈群(王阳、张攀,2018)。在朋友关系方面,数字技术的发展和社交媒体的进步使得自我可以轻而易举扩展自己的社交范围,缔结朋友关系的后备人群规模越大,满足自我偏好的选择性就越强。为满足自我偏好,自我不得不继续扩大搜寻范围,以至于朋友关系的规模越来越大;加之自我精力的有限性,朋友这种“二人关系”属性带来的满足彼此诸多偏好的完整属性越来越弱,只满足其部分偏好的属性越来越强,朋友关系从“完整朋友”被稀释为“部分朋友”。这个结果的极端就是,朋友变成了搭子。这种个体围绕自我需求搭建起的关系网络虽然高效且服务于自我需求,但同样也易碎(杨鑫,2024)。在如今的平台社会之中,个体所遵从的社会秩序似乎不再仅仅依靠制度性

的社会事实所建构和维护,各大商业平台“流量为王”的运作逻辑使得数据成为一种新型的社会知识,掌握了数据的商业平台就像掌握了“普遍的真理体系”,从而构建起新的社会秩序。这种流量却不再是由社会赋予和建构的,而是互联网时代的一个个节点化的个体自发投票和选择的结果。这种以个体为中心进行的资源组织颠倒了传统社会学研究中个体与社会的关系,使得个体而不是社会成为自我的中心。在这个意义上,该书的“自我社会”概念就在贝克等人的“个体化社会”概念上再向前一步,更能代表深度媒介化时代个体与社会关系发生的本质性变化,为洞悉深度媒介化时代社会关系的变革提供了理论工具、拓展了研究思路。

德里达(J. Derrida)用一句“文本之外,别无他物”阐明了他对现实的理解和对文本的重视。他提醒我们,所有的社会现实在本质上都是文本,我们需要对文本进行深入的解读和诠释才能更好地理解社会现实(多伊彻,2021:10)。在此基础上,海德格尔(M. Heidegger)引申出“技术之外,别无他物”阐明了他对媒介的理解,认为在技术这种起整合或框定作用的逻辑之外,并无其他空间。在技术空间中,我们对技术的应对之道应该是泰然处之(贡克尔、泰勒,2019:193);在海德格尔现象学追问的基础上,基特勒(F. Kittler)主张主客体之间的关系由中介建构而成,在建构这种关系的同时,中介也建构了自身作为中间渠道的媒介这样一种位置。由此,基特勒提出“媒介之外,别无他物”的观点,对应主客体二元论,指向媒介的生成性,意图构建一种新的本体论观点。胡翼青教授则认为,“框架之外,别无他物”,指出一切行为皆因框架的生成而生成,即使是精神病患者的反常行为也与框架有关(胡翼青,2023)。而今在这个“连接泛在”的时代,我们甚至也可以说,“关系之外,别无他物”。就像邱泽奇所说的那样:“归根到底,人是在关系中的。”当我们将研究视角从围绕社会的个体转向个体指向社会时就会看到,关系是个体认识社会、理解社会和建设社会的焦点,也是个体认识

自我、建构自我、体验自我、感受自我的焦点,还是自我认识他人、认知社会、建构心灵的焦点(邱泽奇,2023:29)。该书以关系视角为切入点,试图剖析深度媒介化时代复杂的人类社会动态变革的结果。人毕竟是社会的产物,关系似乎可以成为透视复杂社会的刻度尺和理解变动的社会形态运作的观测点,如今我们也必须对关系进行深入的解读和诠释才能更好地理解深度媒介化时代的社会现实。而从"自我社会"中看关系,如今数字社交带来的关系变革,本质上正是从个体围绕社会的社交到个体汇聚关系的社交的转变。

二、群体孤独或心灵幸福:重构社会关系的批判进路

数字技术的进步将人类带入一个更为便利的社交媒体时代。在技术赋能之下,交流逐渐突破物质媒介特别是身体的束缚,使得"缺席的在场"和"在场的缺席"都成为可能(拜厄姆,2020:3—4)。"连接泛在"之下,人们可以通过手机和远在他乡的亲友视频通话,也可以在一场枯燥无聊的会议上"精神出走"。在数字技术带来的无远弗届的便利性之下,雪莉·特克尔在《群体性孤独:为什么我们对科技期待更多,对彼此却不能更亲密?》一书中向社会发出了警世危言:数字技术带来的不仅是经济泡沫,还有情感泡沫。"群体性孤独"关注到数字技术给人类心灵与精神的重创,在此后的学术研究中,特克尔逐渐确立了自己的研究旨趣,那就是用批判的眼光审视数字技术的负面影响,警醒人们在享受数字技术带来的便利的同时,也要了解其坏处。这一观点无疑给正处于深度媒介化进程中的人们带来了很大的启发,邱泽奇教授也是其中之一。受特克尔的启发,邱泽奇教授对"自我社会"和"连接泛在"的思考逐渐有了一套整体的叙事逻辑,从心灵的角度出发进一步系统思考"连接泛在"的本质影响。然而学术研究的过程毕竟是对话的

过程,邱泽奇教授该书另一个重要的学术意义所在,就是通过批判特克尔的“群体性孤独”观点,重新思考数字技术的发展对人类社会关系的本质影响,重构新媒体对人类社交关系研究的批判进路。

频繁使用社交媒体是否会加剧孤独感,这一直是一个充满争议的学术问题。数字技术的发展带来的究竟是一种病态的社交还是更具可能性的社交,这也是该书致力于讨论的话题之一。有学者通过实证研究发现,社交媒体的使用并不一定造成群体性孤独,对孤独感的影响取决于线上社交与线下社交的关系。据 2017 年中国综合社会调查数据的分析结果,频繁使用社交媒体促进了民众与亲朋好友的线下交往和亲密关系,因而对孤独感有明显的缓解作用(许琪、付哲,2023)。这一发现似乎和邱泽奇教授对“群体性孤独”的批判不谋而合。邱泽奇教授认为,“群体性孤独”只是呈现了一部分人在数字社交中的焦虑,并不能刻画数字社交给整体的人类心灵带来的负面影响。这一观点“最大的假设谬误是把现实世界和虚拟世界对立分隔”。由此,作者进行了一连串的发问:生活在身体同场的群体里,是人所必需的吗?身体同场的群体性是社会属性还是自然属性?特克尔认为,身体同场的群体性是自然属性,脱离身体同场的群体性会让人感受到孤独,她因此主张回到身体在场的群体性,通过“重拾交谈”找到“群体性孤独”的解药,对数字技术带来的社会关系变革持警惕、悲观态度。而邱泽奇教授则对此观点进行了再思考,指出身体同场的群体性如果是社会属性,就是可以改变的,不应该被简单列为社会病态。的确,正如作者所发问的那样:“一个普遍存在的社会现象,怎么可以被简单诊断为需要矫正的社会病态呢?”(邱泽奇,2024:16)

除了身体同场,数字同场也是一种新的同场方式。特克尔认为,应该回到身体同场的群体性,通过面对面的交谈实现对“群体性孤独”的疗愈(特克尔,2017:16—21)。“重拾社交,走出永久在线的孤独”,这一观点在如今

深度媒介化时代一切都已然中介化的背景下显然是一种不切实际的主张，这种对于桃花源的向往是一种乌托邦式的幻想，是一种“数字归隐”的避世思想。这种数字归隐往往以失败告终，一次又一次确证着媒介化的深度。因为在深度媒介化时代，人类早已无法脱离媒介，人们就居于媒介之中，与其向往以断开连接的方式重回面对面的交流，倒不如寻找新的同场方式。正如《交往在云端：数字时代的人际关系》的译者所言：“回顾人类历史，几乎每一种全新的社交媒体出现在人们生活中时，我们都难免会产生极端化的应激反应，要么认为它是洪水猛兽，要么认为它是救世灵药……这些简单粗暴的憎恶与赞美，都在人类历史上不断上演，同时也反复被证伪。”（拜厄姆，2020：4—5）因此，邱泽奇教授提出除了“重返身体同场”，还可以通过“数字同场”的新形式实现群体性，这或许就是在特克尔“重拾交谈”的书名之上，他将此书命名为“重构关系”的原因之一。如果说特克尔的主张是一种对身体的复归，那么邱泽奇教授则主张一种自我和关系的复归，也就是“自我社会”。该书第一章就回答了这一问题，并在随后的章节中通过丰富的案例不断印证这一主题。从长远来看，数字社交的重构带来的可能并不是群体性孤独，又或者说这只是对其中一部分的一种可能性结果。乐观地看，数字社交带来新的可能性，带来形塑人类新生活的机会，个体的价值和意义将越来越多地体现于“自我社会”。至此，邱泽奇教授通过与《重拾交谈》《群体性孤独》的对话，站在“自我社会”的背景下数字化技术重构社会关系这一立场，为重新思考深度媒介化时代数字技术尤其是社交媒体对社会关系的影响建立了批判的学术进路。

三、身体同场与数字同场：拥抱数字社交的理性立场

虽然人类对高效率交流的追求史就是不断利用技术挣脱身体束缚的历

史,但是身体在人类的沟通与交流中依然具有十分重要的意义。在传播研究中,身体与技术的关系问题经历了从离身到具身的发展过程(芮必峰、孙爽,2020)。彼得斯(J. Peters)在《对空言说:传播的观念史》中指出,身体是追求理解的方式,纵使交流充满沟壑,但意义就在于身体共同在场的回归。虽然肉体上的“手拉手”并不一定导致精神上的“心连心”,但身体的触碰本身就有意义(彼得斯,2017:376—385)。从视频通话到“元宇宙”概念的火爆再到 Apple Vision Pro 的问世,人们不断通过技术模拟身体在场以缓解交流过程中身体缺席带来的焦虑。在一切都被媒介化的当下,身体仍因为无法被完全数据化而具备不可化约性,人类至少因为肉体的这一特征而保有“连接”还是“断开”的主动权,这使得人类能够从虚拟世界回到现实世界。由此观之,在一切都已然数据化的今天,物质性的肉身作为仍然没有办法被数据化所替代的东西,其受到嫌弃的顽固或许并不是真正的落后,反而是人类逃离技术统治时的唯一退路,也是人类交流最后的归处。但是对于身体同场意义的强调不应遮蔽对于数字同场意义的认可,更不应导致人们用完全悲观的态度看待数字社交对社会关系的重构。因此,邱泽奇教授指出,数字社交也是满足人类群体性需求的有效路径。这一观点在肯定了身体同场的意义所在的同时,又避免了对于身体同场的过度强调所导致的对于数字同场意义的忽视,而这种忽视在数字同场已然成为并列于身体同场的存在方式的时代境况下显然是有些不合理的。

由此可见,长期以来学术界有关身体对传播和交流之意义的强调虽颇有建树,但与此同时,对于身体意义的强调似乎以一种强势的力量遮蔽了其他可能。正如邱泽奇教授指出的:“身体同场固然有其意义,但是其只是满足人类群体性需求的路径之一,而不是唯一路径。”(邱泽奇,2024:16—17)因此,不能简单认为数字社交带来的对身体的远离必然以损伤人们的群体性进而危害人们的心灵体验为代价。相反,在深度媒介化时代,数字同场已

然和曾经的身体同场一样成为一种社会现实,媒介甚至以一种透明且流动的“入身”状态影响着身体的可能性,因此不能将视野停留于身体的意义,数字同场同样具有其独特的意义(李一君、李舒霓,2023)。在这里,邱泽奇教授重构了看待深度媒介化时代人类社交与关系的新立场。不同于以特克尔为代表的学者对数字社交给人类社会关系造成影响的悲观和警醒态度,邱泽奇教授立足深度媒介化时代的特征,以乐观的态度看待数字技术带来的社会关系重构,给以“自我社会”为主题的乐观主义立场通过肯定数字同场的意义与价值缓和了面对数字化浪潮无所适从的人们内心深处的焦虑,让对于技术的社会影响的思考更加理性。

于是我们在该书中看到:何瑜借助电子邮件获得了建立亲密关系的基础,得以将身体隐藏于数字之后开展“码上恋爱”(邱泽奇,2024:32—42);年轻人围绕社交平台聊天记录的“谍战亲密”式恋爱表征着人们运用数字技术重构亲密关系,进而成为重构人类社会关系的一个小小进程(邱泽奇,2024:43—52);许俊熙利用互联网进行语言 cosplay,以满足自己多元的追求与梦想,在数字技术的帮助下发现了更多试演舞台(邱泽奇,2024:94—103);口吃的陈妍通过舞蹈社的微信群减少缺陷暴露的机会,获得跳出缺陷困境的机会(邱泽奇,2024:104—112);不善经营朋友关系的齐学工教授,通过给同事子女进行线上补习激活了自身僵化的社会资产(邱泽奇,2024:113—123);高考失利的程慕欣运用在线社群在工作中突破科层岗位的关系束缚,发现了自己的价值(邱泽奇,2024:157—166);立志永不“鸡娃”的齐薇夫妇在深入日常细节的群体参照中不得不“顺应潮流”(邱泽奇,2024:232—242);双目失明的裴子欣以在线上微信社群听电影的方式构筑起自己部分完美的生活世界(邱泽奇,2024:243—253)。一个个生动而具体的案例,一个个鲜活而立体的人物,都在佐证着作者的观点:社会制度保障个体的日常生活,为生物性生命的维系提供支撑;数字技术又赋予个体连接任意

陌生人的机会,剩下的便取决于个体期望构建怎样的关系,又期待这种关系给自我带来何种体验,进而建构一个怎样的心灵。从这个意义上看,数字化技术影响下的个体似乎并不完全是孤独的,而是有机会热闹的、幸福的和温暖的。至于是通向孤独还是热闹,则完全取决于个体的自我选择。

四、结语

到这里,“自我社会”这一概念的内涵也因为案例分析和理论阐释的交相铺陈而越发清晰起来。在这个数字化技术赋能之下个体完全有机会也有能力“为所欲为”的“自我社会”中,在孤独的另一面,人们需要考量的应该是在面对无限可能和无法积累的社交关系时,需要有一个怎样的心灵空间才能获得幸福(邱泽奇,2024:292)。对于这个问题,邱泽奇教授指出应该建构生活关系的积极性和包容性。由于该书的侧重点是分析数字技术带来的社会关系变革,因此对于这个问题的回答不是主要内容,这里邱泽奇教授只是在全书的延伸中给出了一个简单的答案。当然,如刘海龙教授所言,社会关系的变革并不仅仅是媒体的问题,也是嵌入在新自由主义和消费主义深度发展过程之中的问题。因此,如何在一个更大的语境和视野中理解媒介对人际关系、亲密关系的影响,尤其是对中国式关系的影响,是中国传播学者除了庙堂话题外更应该关注的问题(刘海龙,2023)。在更为广阔的维度上,不少学者从自己的研究领域出发,对于上述问题给出了不同的见解。数字媒体的出现使得公共领域和私人领域的界限发生了混淆和变迁。基于此背景,吴飞教授提出“数字共通”概念,将研究目光从个体的关系拓展到更广阔的关系中,给出了一种在数字化时代重新思考关系的新角度(吴飞、傅正科,2023)。社交媒体时代,在关注自我和关心远方的两极中不断摇摆却忽视了附近的年轻人引起了人类学家项飙的关注,他提出的“重建附近”

也获得了很多年轻人的共鸣(项飙、康岚,2023)。这一观点与邱泽奇教授关于“建构生活关系的积极性”有彼此呼应之处。时至今日,当数字技术以不可阻挡之势嵌入人们的日常生活中,当人们在社交媒体的使用中逐渐感到虽然心灵冲破身体的束缚得以远游,但似乎逐渐迷失了方向时,也不妨从“附近”找到一点改变的力量。

无论是忧虑数字社交给我们带来的群体性孤独,还是对新型社交与新型关系抱以乐观的态度,都是分析深度媒介化时代数字社交本质的一个视角,其作用都在于在思想的交流与碰撞中不断完善人们对这个媒介化时代的认知。从关系的角度来审思深度媒介化时代,可以发现无论何种类型的关系,都不可避免地受到媒介的深刻影响。《重构关系》一书通过对“自我社会”概念的深度诠释,深刻分析了数字化技术影响之下人类社会关系的本质性变革,并为深度媒介化时代的人们如何抵达心灵幸福指出了一条可能的道路,也为以更为理性的学术态度看待新媒体技术对社会关系的重构这一学术议题开了一扇窗。正如潘忠党教授在谈及媒介化趋势时所指出的那样:

> 生活在“媒介化”的社会并不意味着我们成为传媒技术的“臣民”,而是意味着传媒技术成为我们的随从和伴侣,在被界定的场景,我们用它们作为表演的道具;并不意味着我们只能顺从媒介的力量或者掌控媒介者的权力,而是意味着我们可能以传媒技术为资源,改造时刻笼罩、制约着我们的结构和意识形态体系,或至少改造我们与它们之间结构性关联的形态;并不意味着媒介及其技术成为推动历史发展的主角,而是意味着我们可以运用传媒技术建构——创造和再现——更加丰富多彩的历史篇章。(潘忠党,2014)

个体与他人和社会的关系之重构,并不一定以通向群体性孤独为后果;相反,它为人们以此为契机重新建构心灵幸福提供了新的便利。将新媒体的社会关系影响研究从批判的旋涡中抽离,或许是深度媒介化时代媒介已然成为人们的生活空间的现状下更为具有建设性的学术进路。因此,《重构关系》一书的问世不仅让学界重新思考深度媒介化时代数字社交的本质,也让时代潮流下的个体对以何种态度应对深度媒介化时代有了新的思考。无论是在学术研究中,还是在现实生活中,该著作的面世都有较为重要的意义与价值。

参考文献

拜厄姆,2020,《交往在云端:数字时代的人际关系》,董晨宇、唐悦哲译,北京:中国人民大学出版社。

彼得斯,2017,《对空言说:传播的观念史》,邓建国译,上海:上海译文出版社。

多伊彻,2021,《如何阅读德里达》,夏开伟译,北京:北京联合出版公司。

贡克尔、泰勒,2019,《海德格尔论媒介》,吴江译,北京:中国传媒大学出版社。

胡翼青,2023,《框架之外,别无他物》,杨馨:《框架分析:经验组织论》,北京:北京大学出版社。

库尔德利、赫普,2023,《现实的中介化建构》,刘泱育译,上海:复旦大学出版社。

李一君、李舒霓,2023,《流动的具身性:由 ASMR 视频的作用机制再思具身传播》,《现代传播(中国传媒大学学报)》第 9 期。

刘海龙,2023,《媒介技术与人际关系的异化》,《全球传媒学刊》第 6 期。

潘忠党,2014,《“玩转我的 iPhone,搞掂我的世界!”——探讨新传媒技术应用中的“中介化”和“驯化”》,《苏州大学学报》(哲学社会科学版)第 4 期。

邱泽奇,2024,《重构关系:数字社交的本质》,北京:北京大学出版社。

芮必峰、孙爽,2020,《从离身到具身——媒介技术的生存论转向》,《国际新闻界》第 5 期。

特克尔,2014,《群体性孤独:为什么我们对科技期待更多,对彼此却不能更亲密?》,周逵、刘菁荆译,杭州:浙江人民出版社。

特克尔,2017,《重拾交谈》,王晋、边若溪、赵岭译,北京:中信出版社。

王阳、张攀,2018,《个体化存在与圈群化生活:青年群体的网络社交与圈群现象研究》,《中国青年研究》第 2 期。

吴飞、傅正科,2023,《“数字共通”:理解数字时代社会交往的新假设》,《新闻与传播研究》第 6 期。

项飙、康岚,2023,《“重建附近”:年轻人如何从现实中获得力量?——人类学家项飙访谈(上)》,《当代青年研究》第 5 期。

许琪、付哲,2023,《群体性孤独——社交媒体使用对孤独感的影响》,《华中科技大学学报》(社会科学版)第 2 期。

杨鑫,2024,《青年“搭子社交”现象:缘何流行、风险审视与应对策略》,《思想理论教育》第 2 期。

Krotz, F. 2007, “The Meta-Process of ‘Mediatization’ as A Conceptual Frame.” *Global Media and Communication* 3(3).

平台即工厂:数字泰勒主义、劳动裂变激增与基础设施空间

——评莫里茨·奥腾立德《数字工厂》

蔡振华*

摘要:随着 ChatGPT 等人工智能技术的迭代升级,数字平台的智能化趋势备受关注,“机器取代人类”的忧思亦纷至沓来。然而,德国学者莫里茨·奥腾立德在其新著《数字工厂》中却指出,以算法管理为核心的数字平台延续了传统工厂体制中的劳动控制手段,机器非但不会取代人类,反而需要更多劳动力来维持平台资本的弹性积累。借由“数字泰勒主义”“劳动裂变激增”与“基础设施空间”三个核心概念,奥腾立德提出“平台即工厂”这一关键结论:作为组织人力劳动的数字基础设施,物流、游戏、众包及社交媒体平台持续吸纳异质、灵活的海量活劳动,通过独立承包商、计件工资等形式转嫁劳资矛盾,日益形塑碎片化、不稳定的全球劳动空间,其间种族、阶层与性别的不平等机制得以巩固强化。本文认为,奥腾立德敏锐地洞察到数字时代劳动制度的历史延续性,通过强调人与机器互嵌,奥腾立德不仅深刻省思技术替代论的现实根基,同时呼吁研究者关注光鲜技术话语背后被遮蔽的边缘劳动者。

关键词:物流　游戏　众包工作　社交媒体　虚拟移民　零工经济　幽灵工作

1916 年,美国伯利恒钢铁公司(The Bethlehem Steel Company)入驻马里兰州的巴尔的摩,在麻雀角(Sparrows Point)购买钢铁厂、房产甚至农田,历时半个世纪,将这里打造成人口逾 10 万的工业城镇,命名为邓多克(Dun-

* 蔡振华,复旦大学新闻学院。

dalk)。彼时钢铁厂实力甚巨,雇佣 10 万工人,为金门大桥打造钢铁结构部件。第二次世界大战期间,工厂每天能生产一艘船。工厂的这种高效率,与科学管理之父泰勒(F. Taylor)实施的一系列严控劳动的手段密不可分,“泰勒主义”由此得名。而如今,伯利恒公司已成历史,“主厂区的 5 座巨型高炉曾是北美工业制造业实力的象征,如今却成了娱乐区的背景”(奥腾立德,2023:226)。

这一今昔对比似乎暗示了工业制造业乃至工厂的衰落。但若仔细探察,便会发现平常的言语之间却潜藏着值得玩味的“工厂”隐喻。比如,中国当代诸多大学毕业生仍然渴望进入“互联网大厂”。当他们将字节跳动、腾讯、阿里巴巴等前沿平台公司视作“大厂”时,究竟意味着什么?德国学者莫里茨·奥腾立德(M. Altenried)所著《数字工厂》(*The Digital Factory: The Human Labor of Automation*)或可成为讨论的起点。他意图追问:当伯利恒邓多克工厂被亚马逊配送中心替换时,“工厂”真的消失了吗?

奥腾立德目前供职于柏林洪堡大学欧洲民族学研究所和德国融入及移民实证研究所研究院,主要从事互联网政治经济学、数字劳动及城市空间下的平台劳动等研究。自 2022 年此书英文版由芝加哥大学出版社出版以来,他一直笔耕不辍,围绕此书的三个关键概念——“数字泰勒主义”“劳动裂变激增”及“基础设施空间”——建构更精细的理论框架。本文以奥腾立德的《数字工厂》为核心,剖析从“工厂”视角理解物流、游戏、众包工作以及社交媒体的独特价值。

除却“引言”“结论”与“后记”,全书共 4 章,分别考察全球工厂(物流)、游戏工厂(游戏)、分散的工厂(众包工作)以及隐形的工厂(社交媒体)。奥腾立德开宗明义:“本书的中心方法是思考工厂的延续形式,以理解当代数字资本主义。”(奥腾立德,2023:6)依托民族志等方法开展案例研究,作者采访了“亚马逊仓库的工人、柏林机场等其他场所的物流工人、游戏行业和众包工作平台的不同工人群体、社交媒体的内容审核员等”(奥腾

立德,2023:17),前后历时 7 年完成此著。下文将首先澄明“工厂”的当代性,旋即厘清“平台即工厂”这一关键结论背后的概念建构,进而逐章选介核心论据与观点,最后总结作者以“旧”为新思路背后的视角创新与人文关怀。

一、“工人离开谷歌村”与作为系统的工厂

引言开篇,奥腾立德提及艺术家威尔逊(A. Wilson)于 2011 年创作的视频《工人离开谷歌村》(*Workers Leaving the Googleplex*)。该作品左右分屏,显示了两类截然不同的“谷歌人”:左侧三三两两的员工进出主建筑,或搭乘免费单车、豪华摆渡车,或在自助食堂享受美食;右侧的员工则成批离开办公楼,挂着不同于正式员工(白色工牌)的黄色工牌——他们是谷歌颇具争议的“现存图书数字化”项目的“扫描操作员”(ScanOps),这些工人每天凌晨 4 点上班,下午 2 点 15 分准时离开(奥腾立德,2023:4)。很明显,谷歌并不希望黄牌员工与其他雇员聊天,彼时任职于谷歌视频部门一家承包公司的威尔逊,最终因拍摄和采访这些员工而遭到解雇。这从侧面说明,上述判若霄壤的工作境况似乎恰是谷歌意欲轻描淡写的双重现实,这显然可能引起社会各方对公司劳动力结构合理性的质疑。

作者认为,威尔逊的分屏设计无疑参考了两部影片——卢米埃(L. Lumière)的《工人离开卢米埃工厂》(*Workers Leaving the Lumière Factory*)、法罗基(H. Farocki)的《工人离开工厂》(*Workers Leaving the Factory*)——后者“将不同地点、不同历史时期的工人离开工厂的画面与卢米埃工厂的镜头拼接在了一起”(奥腾立德,2023:5)。也许正是创造手法的贴近,启发了奥腾立德关注“工厂”的当代形式。毕竟,“从马克思《资本论》中的英国工厂,到成为整个资本主义时代代名词的福特工厂,一个多世纪以来,工厂也是众多关键经济、社会理论和政治实践的核心”(奥腾立德 2023:6)。那

么,究竟可以如何理解“工厂”?

奥腾立德认为,“工厂”不只是为工人生产劳作而架设的大型建筑,还是“组织和管理生产过程与活劳动力的系统”(奥腾立德,2023:9)。在此,他援引马克思的论述,指出工厂中工人与机器关系的翻转:“在手工业与制造业中,是工人利用工具;而在工厂中,是机器利用工人。”(奥腾立德,2023:9)为把握当下主要由数字平台驱动的劳动力转型,奥腾立德使用了如下三个核心概念——数字泰勒主义、劳动裂变激增、基础设施空间——三者相互嵌套、彼此支撑,共同指向奥腾立德得出的简明结论——平台即工厂。

二、平台即工厂:数字泰勒主义、劳动裂变激增与基础设施空间

笔者认为,奥腾立德将平台视作“工厂”,意在捕捉工厂制度在当下数字平台运作中的适应性。也即强调,经由算法等基础设施架设组织的数字平台,其实隐而不彰地延续了传统工厂的诸多运转逻辑,同时涌现出不同以往的新型劳动趋势。为了说明当前数字平台劳动的制度基础,作者调用了如下三个相互关联的分析性概念。

第一,“数字泰勒主义”(digital Taylorism)被用来说明传统泰勒主义体制的延续性。本文开头提及泰勒制控制劳动的诸种手段,主要包括“计划与执行的分离、任务的分解与标准化以及对工人的精确监督”,它们不仅有助于“工作流程的合理化”,还能“系统性打破工人抵抗的尝试”(奥腾立德,2023:225)。数字泰勒主义并非“20 世纪科学管理的简单重生或延续,而是被用来概念化数字技术如何在新的形式和环境下调动、更新和重组泰勒主义”,它涉及劳动领域一系列的发展与变革,“包括标准化、过程分解、降低技术要求、自动管理和人类运算、算法合作、数字测量以及劳动监督”(奥腾立德,2023:228)。可以说,泰勒主义牵涉的劳动控制在数字平台语境中呈

现出新的特征:监控、引导与评估等各劳动环节皆得到更精细的测量与优化。

奥腾立德指出,数字工厂与产业工厂的关联在于,“二者都属于生产基础设施,都以组织生产过程、劳动分工以及(事无巨细地)控制与约束活劳动力的各类技术为特征”,但由于数字技术所交转的人力物力配置差异,前者“拥有更多的协调性与准确性”(奥腾立德,2023:15—16)。更恰切地说,“任务的标准化、算法管理的手段、对劳动组织过程的监督以及对结果和反馈的自动评估,允许数字工厂通过多种方式接纳形形色色的工人”——“接纳”而非“产生”大批产业工人,被作者视作数字泰勒主义区别于传统泰勒主义的核心所在(奥腾立德,2023:12)。

第二,如果说数字泰勒主义旨在透视劳动控制内部运作的接续与翻新,那么“劳动裂变激增”(multiplication of labor)[①]则重在省察劳动结构与趋势的整体性变迁。梅扎德拉(S. Mezzadra)与尼尔森(B. Neilson)提出“劳动裂变激增”概念以补充“劳动分工”(division of work)概念,具体包括三重意涵:首先,“数字技术和自动化管理的标准化工作程序允许工人的快速吸纳与替代”,使得多元异质劳动者的全球临时合作成为可能;其次,这种临时合作意味着人们的劳动时间与自由时间愈加模糊,个人往往需要从事多份灵活但不稳定的工作;最后,劳动裂变激增内蕴“劳动力地理分布和流动性的明显异化”“性别分工的重构”以及“短期、分包、自由职业和其他形式灵活非常规就业的激增”(奥腾立德,2023:13—14)。可见,劳动裂变激增不只描述出当前劳动力异质、灵活与多重之属性,亦揭示出与其相互映照的全球不平等劳动结构的进一步碎片化与复杂化。

由此,奥腾立德一针见血地指出,数字资本主义“不是通过工厂的终

① 笔者主张将 multiplication of labor 译为“劳动裂变激增”。原因在于,将之译作“劳动力倍增”(《数字工厂》中译本译法)似乎仅意味着劳动者呈指数级增长,但从两位提出者的论述可知,multiplication of labor 强调目今劳动的异质性、灵活性、不稳定性与多重性,因而译作“劳动裂变激增”更能传达数字技术动荡下整体性劳动结构的多层次变迁,其中涵括劳动者数量的增加。

结,而是通过其转型、倍增与泛化来实现”(奥腾立德,2023:229)的。借助数字技术无远弗届的时空辐射力,平台公司得以操持全球化、无间断的标准化劳动进程,它们“能将劳动过程的紧密组织与合同的灵活性或按需而定的劳动形式相结合”(奥腾立德,2023:232)。在下一部分的案例介绍中不难发现,通过诸如零合同制、计件工资等控制手段,劳动速度与强度直接与工资挂钩,借此“劳动剥削所固有的一部分冲突就转移到了个体工人身上”(奥腾立德,2023:234),从而愈发有利于数字工厂资本的弹性积累。

第三,作者采纳伊斯特灵(K. Easterling)的分析方法,侧重分析数字基础设施对“生产空间和劳动力地理分布”的深刻形塑,也即关注基础设施空间的生成。奥腾立德指出:“如果数字技术能让工厂作为一种劳动制度摆脱混凝土的建筑形式,那么数字工厂就可以采取多种不同的空间形式。”(奥腾立德,2023:14)因此,“关注劳动力的迁移、性别构成和其他(空间)分层与碎片化的新旧形式至关重要”(奥腾立德,2023:15)。从全书的讨论来看,奥腾立德特别关注“虚拟移民”背后种族、阶层、性别等因素的“叠变”(intersectionality)效应(吉登斯、萨顿,2023:147—150)。

当然,奥腾立德同样关注数字资本主义的物质性(materiality)面向。他明确指出:“与‘无重量’经济或‘虚拟’经济不同,资本主义的数字转型其实是一个深刻的物质过程。”(奥腾立德,2023:19)此外,他还梳理了“意大利推崇工人阶级利益社会观的马克思主义”(奥腾立德,2023:22)以及“非物质劳动”等理论资源。他表示,自己的研究建基于诸多学者的如下认知:“数字化催生的劳动条件不以自由交流和创造力为标志,而是以降低对工人的技术要求、程序化和控制为特征。”(奥腾立德,2023:24)上述数字工厂的多重面向在全书 4 章渐次铺开、时时呼应。

在此之前,我们有必要梳理上述三个核心概念的内在关联。[①] 笔者认

① 此处笔者与作者奥腾立德进行了邮件联系,确保对核心概念关联的认识没有偏离原意。

为,平台作为“新”工厂,关键在于当前组织劳动的数字基础设施背后以算法管理为核心的技术运作逻辑。该逻辑与“旧”的劳动控制及结构相互勾连,日益形塑出不同于以往的碎片化劳动制度。该制度以数字泰勒主义愈加精准、灵活的劳动过程控制为要义,不断助推劳动裂变激增——个人更可能从事多份不稳定的工作,全球临时合作网络渐趋异质化,其中种族、阶层与性别的不平等机制依然凸显——平台似乎同样承担着“旧”工厂社会空间的功能(奥腾立德,2023:152)。

三、物流、游戏、众包工作与社交媒体中的“工厂”幽灵

第 1 章“全球工厂:物流”说明了在集装箱化与数字化的趋势中,物流行业如何被物质性地架设为一套服务于公司精准控制劳动并确保自身免于承担责任的坚固体系。作者主要围绕亚马逊在布里斯朗(Brieselang)的“订单履行中心”(巨型仓库)展开论述,该中心主要负责存储、分拣与发送各种货物。相比于航运软件、企业资源规划系统(ERP)等基础设施,履行中心里临时员工感受最深的物流工具也许是条码与扫描仪。因为不论是入仓月台、接收区,还是存放区,“都要受到条形码逻辑的控制”(奥腾立德,2023:51)。以最艰难的拣货工作为例,“在扫描仪节奏的指引下,一些工人每次轮班要走超过 20 千米的路程”(奥腾立德,2023:55)。借助“以时间为基础的仓库行动图”专利,扫描仪得以负载倒计时功能,工人必须在计时结束前抵达目的地,一旦偏离关键绩效指标,就要接受领导与区域经理的训话,敦促其保持节奏或加快速度。奥腾立德指出:“‘让每个工人的表现都优于平均水平’,这样不合逻辑的要求就是一个很好的例子,通过看似客观的量化,实现劳动力管理中蕴含的恒定加速逻辑。”(奥腾立德,2023:57)这种“让人类劳动受制于软件逻辑”的设定,“通过颠倒技术与人类劳动力之间的关系,将人力插入复杂的算法机器之中”(奥腾立德,2023:58)。马克思

也曾预言,在机器的自动化系统中,“工人只是大型机器系统间‘有意识的联结’”(奥腾立德,2023:58)。

此外,为了应对诸如圣诞节购物热这样的特殊时期,亚马逊采纳“零小时合同制”(zero-hour contact)及“露营车劳动力”(camper force)项目:前者指“只有在需要时,他们才会被找来工作”;后者“试图安排工人住在露营车里,在全国范围内进行工作”(奥腾立德,2023:64—65)。更进一步,临时合作还被用来“约束和激励那些希望获得永久合同的临时工”(奥腾立德,2023:63),以进一步助推劳动的灵活性。无独有偶,UPS 作为美国最大的私营物流雇主之一,同样采用“远程通信”(telematics)系统、“道路优化与导航”(ORION)算法及传感器等软硬件设备来精准监控“独立承包商”式卡车司机的路线与操作,在计算司机动作的分秒价值的基础上展开深度优化。奥腾立德总结称,数字泰勒主义的一个关键事实是,“劳动的标准化、任务的分解、允许灵活性和劳动力倍增的算法管理”(奥腾立德,2023:62)。但在此运作中,不断生产出一个以劳资争议为核心的相互撕裂与分歧扩大的全球流通体系,其间被努力形塑为“螺丝钉”的工人始终在借助工会、抗议与罢工和上述巨头企业展开抗争(奥腾立德,2023:87—88)。

第 2 章“游戏工厂:游戏”聚焦服务于游戏《魔兽世界》的亚洲数字工厂,借助“金农”“游戏开发和质保数字工人”从事的“无聊、重复、单调、纪律严明的游戏劳动”(奥腾立德,2023:92),揭露与“游乐场”想象截然相反的游戏世界。首先,作者关注游戏内部地下产业的政治经济运作;其次,作者聚焦游戏生产本身展开讨论。“金农”是通过连续游戏来获取道具的工作者,他们可能“一周 7 天,一天 12 个小时”都在和同事打怪(奥腾立德,2023:93)。这一群体大部分是“双重民工”——他们在跨入游戏之前曾是真正的农民,在进入网络空间后仍然招致其他玩家乃至游戏发行商的抵制。这些“金农”需要适应诸如“金农必死”等咒骂与仇视,并时刻面临被玩家举报封号的风险。在此,作者借鉴阿尼许(A. Aneesh)提出的“虚拟移民”(virtual

migration)概念,指出金农面临的双重边缘境况:他们“既是工作室里处于边缘地位的外包劳动力,又是数字领域里因为打金工作备受种族主义攻击的民工”(奥腾立德,2023:107)。

由于涉及流动与“迁徙”,本章更加凸显基础设施形塑空间的维度。奥腾立德认为:“数字技术和基础设施不仅挑战了拓扑经济空间的概念,还挑战了与劳动力流动性、劳动力倍增相关的问题。它们属于正在进行的全球空间异化过程,构成了碎片化的、重叠的、不稳定的地图,并质疑了诸如南和北、中心和外围等固定的分类方式。”(奥腾立德,2023:109—110)比如,游戏工厂中种族化劳动者在网络空间仍然面临现实中的歧视结构,“游戏的空间性激活了以全球不平等为基础、通过各种基础设施实现的复杂价值链”。不仅是“金农”,一群被称为“测试佬”(Testing Monkeys)的质保部门游戏工人同样需要在游戏更新前后的“关键时刻”(crunch time)疯狂加班,其中涉及的性别歧视与劳动剥削已被广泛讨论:关于前者,女性背负照顾子女的重任,通常无法应对冗长且颇具弹性的工作时间,因而被排除在该岗位之外(奥腾立德,2023:122);关于后者,依托“自由精神”与“热情”合理化“无薪加班”的策略无疑是对游戏劳动力的深层剥削(奥腾立德,2023:126)。

第 3 章“分散的工作:众包工厂”特别讲述了众包平台通过计件工资等灵活手段创设的全球异化空间——在此,性别与种族的不平等结构得以再生产。同时,平台公司通过强调技术革新的关键价值来贬抑世界各地众包工人的协作劳动,并将成本累计清算至工人头上。首先,“‘众包工作’泛指通过在线平台外包给大量利用数字设备进行远程工作的人的劳动”,他们通常被视作和前述 UPS 一样的“独立承包商”,平台通过将自己规划为技术公司来避开对这些“没有固定合同的自由职业者”负责(奥腾立德,2023:138)。即便如此,那些肩负照料工作的女性以及发展中国家农村的新数字工人,仍然不得不依赖众包平台来补贴家用或维持生计。虽然这些可以称

作“数字作业流水线”的微型任务通常十分简单,诸如判断图片的文字内容,但这些回报微薄的工作并非常有,必须实时关注平台发布的任务列表。同时,众包工人维权或投诉的可能性极低,他们既然可以被轻松地吸纳进数字工厂的劳作空间,便也可能在一夜之间被扫地出门——排除在平台之外。这是由于,一度被认为过时的“计件工资”形式卷土重来,“消除了操控、监督工人的需要”:“这种合同和工资形式不仅可以为雇主提供灵活性,将停工期、保险和工作设备的成本转嫁给工人,还能作为一种技术,在没有实体工厂和监工的情况下组织劳动过程。”(奥腾立德,2023:54—55)

可以说,以“土耳其机器人”为代表的众包平台几乎无一例外地遵循着“按需劳动、灵活合同和自动化管理的逻辑”(奥腾立德,2023:137)。不过值得一提的是,“土耳其机器人”是 18 世纪一台引发诸多关注的国际象棋机器,其屡战屡胜被认为是机械奇迹,但实际上里面躲藏着一位棋艺高手。也因此,“土耳其机器人已经成为人类与机器、科学与魔法间边界地带的象征”(奥腾立德,2023:144)。奥腾立德以自动驾驶汽车技术的研发为例,指出数十万众包工人注释交通图片的“幕后工作”至关重要,而“围绕自动驾驶讨论的论文总是专注于技术,往往夸大了算法的能力,却忽略了自动化过程中所需的人力劳动”(奥腾立德,2023:144)。与传统泰勒主义的工厂不同,众包平台数字泰勒主义原则支配下的劳动力并非“福特主义意义上的数字大众工人”,而是具有高度异质性、灵活性与不稳定性的劳动力,后者恰是“劳动裂变激增”的关键意涵(奥腾立德,2023:156)。

第 4 章“隐形的工厂:社交媒体”特别关注平台作为基础设施的物质性显现,以及内容审查员在处理社交媒体阴暗面时潜在的负面心理影响。

首先,奥腾立德说明了“社交媒体作为劳动场所的多样性:从北极圈附近的数据中心的保安和技术人员,到硅谷的软件程序员和测试员,再到德国、印度和菲律宾的内容审核员”(奥腾立德,2023:177)。随着所需处理数据量的增加,脸书(Facebook)开始自建数据中心空间:“2013 年,脸书在瑞

典卢利亚的北极圈附近建设了第一座大型欧洲数据中心。寒冷的气候有助于保持服务器的正常运作温度。”(奥腾立德,2023:189)与游戏的物质性相似,数据中心巨大的用电量饱受批评。此外,备用柴油发电机和化学电池的投入也使得基础设施日趋背负环境保护的谴责。由此,社交媒体平台纷纷启动“绿色云计划”(奥腾立德,2023:191),围绕此投入的人力与物力之多恰恰说明了基础设施运作的物质逻辑。

其次,作者指出,“谷歌等搜索引擎不能被视为一种简单的算法,而是由不断变化的物质基础设施、软件和人类劳动组成的混合物”(奥腾立德,2023:183)。奥腾立德在此列举了作为基础设施架构要素之一的算法结构背后不同的利益取径与劳动需求。具体而言,谷歌遵循“网页排名”算法,雇佣评估员对互联网数据检索的层级结构进行评估,“通过改进结果与查询的匹配度来进一步完善算法”(奥腾立德,2023:181)。脸书则遵循“开放脸谱”算法,旨在捕捉用户在脸书外的数据偏好,进而向广告商出售用户的精准画像,其中软件程序员弥足重要。当然,从更广阔的层面鸟瞰,上述大数据及其云计算等基础设施“与国家主权存在着复杂关系”,它们“深刻地嵌入了国家政治地理之中”(奥腾立德,2023:193),这也显示出基础设施与政治经济力量的互构与联合。

最后,奥腾立德指出内容审查尤其复杂,“因为对暴力、色情、幽默和恐怖主义等关键话题的解读,在不同的法律、文化和政治背景下存在很大差异”(奥腾立德,2023:202)。那些分散在柏林、奥斯汀与都柏林的外包民工,通常面临“执行方面的巨大压力”以及过多接触暴力内容的心理风险(奥腾立德,2023:207)。一位受访者甚至坦言:“拉斯维加斯那场可怕的枪击在我看来突然变成了再正常不过的事情。”(奥腾立德,2023:210)这种暴力内容的“脱敏化效应”(desensitization)从长期看可能“泯灭人性”(dehumanization),也即对血腥暴力现实的彻底默然甚至无感。正当这些脆弱的外包员工受困于此时,社交媒体平台可能正在思考:如何“通过相对低廉的

劳动力成本带来额外的收益”,同时免于遭遇“受到创伤的员工提起的诉讼”(奥腾立德,2023:213)?

四、结语

反思基础概念的建构方式,能够更加深入地理解“工厂”隐喻的独特价值。目前传播研究对“平台”的体认,可大致分为技术、媒介与资本三条取径(孙萍、邱林川、于海青,2021),分别强调平台作为基础设备背后的硬软件架构(特别是算法)、借助居间交转连接各种社会资源来显现自身以及作为逐利用具备受资本青睐与支持。若以政治经济学视野观之,数字劳动以及与之相关的零工经济(gig economy)等议题已然位居讨论中心。本文无意梳理该领域的最新面貌,转而思索一个更为基础的问题:这一新劳动形式背后“旧”的根基何在?

有趣的是,这种以“旧”为新的思维方式恰是媒介考古学(media archaeology)的核心特征(施畅,2019),书中亦征引了该领域代表人物帕里卡(J. Parikka)的观点,他亦指出“‘软件工作作为工厂工作’的文化历史”(奥腾立德,2023:127)。笔者认为,奥腾立德恰恰通过历史性回视,揭示了数字平台与传统工厂的深层勾连。他在书中令人信服地指出:“数字资本主义不是工厂的终结,而是其迸发、倍增、空间重构和技术转变。”(奥腾立德,2023:8—9)也因此,作者“避免宣称当代劳动制度的绝对新颖性,反而力图在其中寻找早期劳动制度的延续、重复和重构”(奥腾立德,2023:230)。这种以“旧”为新的整体性视角将笔者带回马克思《资本论》中的精辟透视:“因为工厂的全部运动不是从工人出发,而是从机器出发,所以不断更换人员也不会使劳动过程中断。”(马克思,2018:484—485)如今,数字工厂的人员更换更加灵活巧妙,这也许意味着机器与活劳动更隐秘、更深层的捆绑。

因此,我们需要检视“数字技术与自动化正在消除不需要技巧的程序

化劳动”(奥腾立德,2023:7)等技术替代论观点,其忧思深重,愈发担心智能时代“无用阶层”或将批量涌现(周春林,2023)。不过,正如奥腾立德在书中表明的,自动驾驶技术的算法训练、社交媒体的内容审查,以及游戏开发中副本漏洞的检测,这些看似由机器自动完成的大量工作,实则借由平台分发给世界各地异质的活劳动来合力完成。这些被有意隐藏甚至抹除的人类劳动被格雷(M. Gray)等称为“幽灵工作”(ghost work)(Gary & Suri, 2022)。这一概念力图声张:光鲜夺目的机器与技术背后无数被遮蔽、被边缘化的劳动人民,仍然在努力生活、积极斗争。当然,这些人群日益受到研究者关注。比如,格雷、奥腾立德均提到的内容审核员已得到更细致的关注与理解,罗伯茨(S. Roberts)对商业性内容审查涉及的人群进行了全面、深入的调查,指出网络内容审核员的多元工作模式及其异同(罗伯茨,2023:38—39)。

沿上述思路继续追问,也许不可避免地要论及数字平台的技术治理与伦理运作。巴纳吉(S. Banaji)在反思“创造性”(creativity)的概念建构时指出,不加反思地将创造性视作道德良善、天然可取的规范性界定可能贻害无穷,因为其遮蔽了创造性背后的权力与伦理意涵,人们必须意识到创造性兼具破坏性(Banaji, 2023)。如奥腾立德所言:“许多当代技术的设计首先是为了使别人的劳动——不管是个体工人的劳动,还是更广泛的社会劳动——得到适当的利用和私有化。但这并不意味着,它必须保持这种方式。”(奥腾立德,2023:246)如此观之,以谷歌为代表的技术公司似乎也正在试图放大诸如“智能”技术话语的创造性,这与亚马逊在面临罢工时程式性地向媒体散布自身取得的“自动化”成就如出一辙。我们必须省思:这些技术公司所称的“智能”“自动化”究竟会为谁带来福音?以及,这些创造性的技术话语背后,究竟潜藏着何种尚未被认真审视的破坏性?

参考文献

奥腾立德,2023,《数字工厂》,黄瑶译,北京:中国科学技术出版社。

吉登斯、萨顿,2023,《社会学基本概念》第 3 版,王修晓译,北京:北京大学出版社。

罗伯茨,2023,《幕后之人:社交媒体时代的内容审核》,罗文译,广州:广东人民出版社。

马克思,2018,《资本论》第 1 卷,北京:人民出版社。

施畅,2019,《视旧如新:媒介考古学的兴起及其问题意识》,《新闻与传播研究》第 7 期。

孙萍、邱林川、于海青,2021,《平台作为方法:劳动、技术与传播》,《新闻与传播研究》增刊第 1 期。

周春林,2023,《"过劳人"与"无用人":智能时代的生活危机与超越可能》,《中国图书评论》第 10 期。

Banaji, S. 2023, "Why We Should Take A Second Look at the Politics of Creativity: The Dangers of A Celebratory Mode." in *Creative Provocations: Speculations on the Future of Creativity, Technology & Learning*, Cham: Springer International Publishing.

Gary, L. & S. Suri 2022, "Ghosts in the Machine." in *Critical Perspectives on New Media: Challenging the Neo-Liberal World Order*, Shanghai: Shanghai Translation Publishing House.

计算技术的历史的当代启迪

——专访历史学家由杰夫

由杰夫　叶瀚璋*

在过去的20年里,计算技术已渗透到人类社会的方方面面,从行政管理到虚拟货币,其作用日益凸显。探究计算机和计算技术如何影响人类社会,以及社会如何反作用于技术发展,已成为人文社科领域的重要研究课题。自20世纪70年代起,一些技术史学者开始从社会、经济等多元视角来审视计算机与人类历史的交织。与此同时,计算科学家们也逐渐认识到保存计算机相关历史资料的价值。明尼苏达大学的查尔斯·巴贝奇研究中心(Charles Babbage Institute, CBI)在这一领域的贡献尤为突出。该中心主任由杰夫(J. Yost)是一位杰出的学者,他曾在多个重要岗位上任职,包括《IEEE计算机史年鉴》(*IEEE Annals of the History of Computing*)的编辑(2008—2011)、"计算机历史"(History of Computing)丛书的合作编辑(2019—2021)。目前,他正担任约翰霍普金斯大学"计算与文化"(Studies in Computing and Culture)丛书的合作编辑。在由杰夫和前所长米萨(T. Misa)教授的共同努力下,CBI已经成为美国领先的计算机史研究中心。其档案库中珍藏着对美国计算技术发展产生深远影响的两大公司的资料,这两大公司分别是巴勒斯加法机公司(Burroughs Adding Machine Company)和控制资料公司(Control Data Corporation)。此外,CBI还出版了《界面:计算机和文化论著及评论》(*Interfaces: Essays, and Reviews in Computing and*

* 由杰夫(J. Yost),明尼苏达大学查尔斯·巴贝奇研究中心主任,明尼苏达大学科学技术和医学史研究教授。叶瀚璋,明尼苏达大学科学技术和医学史博士生。

Culture）。在本次访谈中，由杰夫分享了他的研究领域转向、历史对理解和规范计算技术的作用，以及历史学者如何与其他领域学者和计算机专家进行交流等见解。同时，他也揭示了 CBI 崛起为美国最重要的计算技术历史研究中心的秘诀。

一、从汽车史到计算机史

叶瀚璋：非常感谢您能接受这次访谈。您能先向读者简要介绍一下自己的教育和工作经历以及作品吗？

由杰夫：我出生于内布拉斯加州林肯市，并在那里长大。后来，我前往明尼苏达州圣保罗市的麦卡莱斯特学院（Macalester College）深造。大学期间，我开始对各类历史研究产生浓厚的兴趣。临近本科毕业时，我逐渐对技术史产生了特别的喜好。我记得，凯斯西储大学（Case Western Reserve University）是美国首个开设技术史博士项目的学府。尽管之前已有一些科学史项目，但凯斯西储大学的这个项目在技术史领域具有开创性意义。我选择这个项目并非仅仅因为其首创性，更是因为我深受卡罗尔·珀塞尔（C. Purcell）作品的影响，希望能跟随他学习。因此，在 1991 年，我搬到克利夫兰，开始了我的博士生涯。我的研究兴趣广泛，涵盖了社会、文化、知识、商业和组织等多个领域的技术史。我的博士学位论文聚焦于早期汽车工业中的汽车供应网络历史。实际上，我最初对计算机史的研究，是对之前在汽车史领域所做研究的延续。无论是汽车还是计算机，我都从组织史的视角出发，同时融入了一些劳工史和商业史视角。

完成博士学位论文后，我于 1998 年秋天开始在 CBI 担任副所长。其间，我阅读了众多著作，包括比尔·阿斯普雷（B. Aspray）和马丁·坎贝尔-凯利（M. Campbell-Kelly）的《计算机》（*Computer: A History of the Information Machine*）（Campbell-Kelly, Aspray & Yost et al., 2023）一书，并荣幸地

成为该书第 3 版的作者之一。目前,该书的第 4 版将于 6 月底问世。自加入 CBI 并担任副所长以来,我对这份工作充满热情。它不仅涉及行政和管理,同时也是一个研究岗位。我一直在这儿工作,专注于计算机的历史和社会研究。2018 年我升任所长,开始将中心的研究重心转向跨学科领域。我坚信,对计算机和软件历史的研究既是历史研究,也是社会研究。因此,我广泛地与历史学家、社会学家、人类学家以及媒体和传播学者合作。这种跨学科的交流和研究,不仅拓宽了我的研究视野,也为我带来了与不同领域学者和观众互动的乐趣。

叶瀚璋:您为何决定转向计算机史的研究?它与汽车史似乎是两个截然不同的领域。

由杰夫:我一直对 20 世纪以来的技术发展及其与社会史、商业史、文化史和思想史的交织感兴趣。因此,我并未限定自己只研究汽车领域。在研究生阶段,我读到《计算机》这本书,深受其吸引;大卫·博尔特(D. Bolter)的《图灵人:计算时代的西方文化》(*Turing's Man: Western Culture in the Computer Age*)(Bolter, 2023)也让我对文化史、思想史产生了浓厚兴趣。此外,高中时我就曾撰写过关于计算机史的论文,对计算机如何改变世界的历史一直保持着浓厚的兴趣。作为一名技术历史学家,我发现真正有趣的技术并非几个世纪前的老古董,而是 20 世纪那些具有影响力的技术,如汽车、航空、航天和计算机。因此,对我来说,从汽车史转向计算机史的研究是一个自然的转变。随着计算机视觉和人工智能的发展,计算机领域与汽车领域逐渐融合,成为备受关注的领域之一。至今,可以说计算机、人工智能与汽车已经紧密相连,共同推动了技术的进步。

二、跨学科的计算史

叶瀚璋：能否分享一下您对这个领域最新趋势的观察？考虑到它在中国是一个新兴领域，您认为当前该领域的紧迫问题是什么？

由杰夫：自 1998 年进入这个领域，专注于计算机史的研究以来，我观察到了一些显著的变化。当时，相关文献仍然较为稀少。例如，创刊于 1979 年的《IEEE 计算机史年鉴》主要聚焦于“内史”视角，如对第一台数字计算机、第一个存储程序的历史考察。然而，我认为该领域的重要之处在于超越这些内部视角，更多地从社会、政治和组织背景来入手。

这种研究趋势始于 20 世纪 80 年代后期，但真正站稳脚跟并产生重大影响是在 90 年代，尤其是 90 年代后半期。我尤其关注种族、阶级和性别方面的不平等问题，这些问题在我的研究中日益凸显。这些是社会历史的标准研究领域。我认为，在技术史和计算史领域，研究范式的社会史转向相对滞后了大约 10 年，直到 90 年代才真正开始。然而，在过去的 15 年里，计算社会史领域取得了令人印象深刻的进展，相关的方法和理论也得到了显著的发展。因此，我越来越多地将自己的工作视为这一领域内对话的一部分。

2020 年，我与杰拉尔多·迪亚兹（G. Diaz）共同策划并举办了“正义的编码：关于 IT 的全球政治经济中的权力不平等”（Just Code：Power，Inequality，and the Global Political Economy of IT）的活动。此次活动以线上形式进行，吸引了超过 800 人注册，最终有 330 余人参与。无论线上线下，它都堪称科学、技术和医学史领域的一次盛会，汇集了众多人类学家、社会学家、传播学者、媒体研究学者及历史学家。我认为这一群体极具吸引力，目前我正在对活动的论文进行修订和编辑，以形成一本论文集。

2023 年初，我们又举办了“设计自动化：政治、文化和景观，以及学习机器的时代”（Automation by Design：Politics，Culture，and Landscape in An

Age of Machines That Learn)的活动。此次活动聚焦于机器学习及其在社会、商业和政治背景中的发展,继续探讨技术如何影响社会平等的问题,深化了我们对技术本身及其应用的理解。我认为,这对于个人、政策制定者及政府层面做出更明智的决策,以及适当的监管技术都至关重要。

早在 2017 年,我就出版了《运作起来:计算机服务行业的历史》(*Making It Work: History of the Computer Services Industry*)(Yost, 2017)。这本书不仅体现了我对商业和工业领域的兴趣,也深入探讨了该行业的劳工和性别问题。接下来,我们将基于"正义的编码"工作坊的成果撰写一本书。此外,我们还计划出版一本关于设计自动化的专著或期刊特刊。

与不同学者合作极具价值,我参与了众多合作研究项目。我与迪亚兹共同编辑了约翰霍普金斯大学出版社的"计算与文化"(Studies in Computing and Culture)丛书,这是一套以历史学为主要方向的跨学科丛书,涵盖了人类学、社会学等多个学科的作者。我们对此深感兴奋,并已成功出版该系列中的数本佳作,同时还有更多作品正在编辑出版之中。

近期,我们出版了由珍妮特·阿巴特(J. Abatte)和斯蒂芬妮·迪克(S. Dick)编辑的《抽象与具身》(*Abstractions and Embodiments*)一书,该书是一部高质量的论文集(Abbate & Dick, 2022)。接下来,我们计划在本月晚些时候推出安德鲁·利(A. Lea)的《数字化诊断》(*Digitizing Diagnosis*)(Lea, 2023),该书在医学史上具有里程碑意义,并实现了医学史与计算机史的跨学科交融。此外,我们还将出版众多杰出学者在人工智能领域的研究成果,涵盖大型语言模型、模式识别、种族、阶级和性别等多个方面。

我对这一系列图书深感兴奋,因为在 CBI,我有幸结识众多学者,为他们提供帮助、建议,并向他们学习,与他们共事。我热爱编辑和中心主管的工作,并在三个专业期刊的编辑委员会中担任职务,这既是我工作的重要组成部分,也是 CBI 工作的一种延伸。我将自己的职业生涯视为一种使命,既致力于自己的研究,也努力使他人的研究成为可能,并尽我所能地利用研

究所的资源和我自身的能力来实现这些目标。

三、计算技术的历史和当代的对话

叶瀚璋：如今，计算机已经无处不在。了解计算机的历史能够对我们理解现在甚至未来有何帮助？对于社会科学家和公众而言，了解这段复杂的计算机历史有什么好处？

由杰夫：我认为，技术已成为推动变革的强劲动力。它深刻地影响着人们的生活，并催生了一系列与技术相关的社会规范。研究技术的演进历程，可以了解前几代人是如何因应技术变革的。这些历史既充满了激动人心的时刻，也伴有令人忧虑的挑战。普遍来看，年轻一代对迅速的技术变化有着更高的适应性，而部分老年人在面对无处不在的计算机、智能手机和应用程序时则显得较为吃力。因此，回顾历史有助于我们理解技术如何重塑世界，以及我们如何围绕技术发展制定政策、规范与法律。当前，法律相对于信息技术的发展而言存在明显的滞后性，且缺乏充分的监管。例如，对社交网络或人工智能等领域的监管，便是我们面临的重大挑战。

叶瀚璋：面对这些挑战，我们是否需要对技术本身有深入的了解？换言之，鉴于计算机技术的复杂性，作为历史学者，您是否认为我们有必要全面掌握所有的技术知识？

由杰夫：这是一个颇具趣味性的问题。确实，我的一些同事，如托马斯·黑格（T. Haigh）等，他们在本科阶段主修计算机科学。然而，我并非计算机专业出身，未曾涉足编程课程，也未曾有过编程经验。尽管如此，我广泛研究了计算技术的各个方面，主要是为我所从事的口述历史研究做准备。迄今为止，我在不同领域完成了 280 多篇口述历史研究，其中大约三分之二的访谈对象是科学家，包括计算机科学家和工程师。在准备这些口述历史

研究的过程中,我深入学习了编程语言、操作系统、计算机安全模型、人工智能和图形学等方面的知识。但须明确,我并未系统学习过完整的计算机课程。我的研究侧重于商业、社会、文化和知识史的融合。我学习技术知识的程度足以让我分析底层技术,从而辅助我进行这一类型的历史研究。我认为,重要的不仅在于不将技术视为黑匣子,更要关注技术的社会和文化塑造、我们对技术的看法,以及人们如何带来不同的知识和技能。因此,我深入学习了关于底层技术的知识。但同样重要的是,我阅读和研究了劳动史、社会学、性别、种族和民族等领域的知识,这些构成了我研究计算机史的工具包的一部分。

叶瀚璋:正如您之前提及的,跨学科的计算机史研究能为决策提供更丰富的信息。然而,这个领域的历史学家和社会科学家是如何与计算机专家及决策者进行实际交流的呢?

由杰夫:举例来说,我的一些同事会参与国会听证会,直接与政府沟通。此外,如阿隆德拉·纳尔逊(A. Nelson),她被任命为信息技术不平等问题的高级顾问。纳尔逊女士是普林斯顿高等研究院的学者,她在该领域具有显著影响力。在人工智能和大型语言模型领域,也有许多学者——如华盛顿大学的计算语言学家艾米丽·本德(E. Bender)、谷歌前伦理人工智能学者蒂姆尼特·格布鲁(T. Gebru),后者目前正在经营她与他人共同创立的分布式人工智能研究中心(DAIR)——成为公众关注的焦点,并积极参与政策辩论。尽管并非所有历史学家都能获得这种待遇,但我们都在努力促进领导者利用这种理解来尝试制定政策。因此,我认为这些真正令人印象深刻的学者的工作是至关重要的。他们有效地利用了更广泛的社群资源,开展了更广泛的对话,我认为这是非常有益的。

与此同时,我认为当前的主要挑战在于公司所掌握的巨大权力。众多公司、大集团以及行业内存在诸多强势的游说团体。然而,由于技术的日新

月异,政府在监管方面面临极大困难。以美国为例,社交媒体、人工智能和隐私等信息技术领域确实缺乏足够的监管。尽管欧洲在这方面有所进步,但技术监管依然面临重大挑战,且往往存在滞后现象。这些公司往往利用各种手段绕过现有法规,过度追求股东价值和利润最大化,而忽视了对公共利益的关注。

叶瀚璋:您与公众互动的经验能否分享一下?我们虽多聚焦于学术界的话题,但您还涉足面向公众的博客写作。

由杰夫:我坚信与公众沟通的重要性。参与"正义的编码""设计自动化"研讨会的人群虽以学者为主,但他们同样属于广大公众。2020 年,即"正义的编码"研讨会举办之年,我与档案管理员阿曼达·威克(A. Wick)合作推出了一份面向公众的学术短文期刊——《界面:计算机和文化论著及评论》。我们希望将其打造成一本技术类杂志,强调它的可读性。相较于学术期刊,该杂志的文章篇幅更短,图文并茂,注重可读性,避免使用过多的专业术语。我们对学术界和更广泛的社区对该期刊的积极反应感到非常满意。

你提到了我的博客,那就不得不提我对加密货币领域的兴趣。2020 年,随着数亿人持有比特币或其他加密货币,我愈发关注这一领域。2012—2015 年,我与米萨共同研究了计算机安全历史,这让我对密码学及其在现实世界中的应用产生了浓厚兴趣。此外,我对隐私问题的关注始于研究生期间的第一篇论文。谈及区块链,比特币的发起者——一个赛博朋克团体,尤其吸引了我的注意。尽管比特币创始人的身份仍是个谜,但关于其身份的猜测众多,其中一些人认为它主要是由一群我持反对态度的自由主义右翼政治倾向者推动的。然而,也有许多人专注于其去中心化的特性,并运营各种区块链项目以服务于社会公益,这些项目往往展现出左翼倾向。这种二分法正是吸引我的地方。为此,我开设了一个博客,专注于研究区块链技

术的社会历史和社会学。我撰写了涉及种族、性别、政治经济学的文章,这些文章在公众中得到了良好的反响,也引起了学术界的兴趣。甚至我的一些同事也使用这些文章作为本科生的教学材料,这让我感到非常欣慰。为了更自由地表达我的政治观点,我决定将这些内容放在我的个人网站上,而非明尼苏达大学或 CBI 的服务器上。

四、CBI 的诞生及其馆藏

叶瀚璋:正如您所说,CBI 对您和众多致力于该领域研究的学者而言都意义非凡。您能为我们简要介绍一下这个研究所吗?

由杰夫:成立 CBI 的构想,是在它正式在加州成立前的数年间逐渐形成的。这一构想的推动者包括计算机行业的退休人士,尤其是我们的主要创始人欧文·托马什(E. Tomash)和阿黛尔·托马什(A. Tomash)。他们与工业界的朋友合作,并获得了学者和策展人的宝贵建议。在诸多建议中,常春藤盟校、史密森尼学会以及欧文·托马什接触的其他历史学家提出,若要在计算机历史上留下深刻影响,就应在大学内建立一个中心,该中心应包含图书馆和博士项目两大核心部分。欧文·托马什采纳了这一建议,并在全国范围内进行招标。最终,明尼苏达大学凭借其强大的科学技术史课程、研究型图书馆,以及对博士项目和学院的浓厚兴趣与坚定承诺成功中标,CBI 因此落户于明尼苏达大学。此外,我们的图书馆在美国研究型图书馆中位列前十五,尤其在档案和特殊馆藏方面,不是前五就是前十。就设施而言,我认为我们的图书馆可能位居前五。因此,我们现在拥有一个专门的特殊馆藏图书馆大楼。

我们现在所在的安德森图书馆,建于 20 世纪 90 年代后期,2000 年正式启用。该建筑的三分之二位于地下,为 CBI 及大学档案馆等档案单位提供了可容纳数个足球场大小的可调节温度的储藏库,用以存放大量纸质记录

和照片。我们充分利用研究项目和档案的协同功能,实现了资源共享与互补。自 CBI 创立之初,历史学家亚瑟·诺伯格(A. Norberg)和比尔·阿斯普雷就积极投身于口述历史项目,他们不仅参与展示和出版,还建立了广泛的联系网络,为馆藏的发展创造了丰富机会。此外,我们还拥有一个充满活力的口述历史项目,这一项目极大地促进了我们关系网络的拓展,并帮助我们收集到世界上关于计算机、软件和网络历史的一些重要材料。自 1980 年 CBI 开始运作起,这一传统便持续发挥作用。我加入 CBI 后持续与多位所长合作,共同推动这一项目的深入发展。

我所做的一个显著改变在于,明确而系统地将我们的研究范围扩大至计算、软件和网络的跨学科领域,从而开辟了全新的研究领域。同时,我前所未有地利用社交网络来推广 CBI,强化了社交网络在推广出版物、研究计划和资源方面的作用。具体来说,我通过社交网络,如推特(Twitter,现为 X)、Mastodon、LinkedIn、Threads 和 BlueSky 等平台,积极推广了我们的研究成果、活动以及出版物等。这一举措不仅扩大了我们的影响力,还使我有机会与历史学界以外的众多社会科学家建立联系。

叶瀚璋:您能分享一下 CBI 藏品的具体情况吗?据我所知,这是非常独特且珍贵的收藏。

由杰夫:当然,正如我之前所提及的,CBI 从成立之初便开始进行收藏工作。目前,我们拥有超过 300 个收藏系列,规模大小不一。其中,我们特别关注两个主要的企业资料收藏系列,尽管它们并非我们唯一的公司收藏,但其因规模庞大而显得尤为突出。这两个系列分别是巴勒斯加法机公司的资料和控制资料公司的资料。前者于 19 世纪 80 年代作为计数器公司而起家,并在 20 世纪 50 年代涉足计算机领域。我们收藏的这些记录,涵盖了该公司从创立之日起到 20 世纪 80 年代的全部历史。控制资料公司于 1957 年在明尼阿波利斯成立,直至 1992 年因财务困境进入破产管理,最

终被拆分。其中一部分作为 Ceridian 公司得以幸存。公司的重组和解散往往为保存这些庞大的公司记录提供了机会。通常,公司希望维护和控制自己的记录,但在某些情况下,如巴勒斯加法机公司在 80 年代与 Sperry Univac 合并成立优利系统(Unisys)时,由于新公司的成立、名称的变更以及维护档案的高昂成本,其选择将跨越一个世纪的巴勒斯加法机公司资料作为礼物赠送给我们。经过整理后,这些资料形成了一个包含 800 个档案盒的庞大收藏。

控制资料公司的档案规模也相当可观。在该公司接近倒闭时,我们得以接收其档案库。我们同样收藏了计算机协会(ACM)的档案。ACM 作为一个蓬勃发展的专业组织,汇聚了约 100 000 名计算机专业人员和科学家。通常,非营利组织在自身运作正常时更倾向于捐赠其档案,而这种情况在公司中则较为罕见。然而,仍有一些公司会在特定时间选择捐赠档案。例如,福特汽车公司在 50 年后向亨利·福特博物馆和格林菲尔德村图书馆捐赠了其档案,我相信他们现在仍遵循这一做法,即所有 50 年前或更久远的资料都会被移交至档案馆。这是一种常见的方式,旨在确保与公司利益不直接相关的产权或私人信息不被公开。我们的馆藏类型丰富多样,包括国家标准局收藏中一些并不广为人知的灰色文学作品,以及数以万计的灰色文献报告。此外,我们还收藏了行业协会的档案集,如 ADAPSO,这是一个支持软件和服务行业等行业发展的组织,其档案记录以及两份由主要 IT 行业委托进行的市场研究报告均为我们所拥有。著名计算机科学家的论文,如珀利斯(A. Perlis)、威尔(W. Ware)的作品就在我们的收藏之列。在我担任所长期间,我们的研究团队共同努力,积极收集各类档案记录,以提升人们进行社会史研究的能力。这些与计算机相关的收藏,涵盖了劳工历史、性别、工会、自动化等多个社会议题。同时,我们一些较老的馆藏也为研究社会历史提供了宝贵资料。

叶瀚璋：巴勒斯加法机公司、控制资料公司等通常在合并或倒闭时捐赠资料，不过，ACM 这样的案例依然存在。我想知道，他们是否只是捐赠了他们认为可以公开的资料呢？

由杰夫：我认为，ACM 的大部分组织办公室记录都已捐赠给我们。这个组织是志愿者结构，拥有大约 30 个不同的兴趣小组，我认为这是一个相对完整的收藏。关于具体细节，档案管理员威克可能会比我更了解。我记得有一份协议规定，ACM 会定期对这个收藏进行补充。他们捐赠的东西可能不包括最近 5 年左右的材料，因为这部分记录他们还在积极使用。但我认为他们有继续补充捐赠材料的计划，这不会像福特汽车公司那样滞后 50 年，可能只是滞后 10 年左右。

叶瀚璋：是 CBI 先联系了他们，还是他们联系了 CBI 来捐赠？

由杰夫：我们的工作人员历来都是 ACM 历史委员会的成员，现任档案管理员威克近期才加入该委员会。过去几年，我曾领导 IEEE 计算机学会的历史委员会；但成为所长后，我认为加入 ACM 历史委员会更为合适，因此我辞去了 IEEE 的职务，并希望在那里以普通成员的身份参与。这一决策已持续 6 年。ACM 历史委员会十分活跃，并为 ACM 提供关于历史问题的建议。事实上，自 ACM 历史委员会成立以来，CBI 的工作人员就与其保持着联系，这种长期联系使我们成为接收和存储其记录的自然选择。

叶瀚璋：所以，你们一直有长期的合作。

由杰夫：确实。我想强调的是，委员会中的志愿者角色、编辑委员会中的领导角色以及外部组织中的委员会角色，一直是 CBI 及其外部扩展的核心要素。目前，我正致力于一个项目，该项目涵盖一项活动的策划以及与美国国家工程院相关的多篇出版物。这一志愿服务对于促进学群间的交流与合作、发掘收藏和研究机会至关重要。社会网络对于研究所而言具有极高

的价值,这也是我热衷于此并发现其专业益处的原因。

叶瀚璋:我们谈到了研究所的历史,也谈到了它与外部组织的联系。CBI 接待了许多学者来这里从事收藏工作或他们自己的项目。您能和这些访问学者分享一下您的经验吗?

由杰夫:CBI 自创立之初便肩负两大使命,即口述历史(开展研究)与藏品利用(发布研究成果)。我们设立了托马什奖学金计划,最初简称为 CBI 奖学金,并将所有获奖者尊称为“托马什研究员”。多年来,我们持续颁发这一奖项,目前已有约 44 位托马什研究员,包括比尔·阿斯普雷、珍妮特·阿巴特等该领域的杰出领导者。此奖学金计划对获奖者助益良多,不仅资助了他们的研究生研究,也助力了他们的论文撰写。众多获奖者在此投入大量时间,利用藏品资源,并接受所长和副所长的指导。在职业生涯中,我有幸指导了多位这样的研究者,并从他们日后的成就中受益匪浅。此外,我还积极推动了大学的跨学科博士研究员(IDF)项目。在我担任所长期间,我们在获得 IDF 研究员方面取得了显著成功。IDF 奖学金价值 25 000 美元,面向中心主要领域以外的申请者。尽管该项目竞争激烈,通常仅有约四分之一的申请人能获得资助,但我们的成功资助率却高达约 80%。过去 5 年里,我们成功吸引了大约 6 名 IDF 研究员。除了上述奖学金项目,我们还接待了国际访问学者。与这些海外学者共同工作、学习和交流,无疑为我们带来了丰富的经验和视角,这是一段非常宝贵的经历。

叶瀚璋:您能和我们分享一下您与访问学者合作时最难忘的时刻或记忆吗?

由杰夫:与跨学科研究员,包括与托马什研究员合作,具有深刻的意义。对此我深感愉悦,并从众多同事那里汲取了丰富的知识。特别值得一提的是,我与迪亚兹的合作尤为密切,他既是我的研究员,又是我亲密的同事和

朋友。自他成为托马什研究员以来,我们在多个项目上携手合作,合作频率远超其他同事。我们之间的合作顺畅而富有成效,对计算社会历史和跨学科社会研究怀有共同的热情与奉献。我们曾共同策划多项活动并出版丛书,他在 CBI 的驻留时间或许超过了其他任何一位托马什研究员,我与他逐渐发展成一种紧密的学术伙伴关系。这一关系对于 CBI 而言,在多方面都具有极大的价值。我接任所长一职时,CBI 进行了重组,不再设立副所长职位。由于我是全职所长,没有本科教学职责,因此有更多的时间和精力来专注于研究和合作。在参与活动和项目时,与多方合作尤为重要。例如,在最近的“设计自动化”活动中,我与来自科罗拉多大学的柯勒特·佩罗德(C. Perold)、郑芳芳以及迪亚兹展开了紧密的合作。

叶瀚璋:您和其他访问学者还有联系吗?

由杰夫:我与许多人都保持着频繁的联系。他们中不乏该领域的佼佼者,如比尔·阿斯普雷,自我涉足计算机历史领域起便与之相识,并且成为我的导师。在早期阶段,比尔一直是我的良师益友,我们共同合作开展了多个项目,至今已逾 20 年。还有像保罗·塞鲁齐(P. Ceruzzi)这样的杰出人士,以及一些在我来之前就加入了 CBI 的学者。在加入 CBI 后不久,我便有幸结识了这些学者。在这漫长的岁月里,我有幸结识了这些学者,并一直与他们保持联系。他们为 CBI 的发展做出了巨大贡献,同时也在我的职业生涯中给予了极大的帮助。当然,我也尽我所能为他们提供支持,这种互助关系非常宝贵。

五、为亲历者留下口述历史

叶瀚璋:您能分享一下开始一项口述历史工作的流程吗?对于与科学家沟通的方式,您有什么特别的建议吗?

由杰夫:CBI 创始人诺伯格具备前瞻性的眼光,在创始阶段就启动了历史资料收集和口述历史工作。在加入 CBI 之前,他已经积累了丰富的口述历史工作经验,这对 CBI 的发展起到了积极的推动作用。诺伯格聘请了热衷于口述历史的阿斯普雷,随后是米萨和我,我们同样对口述历史充满热情。我个人已经完成了近 300 份口述历史记录。

历史学家对口述历史的态度非常谨慎,原因在于其本质乃基于受访者的主观记忆与观点,同时受访谈者提问的主观性影响。尽管如此,这种“动态”特性使得口述历史成为研究的有力辅助。通常,历史学家更倾向于依赖档案研究,但口述历史能在档案不足时提供必要的补充,尤其在某些无法仅凭档案完成的项目中。例如,我在研究计算机中的性别问题时,发现关于性别与计算机服务行业的记录几乎空白。为了填补这一空白,采访计算机服务行业中的女性领导者显得尤为关键。幸运的是,我拥有相应的社会网络和人脉,使得我能够采访到计算机服务领域的杰出女性专业人士和企业家。这些口述历史不仅捕捉了计算机历史中被遗忘的声音,还展现了边缘化个体的故事,为社会历史研究提供了重要资料。同时,我们的口述历史项目也涵盖了与杰出计算机科学家、图灵奖获得者的合作,其范围之广可见一斑。

口述历史承载着多重目的。在进行口述历史研究时,一个关键技巧是尽可能收集多样化的视角,通过多方比对来确保材料的可信度。CBI 的口述历史项目往往与某个受资助的研究项目紧密相连,这对于 CBI 的财务状况及研究所与项目的声誉均至关重要。我们在争取资助方面取得了显著成就,特别是来自美国国家科学基金会(National Science Foundation, NSF)的资助。我有幸担任 NSF 五个项目的首席研究员,这些项目多为 3 年期,资助金额介于 20 万—50 万美元之间。这些受资助的项目不仅为聘请研究生作为研究助理提供了机会,更将口述历史作为项目的重要组成部分。在准备口述历史项目时,每个人都需要进行独特的准备工作,这基于过去对特定

领域的采访经验和对该领域专业知识的积累。例如,我曾与米萨共同负责或主导了关于软件历史、计算机安全历史、NSF 历史及网络基础设施 Fast Lane 的项目。目前,我正致力于另一个关于安全和隐私的项目。作为历史学家,我持续学习相关知识,这些学习基于每次采访的积累,而非一次性完成,从而确保采访质量得到提升。例如,若我计划采访某人关于计算机体系结构的问题,那么与其讨论编程语言则显得相对不那么必要,效率亦会降低。

叶瀚璋:访问一般是您自己做还是雇人做?

由杰夫:我几乎独立完成了所有口述历史工作,这些成果大多源自我参与的项目,但也有例外,比如我负责的关于软件历史的项目,我仅完成了四分之一到三分之一的工作。为补充这部分缺失,我们聘请了博士后学者菲利普·弗兰纳(P. Frana),他目前执教于詹姆斯麦迪逊大学(James Madison University, JMU)。他出色地完成了该项目剩余三分之二的部分。值得注意的是,这是 CBI 历史上唯一一次,大部分口述历史工作不由所长或副所长直接完成。此外,我们也有一些口述历史资料是通过外部捐赠获取的,对于这类捐赠提议,我们总是持开放和欢迎的态度。这些捐赠的口述历史资料通常由资深历史学家精心整理,因此其质量完全符合我们的标准。人们愿意捐献口述历史,这无疑是件值得称赞的事情。对于捐赠者而言,这也是一种双赢的选择。

叶瀚璋:要完成这些口述历史项目,肯定要付出巨大的努力。

由杰夫:确实,这是一项极其繁重的任务。在新冠疫情之前,我们主要进行的是面对面的口述历史记录。随着新冠疫情大流行,一些领先的项目开始尝试通过视频会议进行,我发现这种方式运作得相当好,甚至更加高效。现在,我仍会偶尔进行面对面的口述历史工作;但大多数情况下,我认

为 CBI 将会通过 Zoom 进行资料收集。这种方式极大地简化了工作流程，也降低了我出差的频率。如今，作为中心唯一的历史学家，尽管我居家办公的时间增多，但我仍能通过 Zoom 进行工作，并在研究人员来访时与他们见面。我始终积极参与对研究人员有益的会议，因此减少出差次数对我而言是极大的便利。过去我确实喜欢前往硅谷等地出差。但如今，我很高兴出差的次数有所减少，我认为通过视频会议进行的口述历史工作并未受到太大影响。

六、CBI 模式

叶瀚璋：接下来，我想提出一个假设性的问题：如果有人询问您如何创办类似 CBI 这样的一个机构，您会给予他哪些建议？

由杰夫：若有人计划创办一个研究所，我强烈推荐他们参考 CBI 模式，它已被证实是极为有效的。这种模式强调研究项目和图书馆之间的紧密合作，这种合作带来了诸多益处。同时，我认为口述历史不仅本身就是一种宝贵的资源，还能为馆藏发展创造更多机会。我建议研究所在创立之初就建立这样的伙伴关系。此外，我认为拥有访问学者对研究所的发展至关重要。我个人非常支持跨学科的研究模式，因为社会学家、人类学家、传播和媒体研究学者、地理学家等领域的学者都能为历史研究提供独特的视角和丰富的资源。他们的加入不仅丰富了各自的研究项目，也为历史学家的工作带来了新的方法和思路。同时，跨学科研究还能够帮助我们与制定政策和有影响力的人建立联系，使我们在参与对话时更加得心应手。

叶瀚璋：关于 CBI、您的个人经历、学术界或该领域的其他方面，您是否有什么想要补充的，或是对读者说的话？

由杰夫：在我的职业生涯中，我有幸与众多杰出的人才共事，并师从非

凡的导师,如米萨、马奥尼、阿斯普雷等。与迪亚兹、莱特(J. Light)、尼尔(D. Nye)、比尔以及科塔达(J. Cortada)等高级研究员的合作经历同样令人难忘。能够与他们共事并汲取他们的智慧,对我而言是一种极大的乐趣。我为自己能成为科学、技术和医学史项目的一分子而感到骄傲,与来自世界各地的研究生一同进行研究,是我职业生涯中的一大亮点。我特别喜欢在论文委员会的工作,这使我深感有意义。我非常感谢这些机会,让我能够成为对人类有贡献的基础建设的一部分。这份工作充满了价值,我很荣幸能在这里工作 25 年。

叶瀚璋:对于希望了解计算机历史的初学者,您认为哪些书籍对他们有帮助?

由杰夫:在我个人的学习经历中,坎贝尔-凯利和阿斯普雷的《计算机》一书对我影响深远。2014 年,我有幸与恩斯明哥(N. Ensminger)等三位同事共同参与了该书的修订版编写工作。这本书原本就是作为一本介绍性文本而撰写的,如今已被翻译成中文、日文等多种语言,并在众多课程中广泛使用。鉴于其广泛的认可度和学术价值,我强烈推荐这本书给初学者。在第 4 版中,我又与郑芳芳和迪亚兹这两位同事共同努力,为该书增添了诸多新内容,尤其是关于计算的社会背景部分。

我的前同事麦地那(E. Medina)的《控制论革命者:阿连德的智利的技术与政治》(*Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*)(Medina, 2011)是一本内容非常丰富的书,获得了多个图书奖项。阿巴特的《重编码性别》(*Recoding Gender*)、《发明互联网》(*Inventing the Internet*)也是我非常喜欢的两部作品(Abbate, 1999; Abbate, 2012)。诺贝尔(S. Nobles)的《压迫的算法》(*Algorithms of Oppression*)(Nobel, 2018),本杰明(R. Benjamin)的《技术之后的种族》(*Race after Technology*)(Benjamin, 2019)等作品,虽然不直接属于历史领域,但同样值得一读。莱特的

《从战争到福利》(*From Warfare to Welfare*)(Light, 2005)是一部令人惊叹的作品,主要探讨了兰德公司及其不断变化的使命,以及军事和军事主义模式如何以有害的方式被用于城市中的社会问题。这本书对我个人产生了深远的影响,也是我职业生涯中最喜欢的作品之一。哈佛大学社会学者雷雅雯(Y. Lei)的《有争议的公共领域》(*The Contentious Public Sphere: Law, Media, and Authoritarian Rule in China*)(Lei, 2017)一书同样值得一读,她在"正义的编码"工作坊中介绍了这本书,并为《正义的编码》一书撰写了一篇论文。

参考文献

Abbate, J. & S. Dick 2022, *Abstractions and Embodiments: New Histories of Computing and Society*, Baltimore: Johns Hopkins University Press.

Abbate, J. 1999, *Inventing the Internet*, Cambridge, MA.: MIT Press.

Abbate, J. 2012, *Recoding Gender: Women's Changing Participation in Computing*, Cambridge, MA.: MIT Press.

Benjamin, R., 2019, *Race After Technology: Abolitionist Tools for the New Jim Code*, New York: Polity.

Bolter, J. 2023, *Turing's Man: Western Culture in the Computer Age*, Chapel Hill: University of North Carolina Press.

Campbell-Kelly, M., W. Aspray & J. Yost et al. 2023, *Computer: A History of the Information Machine*, New York: Routledge.

Lea, A. 2023, *Digitizing Diagnosis: Medicine, Minds, and Machines in Twentieth-Century America*, Baltimore: Johns Hopkins University Press.

Lei, Y. 2017, *The Contentious Public Sphere: Law, Media, and Authoritarian Rule in China*, Princeton: Princeton University Press.

Light, J. 2005, *From Warfare to Welfare: Defense Intellectuals and Urban Problems in Cold War America*, Baltimore: Johns Hopkins University Press.

Medina, E. 2011, *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*, Cambridge, MA.: MIT Press.

Noble, S. 2018, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press.

Yost, J. 2017, *Making It Work: History of the Computer Services Industry*, Cambridge, MA.: MIT Press.