智能社会研究

（双月刊）

2024 年
第 3 期
2024 年 5 月 10 日出版

2022 年 11 月 10 日创刊

总第 10 期

目　次

论文

自由与认同：数字游民文化与本土化社会实践研究

——基于西南某地的田野调查 …… 孙艺珂　周承磊（ 1 ）

智能时代劳动范式的数字转型

——基于政治经济学批判的视角 …… 周光港（ 18 ）

政治经济学批判语境下数字异化的深层逻辑与扬弃进路 …… 魏金鹏（ 37 ）

数字时代的社会自我

——破解碎片化自我困境的路径探索 …… 高思蓉　王化起（ 52 ）

赛博格工人、“科技恶魔”与金币农夫

——东方主义遭遇数字时代 …… 何祎金（ 66 ）

研究报告

中国个人信息保护经验研究的方法论反思

——欧盟经验的启示

…… 张月明　李汶龙　李汶锴　李子豪　李傲兰（ 85 ）

理性、权力与生态

——平台研究相关中文文献述评 …… 袁方杰（114）

译文

气候变化与人工智能政治路径

——人工智能和人类能动性下的技治主义民主困境

…… 马克·科克伯格　亨里克·塞特拉　著

李　帅　李　芳　译(138)

归因法:网络攻击来源归因规则

…… 德尔伯特·特兰　著　裴　轶　强心语　译(160)

书评

数字技术支配下的生活世界

——读杰米·萨斯坎德《算法的力量:人类如何共同生存?》

…… 王国伟(203)

CONTENTS

THESIS

Freedom and Identity: A Study on Digital Nomad Culture and Localized Social Practices Based on Fieldwork in the Southwestern Region of China ····· Sun Yike, Zhou Chenglei(1)

Digital Transformation of the Labour Paradigm in the Age of Intelligence: A Perspective Based on the Critique of Political Economy ····· Zhou Guanggang(18)

The Deep Logic and Sublation Approach of Digital Alienation in the Context of Political Economy Criticism ····· Wei Jinpeng(37)

The Social Self in the Digital Society: Exploring Paths to Resolve Fragmented Self Dilemmas ····· Gao Sirong, Wang Huaqi(52)

Cyborg Labor, "Devil of Science" and Gold Farmer: An Encounter Between Orientalism and Digital Times ····· He Yijin(66)

RESEARCH REPORT

Methodological Reflection on the Study of China's Personal Information Protection Experience: Insights from the EU's Experience ····· Zhang Yueming, Li Wenlong, Li Wenkai, Li Zihao, Li Aolan(85)

Rationality, Power and Ecology: A Review of Chinese Literature of Platform Research ····· Yuan Fangjie(114)

TRANSLATED TEXTS

Climate Change and the Political Pathways of AI: The Technocracy-Democracy Dilemma in Light of Artificial Intelligence and Human Agency
..... written by M. Coeckelbergh, H. Sætra; trans. by Li Shuai, Li Fang (138)

The Law of Attribution: Rules for Attributing the Source of Cyber-Attack
..... written by D. Tran; trans. by Pei Yi, Jiang Xinyu (160)

BOOK REVIEW

The Living World Dominated by Digital Technology: Reading Jamie Susskind's *Future Politics: Living Together in A World Transformed by Tech* ..... Wang Guowei (203)

自由与认同:数字游民文化与本土化社会实践研究*

——基于西南某地的田野调查

孙艺珂　周承磊**

摘要:随着全球化和数字技术的迅速发展,数字游民这一群体在中国逐渐壮大。他们试图摆脱传统社会的束缚,在身份认同、整体自主性和社会互动等方面建构自己的生活哲学,其社会实践可被视作一场大型生活实验。社交媒体等载体孕育和强化了数字游民的身份认同与集体归属感,他们解构传统工作观念,秉持新工作主义理念,追求工作与生活的平衡。数字游民的地理套利行为一方面能够促进乡村振兴,另一方面也可能改造传统乡村生态。数字游民的发展尚处于探索期,其结局是找到人生的旷野还是回归主流,仍未可知。

关键词:数字游民　远程工作　地理套利　共居社区　乡村振兴

一、研究缘起与文献回顾

目前,越来越多的人开始探索新的生活和工作方式。数字游民作为一种新兴的社会群体,正成为这种探索的代表。他们冲破传统的职业束缚,选择了一种灵活的、不受地理限制的工作方式,以此应对现代职场的高压和不确定性。本文旨在探讨数字游民如何寻找新的身份认同,他们的生活方式如何

* 本文系2023年国家社会科学基金一般项目"逆城市化移民在西南边疆的媒介实践研究"(项目批准号:23BXW103)的阶段性研究成果。

** 孙艺珂,中国传媒大学广告学院。周承磊,中央民族大学民族学与社会学学院。

反映出个体对工作和生活平衡的追求,并与传统社会结构相互作用。

“数字游民”(digital nomad)描述的是那些利用数字技术进行远程工作,并选择经常更换居住地或在世界各地旅行的人(Thompson, 2019),其与背包客的主要区别在于工作和生活方式。背包客通常是在旅行期间休假或暂停工作,而数字游民在旅行的同时工作,利用数字技术和互联网获得收入(Stumpf, Califf & Lancaster, 2022)。数字游民与传统远程工作者也有所不同:传统远程工作者一般为他雇者,拥有定期收入,倾向于在一个地点生活,往往缺乏社交活动(Matos & Ardévol, 2021);数字游民则追求更高的自主性,希望能够自主选择自己的工作项目、工作时间和居住地。数字游民主张,工作在多大程度上受到监督,比从这种受监督的劳动中领取报酬的多少,更能反映真实的社会等级(Cook, 2023)。

德勒兹(G. Deleuze)、萨特(J. -P. Sartre)和马尔库塞(H. Marcuse)的理论,为数字游民的生活方式提供了思想基础。德勒兹倡导的流动性和去中心化理念(Yang & Yang, 2009),萨特存在主义理论中的个人自由与责任(Glendinning, 1999),马尔库塞对一元化思维和消费文化的批判(Peñaranda, 2013),共同构成了数字游民价值观的哲学框架。这一框架鼓励个体超越传统模式,通过行动进行自我定义,同时利用技术解放潜力,探索自由和灵活的生活与工作方式。

追溯西方数字游民的历史源流,有两个脉络:工作模式上,背包客逐渐转变为收入型旅游博主,利用数字技术和社交媒体进行内容商业化,互联网远程工作者的兴起也实现了工作与旅行的融合(Makimoto & Manners, 1997);价值取向上,贝格曼(F. Bergmann)的新工作主义理论、费里斯(T. Ferriss)的《每周工作四小时》和格雷伯(D. Graeber)对现代工作的批判,共同塑造了数字游民对新工作和生活观念的接受和实践。

既有研究涵盖了数字游民的多重身份认同、选择居住地点的影响因素、生活方式实践、与传统环境的互动模式等多个方面。这些研究采用质性研究

方法,对不同国家和地区的数字游民进行调查,揭示这一群体的特征和行为模式。

数字游民的身份认同是多元的。已有研究发现,社交媒体关于数字游民的叙事有四种:浪漫主义者,强调数字游民生活的自由和灵活;现实主义者,强调数字游民生活的实际难处;机会主义者,主要关注数字游民生活方式的商业化机会;理想主义者,认为数字游民生活是人生的最佳状态(Ehn, Jorge & Marques-Pita, 2022)。这些类型化的叙事塑造了数字游民的集体认同。数字游民的身份认同也是流动的,他们通过符号性消费建构自己的形象和流动主体性。对他们来说,这种生活方式既提供了身份转换的机会,也蕴含着自我实现的意义,是人生阶段的一个过渡。但是,这种生活方式也存在财务风险和永久边缘化的可能(Louis & Mielly, 2023)。

在选择居住和工作地点时,数字游民重视目的地的自然文化资源、基础设施条件、生活成本以及信息网络,这也是一些热门数字游民聚集地的关键吸引因素(De Almeida, Correia & De Souza et al., 2022)。数字游民对联合工作空间的高质量服务和设施有着浓厚兴趣,这凸显了他们在选择工作地点时对环境的特殊关注,而远程工作的增加也为他们创造了新的机遇(Orthaber & Zupan, 2021)。数字游民试图在工作和生活之间找到平衡,有研究提出了数字游民个人创业的阶段性模型(Louis & Mielly, 2023)。此外,克罗地亚政府为吸引数字游民而采取的政策措施,以及公司管理层在整合数字游民时面临的挑战,均表明了主流社会对这一群体需求的理解与适应(Gorliński-Kucik, 2021)。

中国的数字游民受传统文化和个人主义影响,有独特的发展路径。在实践数字游民生活方式时,他们无法享受到与之匹配的全球化工作机会、远程工作政策和高社会接受度。因此,从概念旅行的视角开展数字游民本土化研究有其必要性(姚建华、杨涵庚,2023)。

二、从身份构建到生活哲学

(一) 数字游民群体画像

本研究对 98 位数字游民进行了问卷调查。结果表明:数字游民群体性别分布平衡,男性占比为 54%,女性占比为 46%;普遍年轻化,25—34 岁的数字游民占比为 57.14%,这一阶段正是职场黄金年龄;大多拥有本科及以上学历(95.92%),其中本科学历占比为 68.37%,硕士学历占比为 27.55%,教育资本积累程度较高;普遍认为自己“不受人管辖,思想独立”(78.57%)、“怀有强烈的求知欲和好奇心”(56.12%);成为数字游民的首要原因是追求自由和独立(96.94%),其次为对传统工作方式不满(59.18%)、追求更高收入(26.53%)。数字游民追求高度的个人自由和灵活的工作模式,主要涉及互联网相关行业及创意产业,可见他们在新兴产业中的活跃度和在传统行业以外的开阔视野。

(二) 身份塑造的双重路径

1. 跨越边界的自我表现

数字游民通过社交媒体平台展示旅行经历和生活实践,以建构和巩固身份认同。在社交媒体上,最能代表数字游民的画面应当是:迷人的自然风光做背景,笔记本电脑屏幕做前景,画面的其他部分不经意间展示出本地特色美食或其他标志性地理符号。这种自我表征机制强化了他们对自身角色的认知和自豪感,并以此为“暗号”连接全球数字游民,强化自我身份感。特别是在多元文化背景下,这种展示强调了他们作为全球公民的身份特征。美好生活的媒介展演既是对所选生活方式的自我肯定,也引发了围观者对数字游民生活的向往和传统工作模式的反思。

数字游民的身份认同具有流动性与多样性特点。数字游民在社交媒体的虚拟空间中叙述个人经历,进行视觉呈现,构建了一个多维度、动态发展的个人形象。他们在自我表征时频繁使用不同标签,涵盖生活方式、职业兴趣及文化经历,反映了身份的流动性与多样性。例如,李明(化名,下同)的个人介绍中有自由撰稿人、环球旅行家、摄影爱好者、环保志愿者等多个标签。他展示了在不同国家的旅行经历,如在泰国拍摄日出照片、在西班牙参加当地的节日庆典,以及在印度参与环保活动。在职业社交平台上,李明强调他自由撰稿人的职业身份,分享他为多个国际杂志撰写的文章的链接。通过这些不同的标签和社交媒体平台,李明的自我呈现出流动性与多样性。

2. 共居社区中的群体共鸣

相较于西方的数字游民,中国数字游民更倾向于通过社群获取社会支持和认同感。长期维持数字游民生活对个人生存能力的要求很高,而参与在线社群则是一条寻找可持续生活方式和职业的路径,也是一种减轻经济压力、提高职业成功率的策略。在数字游民面对更多挑战和不确定性时,共居社区为他们提供了必要的支持和资源。

数字游民共居社区是一种新兴住宿模式,整合了居住和办公的功能。社区内一般配备了完善的办公设施,有相应的社交活动区域,以满足成员独立工作的需求及与其他数字游民的交流合作。此外,共居社区还是个体自我发现和集体身份构建的重要场所。每一位成员在申请入住时都需要提交一份答卷,社区主理人基于申请人的教育水平、职业经历、兴趣爱好等背景筛选合适的候选人,以保持社群成员间的异质性。筛选出的数字游民个体在社区中极易获得群体共鸣,社区中的社交互动则进一步强化了自我与社区集体之间的联系。例如,A 社区每周一晚上举行“思享厨房”活动,社区成员自我介绍并相互认识;B 社区每周六举行人工智能(AI)工作坊,教零编程基础的数字游民学习使用 AI 工具;C 社区不定期举办小众电影观影分享活动;D 社区每周三下午举行小型分享会,分享人互相交换技能和需求。

在共居社区中,我们抱着一种开放和坦诚的生活态度。每当有新成员加入,我们都热情地欢迎他们,认为他们带来的一切都是对我们社区文化的丰富。我们之间的交流始终坦诚直接,大家都乐于分享自己的想法和需求。如果你需要帮助,或者有能力帮助别人,我们总是愿意坐下来讨论,寻找最佳的支持方式。这种互相支持和坦诚沟通的氛围,是我们共居社区的核心特质。(D 社区主理人晓薇)

我们的共居社区每周五都会举办一次名为 Foodie Friday 的活动,邀请所有伙伴参与。这是一个社交和学习的机会,大家在这里相互认识,可能会一起学习制作咖啡,或者讨论数字游民生活相关的话题。同时,我们也会向大家介绍并请他们尝试各种植物性食品。比如,你可能会看到一些不常见的水果的独特吃法,有时会做水果吐司,或者把香蕉做成奶昔。这些活动不仅丰富了我们的生活,也增进了社区成员之间的了解和交流。(C 社区主理人雨欣)

(三) 生活哲学的五个维度

数字游民的生活哲学不限于单维度的自由,更关注在经济、时间、兴趣爱好、居住地选择以及社交关系等多个生活领域中的整体自主性。数字游民追求的自由可分为五个维度:经济自由、时间自由、志趣自由、居住自由和情感自由。经济自由包括收入自由和支出自由,前者指想赚钱的时候就能赚到钱,后者指不会因为金钱的匮乏而无法实现自己的目标。时间自由并不是无所事事地“躺平”,而是有自主分配时间的权利。志趣自由是数字游民最典型的特征,也是最容易实现的目标。数字游民摆脱传统价值观的束缚,自由选择喜欢的事情。居住自由即根据自然、人文、社交、物价、签证等因素选择“游”到哪里。情感自由经常与社交资本相混淆,本调查显示有 61 位数字游

民感到孤独,有情感支持和归属感方面的需求,他们渴望与他人进行深度交往。

李华是一位远程工作 8 年的 IT 工程师,通过在线项目协作可以在全球任何地方赚取收入。他为了维护客户关系,工作的时候比在办公室里更加投入。王芳这样的自由撰稿人可以根据自己的生活节奏自主安排工作,选择在清晨创作,下午休闲,从而平衡工作与生活。旅行博主兼旅拍摄影师张强,能够将个人爱好与职业相结合,在追求个人价值的同时谋生。赵怡然选择在大理、清迈、巴厘岛、里斯本等地生活和工作,探索多元文化。区块链从业者经常参加国际黑客松或快闪城市活动,在活动期间共居共创,建立全球性的社交网络,遇见志同道合的朋友,丰富个人和职业生活。

三、从工作自由到价值追求

(一) 工作与生活的平衡

1. 新工作主义思想风暴

数字游民视新工作主义(new workism)(Bergmann, 2019)为圭臬。这一理念的核心在于重构工作的意义和方式,强调在工作选择和方式上的个人自主性,并寻求工作与生活更好的平衡。这一理念鼓励工作模式创新,如远程工作和自由职业,反对传统集中式、等级化的工作结构,倡导更分散、去中心化的工作环境。蒂莫西·费里斯(Ferriss, 2007)提出了“设计生活方式”的理念,即根据个人愿望和目标主动设计生活,而不是完全被工作所左右,同时利用地理套利来提升生活质量,并通过时间和财务管理实现更多的自由和独立。人类学家大卫·格雷伯(Graeber, 2018)则从批判的角度出发,认为现代社会中大量的工作毫无意义,这些工作往往使人感到自己的劳动毫无价值。他对传统的工作伦理提出了疑问,认为工作不应该仅仅是为了工作本身,而

应该是实现个人价值和为社会做出贡献的手段。

2. 生计转型各显神通

数字游民的生计方式多种多样,基准线是自给自足。从全职工作到自由职业,从股票投资到资产管理,各式各样的途径共同构成了他们经济生活的多元图景。在这些方式中,自由职业者更受益于互联网零工经济的兴起。他们最具有数字游民的特点,通过网络平台接受各式各样的零工任务。远程工作的数字游民,其职业生涯相比之下更为稳定。然而,稳定并不意味着轻松。为了保持这种稳定,他们必须投入相应的时间和努力,这往往意味着必须在工作与生活之间找到一种微妙的平衡。由于远程工作的特性,工作与生活的界限变得模糊,二者常常交织在一起。

数字游民的经济来源主要有四种:劳动收入、流量收入、创意收入和资产收入。首先是劳动收入。远程工作者,以工资维持生计。佟彤是一名时尚杂志编辑,供职于上海某公司。她和母亲住在一栋白族村民的院子里,定期出差,其他时间远程办公,每个月有稳定的收入,上海的高薪资水平和×市的低生活成本,让佟彤的幸福指数飙升。比起都市的快节奏生活,她在×市可以用半个月的时间参加戏剧社,每周参加摇摆舞会,给其他舞者拍照,并获得“百万摄影师”的赞誉。其次是流量收入。志成原是一名企业销售人员,爱好摄影,新冠疫情期间来到×市,成为旅拍博主。旅游旺季的时候,粉丝受其吸引成为客户,同时他还获得内容流量本身带来的广告商单。再次是创意收入。明炫曾在海外学习戏剧,他在×市发起疗愈喜剧工作坊,以戏剧的方式帮助参与者摆脱生活中的尴尬,工作坊报名费为 298 元/人。最后是资产收入。×市有丰富的旅游资源,有些数字游民通过经营民宿、酒吧、创新教育学校等方式获得收入,也有 Web 3.0 从业者从虚拟资产交易中获得收入。刘健是一位经验丰富的数字游民,早期生活在清迈,曾开过一间网红咖啡馆,买咖啡的人半夜还在排队。新冠疫情期间,刘健来到×市,在一条主干道边开了一家酒吧,同时为数字游民共创活动提供场地支持。

3. 金钱焦虑与精神自由的考量

数字游民脱离一线城市紧张的工作节奏,来到×市难以维系之前的收入水平,“即使原工作模式转入远程办公模式,在薪资水平和职位晋升方面也会受到一定程度的影响”(数字游民 Tina)。有相当一部分数字游民脱离原来的工作岗位之后能凭借原工作技能成为自由职业者,或者放弃原来的工作内容而开启新的事业。小薇曾经在外企和商学院工作,开始数字游民生活后决定将人生教练(life coach)作为自己的新职业。尽管人生教练在国内有一定的市场需求,但了解并愿意为之付费咨询的人寥寥无几。在新事业起步阶段,小薇的收入水平可能会大不如从前,甚至需要其他收入补贴。

在数字游民群体中,有相当一部分间隔年(gap year)学生或刚刚毕业的应届生,没有工作经验和积蓄,他们的收入来源主要是父母补贴,因此数字游民的生活在带给他们精神自由、生活愉悦的同时也带来了金钱焦虑。2000 年出生的李磊本科毕业后来到×市,他本人从未上过班,凭着对电影的一腔热爱,又恰逢 AI 绘图工具技术成熟,受惠于零工经济,他通过网络接单为客户制作 AI 动画。工作一段时间后,他认为还是需要一份稳定的收入,因此在旅游商业区某酒吧谋得一份工作,以此缓解自己的金钱焦虑。自称数字无业游民的陈曦正处于科研间隔年,因为经济不独立而经常与父母产生摩擦,争执的核心是父母希望他早点毕业找份工作。陈曦的男朋友在某世界 500 强企业工作,在陈曦的鼓动下准备辞职成为数字游民。由于工作年限短,积蓄只够两人生活几个月,辞职之后的生计问题也让陈曦的男朋友有些焦虑。

数字游民经常有金钱焦虑,但他们并不会因此而放弃这种生活方式。陈曦认为她和父母、男朋友的矛盾点主要在于金钱,而她本人从来没有赚过钱,也没有赚钱的经验,对自己能否赚到足够的钱没有信心。她希望通过做自媒体来寻找一些变现的机会,抓住一切机会寻求经验。本硕博一路连读的小雅没有工作经验,在同龄人都成家立业的年纪,自己还无法实现经济独立,无法赚钱是当下生活状态唯一的缺憾。有趣的是,李磊、陈曦和小雅都没有因此

终止数字游民的生活状态转而去工作,陈曦甚至还劝说男朋友辞职成为数字游民。由此可见,金钱焦虑似乎并不足以改变他们的生活志趣,于他们而言这不过是人生探索期的阵痛,他们对走出金钱焦虑充满信心。“我说的‘希望被资本收买’也就是口嗨一下,要是真被资本收买了,我可受不了。”(数字游民小雅)

(二)追求内在价值的新游牧主义

数字游民重构生活和工作模式,采取远程工作、自由职业等灵活多变的工作方式,以游牧的生活方式对抗创造物异化和创造过程的压力。“游牧的核心在于对在场的超越。我们所推崇的游牧主义,是勇于打破系统、结构、疆界的束缚,是对边界和可能性的探索,是对创造性、想象力的思考。”(数字游民映雪)不同于“躺平”,×市的数字游民强调自我实现,超越职业成就的界限,探索个人兴趣和生活方式,从而减轻社会规范和期望带来的心理压力。他们追求内在满足而非外在奖励,注重工作和生活的本质意义,这种内在导向的生活方式促使他们真实地表达自己的需求和个性。笔者在田野调查中发现,数字游民的内在自我区别于荣格(C. Jung)的意识自我(ego),并非只关注自己的感受,他们将社会公共价值纳入自我实现的范畴。例如,小晴曾是一名公办中学的数学老师,跟随某公益组织来×市支教,服务期结束后继续留在×市。

数字游民通过社区和支持网络,寻获精神和情感上的互助,以共同体的形式抵御社会压力,增强内在自由和自主感。他们旅行体验多样生活,打破单一文化或观念的束缚。他们摆脱了 20 世纪五六十年代美国社会学理论中“组织的人”和“一维的人”的境遇,展现了从“他定型”的人向“内定型”和“普洛透斯式”的人的转变(科恩,1986)。这种变化反映了个体思考自我存在的象征概念的稳定性,以及对新文化象征的接受。“我前 7 年在中国最好的大学读书,从小一直是别人眼里的好学生,我想撕掉‘好学生’的标签,找到自己。”

(数字游民黑马)“我想对生活有更多的掌控,撕去外界的标签,进行注意力管理,将注意力集中在当下,减少外界对自己的影响。”(数字游民 Candy)

四、从地理套利到乡村共栖

(一) 流动性液态生活体验

数字游民的生活方式体现了鲍曼(Bauman, 2000)所描述的“液态现代性”,即个体生活方式、社交关系和职业路径的灵活性和不确定性,而且在身份认同的构建和社交关系的形成上揭示了全球化和数字化背景下的挑战与适应策略。在不断变化的地理和文化环境中,不同的居住地和社交圈子,塑造了一种流动和多样的身份认同。吉登斯(Giddens, 1991)的自我身份理论指出,这种不断变化的环境促使个体进行自我反思,形成多元化和灵活的身份。在这个过程中,文化适应能力成为数字游民身份构建的重要组成部分,他们吸收不同的文化元素来形成独特的个人身份。

不同于熟人社区,数字游民社区的高度流动性难以让数字游民建立持久关系,但他们在不同社区的短暂停留中形成了一种新的社区归属感。这种基于共享经历和价值观的归属感可能不如传统社区稳定,但提供了重要的社交支持和认同感。

每次回去都会发现之前刚刚认识的人已经不在了,社区里都是不认识的人。(数字游民佳倩)

×市的流动性特别大,有种相忘于江湖的感觉,即使约个饭都不一定能约到,今天在这里,明天就不一定了。(数字游民杨雪)

热而且疲惫,我称之为“热疲惫”风格。新冠疫情前我在西班牙考雅思,口试的时候大家在另外一个房间里等候,居然没人紧张,都和我热烈

地聊起来,毫无顿挫、敏感、腼腆感,聊的时候都和我热烈地目光接触。道别的时候和老朋友一样,这就是“热融感”。我在国内遇到的唯一真诚又不疲惫的中国人,就是那位希腊人长相的笛子手。(数字游民程鑫)

(二) 在地性融合共生

数字游民的流动性与在地性相互交织,反映出个体与地方、全球与本地之间的复杂关系。他们的地理套利策略对当地社区产生了经济、文化、社会等多方面的影响,他们在享受全球流动的自由的同时,也寻求与特定地区的文化和社会环境建立联系。全球流动性理论认为,现代社会中的个体不仅在全球范围内流动,同时也与特定地方产生紧密联系(Withaeckx, Schrooten & Geldof, 2015)。

数字游民的地理套利策略的实质是在不同地区间平衡生活成本和收入潜力,这种现象在经济地理学中被广泛讨论。数字游民在特定地区的集中可能对当地社区产生经济和文化影响,如里斯本当地居民就曾抗议数字游民影响了当地物价(Xiao & Wan, 2023)。地理套利在当地人眼里是新殖民主义,对此当地人也有自己的回应方式,涨价是最直接的表现。在×市的菜市场买菜,用不同语言获得的报价略有差异,说普通话的最贵,当地方言次之,当地民族话最便宜。

很多数字游民目的地的承接能力有限,一下子涌入那么多人对当地人造成了很大影响,数字游民太多会挤占当地人的公共资源。所以,要以去中心化流动的适应性,来克服集中化脆弱的稳定性。(数字游民杨雪)

我知道买菜是贵的,但我从不讲价,因为我本来就在地理套利,我愿意让他们赚一些,一共就几块钱,其实也贵不到哪儿去。(数字游民黄杰)

数字游民社区更加重视与当地社区的文化融入。D 社区位于大湾村,地处热门景点和城区之间,这里的村民对待外来人态度谨慎,甚至出租房屋时也不会像景区里的村民一样在大门上直接挂牌。主理人李婷向村民介绍美好生活的社区理念,邀请村民到社区喝茶,与村民建立情感连接,获得了在地村民的信任,村民甚至将自家房源委托给她出租。此外,A 社区、B 社区的一些举动,也反映了数字游民社区对融入当地文化的努力。

> 我想把这个村子做成一个共同的社区。我们把所想要的美好生活状态,通过我们的行为带给他们。现在我要是出去一趟,回来时手上的东西可能都拿不下了,特别是蔬菜,当地村民都一筐一筐地送给我。他们一有空也会到我这边来喝茶聊天,了解我们所做的事,甚至他们也愿意参与。(D 社区主理人李婷)

(三) 数字游民与乡村振兴

1. 数字游民业态的可持续性探索

《经济学人》曾探讨数字游民是否会从小众生活方式转为主流生活方式(Levels, 2020),说明了这一现象对社会结构与互动模式变革的潜在影响。海外市场对数字游民这一新兴趋势的敏感反应,体现在对此群体商机的迅速捕捉和相应商业模式的形成上。多个平台,如 Digital Nomad Academy、Nomadlist、Meetup、Remoteyear 及 DNX 数字游民大会等,通过提供工作、旅居信息、社交需求解决方案以及远程工作地点服务等方式,支持并引导了数字游民生活方式的发展。曼谷、巴厘岛等地的数字游民聚集区及共居社区、共享办公空间的出现,不仅证实了数字游民作为一种生活理念的影响力,也展示了其作为一个可商业化群体标签的潜力。信息技术、零工经济的发展促成了这种新型数字生活方式的兴起,而商业领域的活跃参与则进一步扩展和巩固

了这一趋势。

相对于海外,中国的数字游民业态还在起步阶段,只零星出现了几处配套服务。在×市,将“数字游民”作为标签的共居社区有4个,共创空间有2个,共享办公空间有2个。共居社区成员并不完全是数字游民,还有间隔年的学生和处于倦怠期的职场新人。如何持续运营下去,是数字游民社区面临的共同问题。共享办公空间中办公的人很少,某精心装修的共享办公空间开业时,会员价每周179元,非会员价每天39元,但3个月后只要点9.9元一杯的咖啡即可办公一整天;由于进来办公需要脱掉鞋子,老板还给每位客人送了一双袜子。“老板在外地,也不缺钱,已经放弃这里了,现在纯做公益。”(共享办公空间员工李浩)

最早将“数字游民”概念带到×市的E共享办公空间在场地到期后没有续租;B社区进行了新一轮的融资,扩大了住宿容纳量;D社区目前的主要收入来源是房屋租金;A社区在探索可持续的商业模式,早期通过小额众筹获得启动资金的A社区,更加带有试验的底色。“之前有一次开会大概讲了一下运营情况,我听下来感觉是没盈利,且负责运营的同学还往里投了很多,之后应该会半年公示一次财务情况。”(A社区员工桃子)

2. 数字游民的短暂停留与长期影响

×市的乡村同样面临空心化问题,数字游民的涌入是对这些空间的再填充,他们为乡村带来了新的活力和资源,试图建立新的生活和工作方式。×市政府引入数字游民考察团,走访各个县,进行实地调研,组织县级政府和考察团成员直接对话,一方面向数字游民展示各个县的人文和地理特色,另一方面希望数字游民代表可以为县级乡村振兴发展提供建议。作为一个具有全球视野、灵活工作模式的群体,数字游民的到来可能为乡村带来新的发展机遇。

数字游民的核心特征在于“游”,在于流动,因此数字游民的进入并不能完全解决乡村空心化的问题,他们与乡村的关系往往是短暂的、不稳定的,长

期的、稳定的效果仍然是一个未知数。“我已经驻村 2 年了,我们有几个老院子正在改建,我们村非常缺年轻人,现在来的人一般住 1 个月左右就走了,最多 3 个月,我们希望他们可以住得更久一点,欢迎他们成为新村民。”(驻村干部张勇)这种短期的、不连续的关系,可能导致乡村社会、乡村生态的进一步断裂,在这种背景下,乡村地区的发展和转型可能是一种重构,与原有的乡村生态有所不同。“动态的‘切入与接替’带来的是一种辩证的后果,即在否定原有乡村生态的同时,创造出新的生活方式和社会结构。”(胡翼青,2024)数字游民与乡村的互动促进了乡村新生态的建立,融合了现代化元素、技术创新和全球化影响。

五、总结

面对传统价值观和就业压力,数字游民在竞争激烈、充满不确定性、多元化的社会环境中,做出了一种新的选择,跨越性地实践了一种替代性的生活方式。他们的生活安排基于地理流动性和经济自主性的综合考量,其中包含了理想的居住环境、灵活的职业安排以及对社交和社区支持的需求。这种生活方式强调工作与生活平衡的重要性,并倾向于精神上的满足而非物质上的奢华。因此,关于“数字游民”的定义,不能简单地从经济收入单方面衡量,而应当包含生活与工作平衡可持续、与消费主义脱钩、拥有身份认同和社区感。

某种程度上,数字游民代表一种非主流的文化符号。然而,也有人拒绝这样的标签,他们认为这个标签下缺乏实质内容,或是对这种生活方式的真实性和持续性持怀疑态度。潜在的数字游民群体对数字游民的想象往往源自社交媒体上的美好描绘,但实际上这可能只是一种表象,隐藏着早期数字游民通过社群和媒体影响力对后来者进行某种形式的“收割”,从而推动“数字游民”概念的普及和商业项目的发展。在中国,数字游民与其说是一种生活方式,不如说是一场生活实验。

参考文献

胡翼青,2024,在南京大学人文社会科学高级研究院“数字游民的理论探索与本土化实践”学术沙龙上的发言。

科恩,1986,《自我论:个人与个人自我意识》,佟景韩、范国恩、许宏治译,北京:生活·读书·新知三联书店。

姚建华、杨涵庚,2023,《西方数字游民研究述评与中国启示》,《中国青年研究》第 11 期。

Bauman, Z. 2000, *Liquid Modernity*, London: Polity Press.

Bergmann, F. 2019, *New Work New Culture: Work We Want and A Culture that Strengthens Us*, Alresford: John Hunt Publishing.

Cook, D. 2023, "What Is A Digital Nomad? Definition and Taxonomy in the Era of Mainstream Remote Work." *World Leisure Journal* 65(2).

De Almeida, M., A. Correia & J. De Souza et al. 2022, "Digital Nomads During the COVID-19 Pandemic: Evidence from Narratives on Reddit Discussions." 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD).

Ehn, K., A. Jorge & M. Marques-Pita 2022, "Digital Nomads and the COVID-19 Pandemic: Narratives about Relocation in A Time of Lockdowns and Reduced Mobility." *Social Media+Society* 8(1).

Ferriss, T. 2007, *The 4-Hour Workweek: Escape 9-5, Live Anywhere, and Join the New Rich*, New York: Crown Publishing Group.

Giddens, A. 1991, *Modernity and Self-Identity: Self and Society in the Late Modern Age*, Stanford: Stanford University Press.

Glendinning, S. 1999, *The Edinburgh Encyclopedia of Continental Philosophy*, Milton Park: Psychology Press.

Gorliński-Kucik, P. 2021, "On Liberatory Strategies of Digital Nomads." *Er(r) go: Teoria-Literatura-Kultura* 1(42).

Graeber, D. 2018, *Bullshit Jobs: A Theory*, New York: Simon & Schuster.

Levels, P. 2020, "The Future of Remote Work." https://levels.io/the-greatest-migration.

Louis, D. &M. Mielly 2023, "People on the Tweets: Online Collective Identity Narratives and Temporality in the Lebanese Revolution." *Organization* 30(1).

Makimoto, T. &D. Manners 1997, *Digital Nomad*, Hoboken: John Wiley & Sons.

Matos, P. & E. Ardévol 2021, "The Potentiality to Move: Mobility and Future in Digital Nomads' Practices." *Transfers* 11(3).

Orthaber, S. & S. Zupan 2021, "Translators as Digital Nomads: Constructing Professional Identity Online." in *Mobilität-Wirtschaft-Kommunikation: Wie die Mobilität von Unternehmen, Personen, Kapital, Waren und Dienstleistungen die Kommunikation verändert*, Wiesbaden: Springer Fachmedien Wiesbaden.

Peñaranda, G. 2013, "New-Materialism and Reification in the Infoproduction Era." *Communication +1* 2(1).

Stumpf, T., C. Califf & J. Lancaster 2022, "Becoming A Digital Nomad Entrepreneur: A Process Theory." *E-Service Journal* 13(3).

Thompson, B. 2019, "The Digital Nomad Lifestyle: (Remote) Work/Leisure Balance, Privilege, and Constructed Community." *International Journal of the Sociology of Leisure* 2(1-2).

Withaeckx, S., M. Schrooten & D. Geldof 2015, "Living Across Borders: The Everyday Experiences of Moroccan and Brazilian Transmigrants in Belgium." *Crossings: Journal of Migration & Culture* 6(1).

Xiao, J. &Z. Wan 2023, "Nomads in A City of Their Own Making: A Portrait of Migrant Workers in Shenzhen, 1980-2000." *Inter-Asia Cultural Studies* 24(3).

Yang, H. &S. Yang 2009, "A Study on the Blurring Boundary Phenomena Expressed in Complex Fashion Space of 21th Century: Focusing on the Theories of Gill Deleuze and Felix Guattari." *The Research Journal of the Costume Culture* 17(4).

智能时代劳动范式的数字转型

——基于政治经济学批判的视角

周光港*

摘要:机器体系的自动控制与独立分化催生出数字时代虚实共存、开放互联、多元集中的平台化劳动资料,劳动者在全景化、流动化的监控体系中被调制为资本掌控的数字傀儡。资本对人类智力和社会属性的剥削需求,使得借助网络联结起来的个体转化为数字劳工,扩展了总体工人的界限。工业劳动中智力因素的消解也逐渐外化出专业化脑力劳动者,将工厂中的机器看管活动转化为网络媒介中具备生产性、社会性、主体性的认知交往,并培育出主体间的通用智能,赋予劳动解放以数字潜能。劳动范式数字转型的政治经济学批判揭示了资本、劳动、技术的共同在场以及劳动解放的时代际遇,彰显了马克思劳动价值论的不朽价值。

关键词:劳动范式　人工智能　数字劳工　政治经济学批判

数字化革命以来,半导体制造的微型化、集成化趋势愈加凸显,纳米计算、自旋计算、量子计算等算法系统飞速迭代,推动着新一轮机器智能革命和互联网发展的浪潮,人类正式步入实体数字化、万物互联的全新场域。人工智能的通用性、适应性、创造性在提高人类收集信息和处理信息的能力的同时,也深刻地改变了机器大工业生产以来的劳动范式。第 52 次《中国互联网络发展状况统计报告》数据显示,在中国,“至少掌握一种初级数字技能(能够使用数字化工具获取、存储、传输数字化资源)的网民占比 86.6%;至少掌握一种中级数字技能(能够使用数字化工具制作、加工、处理数字化资源)的网

* 周光港,吉林大学马克思主义学院。

民占比 60.4%”(新华社,2023)。机器的智能变革正在不断重塑劳动体系,在改变劳动者存在状态的同时,突破了大工业时代社会劳动力的技术限制。

然而,对这种与人工智能适配的劳动形态,仍旧存在诸多观点上的分歧。技术乐观主义者如保罗·多尔蒂(P. Daugherty)和詹姆斯·威尔逊(J. Wilson)认为,机器智能的发展将会塑造一种人机协作的全新框架和理解方式,人机融合对空白领域的填补和人类相对劣势领域的增强“以更人性化的方式重构业务流程,并让人们拥有超强的能力以实现高效工作”(多尔蒂、威尔逊,2018:197—198);而在以尤瓦尔·赫拉利(Y. Harari)为代表的技术悲观主义者看来,一个比人类更了解人类的“超人工智能”只会演化出统治人类的利维坦,大多数人将沦为“无用阶级”,“这些人对社会毫无用处,整天活在现实与虚幻之间”(赫拉利,2017:296)。在技术发展的未知性面前,也有学者试图规避人类与技术这一根本性问题,将智能看作新的“物自体”,认为“也许冥冥宇宙间有某种原因,会使它更倾向于走向有利于人类的结局”(蔡斯,2017a:174),主张回到主观个体的行为选择之中来把握人工智能的发展动向,但这种对宏大叙事的后现代排斥实际上将问题又推回历史的彼岸。当代西方学者的观点为我们理解智能时代劳动范式提供了借鉴,但他们却忽视了存在于技术与劳动之间的资本维度。“知识和技能的积累,社会智力的一般生产力的积累,就同劳动相对立而被吸收在资本当中,从而表现为资本的属性。”(马克思、恩格斯,2009a:186—187)马克思指出,社会普遍知识的资本占有成为资本主义生产方式的一般属性,这一点在当代西方数字资本主义中并没有发生实质性的变革。因而,基于马克思政治经济学批判的理论视角来分析数字时代资本宰制下的机器智能演进以及劳动范式的数字化转型,有助于破解数字化资本增殖逻辑的拜物教幻象,从而科学认识智能机器时代人机关系的当下嬗变,探求数字技术加持下人类解放的劳动路径。

一、从机器体系到网络平台:劳动资料的数字塑形

劳动资料作为劳动个体传递体力、外化智力的中介,是劳动过程的核心构成与时代表征,反映了数字社会转型过程中劳动范式数字化的器件形塑。早期工场中受限于手工操作的通用工具,在资本控制的分工进程中被拆解重组,实现了器件层面的机械化和系统化变革。“劳动资料经历了各种不同的形态变化,它的最后的形态是机器,或者更确切些说,是自动的机器体系。”(马克思、恩格斯,2009a:184)马克思认为劳动资料的物质属性与社会属性之间相互作用,推动劳动资料发生了革命性转换,工人所面对的不再是单个工具,而是涵盖工具机、动力机和传动机在内的一整套机器体系。

局限于工具的生产时代是前大工业时代。在工场手工业时代,“范围有限的知识和经验是同劳动本身直接联系在一起的,并没有发展成为同劳动相分离的独立的力量”(马克思、恩格斯,2009a:357)。脑力与体力都集中在劳动者身上,马克思形象地称之为“手和脑还没有分离”。知识以局限性技巧的形态存在,并随着手工劳动施加于产品之上,劳动者通过对技艺的掌控在生产过程中把握着部分主动权。而在机器大工业时代,“使用劳动工具的技巧,也同劳动工具一起,从工人身上转到了机器上面”(马克思、恩格斯,2009b:483)。工人所熟稔的手工操作被工具机取代,简单体力输出作用被动力机取代,工人群体中世代沿袭的生产技艺也因为劳动载体的废除而消弭,工人彻底沦为了一种机器零件,丧失了独立地位。复杂工艺中蕴含的工人的智力因素得以独立化与外化,成为由技术和研发人员等脑力劳动者专门负责的职业,并固定于工厂之外。劳动资料的机械变革割裂了与传统手工劳动的原初统一,解放了原本由人手限制的生产效率,极大地提高了社会生产力。而手工劳动的智力因素被固定到机器之中,成为工业革命发生的智力支撑。尽管机器本身就是人类智力发展的高度凝结,但机器并非智力的直接来源,不能

自主完成脑力工作。在智力层面,机器仍具有工具属性,是一种“智力工具”。

然而劳动资料的演进并没有停留于机器对体力劳动的简单替代,而是在两个方向继续演化,一个是新的机器构件的形成,另一个是机器构架的独立化。自瓦特(J. Watt)借助离心调速器改良蒸汽机以实现转速自动控制以来,作为一种独立元件的控制机雏形就已初步显露,无论是纸张计数器还是键控穿孔检验器都是为了部分实现生产过程的自动控制,即“根据来自直接作业机械装置以外的信息对机器进行控制”(布雷弗曼,1978:169),控制机很快成为机器系统中的必要构成。就机器构成发展来讲,直到1969年可编程逻辑控制器(PLC)问世,机器生产的控制才得以脱离劳动者的外部直接管理,进入自动化生产阶段。“可编程逻辑控制器就是原始的计算机”(蔡斯,2017b:14),其本身就是打孔控制技术与晶体管技术结合的产物,构成了生产领域的机器大脑。就技术逻辑来说,正是雅卡尔提花机中的打卡控制系统促使巴贝奇(C. Babbage)构想出了计算机的前身分析机,而且这种打卡控制的想法被图灵(A. Turing)用以阐发现代计算机的数学概念(胡翌霖,2020:114)。控制机革命不仅生成了机器的控制器官,而且塑造了计算机这一外在于人脑的机器大脑。

另一个重要趋势是机器构架的独立化。大工业时代机器体系各组成部分的相互作用“只有在整个机器体系由一台自动机推动的机械工厂中才得以完成”(马克思、恩格斯,2009a:329),但伴随着资本扩张所带来的远距离信息和能量传输需求,动力机、传动机和工具机不再局限于工厂空间内部,而是趋向离散化、专业化,形成了发电厂、输电网络、车间、公司等专门场所。更重要的是,受益于电力革命,包括电报、电话在内的远距离信息传输技术开始迅速发展起来,“新信息传递工具的发明及其广泛使用,大大延伸和扩展了人类的感觉器官和神经系统”(陈万球,2023),并经由电子产业催生了互联网这一革命性成果。以数据处理为核心的计算机和以数据流传输为核心的互联网,构成了人类外部的大脑器官和神经网络,开启了数字革命的新浪潮。

数字化革命最突出的特征就是信息数据日益占据产业发展的核心地位，作为数字时代的战略性资源，数据取得了与工业时代的石油资源相当的地位，构成了数字时代机器运行的基础。凭借“能独占、提取、分析和使用记录下来的日益增加的数据量”（斯尔尼塞克，2018：49）的功能，平台晋位为数字经济的基础设施，成为互联互通的中介环节。

首先，相比手工业和工厂中实体性的劳动资料，网络平台的劳动资料具有虚实共存的特征。依托于计算机和互联网，网络平台是实体终端、数据链条和云端数据共同构成的信息体系，在职能上形成了信息收集、信息传输、信息储存、信息处理和信息反馈的协动体系。平台开辟了与工具链并行的数据流，通过重构劳动者与劳动资料、劳动对象的组合方式提高了生产效率。其次，平台具有开放互联的特点。如果说机器体系通过产品生产的顺次工序聚集，独立化的机器工厂通过电力线路远距离结合，那么网络平台就处于信息交流的中心节点，网络生态中的空间优势赋予了它在多维度上联结各终端的能力。平台企业借助红外、电磁、震动等传感手段，拓展网络爬虫、Cookie 授权和信息交易等数据获取渠道，并辅以深度学习等先进算法来提高处理和控制信息的能力，通过监控网络多元主体之间的数字活动，在数字经济中占据垄断地位。最后，平台具有多元集中的特点。与工厂变革过程中实现机器设备的集中化使用不同，平台是建立在去中心化网络基础上的再集中，这种区别于传统集中组织的劳动控制塑造了弹性时间、免费服务与自由选择的新自由主义意识形态，削弱了平台劳动者和平台用户的受压迫感。但是隐藏在表面自由背后的是更为精巧且深入的剥削，零工经济中生产界限的模糊无限制提高了劳动强度，劳动者同时失去了制度保障下的福利措施，而劳动者与用户的信息又按照服务条款为平台所有，“平台正是其所建构的生态系统的绝对中心”（吴静，2002）。

二、从总体工人到数字劳工:劳动主体的数字扩散

在《资本论》“分工与工场手工业”一节中,马克思着重分析了工场手工业中“局部工人”与“总体工人”的辩证关系。分工基础上的协作拆解了原本独立化的手工业工人的劳动,按工场内部生产工序安排工人终生重复一种局部操作,使得“他们的片面活动现在取得了一种最适合于狭隘活动范围的形式”(马克思、恩格斯,2009b:391),这种碎片化操作在工人身上的固定化最终塑造出一种“局部工人”,而各个生产环节的局部工人组合成为工场内部活的总机构,即“总体工人”。总体工人首先标志着对局部工人的高效率的整体运用,这种整体性的效率既体现在总体工人对局部工人片面发展出的专业能力的吸收运用上,还表现在总体工人赋予各局部工人一种量的规定性上,形成了各局部工人之间有机协调的比例关系。总体工人的另一重表现是,相比于早期手工业中工人的独立属性,各局部工人组合为一种社会结合劳动,这在工场手工业时代表现为工场内部的社会化联合,而机器大工业对生产领域的扩展则使得总体工人的社会属性充分体现出来。针对大工业社会涵盖管理者、设计者以及技术人员在内的全新劳动主体的生产性,马克思(马克思、恩格斯,2009b:582)指出,“只要成为总体工人的一个器官,完成他所属的某一种职能就够了”。在《1863—1865 年经济学手稿》中讨论生产劳动与非生产劳动时,马克思(马克思、恩格斯,2009a:524)还预测了所有使资本实现增殖的服务活动的生产劳动性质,“所有服务的执行者都转化为雇佣工人,从而都具有这种与生产工人相同的性质”。马克思对总体工人的指认,针对的是大工业生产中脑力劳动与体力劳动分离的情况,以及脱离直接劳动的工人数量不断增加的趋势,以便及时调整生产劳动者的范围。

尽管马克思对总体工人的界定昭示着非体力形态劳动者的出现,但新兴劳动主体的形成随着后工业社会的来临与福特制度的瓦解而愈发重要,数字

时代的到来将劳动主体变革置于学者议题的中心。哈特(M. Hardt)、奈格里(A. Negri)在对非物质劳动具体形式的分析中,阐发了非物质劳动的数字劳动维度,指出数字劳动的非物质表现为通信技术与工业生产相互融合塑造出的生产数字符号乃至交往关系的劳动形式:“不断地广泛应用计算机已逐步趋于对劳动实践和关系以及所有的社会实践和关系进行重新定义。”(哈特、奈格里,2008:284)克里斯蒂安·福克斯(C. Fuchs)则基于数字产业的全球布局推衍出一个涵盖数字产业物质与非物质全领域的数字劳动主体概念,“涉及数字媒体技术生产和内容生产中的所有活动”(福克斯,2020:458),数字化总体工人扩展到数字产业全球链条上的一切被压迫者。如果说总体工人是大工业时代劳动主体的必然状态,那么资本数字化转型所催生的劳动主体就必然符合平台企业的扁平结构。结合诸多学者对全新劳动模式的指认,可以发现数字劳工就是数字化增殖逻辑裹挟下的劳动主体。

劳动与资本的辩证逻辑,要求劳动者与资本之间保持一种动态互斥的矛盾状态,无产阶级的全新主体就是被数字资本逻辑支配的劳动者。区别于传统工业社会的两极支配格局,当下的资本支配呈现出一种网格状的中心辐射结构,以掌握数据信息中心的多个平台企业为核心向四周辐射,并在每个信息结合点上汇集。在这种以信息流动为核心的支配体系中,无产阶级的数字肉身,即平台用户,成为资本增殖的数字维度上的剥削对象。这种新形态数字劳工的主要劳动形式就是数字劳动,与数字劳动形式同时生成的还有劳动者的全新主体性的塑造,“这一主体性主要在知识、交往和语言的领域内发挥作用”(哈特、奈格里,2008:31)。

作为数字时代的劳动主体,数字劳工是总体工人在数字化时代的社会扩散,在多个方面呈现出区别于机器大工业时代总体工人的新表征。首先,在劳资关系上,可以发现数字劳工与总体工人的不同之处在于,后者需要进入交换体系中才能为资本服务,但数字劳工却不需要明确的雇佣关系存在就可以被资本利用。资本借助平台这一实质从属的全新形式,将数字媒介使用者

的生命活力、注意能力、精神潜力、社会合力全都纳入资本增殖逻辑之中。继总体工人在机器大工业时代拓展到脑力劳动者和服务工作者之后,数字平台的信息剥削使其进一步将非雇佣劳动者纳入进来。其次,在劳动者之间的分工体系方面,资本通过将数字劳工片面化发展为生产生产性知识的劳动者与生产消费性信息的劳动者,进一步在智力层面将劳动者禁锢在不同的数字囹圄之中,受限于彼此之间庞大的知识鸿沟与技术差距,知识生产者之间的互相认同更为困难。最后,数字劳动的非物质性使得劳动者不仅生产出作为实体产品的信息,还生产出一种同意制造机制,即自我服从的意识形态。与工业时代资本家凭借有形的监工或者无形的工资罚扣制度来约束工人不同,数字劳工沉浸在资本提供的免费服务中,以产消者取代了自身劳动者的认知,劳动主体性转化为协助资本增殖的主体性。但同时,数字劳动者的意识形态生成性将改造自身的自由本质进一步揭示,异化了的主体自为已经为最后的解放提供了条件。

三、从机器看管到认知交往:劳动活动的数字编码

“机器不是使工人摆脱劳动,而是使工人的劳动毫无内容。”(马克思、恩格斯,2009b:487)马克思在分析直接生产过程中机器对劳动者体力和智力的排斥效应时,指出在机器所蕴含的一般社会知识面前,手工业工人赖以为生的技巧经验全部失效,只能作为有意识的肢体服务于机器。劳动者只是负责看管机器,保证其正常运转和防止其出现故障,“表现为人以生产过程的监督者和调节者的身份同生产过程本身发生关系”(马克思、恩格斯,2009a:196)。

一方面,生产过程内部劳动者体力因素与智力因素的彻底驱逐决定了机器看管活动的非人属性,甚至很难将其称为一种劳动。与劳动工具发生直接关联的劳动活动确证着工人的主体地位,工具成为劳动者挥洒自己体力、运用自己智力的物质载体,在工场手工业中工人仍旧凭借其技艺的独占地位与

分工原则抗衡;但是大工业时代,“机器则代替工人而具有技能和力量,它本身就是能工巧匠,它通过在自身中发生作用的力学规律而具有自己的灵魂”(马克思、恩格斯,2009a:185),工人只是单纯作为耗材被添加进生产过程中,负责给机器打下手。资本在将工人解放为所谓“自由劳动者”之后,再一次将工人抛入更为沉重的物化之中。另一方面,大工业的发展瓦解了生产领域内部劳动者的智力因素,同时又将生产领域拓展到工厂外部。与机器照看活动相对应的,是工厂外部脑力劳动者的开发、设计、检验、调节、运维等一系列被扩充进生产过程的智力活动,可以说外部脑力劳动者的被剥削直接造成了工厂内部劳动者的自我丧失。智力活动的资本化扩展了生产劳动的范围,提高了实现资本增殖的剥削基数。

伴随着机器自动化进程不断地削减工厂内部的劳工数量,交往数字化也将外部脑力劳动者的活动复杂化、情感化、社会化,驱使单纯脑力劳动转化为直接改变情感心理的认知交往活动。在分析完机器劳动使得工人转化为外在于工厂的监督者和调节者后,马克思(马克思、恩格斯,2009a:196)旋即指出:“关于机器体系所说的这些情况,同样适用于人们活动的结合和人们交往的发展。”也就是说,人们在工厂劳动过程中的交往被挪移到工厂外部的社交场域中实现,特别是随着数字媒介时代的到来,脑力劳动被塑造为一种纷繁复杂的认知交往网络。

首先,数字时代的认知交往具有物质生产性,以人与物之间的生产交往为基础并反过来对生产活动加以调节。在《德意志意识形态》中,马克思通过揭示费尔巴哈等一批德国哲学家的唯心主义本质,廓清了意识形态的物质根基,指出“发展着自己的物质生产和物质交往的人们,在改变自己的这个现实的同时也改变着自己的思维和思维的产物”(马克思、恩格斯,2009c:525),提出了历史实践基础上的交往观,明确了物质交往之于精神交往更为根本的地位。马克思对机器大工业时期生产领域扩展的研判,进一步确证了精神交往所具有的生产性质,因而在发达资本主义时期,“作为生产者出现的,是社会

活动的结合”(马克思、恩格斯,2009a:200)。数字资本主义时代主客体之间的数据交融塑造了全时在线、人机互联、虚实共生的多元数字症候,工厂劳动与社群交往之间的空间界分逐渐消弭,认知交往愈发获得生产的性质。

其次,数字时代的认知交往是一种高度社会性的活动。这一点为当代西方马克思主义者所注意,并将其曲解为“交往对劳动的替代”。哈贝马斯(1999:49)就将劳动规定为“工具活动”,将交往规定为“以符号为媒介的相互作用”,并在此基础上确立起交往行为相对于劳动行为的优先性,这就忽略了劳动本身的交往属性,颠倒了物质生产与精神生产之间的关系。因而,哈贝马斯将资本主义的异化现象解读为作为劳动领域的“系统”对作为交往领域的“生活世界”的殖民化,但数字时代的交往构建绝非生产领域对纯洁无瑕的交往领域的入侵,而是这种社会性、生产性的交往互动被资本所宰制,是资本对发展起来的社会劳动力的非法篡夺。区别于传统主客二分的思维,认知交往的数字构型表现为主体借助数字中介于诸主体、诸客体之间流动,主客体的“虚体”存在构成了数字网络内部的联结点,其独特之处在于数字交互并非主体与客体之间的单向映射,而是交往内部各单元的去中心化组织,“它是一种网络,而这种网络依赖于各个节点”(蓝江,2022:31),这彰显出马克思社会性概念的原则高度。

最后,数字时代的认知交往是具有高度主体性的活动。依照生产组织变迁的理路,传统大规模制造业中的产品冗余和创新迟滞等因素导致福特主义生产模式的效率峰值逐渐衰减,而精益生产的弹性模式带来了高效率信息传递需求,重塑了精细化的管理单元,生产劳动的主体性得以凸显。在认知资本主义者、自治主义马克思主义者等看来,数字交往的认知属性应当被理解为“非物质性”,在这种非物质劳动中,“集体智慧成为主要的生产要素,成为财富和价值的真正实质”(Boutang, 2011: 30)。一些西方学者以劳动的非物质倾向来质疑马克思劳动价值论的当下合理性,但实际上,隐藏在知识生产、创新才智等劳动精神维度背后的是劳动者主体性的彰显。所谓“非物质劳

动”就其本质来讲是一种虚拟实践,“这种实践性是有意图、有方向的,是朝向人与社会之间的行动的,也是朝向人自身的改变的”(杜骏飞,2021),是被数字交往所解放出来的人的主体性本质。

四、从实际从属到算法监控:劳动控制的数字调制

“作为劳动过程和价值增殖过程的统一,生产过程是资本主义生产过程,是商品生产的资本主义形式。”(马克思、恩格斯,2009b:230)马克思指出,资本主义商品生产的过程不仅是劳动过程,更是实现价值增殖的过程。在产品生产向商品生产的转化过程中,在场的并不仅仅是劳动者、劳动资料,甚至不主要是劳动者、劳动资料,真正主宰二者的是资本。正是资本对劳动力以及劳动资料的占有,才使得一无所有的自由劳动者和归资本家所有的劳动资料得以组合起来。马克思将资本家对劳动能力的购买与消费意义上的劳动从属称为形式上的从属。这种形式从属除了确立起资本对劳动的统治关系外,在工艺维度和生产方式上与前资本主义生产并没有区别,因而资本只能通过攫取绝对剩余价值来实现增殖。资本为了更高效地剥削剩余价值,借助协作生产、分工组合、技术改造、自动机器等方式不断改进制造工艺和生产方式,释放出生产过程中的社会生产力,推动劳资关系从形式从属向实际从属转换,“劳动对资本的实际上的从属是在创造与绝对剩余价值不同的相对剩余价值的一切形式中发展起来的”(马克思、恩格斯,2009a:383),这一过程恰恰是相对剩余价值生产的形成过程。

实际从属的形成标志着资本对劳动创造出来的社会生产力的增殖化占有,是资本对劳动控制的日渐深化。首先,这一过程是活劳动与死劳动的敌对过程,是死劳动支配活劳动的过程,即“活劳动是替积累起来的劳动充当保存并增加其交换价值的手段”(马克思、恩格斯,2009c:726)。活劳动作为劳动者自身生命力量的对象化实现,作为人体力与脑力的支出和消耗,是即时

改造外部世界的过程;而死劳动作为物化劳动,作为凝固状态的活劳动,往往以劳动产品或者劳动资料的形式表现出来。机器体系使得“对象化劳动在劳动过程本身中与活劳动相对立而成为支配活劳动的力量”(马克思、恩格斯,2009a:185—186)。活劳动由作为一种创生性的原发力量转化为机器这一物化产物,工人当下的活劳动被物化劳动所控制。考虑到对象化劳动本身也是工人既往创造对象的实体化,那么死劳动、活劳动之间的对立就成为工人自己反对自己的过程,是资本借助工人的过去来否定工人的当下,反映了资本支配下劳动者的时间悖谬。其次,实际从属的形成也标志着工人主体性的消解,即从手工业工人之间的有机合作向服侍机器的肉身零件之间的机械联结转换。如果说在工场手工业时期,工人借助自身的独立地位和集体的生产智慧取得了一定议价实力,那么机器操作的细碎化则加剧了工人劳动过程的异化,工人彻底沦为机器的零件。机器对社会劳动的吸纳使得工人的知识优势彻底沦丧,成为单纯体力劳动的输出者,劳动者的独立主体性地位最终丧失。最后,从形式从属到实际从属的演进过程也与资本对劳动的监控活动息息相关。与资本对劳动形式上的吸纳相对应的是资本对工人的直接监视,这既是为了防止工人破坏资本所有的劳动条件,也是为了高效地消费所买到的劳动力商品,“因而是在资本家的监督和管理之下进行这种消费”(马克思、恩格斯,2009a:372)。但是伴随着生产过程中机器的运用,工人对劳动环节的实际从属使得资本家开始脱离直接生产过程,采取一系列包括罚金簿、计件工资在内的间接措施来对工人进行监管,进一步完善了工作场所的监控体系。

伴随着福特主义巨量生产催生的巨量消费需求,资本亟须将休闲、消费、生育乃至整个再生产领域囊括到监控的视域之中,以确保剩余价值的成功转化,这揭示了内在于资本的原始冲动,即疯狂地寻找实现剩余价值最大化的方式。当下资本则通过将数字技术纳入劳动支配的进程中实现了实质吸纳的数字转化,将机器体系向着数字平台推进,劳动控制的大数据耦合催生了监控方式的算法革命。

如果说资本的一般智力捕获是通过对脑力劳动者的控制和机器的占有来实现的,那么数字时代的劳动控制方式就转化为无处不在的电子网络对个体全部细节的监视控制,即德勒兹(2014:196)所说的“调制”(modulation),“控制是一种调制,像一种连续的、每时变化的、自动变形的造型,或像一种每个网孔点都在变化的网筛”。数字网络对实质从属的改造就是一个调制过程,即在不改变原有控制效果的条件下,将全新的数字控制机制融入进去,从而达成一种数字资本与实质从属的榫合结构。

具体到劳动者的实际控制,一方面,调制过程需要对异质性的个体进行渗透,“曾经看似稳固固定的监控,如今已变得更加灵活机动,渗入并蔓延到许多生活领域”(Bauman & Lyon, 2012: 3),构成了数字时代资本监控的核心表征“流动”,数据监控突破了闭路电视系统的固定视角,呈现出去中心化、解辖域化的弥散趋势,借助移动设备的全感官信息窃取延伸了日常生活领域的资本权力触角。另一方面,调制又能够通过量化处理、信息屏障、价值涓滴、记忆重塑等方式对个人进行控制,将资本意志贯穿于个人心理层面,一个全知全能的数字虚体裁决着个人的所思所想所求,这构成了数字时代劳动控制的算法监控逻辑。可以发现,劳动者与资本之间的关联变得更加稳固且隐蔽,劳动对资本的依附性更强,由原本劳动力恢复与维持的成本需要资本支付,转化为劳动力的精神需求需要资本予以满足。在约迪·迪恩(Dean, 2012: 129)看来,以认知交往为操纵对象的监控方式实现了资本对人的劳动能力、情感心理乃至社会关系的直接剥削,“交往资本主义攫取了社会实存,将其私有化,并试图将其货币化”。

五、从一般智力到通用智能:劳动解放的数字赋能

机器大工业废除了工场手工业中的分工原则,机器以固定资本的形态将生产力从人手的局限中解放出来,“固定资本的发展表明,一般社会知识,已

经在多么大的程度上变成了直接的生产力”(马克思、恩格斯,2009a:198)。可见,马克思将一般智力的发展看作直接生产力的进步,而这种生产力却转而为资本所挟制。一般智力是在直接生产过程中智力因素的剥离过程中逐渐形成的,“这个分离过程在简单协作中开始,在工场手工业中得到发展,在大工业中完成”(马克思、恩格斯,2009b:418),机器作为固化的一般智力成果彻底完成了这一分离过程。但一般智力除了被机器所吸纳外,还通过资本的教育因素赋予外在于直接生产过程的脑力劳动者,他们构成了一般智力的活劳动载体。因此,一般智力的运作具有物化产物和对象化活动双重载体。马克思同时清醒地认识到,资本主义生产关系下技术的发展最终也逃不脱成为资本剥削帮凶的宿命,作为社会知识的物化凝结,机器以固定资本的形态站在了劳动者的对立面。从这个角度上说,一般智力内蕴着活劳动与死劳动、活知识与死知识的对立。

但一般智力同样具有主观与客观双重向度的解放效用。首先,资本主义在划分了必要劳动与剩余劳动的基础上确立了自己的增殖逻辑,压缩必要劳动增加剩余劳动的内在驱力——对相对剩余价值的追求——使得资本被迫进入强制进步的循环,资本对一般智力的吸纳使其在压缩必要劳动方面完成得过于优秀,最终只能以终结自己来完成自身,为劳动解放确立了客观的现实条件,“以交换价值为基础的生产便会崩溃,直接的物质生产过程本身也就摆脱了贫困和对立的形式”(马克思、恩格斯,2009a:197)。其次,生产和生活资料的高效率生产在不断缩减必要劳动的同时,也为劳动者发展自身提供了时间可能,劳动者素质的发展又成为生产率提高的动因。与此同时,资本主义多样化社会分工的生产需求也使得资本家致力于培养出全能的劳动者,“用那种把不同社会职能当做互相交替的活动方式的全面发展的个人,来代替只是承担一种社会局部职能的局部个人”(马克思、恩格斯,2009b:561),这就为工人个体的发展创造了外在动力,从主观维度凸显了一般智力的解放性。相比起机器大工业时代一般智力的解放可能,资本增殖的数字技术耦合

和无孔不入的劳动渗透似乎已经将资本权力毛细血管化,消解了一般智力的解放潜能。但是,一般智力并没有停留于机器大工业的双重存在,而是进阶为一种“通用智能”,赋予劳动解放以全新的生命力。

“生产方式的变革,在工场手工业中以劳动力为起点,在大工业中以劳动资料为起点”(马克思、恩格斯,2009b:427),数字时代劳动资料的平台化进展表现为智能机器依托于万维网、计算中心存储和处理云端数据。平台资本主义借助大数据挖掘来提高算法效率,改进软硬件配置,更新平台架构,提高信息联结效率,并以服务条款绑定的方式来进一步掠夺信息,反过来强化自身技术优势,“在硬件、数据库组织和网络基础设施方面的进步,都为在速度和洞察力方面获得了比竞争对手更多的优势,发挥了重要作用”(斯尔尼塞克,2018:114),正是数字平台对用户大量数据的圈占和攫取,为智能机器的深度学习提供了养料,并最终衍生出一种云端智力,即通用智能。

通用智能一般被理解为一种前沿数字技术,其机器本质往往为人工智能、脑机接口、虚拟现实等术语所遮盖,而在“劳动—技术—资本”的辩证逻辑中考察通用智能,则可以总体上把握通用智能概念的三重维度。

第一,通用智能是机器智能发展的最新成果。机器体系的控制机革命使得具身性的脑力活动转化为异在的“机器大脑”,将智力从个体性智力之中解放出来,以外部机器的方式进行整合,推动了脑力劳动的自动化与社会化。同时,“智能自动化生成的‘机器神经网络’不再是‘传导’而是可以‘代替’人脑及其产生的智能”(刘方喜,2022:75)。对狭义人工智能(artificial narrow intelligence,ANI)来说,机器只是对特定领域问题的高效计算和处理,这是人类专门设计的智能,部分完成了对人脑记忆存储和理性运算能力的替代,但是并没有发展出自主性和通用性。而以 ChatGPT、盘古大模型、文心一言等大模型为代表的通用人工智能(artificial general intelligence,AGI)则试图真正地复制人脑认知框架,构造出一种人类思维的机器理解方式。当下机器智能的全领域运用将通用智能的美好构想现实化,在人脑外部培育起强大的智力

来源。

第二,通用智能标识出数字时代劳动者主体性的存在状态。在维尔诺(2017:82)看来,后福特主义时代的一般智力“首先体现着活生生的主体的交往、抽象思考和自我反思”,通用智能成为劳动者的数字存在状态。通用智能存在于每个劳动个体之中,维尔诺认为新的劳动方式可以生产一种共同性,这种共同性之于劳动者意味着共同生产、共同成果、共同享用,这种共同性的指认是在数字时代对马克思一般智力的化用,也是对主体间性智力内容的强调,彰显了普遍智能下劳动者的主体性存在。自治主义马克思主义者试图宣扬一种劳动解放的能动路径,即通过个体对一般智力的掌握和社会化运用以发扬劳动的主体性,进而终结资本统治。

第三,通用智能揭示了数字时代劳动解放的全新路径。西方左翼学者对通用智能的主体性解放路径进行了一番探索,但却忽略了通用智能的客体性以及通用智能资本主义运用的矛盾性。数字资本支配下的一般智力存在于一个物质整体之中,一个由数字端口联结而形成的网状结构,表现为人与人、人与物之间的互联互通,也可以称作一种百科式的集体智能,这就使得解放关乎联结着的每一个人、每一个事物。脑力劳动者必须通过当下的网络体系实现与体力劳动者的大联合,这种联合不单是人数上的堆叠,更是信息权力的争夺,是平台与平台的斗争。一方面,要隔绝来自资本平台的信息操控,这既是割断信息资本毛细血管般的权力触角,也是网络意识形态阵地的争夺战,因此更需要每个数字劳工的合作,构建相互之间非异化的联结方式;另一方面,要占有数字寡头的生产资料,实现数字基础设施公有化、社会化,这是数字时代社会主义构建的基本原则。

可见,通用智能的资本挟制必然激化数字资本主义的基本矛盾,即数字信息的公有本性与资本寡头占有之间的冲突。正是这些新元素,使得通用智能就其内在逻辑而言具备挣脱资本逻辑的可能性。而网络时代劳动者的联合彻底扬弃了无产阶级的分裂状态,在终结了资本主义私有制后,劳动者得

以通过体力的挥洒和智力的锻炼来发展自身的总体智能,“劳动已经不仅仅是谋生的手段,而且本身成了生活的第一需要”(马克思、恩格斯,2009d:435),最终实现人的存在意义上目的与手段的统一。

六、结语

在对机器掌控劳动背后的资本原则的政治经济学批判中,马克思运用“异化劳动”“一般智力”“固定资本”“实际从属”等核心概念分析了机器体系形成过程中劳动过程的境遇转换,通过对劳动与生产概念异质性的指认,辨识出大机器时代劳动形态的根本性变革,并基于大工业生产探讨了劳动解放的技术可能;同时指出,终结资本的无产者既在资本对所需工人的有意塑造中汲取技能,又在生产过剩的危机循环中获得斗志,进而生成超越资本原则的主体性条件,劳资关系的辩证互构揭示了大工业时代劳动范式的大转型。

延续马克思政治经济学批判的理路,相较于机器大工业,数字社会中的资本运作看似摆脱了一系列桎梏,虚构出“无摩擦资本主义”的幻梦,但是结合数字化劳动过程的当下境遇,可以发现资本主义内在矛盾并没有消弭,资本主义与数字技术的双向耦合对人的社会性本质的增殖化吸纳直接侵犯了个体存在的独立性,劳动者陷入更为深重的数字困境。但与此同时,数字时代的互联互通也必然带来劳动者多元维度的融合,万物互联驱散了笼罩在虚体之上的物性,重新唤醒了主体间性的解放脉冲,掌握了数字技术的数字劳工必然会进一步彰显自身的主体性存在,打破数字劳动的异化束缚。劳动范式的数字转型揭示了资本主义的数字形变和逻辑转轨,开显了劳动解放的当代际遇,再一次证明马克思劳动哲学以及劳动价值论在数字时代的现实生命力。

参考文献

布雷弗曼,1978,《劳动与垄断资本:二十世纪中劳动的退化》,方生等译,北京:商务印书馆。

蔡斯,2017a,《经济奇点:人工智能时代,我们将如何谋生?》,任小红译,北京:机械工业出版社。

蔡斯,2017b,《人工智能革命:超级智能时代的人类命运》,张尧然译,北京:机械工业出版社。

陈万球,2023,《人工智能时代的劳动方式变革图式》,《武汉大学学报》(哲学社会科学版)第 1 期。

德勒兹,2014,《哲学与权力的谈判》,刘汉全译,南京:译林出版社。

杜骏飞,2021,《数字交往论(1):一种面向未来的传播学》,《新闻界》第 12 期。

多尔蒂、威尔逊,2018,《机器与人:埃森哲论新人工智能》,赵亚男译,北京:中信出版社。

福克斯,2020,《数字劳动与卡尔·马克思》,周延云译,北京:人民出版社。

哈贝马斯,1999,《作为"意识形态"的技术与科学》,李黎、郭官义译,上海:学林出版社。

哈特、奈格里,2008,《帝国:全球化的政治秩序》,杨建国、范一亭译,南京:江苏人民出版社。

赫拉利,2017,《未来简史:从智人到智神》,林俊宏译,北京:中信出版社。

胡翌霖,2020,《人的延伸:技术通史》,上海:上海教育出版社。

蓝江,2022,《一般数据、虚体与数字资本:历史唯物主义视域下的数字资本主义批判》,南京:江苏人民出版社。

刘方喜,2022,《脑工解放时代来临:人工智能文化生产工艺学批判》,杭州:浙江工商大学出版社。

马克思、恩格斯,2009a,《马克思恩格斯文集》第 8 卷,北京:人民出版社。

马克思、恩格斯,2009b,《马克思恩格斯文集》第 5 卷,北京:人民出版社。

马克思、恩格斯,2009c,《马克思恩格斯文集》第 1 卷,北京:人民出版社。

马克思、恩格斯,2009d,《马克思恩格斯文集》第 3 卷,北京:人民出版社。

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。

维尔诺,2017,《诸众的语法:当代生活方式的分析》,董必成译,北京:商务印书馆。

吴静,2022,《总体吸纳:平台资本主义剥削的新特征》,《国外理论动态》第 1 期。

新华社,2023,《10.79 亿网民如何共享美好数字生活?——透视第 52 次〈中国互联网络发展

状况统计报告〉》,https://www.gov.cn/yaowen/liebiao/202308/content_6900651.htm。

Bauman, Z. & D. Lyon 2012, *Liquid Surveillance: A Conversation*, Malden: Polity Press.

Boutang, Y. 2011, *Cognitive Capitalism*, E. Emery trans., Cambridge: Polity Press.

Dean, J. 2012, *The Communist Horizon*, London: Verso Press.

政治经济学批判语境下数字异化的深层逻辑与扬弃进路

魏金鹏*

摘要:在数字时代,数字资本脱颖而出,成为引领传统资本的新形式。数字资本将数字劳动的新形式推广到全球范围,使处于数字异化状态下的劳动者陷入困境,自觉或不自觉地受到压迫。在数字劳动中,数字劳动者处于被剥削的地位。对于数字异化状态,存在论视角的批判显然不足以使之得以扬弃。目前,从政治经济学的批判角度来分析数字异化,评析几种异化扬弃状态,凸显马克思主义对异化状态的扬弃精神,通过具体的、革命的实践行动扬弃数字异化显得十分重要。

关键词:马克思主义　政治经济学批判　数字异化　扬弃

当今,快速发展的数字技术已将人们带入全新的数字时代。5G、物联网、元宇宙、Web 2.0、ChatGPT 等技术名词层出不穷,我们几乎来不及欣赏每一个科学的杰作,某项技术就已成为明日黄花,似乎技术的发展无法等待人们和技术产品之间的“共鸣”。一切似乎都如此不受掌控,但正如哈特穆特·罗萨(2022:10)所说:“但是,生命之活力、情感上的触动,与真正的经验,都是在与不受掌控之物的相遇中形成的。”在数字时代,我们与技术产品的关系发生了新的变化,呈现出新的形态。在数字时代的浪潮中,我们越想抓住解放的浮木,就发现自己越是深陷于大数据的海洋,成为一片数据浮萍。在数字风暴中,人们渴望解放,却无力地在风暴中哀叹。毋庸置疑,我们已经陷入了数

* 魏金鹏,东北师范大学马克思主义学部。

字异化之中,因此揭露数字异化的神秘面纱,使人们从数字异化的状态中解放出来,成为当务之急。

一、数字异化的历史根基与脉络

(一) 数字异化的理论资源

就像印刷技术粉碎了教士对知识的垄断,火药摧毁了骑士的城堡一样,当今数字技术也以轰轰烈烈的方式展开了自己的变革。数字时代已经来临,新形式的异化状态也随之出现。在此,我们需要回顾异化观的理论源头。首先,在黑格尔的意义上,自我具有纯粹的否定性,异化是人的客体化活动,它本身就是自身实践活动的结果。然而,黑格尔(2013:74)所指的异化是精神的异化:“即是说,抽象的东西,无论属于感性存在的或属于单纯的思想事物的,先将自己予以异化,然后从这个异化中返回自身,这样,原来没经验过的东西才呈现出它的现实性和真理性,才是意识的财产。”在黑格尔的论述中,异化是精神自身的异化,是精神自身的运动。通过这个运动,精神确认其实存,以实现黑格尔哲学体系中的“现实”。

异化观念在青年黑格尔派中得到了发扬。与黑格尔的现实观不同,费尔巴哈的现实观更加“现实”。费尔巴哈(2010:1)揭示了人的宗教异化:“至于那异于人的本质、不依靠人的本质的实体,亦即那不具人的本质、人的特性、人的个性的实体,真正来说,不是别的东西,就是自然。”按照费尔巴哈的观点,人的本质就是其自然属性,人就是人自身,不是上帝创造了人,而是人创造了上帝。然而,人却反过来受到自身创造物的制约。费尔巴哈高举人文主义旗帜,看到了人的实践活动,但并不理解这种实践活动,他将人的实践活动视作一种消极的机械运动,而非人类自由自觉的活动。

马克思扬弃了费尔巴哈的旧实践观,在吸收黑格尔异化理论的基础上,

对黑格尔的辩证法进行了彻底改造,逐步形成了辩证唯物主义的异化观。在《1844 年经济学哲学手稿》中,马克思提到了人的四重异化,即人与劳动产品异化、人与劳动活动异化、人与人的类本质异化、人与人相异化。此时,马克思的异化思想尚未成熟,在很大程度上是在哲学意义上使用异化概念的。然而,试图将政治经济学问题形而上学化的努力最终将是徒劳的,永远无法接触社会实际。正如马克思(马克思、恩格斯,2012:223)所说:“谁用政治经济学的范畴构筑某种意识形态体系的大厦,谁就是把社会体系的各个环节割裂开来,就是把社会的各个环节变成同等数量的依次出现的单个社会。”随着马克思逐渐清算自己的青年黑格尔派思想,他将带有形而上学色彩的“异化”概念发展为对劳动者的异化劳动批判,并在此基础上扩展到对资本主义生产方式的整体批判。资本主义社会的社会关系是处于异化状态中的,生活在资本主义社会关系下的人也是处于异化状态中的。

(二) 数字异化的新形式和新特点

数字异化正是在传统的异化观点基础上发展起来的,它既具有传统异化观的特征,又呈现出数字时代的新形式,表现出新的特点。数字异化是当代异化劳动的一种新形式,表现为人作为实践主体在数字劳动过程中创造劳动产品,但这些产品却被数字生产资料的所有者所占有。在数字时代,人的自由自觉活动变成了异化的活动,被外在力量反过来控制着自身。“手推磨产生的是封建主的社会,蒸汽磨产生的是工业资本家的社会。”(马克思、恩格斯,2012:222)现代社会是科学技术迅速发展的社会,科技的进步推动了社会生产力的提升,劳动者的劳动形式发生了巨大变化,实践活动也展现出新的特征。虚拟现实技术创造了元宇宙虚拟空间,为人类的创造活动提供了新的空间;万物互联打破了传统的时空界限,劳动者不再局限于传统的办公室和车间;数字技术蓬勃发展,使得数字劳动成为一种新的劳动形式。与传统劳动相比,数字劳动具有以下新特征:灵活性、互动性和逃逸性。

首先,数字劳动具有灵活性。数据是数字劳动所需的必要原料,而数据并没有物理意义上的体积。因此,在工厂或办公室以外的地方,也可以使用数字技术进行数字产品的生产,数字劳动者的工作场所范围得到了扩展。借助技术手段,劳动者可以在家中、游乐场、餐厅等任何地方进行数字劳动,只要能够通过网络连接,传统的劳动场域被彻底打破了。

其次,数字劳动具有互动性。与传统劳动模式不同的是,在数字劳动中,劳动者不再被迫服从劳动工具,而是享有更多的“自我掌控”能力。他们能够通过自身的劳动改变劳动活动的形态。数字劳动者所使用的工具往往是日常生活中必不可少的,如智能手机等。因此,每个人在使用这些工具的同时,也将自己创造的数据反馈到大数据集成机制中,共同塑造数据共同体。

最后,数字劳动具有逃逸性。数字劳动的一种形式是“玩劳动”,通常在人们享受休闲生活时进行,如玩电子游戏、浏览娱乐视频或进行网购等。在这些活动中,人们实际上就是在进行数字劳动。数字资源的所有者将数字劳动包装为“解放”手段,劳动者在不经意间进行劳动,却未能察觉自身被剥削的状况。

这些新特点掩盖了数字劳动在资本介入下的剥削性,以及数字劳动与传统劳动形式之间的共同特点。然而,依然存在着劳动者与自身劳动产品、劳动手段的异化,以及剥削者与被剥削者之间的类本质的差异。这使得数字异化成为人们日常生活中潜藏的一部分,人们既无法完全享用自己生产的数据产品,又无法意识到自身的异化状态。正如蓝江(2022:118)所指出的:“产品中所凝结的工人的生命力被资本家占有,由此,人与人之间的关系发生了异化,一部分人成为另一部分人的统治者和压迫者。”

(三)数字异化的动能和限度

数字劳动以一种新型的“解放”方式渗透到人们的生活中,凭借其灵活性、互动性和逃逸性的特点,使人们“自愿地”认同并投入这种劳动方式中,夜

以继日,乐此不疲。然而,数字劳动并未改变其异化劳动的本质。正如周光港、赵海月(2022)所指出的:“数字劳动始终是资本用以实现自身增殖的手段,这就是数字资本主义条件下的数字劳动的内在本质。”数字劳动的异化是数字异化的一个重要方面,在人们陷于数字异化的时代,互联网精神荡然无存。本应以共享、合作为目标的互联网,在资本介入下成为资本扩张的工具和爪牙。数字剥削者(互联网平台)以剥削数字劳动者(大众)为目标,导致数据的资本化。在数字劳动中,数字产品越多地由数字劳动者生产出来,数据的资本化就越深入,进而更加促进数字产品的生产。更重要的是,数字劳动者与数字剥削者以及生产出的数字产品之间存在不公平的分配关系。尽管数字劳动者是数字产品的创造者,却无法完全占有和享用自己生产的数字产品。数字产品被数字剥削者通过技术手段(用户协议、大数据技术)占有,使得本应属于数字劳动者的数字产品成为与他们相异的力量。

随着数据资本化的进一步发展,数字剥削者将加强对数据的控制,引入更科学的算法,以更多地占有数字劳动者生产的价值,并进一步扩大已有的数据。正如马克思(马克思、恩格斯,2012:648)所指出的:“和任何其他一定的生产方式一样,把社会生产力及其发展形式的一个既定的阶段作为自己的历史条件,而这个条件又是一个先行过程的历史结果和产物,并且是新的生产方式由以产生的既定基础。”在数字时代,社会生产力水平相较于以前有了巨大提升,这为数字资本主义生产方式的发展提供了客观条件。在数字劳动中,数字技术是不可或缺的要素,而数字技术所依赖的数据则是最重要的生产资料。因此,数字剥削者拥有更多的数据,就能更轻松地占有更多的数字产品,进一步推动数据的资本化。如果数字劳动者能够完全占有和享用个体创造的数字产品,而不被数字剥削者无偿占有,那么这种劳动就不再是剥削的数字劳动,也不存在异化。然而,数据的资本化使生产过程中的生产资料被数字剥削者所占有,数字劳动者与其生产资料分离,加剧了数字异化程度。为了在数字劳动过程中剥削更多的价值,数字剥削者更倾向于将数字生活方

式普及到整个社会。当社会的所有成员都被纳入数字劳动过程中,新的社会关系得以建立,数字剥削者对数字劳动者的奴役达到顶峰,而数字劳动者却对此浑然不知,这就是处于数字异化下的社会关系。

二、数字异化的发生逻辑

(一) 数字化生活的扩展和失序

数字异化的发生具有时间维度的连续性和空间维度的延展性。在前数字时代,数字技术的发展水平尚未能够完全改变人类的生活。随着数字技术的进一步成熟,生产方式也发生了变化。与数字技术相适应的是数字生产关系,数字剥削者为了获取更多利润、占有更多数据,不惜将数字劳动推广到社会生产的各个方面,由数字技术创造的数字空间也被纳入数字大生产的范畴。资本化的数字空间打破了传统的劳动时空观念,改变了传统的劳动形态,塑造了新的异化关系。

因此,在数字时代,人们进入了这样一种状态:在劳动者主体的感觉中,人的数字活动表现为对数字产品的占有和享受;但随着数字活动创造新的价值,而这些新的价值却无法被劳动者所使用,便产生了一种对数字生活的无力感和茫然。数字资本的“幽灵”令人深深“着魔”。诚然,数字技术给我们的生活带来了新的变化。可以说,在当前的数字时代,离开数字技术就很难正常生活。表面上看,人们似乎是主动拥抱新技术,而新技术在发挥自身效用的过程中造福人类。然而,卢卡奇(1999:163)指出:“要使资本主义生产完全产生效果的前提成为现实,这种变化过程就必须遍及社会生活的所有表现形式。”从数字剥削者的角度看,大众是被迫接受数字生活的,使资本增殖成为数字剥削者的“神圣使命”。“资本增值逻辑对人类数字生活的全面侵入,使得‘时间节约型经济’的数字泰勒主义模式日益稳固,个体的居家生活、休闲

娱乐、沟通交往、旅行交通等各个方面都被其窥探与控制，催生出人类生活维度的数字异化。”（成龙、王楠，2022）

大数据监视和控制了我们生活的各个方面。数字技术本应为我们的生活服务，提高我们的生活质量，让我们拥有更多自由时间，但它却成为我们的影子，与我们时刻相伴、如影随形。它随时准备将我们吞噬到数字黑洞中，我们逐渐失去了对美好数字生活的掌控。沉醉于其中时，我们往往察觉不到自身的异化状态。资本的无限扩张在数字空间中不断升级，市场秩序成为数字空间的绝对秩序。数字技术的发展将人与人交往的空间拓展到虚拟世界。丹·席勒（2001：2）指出：“事实上，电脑网络空间远未将我们带入一个高科技伊甸园，相反，市场体系正在借助其中为人熟知的作用迅速将其征服。”资本通过无处不在的计算机网络拓宽了市场的有效范围，在数字技术和占有数据资料的基础上，创造了作为市场交易中介的网络平台，形成了新的生产方式。在这个平台中介的基础上，出现了新的资本形态——数字资本。

（二）数字资本的产生与聚集

在资本主义发展史上，资本在传统产业资本的基础上发展出了金融资本。传统的资本形态依靠商品作为价值载体来实现资本的增殖，成为资本的物质基础。正如吴欢、卢黎歌（2023）所指出的：“抽象意义上的资本价值生产和增殖过程必须以商品这一物化的价值载体来实现，价值载体即商品的数字化是数字资本形成的物质基础。”在数字时代，数字产品充当资本运转中的商品角色，承载了数字劳动者的无差别一般劳动。从数字剥削者的角度来看，如果数字劳动者生产的数字产品不能为数字剥削者创造利润，那么数字剥削者期盼的资本增殖就无法实现。

数字产品作为价值载体，是数字劳动者在数字劳动过程中创造出来的。然而，我们不能仅仅将数字劳动看作单纯的商品生产过程。正如马克思（马克思、恩格斯，2009：485）强调的：“资本不是物，正像货币不是物一样。”资本

不是一般的物,而是其背后的社会关系。数字劳动的过程也是生产新的社会关系的过程,新的生产方式形成了新的社会关系,其中包括数字剥削者对数字劳动者的剥削关系,即数字资本。数字资本作为新的资本形式,在数字时代诞生并具备额外的权力。在商品流通领域,数字资本的生产方式引导着商品生产者摆脱市场的盲目性,使产业资本家的产品更好地适应消费者的需求,并降低其获取市场反馈的成本。在金融领域,数字资本的运算能力为金融资本家提供了更加有效的投融资机制,增强了金融资本的抗风险能力,使金融资本以更好的状态投入流通之中,实现资本增殖。

然而,数字资本本质上依然是资本,依然带有旧的生产关系性质,数字剥削者对数字劳动者的价值剥削的实质没有变,数字资源的私人占有没有变。数字资本站在资本形态的塔尖,引领着传统的产业资本和金融资本。相比这两种传统的资本形态,数字资本同样有着无限增殖的需求,甚至比它们更加需要自身的增殖。在数字时代,没有流量就等于死亡。缺乏作为生产资料的数据,数字资本的生产关系就无法进行再生产,也无法达到资本增殖的目的。因此,数字资本必须将每个自然人都转化为数据人,通过算法机制达到对每个个体信息的掌握和控制。正如蓝江(2022:229)所言:“这种再生产就是通过算法治理来实现的,算法实现的基础就是不断地将现实社会中的个体再生产为数据人,并嵌入到巨大的算法治理的框架中。”我们不再是我们自身,而成为算法中的一串数字。为了完成对数字劳动者的剥削,数字资本不仅需要将所有的时空维度都纳入算法机制框架中,还需要让数字劳动者在不知不觉中沉溺于数字异化的状态。

(三) 数字技术意识形态的形成

在现代背景下,科学技术与资本的结合更加紧密:一方面,科学技术的发展为资本扩张提供了效率更高的技术手段,而资本扩张带来的利润又反哺科学技术的进一步发展,使得科学技术发展与资本扩张互相促进;另一方面,科

学技术的发展不仅为资本提供了新的技术手段,还在社会意识方面形成了新的强力影响,即一种“有效等于有价值”的意识形态,它以工具理性遮蔽了价值理性,将有效性与价值联系起来。这种畸形等式不断扩展,即“技术越有效就越有价值”,成为人们的共识。同时,资本的扩张也不仅仅是为科学技术的发展提供资金支持,资本依靠其“全民化”“普遍化”的运作逻辑,使得新的科学技术得以渗透到人们的日常生活中。

表面上看,这种资本和技术的双生是一种良性循环。数字技术为现代社会的人们带来各种便利,使人类迈入数字时代,人们的生活水平日益提高,数字技术为人们创造了更多的空闲时间,似乎人的解放就在眼前。然而,当我们深入分析这一体系的核心,就会发现我们忽视了技术、数据和资本都由数字资本的既得利益者所控制。作为数字生产者的劳动者不仅无法完全享有自己创造的劳动产品,还在意识形态上完全受到资本和技术联盟的操控,陷入深深的异化状态。现代科技企业普遍倡导人们过上“科技生活”,打出“科技引领未来生活”的口号,将产品技术不成熟的问题归结为产品使用者不能适应新的技术手段。例如,智能汽车驾驶事故被科技公司归咎于司机操作不当,网络平台上用户信息泄露被科技公司解释为“用户风险意识不强”,科技公司的新操作系统无法满足用户需求却拒绝接受任何反馈。

不仅科技公司对用户进行精神奴役,用户本身也在进行自我精神奴役。正如马克思(马克思、恩格斯,2009:47)所言:“在资本主义生产占统治地位的社会状态内,非资本主义的生产者也受资本主义观念的支配。”种种现代化过程中产生的科技崇拜现象,本质上是一种前反思的科学实在论。它表现为对科学技术的盲目崇拜,科学技术在科学崇拜者眼中成为权威,并进一步演变为一种意识形态。数字资本渗透着数字技术意识形态的强权,对技术的解释具有绝对权威,并基于此构建了庞大的意识形态机器,以压制任何试图反抗技术权威的行为。人们在数字技术意识形态下的任何反抗,无论正当与否,都会被谴责为“拒斥现代生活”。数字技术意识形态通过将自身与现代社会

紧密联系,以实现其“统治”。然而,数字时代的便利尚未普惠全体大众,数字异化的达摩克利斯之剑已经悬挂在我们每个人的头上。人们在肉体、精神和社会关系上全面陷入数字异化的危机之中。

三、扬弃数字异化的进路

(一) 同技术讲和:加速主义批判路向

在现代,人们的生活节奏被绑上“火箭”。我们时常在生活的奔波中感到眩晕,仿佛时间被加速了。我们不得不追赶社会的速度,因为只有跟上发展的脚步才有可能生存。然而,单向的时间本身是不可能加速的,不存在按下快进键的造物主。数字生活状态下,技术的进步给人们带来了更多的选择和机会。人们生活的“幸福标准”每天都在发生变化。各种新奇的刺激迫使人们主动或被动地投入对外部世界的索取中。人们对社会加速和时间“减少”的恐惧,恰恰是因为人们将自身的时间投入单向度活动中,失去了反思和内在的批判性。人们完全为生活的各种具体事件所宰制。

因此,加速主义批判路向认为人陷入数字异化的根源在于人的时间观念的改变。技术的快速发展裹挟着人,消除数字异化需要人们主动摆脱技术加速的社会节奏,唤醒数字劳动者的主体意识,拒绝加速化社会潮流的裹挟。可以利用各种具体的方法,如减少数字阅读、降低网络交往强度、正确对待数字产品等,让自身的时间“慢下来”。正如罗萨(2018:101)所说:“因为,有一种社会意识隐藏在主流的自由主义的自我自由感背后,它以压倒性的力量将一切推往反方向。”

因此,在慢节奏的生活中拥抱数字技术,唤醒隐藏在主流文化下的真正的“慢生活”内核,就是加速主义批判路向扬弃数字异化的道路。当然,加速主义批判路向承认田园牧歌式的生活不可返回。然而,这一扬弃数字异化进

路的缺陷在于,将数字异化问题看作一种简单的生存论问题。它仅仅将人的异化看作一种自我异化,缺少劳动异化的政治经济学批判向度,未能认识到人陷于数字异化状态是数字技术和资本逻辑联姻的结果。如果不能将人从被剥削的数字劳动中解放出来,仅仅试图同无序扩张的资本逻辑和加速的社会节奏“讲和”,就会在与数字剥削者的斗争中成为妥协派。

因为在加速主义批判路径的语境下,人们数字异化状态的根源在于人们没有正确处理自身与数字技术的关系。如果人们能够唤醒自身的自由自主意识,掌控数字技术的发展,就会迎来解放。加速主义批判路向准确地看到了技术进步带来的社会加速的现实,却将社会加速的根本原因看作人的自主性的丧失,退化为一种失去锋芒的文化批判。这将会进一步掩盖数字资本的剥削本质,使人的解放活动笼罩在数字意识形态的阴霾下,无法从根本上消除数字时代人的异化状态。加速主义批判路向搞错了斗争的对象,就像堂吉诃德骑士冲向风车一般。他始终认为自己在与怪物交战,却无法看到捶打自己的不是怪物,而是风车。

(二) 以技术治技术:数字资源建构路向

数字资本的增殖和扩张就像不断旋转的线圈,数字技术和资本紧密地结合在一起,飞速旋转,形成正负两极。既然数字技术和资本逻辑的结合已经成为事实,数字资本的无序增长也已经发生,就有学者提出通过技术改良的手段来建立完善的资本市场规范,进一步优化数字资源分配手段,使数字技术的发展始终在人的掌控之下。这就是另一条扬弃数字异化的路径——数字资源建构路向。数字资源建构路向认为,人们陷入数字异化状态的原因是数字技术的野蛮增长。因此,要消除数字异化,就需要人的介入。通过规范数字资源的生产和运行等技术步骤,破除数字异化的负面影响,使数字技术真正为人所用。例如,意大利自治学派的思想家乐观地将一般智力的生产看作通向共产主义的一种可能路径。他们认为非物质生产的产物是主体性,这

种主体性可以打破无产阶级的孤立状态,引导无产阶级走向团结,为彻底摧毁资本主义世界积蓄力量。然而,“非物质劳动以信息、智力、知识等要素产出,标志着资本建构了与自己相适合的生产方式。正是在特定的社会关系即资本主义生产关系中,人和劳动、生产要素建立起联系”(赵小宇,2023)。

数字劳动与非物质劳动一样,通过一般数据的生产、资金注入下的技术发展以及不断积累的数字知识和先进算法等要素,使数字资本构建了与自身相适应的生产方式。而这种特定的生产关系就是数字资本的生产关系。只有在资本主义生产关系的条件下,才能诞生与数字资本相适应的生产关系。因此,数字资源建构路向的缺陷就在于:虽然这一路向认识到了资本逻辑无序扩张的问题,但没有试图断绝资本逻辑和数字技术的紧密联系。在理论层面没有意识到私有制在资本逻辑运行中的重要地位,在方法层面仍然主张“以技术治技术”。技术问题从来都不只是简单的科学问题,而是与政治、经济、社会文化等多个方面相互关联的问题。倾向于“以技术治技术”的思潮往往助长技术的统治,并存在很大的缺陷。正如哈贝马斯(1999:99)所说:“一方面,它假想了技术进步的内在强制,而技术进步把这种虚假的独立性仅仅归功于在技术进步中起作用的社会利益的自发性;另一方面,这种模式以解决技术问题和实践问题中的根本不可能存在的合理性的连续性为前提。”在数字技术和资本的结合中,数字技术将自身的进步归功于资本的注入,这种解释掩盖了广大被剥削的数字劳动者创造的一般数据的作用。技术统治的手段无法消除数字异化现象,反而使人更加依赖技术,助长了数字技术和资本联合的气焰,还可能加重数字异化的问题。要扬弃数字异化,就必须准确地看到真正需要消除的对象,即私有制基础上的资本主义生产方式本身。

(三)真正的革命实践:资本逻辑剔除路向

综上所述,无论是存在论视域下的加速主义批判,还是从控制论角度出发的技术统治,都无法真正地消除人的数字异化状态。要扬弃异化状态,我

们应当回到马克思的视角。一方面,马克思充分肯定资本主义生产方式带来的强大生产力。虽然马克思的时代不存在数字技术,但他很有远见地预见到了技术进步对生产力发展的强大推动作用。然而,另一方面,“资本作为孜孜不倦地追求财富的一般形式的欲望,驱使劳动超过自己自然需要的界限,来为发展丰富的个性创造出物质要素,这种个性无论在生产上和消费上都是全面的,因而个性的劳动也不再表现为劳动,而表现为活动本身的充分发展,而在这种发展状况下,直接形式的自然必然性消失了;这是因为一种历史地形成的需要代替了自然的需要”(马克思、恩格斯,2009:69)。

在资本主义的运行逻辑中,生产的目标不是满足人们的生活需求,而是超越自然需要,成为一种社会需求,即资本家对资本增殖的追求。然而,持续的资本增殖导致劳动者与其劳动产品的进一步分离,进一步加深了劳动者的异化状态。因此,要扬弃异化状态,就必须将人从资本增殖的“着魔”状态下解放出来。这意味着彻底扬弃整个资本主义生产方式,而不是与资本妥协,这只会纵容资本的扩张。同时,仅仅依靠技术管制只会陷入工具理性统治的泥沼。要扬弃数字异化的资本逻辑,我们需要采取一种剔除资本逻辑的路向。

资本逻辑剔除路向直指资本主义生产方式的症结所在——私有制。私有制基础上的数字劳动,造成了数据资源的占有者凭借对数字生产资料的控制而无偿占有了价值生产者创造的剩余价值。数字劳动者的劳动活动从人的自由自觉的活动转变为一种异化活动,一种被迫的、消磨生命的劳动。数字劳动者无法完全占有其劳动成果,在劳动过程中还需要创造剩余价值以达到数字剥削者追求资本增殖的目标。劳动活动本身并不是人的异化根源,劳动者与自己生产的劳动产品的分离才是异化的开始。若要扬弃数字生活中人的不自由状态,扬弃数字异化,就必须从根本上剔除以私有制为基础的资本主义生产方式,实现共同生产。正如马克思(马克思、恩格斯,2009:67)所言:“如果共同生产已成为前提,时间的规定当然仍有重要意义。社会为生产

小麦、牲畜等等所需要的时间越少,它所赢得的从事其他生产,物质的或精神的生产的时间就越多。”在共同生产条件下,数字资源归大众所有,数字劳动者可以完全占有和享受自己创造的全部数字产品。数字劳动不再是具有剥削性质的异化劳动,人们也能够从数字异化的不自由状态回归自然状态。因此,只有在数字资源所有制领域进行彻底的、革命性的资本逻辑剔除活动,才能真正将人们从数字异化的束缚中解放出来。

四、结语

马克思的时代没有数字技术,也不存在数字异化。然而,马克思在当时的社会背景下准确地预见了我们如今身处的数字时代的情况。只要广大劳动者仍处于异化状态,只要人类尚未完全解放,只要压迫和不平等仍然存在,马克思的思想就不会过时,革命的实践就不会过时。正如马克思(马克思、恩格斯,2012:9)所说:“批判的武器当然不能代替武器的批判,物质力量只能用物质力量来摧毁;但是理论一经掌握群众,也会变成物质力量。”面对数字时代下的异化状态,我们更应坚定对马克思主义的信心,运用马克思主义的理论,实践马克思主义的原则。只有通过这样的实践,我们才能扬弃不自由的异化状态,最终实现自由人的联合体。

参考文献

成龙、王楠,2022,《数字异化:表象、根源及其破解途径》,《中国地质大学学报》(社会科学版)第 6 期。

费尔巴哈,2010,《宗教的本质》,王太庆译,北京:商务印书馆。

哈贝马斯,1999,《作为“意识形态”的技术与科学》,李黎、郭官义译,上海:学林出版社。

黑格尔,2013,《精神现象学》,贺麟、王久欣译,上海:上海人民出版社。

蓝江,2022,《一般数据、虚体与数字资本:历史唯物主义视域下的数字资本主义批判》,南京:江苏人民出版社。

卢卡奇,1999,《历史与阶级意识》,杜章智、任立、燕宏远译,北京:商务印书馆。

罗萨,2018,《新异化的诞生:社会加速批判理论大纲》,郑作彧译,上海:上海人民出版社。

罗萨,2022,《不受掌控》,郑作彧、马欣译,上海:上海人民出版社。

马克思、恩格斯,2009,《马克思恩格斯文集》第 7、8 卷,北京:人民出版社。

马克思、恩格斯,2012,《马克思恩格斯选集》第 1、2 卷,北京:人民出版社。

吴欢、卢黎歌,2023,《数字资本与剩余价值规律的新表现——数字经济时代马克思资本理论的再阐释》,《大连理工大学学报》(社会科学版)第 2 期。

席勒,2001,《数字资本主义》,杨立平译,南昌:江西人民出版社。

赵小宇,2023,《非物质劳动批判:回到马克思的劳动理论与现实》,《黑龙江社会科学》第 1 期。

周光港、赵海月,2022,《数字资本主义下数字劳动的生成逻辑、模式辨识与现实启示》,《山东工会论坛》第 4 期。

数字时代的社会自我

——破解碎片化自我困境的路径探索

高思蓉　王化起*

摘要：自我的形成过程、自我与社会的关系一直是社会学研究的重点。数字社会中，个体自我的独立性不断增强，但这也伴随着个体心理危机，以及个体化、个体社会性等议题中的理解偏差。在阐述个体自我论和社会自我论两种研究取向的基础上，本文以符号互动论的代表性人物米德的社会自我论为切入点，强调自我的形成需要有他者的在场，在特定的社会情境和社会互动中产生。本文分析了数字化社会中社会互动形式的变化，认为数字技术的普及使得个体社会关系和自我认同呈现出原子化、碎片化的趋势，但这种趋势并不意味着个体的形成不带有他者经验，而是说明社会组织本身的分散性。为在分散的数字社会中加强碎片化自我的建设，本文提出两种路径：一是借助项飙的“附近”概念，重新组织外界他者经验；二是利用霍耐特的三种承认模式——爱、法律和团结，在三种承认模式中寻求更高层次的共同体认同。

关键词：社会自我论　米德　数字社会

一、引言

“自我”是一个古老的话题。在哲学领域，尼采、弗洛伊德、海德格尔等学者都对自我的构成以及人的本体性进行了讨论（魏波，2023）；在社会学领域，也一直有着社会自我论和个体自我论、个人身份和社会身份的争论（Parkes，

* 高思蓉，云南大学民族学与社会学学院。王化起（通讯作者），云南大学民族学与社会学学院。

Schneider & Bochner, 1999)。人是社会性的动物,但是人们对自我塑造中的"社会性"的理解却存在争议(王晓升,2019)。自我如何产生,它的构成要素是什么,存在形式是什么?要回答这些问题,就离不开对自我—社会关系的考量,就离不开对自我形成的社会过程的考察。

不难发现,对"自我"的理解随着社会的发展而发生变迁。在社会发展进程中,现代人的生活逐渐出现个体化转向,个人独立性不断增强,在人格崇拜的基础上,现代个体逐渐树立起人格神圣的观念,但这也影响了现代人的自我塑造,对人的社会性的片面理解使得现代人产生了心理危机(杨锃,2019)。在现代性背景下,对自由和解放的追求往往伴随着个人的物化和原子化,个体逐渐成为被看待和对待的碎片化对象。因此,如何重塑自我、重建社会关系便成为一个重要的时代命题。

当代社会是一个快速变革的社会,数字技术的发展和普及正深刻改变着人们的互动形式——社会网络也影响着人们的自我形塑(Dufva & Dufva, 2019)。数字利维坦对个体化社会的生存基石造成了极大冲击(郧彦辉,2015)。一方面,数字技术推动了社会社群的形成和集体身份的建构(赵一璋、王明玉,2023);另一方面,数字技术的发展也使得社会关系碎片化,社会逐渐分裂。在这样的背景下,一场关于自我定义和社会互动模式变化的革命正在逐步展开。

现代数字社会的自我是如何建构的?如何在碎片化时代重塑自我?作为符号互动论的代表性人物,米德(G. Mead)继承了形式社会学的传统,关注在抽象的社会互动形式中自我的形成与发展(Morgan & Schwalbe, 1990)。本文将从米德的社会自我论出发,结合项飙的"附近"概念和霍耐特(A. Honneth)的承认理论进行论述,讨论在现代社会重新组织自我的可能性。

二、关于"自我"的讨论:社会自我与个体自我

在思想史上,关于"自我"的理解经历了漫长的变革。西方哲学史上,个

体自我观有着漫长的历史,从笛卡尔的“我思故我在”到洛克描述的自然权利、自由意志,自我往往被理解为孤立的存在(高新民,2020)。人们对人之社会性也往往从外在性方面去理解,将人和人的社会关系看作产生于终结斗争的需要,将人类共同体看作孤立主体的组合(王晓升,2019)。在前现代,自我被理解成社会秩序中一个不可缺少的部分,尚未成为独立的个体。随着西方社会的现代性转变,个人生活和私人生活分离,商品经济结构导致人格的商品化,自我构建由此出现危机。由于现代自我认同的种种困境,在后现代,在“自我”概念的理解上又发生了回归古典共同体社会的转向,强调内在的关系实在性(袁祖社,2010)。

这种自我观,就是一种社会自我论。社会自我论的思想源头可以追溯到“美国心理学之父”威廉·詹姆斯(W. James)(张文龙、高新民,2021),在其自我结构的人格理论中,他对纯粹自我、经验自我、物质自我以及社会自我进行了分析,承认身体在认知过程中发挥的重要作用。这种具身认知对身心一体观点的支持,对后来的心理学家产生了极大影响(董晶晶、姚本先,2017)。在詹姆斯对自我进行划分的基础上,库利(C. Cooley)提出了“镜中我”的概念,对自我的形成进行了多向度的探讨(周晓虹,2008)。库利指出,自我的形成是一个动态的过程,不是个体内部的产物,而是在与他人的交往和行动中不断演化和塑造的。个体在与他人的交往中,通过观察、沟通和互动,逐渐理解他人的想法、情感和态度,同时也意识到他人对自己的看法和评价。这种互动和反馈不仅帮助个体更好地理解自己的身份和角色,还促使个体对自己的认知和行为进行调整和修正。库利之后,米德继承了符号互动论的传统,发展并完善了社会自我论(柯泽、宋小康,2021;Gillespie, 2005)。与强调个体自我的个体心理学不同,米德的社会心理学是行为主义范式的,把个体置于社会过程中来研究,用社会整体有组织的行动来解释个体活动,更关注社会对个体经验的影响。相对于华生(J. Watson)将心理现象归为类似条件反射的生理机制,米德认为可以否认心灵实体的存在,却不能否认它在其他意

义上的存在,心理行为可以用非心理行为进行解释。华生所持的个人自我理论与米德提倡的社会自我理论不同,个人自我理论从个体自我引出社会过程,但这种理论无法解释心灵、自我的产生;社会自我理论则用生理关系解释逻辑上为先的东西——社会行为过程的存在,再通过个体在其中的经验性参与引申出个体自我。

三、米德关于社会自我的讨论

那么,在米德那里,社会自我是如何产生的呢?这主要和表意姿态进入会话、个体对社会外部姿态会话进行内在转化以及"客我"与"主我"共同出现在经验中有关。

(一) 表意姿态进入会话

社会自我产生的第一层面,是表意姿态进入会话。不同于生物对外界刺激做出的本能反应,人类之间的姿态的会话还包含情绪的表达,姿态是"引起另一方的顺应反应的那些动作状态"(米德,2005:35),这意味着姿态"符合第一个人经验中的一种意义且在第二个人那里唤起那个意义"(米德,2005:36)。姿态会话的内在化,说明姿态成了一种表意的存在。

在此基础上,表意姿态进入会话,与语言这种有声姿态作为一种社会性刺激具有特殊性,以及人类之间有复制有声姿态的倾向有关。一方面,通过有声姿态的复制,个体能对他自己的刺激做出和别人一样的反应;另一方面,有声姿态的复制能创造具有普遍意义的空间,代表了机体对某些刺激做出的与其他机体共通的反应,从而使语言运载一系列在不同个体经验中具有相同意义的符号。由于"意义的逻辑结构要到姿态与顺应反应以及与特定社会动作结果的三重关系中去找"(米德,2005:63),因此在一组姿态的会话中,当一个姿态既可以向他人表示,又可以向做出动作的有机体本人表示的时候,意

义就产生了。

（二）对外部会话的内在转化

社会自我产生的第二层面，是对社会外部姿态会话进行个体内在经验性转化（心灵的产生）。在姿态会话的基础上，人能借助姿态提供的过程在他身上引起在其他人身上会引起的反应，并向自我指出尚未发生的行动的刺激，通过对该刺激进行分析，按照该刺激指令确定整个行动（正如前面提到的，这种唤起反应与语言的本质有关）。简单来说，就是“一个人知道其他人对他说的话的意义”（米德，2005：86）。

在这个过程中，心灵逐渐浮现出来，因为这种反思智能是心灵的一大特征：通过在社会行动和姿态会话中产生的符号，人类将相关对象的特征和对象本身分离出来，并根据中枢神经系统的隐含反应机制，“能够在一个已经开始的动作实际结束之前隐含地检验各种可能的结果”（米德，2005：93）。在这样一种机制中，有机体将自己置于同对象的一种关系之中，通过体内存在的感受性结构使得某些对象开始相对于有机体而存在，社会经验进入个体经验中，通过反射，有机体“能够对他自己采取他人所取的态度”（米德，2005：105），心灵——这种“以成套的符号为中介”的一种“有机体与情境之间的关系”（米德，2005：98）便产生了。社会自我的产生，也意味着完整自我的出现。

（三）社会自我的出现：主我与客我

米德认为的完整的自我，包括主我和客我：他人的态度被组织起来构成了客我，而有机体“对他人态度的反应”构成了主我，客我与主我出现在经验中，便构成了人格（米德，2005：137）。自我的发展有两个阶段：在第一个阶段中，自我在由个体和其他个体构成的特定社会姿态及其相互作用中形成（正如我们在上文分析的一样）；在第二个阶段中，只有当自我对他所属的社会群体参加的社会活动采取了有组织的共同体（泛化的他者）的态度，完全的自我

才算形成,社会性自我也由此形成。

在这个过程中,个体的自我组织和形成是与其所处的社会情境密切相关的。完整的自我其实反映了“作为一个整体的社会过程的统一性与结构性”(米德,2005:113)。因此,个体的自我不是孤立存在的,而是在人与人的关系中、在社会情境中得以实现和展现的。个体的社会性自我是在社会互动和交往中逐步建构并完善的。

(四)自我认同与共同体

在自我与共同体的关系中,个体的行动往往受到群体中其他个体的态度和社会控制的影响。在正常情况下,个体的主我能够将自我用作实现共同体公共事业的手段,这也是“他对他作为共同体一员所具有的价值作出情感反应的根源”(米德,2005:157)。当个体感到自己作为共同体一员时的获得感较强时,他也更容易对该组织的价值和目标产生情感反应,从而更加乐意为共同体的利益和目标而努力。

然而,当个体过度强调自身权利而忽视了共同体利益时,“‘主我’便是与‘客我’相对的支配因素”(米德,2005:165)。这意味着个体过度强调个人利益和权利,而忽视了群体的整体利益和目标。在这种情况下,个体可能会与共同体产生冲突,而主我则成了个体与群体之间权力斗争的表现。因此,当个体在某一共同体中受到阻碍时,他需要一个新的情境,在新的情境中获得自我表现的机会。在那个新的情境中,“整个共同体都在做同一件事。以前存在的压抑消失了,他与共同体融为一体,共同体与他融为一体”(米德,2005:172)。这意味着个体的主我能够在其中得以发挥和表现,而不仅仅是因循守旧地受制于现有的客我。实质上,这一过程在某种程度上是求助于一个逻辑意义上更大的共同体,即个体试图通过融入更广泛的社会群体或社会环境来实现自身的价值和自我表达,从而寻求更大的发展和实现。事实上,一个健康的共同体不仅需要秩序和规范,“还必须为个体本身的表现留下余

地,必须为这种表现提供手段”(米德,2005:174),从而使新的共同体组织成为可能,让主我的价值更好地实现。

在今天,米德社会自我论的构建不仅有实际的案例支持[比如有学者使用来自 56 个自然社区的网络数据验证了“自我”概念灌输中兼容内化模式的作用(Yeung & Martin, 2003)],其提出的寻求更大共同体的设想,也为我们后续处理当代数字社会之下的个体碎片化现象提供了解决思路。

四、数字社会互动形式的变化

由上面的分析可以看出,米德认为自我的形成需要有他者的在场,在特定的社会情境和社会互动中产生。通过表意姿态进入会话的符号的基础和中枢神经系统记忆的生理基础,个体在一定的社会情境中完成自我的形塑与自我价值的实现。社会情境由意义符号构成(苏振东,2011),现代社会数字化的转向,使得社会情境发生了哪些改变,自我的形成和实现又相应地发生了哪些变化呢?

数字社会具有信息数字化、信息计算、数字智能化的特征(王天夫,2021),这直接改变了现实生活中人们的人际互动模式和社会关系建构方式。在现代城市公共空间中,我们面对着一个无处不在的、移动的新媒介环境(Drucker & Gumpert, 2012)。在这样的环境中,新媒介通过嵌入式和移动媒体为那些进入的人提供了一个连接和分离的组合,开展更多远程合作(王天夫,2021),在身体缺席的情况下在场。表面上看,数字技术的发展减少了空间对人际交往的限制,人们可以在虚拟媒体中发生互动(Madouni, 2020),人与人的联系更加频繁了;但实际上,数字技术的普及也导致了个体之间社会关系边界的逐渐消解。在传统社会中,人们的社会关系往往受限于地理位置和社会身份,个体所处的社会角色和身份边界比较清晰。然而,在数字社会中,个体可以在虚拟空间中扮演多种身份和角色,社会关系的边界变得模糊

不清。这种模糊的社会关系边界使得人们在建立和维系社会关系时缺乏方向性和稳定性,个体之间的联系变得更加随意和碎片化。线上线下交往方式的模糊、社会关系边界的消解使得人际关系走向淡薄,个体走向消解(赵一璋、王明玉,2023),个体原子化、无缘社会的声音不绝于耳。

但是,现代社会中碎片化、原子化的自我,并不意味着个体自我的形成与他者的经验相脱离。米德的自我形成理论强调了个体与外部社会的互动和内在经验性转化的过程,这构成了完整的社会性自我形成的基础。碎片化的自我实际上反映了外部社会姿态的分散性。在传统的行动模式中,行动主体直接介入并产生对客体的影响,整个过程相对较为连贯和稳定(陈氚,2021)。然而,在信息时代的行动模式中,信息的传播与主体的分离导致了信息作为主体产生影响的情况。信息的流动性和不确定性使得社会行动更加碎片化,个体的行动往往不再具有明确的连贯性和稳定性,而是受到各种信息和因素的影响,呈现出多样化、分散化的特征。

因此,信息时代中社会行动的碎片化与传统行动模式的差异在于,信息的分散性和不确定性使得个体的行动更加多样化和灵活化。这种碎片化的行动模式也反映了现代社会的复杂性和多样性,个体需要在不断变化的信息环境中适应和应对,以实现自身的目标和价值。

数字连接作为一种元关系,不仅改变了个体—社会的组织逻辑,而且泛在连接作为自变量对整个社会体系产生了机制性的影响,这包括政治、经济和社会关系等因素形塑的社会分化机制(邱泽奇,2022)。因此,社会主体自我的变迁反映的是社会的变迁,碎片化自我折射出社会关系的分散化。按照贝克(Beck & Beck-Gernsheim, 2001)对个体化三重维度——脱嵌、传统丧失、再嵌入——的分析,个体脱嵌意味着个体与传统的社会结构、价值观念逐渐疏离,个体的社会关系和身份认同遭受到了冲击。在数字社会不断发展的过程中,个体感受到的原子化、碎片化、孤立的痛苦,实际上来源于在个体脱嵌中对自我的错误理解——忽视了米德描述的自我产生的他者的在场、自我形

成的社会性、社会姿态分散以及对自我组织再嵌入的缺失。因此,在社会变迁和转型的过程中,个体需要重新理解和适应社会的变化,并寻找新的社会支持和认同。然而,个体往往缺乏对社会变迁的正确理解和应对策略,导致个体与社会关系的疏远和孤立。我们需要进一步思考的是,社会变迁本身是有方向的、积极的,在社会转型的过程中,我们如何给予个人以更多的支持。

五、碎片化自我的新组织方式

由于传统社会向现代数字社会转型这一客观过程不可逆转,因此我们需要寻求这一变迁中碎片化自我的新的组织方式。按照米德的分析,自我的组织需要外部的他者经验的进入,而自我的实现需要在共同体中完成,当某一共同体受阻时,个体需要一个新的、逻辑意义上更大的共同体。

因此,面对碎片化的自我,我们首先要承认自我形成的社会性,重新组织外界他者的经验。在这里,我们可以借助项飙的“附近”概念进行分析。项飙(项飙,2022)指出,在现代的中国城市社会,人们往往只关注到极近(自身)和极远(想象的世界),而忽视了中层空间。中层空间的丧失,意味着个体在组织外界他者经验时面临着不完全性。因此,要重新组织他者经验,我们就需要寻回消失的“附近”。这种寻回,从物理空间意义上讲,可以是关注附近的生活场域,通过“最初 500 米”开启我们与外部世界的联系。通过重视附近的生活场域,个体可以重新建立起与周围社会的联系,并逐渐建立起归属感和社会认同感。从抽象空间来看,为了满足社交需要,个体可以通过“位置信息的分享与实践”实现超时空的空间连接,打造共同在场感(姚亦恬,2023),在抽象空间中建立起超时空的社会关系,从而扩展社会交往的范围和方式。

可以看到,这种对“附近”空间的寻回方式,抛弃了传统“边缘—中心”的

叙事方式,力求创造一个切实可感的、可以自我掌控的场域,在具体的社会情境中与他人发生互动,感知他人的经验,与他人进行会话。在这样一种情境中,自我的实现也是向内寻求的。正如项飙一直主张要把自己作为方法,从自我经验出发去理解世界一样,在“附近”的场域中发展自我,组织他人的态度构成客我,从自己出发对他人态度做出反应以建构主我,主我与客我在这个循环中交织互动,完整的自我也就渐渐形成了。

其次,由于个体在旧的共同体(碎片化的外界姿态)中组织自我、实现自我受到阻碍,因此他只有寻求一个更高层次的共同体范围,才能改变现有的社会秩序,完成自我的重新塑造。那么这个“更高层次的共同体”需要如何寻找和打造呢?我们可以从法兰克福学派第三代掌门人霍耐特的承认理论那里寻求答案。

霍耐特吸收了黑格尔的观念,认为一个主体的自我认知必须得到另一个主体的承认,强调个体之间相互关系的重要性,以及在社会互动中个体自我认知的形成和发展。不过,霍耐特并没有止步于黑格尔的抽象思辨,而是借助米德的实用主义思想,将这一理论观念置于现实经验的框架中进行深入研究。在黑格尔和米德的基础上,霍耐特(2005)总结出了三种主体间的承认模式:爱、法律和团结。在“爱”的承认中,主体在亲密关系中获得情感支持,互相确认需求,互相承认,形成相依为命的关系;在“法律”的承认中,由于法律的普遍性,个体权利与具体的角色期待分离,个体“把自己的行为看作对他们的独立性的普遍被尊重的表达”;在“团结”的承认中,社会伦理具有多元价值倾向,个体的价值能在社会重视中得到充分实现。承认形式的个体化,意味着个体实践的自我关系从过去的集体团结演变为现在的独立自主。

具体来说,在寻找“爱”的共同体的时候,个体可以通过建立亲密关系和情感支持网络来弥补社会碎片化带来的孤独和失落感。在这种关系中,个体与他人相互确认需求,形成互相承认的关系,从而获得情感上的支持和安全感。在“法律”承认中,个体可以通过法律的规范性、普遍性原则来确保自己

的权利和地位得到尊重。法律作为一个普遍性的规范系统,可以帮助个体超越片段化的社会关系,保障个体的自由和权益。通过法律承认,个体可以重建对社会秩序和公正的信任,从而塑造自己的社会性自我认同。在社会团结中,个体可以在多元化的社会伦理中寻找归属感和认同感,通过参与社会团体和群体活动,个体可以体验到社会认同的力量和意义,从而重新塑造自己在社会中的角色和价值。

总的来说,面对碎片化自我的重新组织:一方面,个体可以寻回消失的“附近”,通过会话理性重新组织外在他者的经验;另一方面,个体可以通过爱、法律和社会团结等不同形式的承认来重新塑造社会性自我。这些将有助于个体在碎片化的社会中找到自己的位置,并重建社会性自我认同和关系网络。

六、总结与讨论

通过对社会自我论和个体自我论的阐述,以及对数字社会互动形式变化的分析,我们可以看到现代社会中自我形成和社会互动模式发生了深刻的变革。传统的共同体关系逐渐被数字化连接所取代,个体的社会关系和自我认同出现了碎片化、原子化的趋势。然而,这种变革并非单纯消极的,而是为个体重组自我、重塑自我提供了新的机遇和挑战。

为了寻找碎片化自我的组织方式,本文分析了米德的社会自我论。米德强调在社会自我的形成中,表意姿态进入会话、个体对社会外部姿态会话进行内在经验性转化的重要性,通过对他人态度的组织与反应,包含主我与客我的完整自我得以形成。自我的形成需要在社会情境中完成,自我的实现同样需要在社会共同体中打造。

因此,面对数字社会中个体感受到的原子化、碎片化、孤立的痛苦,我们应该关注自我产生过程中他者的在场、自我形成的社会性、社会姿态分散这

些客观事实,以及对自我组织的合理再嵌入。

首先,个体需要意识到自我形成是一个社会性过程,依赖于外部他者的经验和社会情境的影响。在数字社会中,重新组织外界他者的经验变得至关重要,个体需要通过关注身边的生活场域、利用位置信息分享等方式来重建与外部世界的联系,找回消失的“附近”,以实现自我在社会互动中的价值。其次,个体需要寻求更高层次的共同体认同,以完成自我在社会中的重新塑造。这种更高层次的共同体可以是霍耐特基于爱、法律和团结的承认模式,通过这些模式,个体可以在亲密关系、法律框架或社会伦理中找到对自我的认同和实现。这种个体化的承认形式使得个体不再仅仅依赖于社会团结,而是能够独立自主地实践自我关系。

总的来说,数字社会的到来带来了个体社会关系和自我认同的深刻变革,但同时也为个体提供了重组自我、重塑自我的机会。通过重视外部他者的经验、寻求更高层次的共同体认同,个体可以在数字化连接的社会中找到自我实现的新路径,从而应对碎片化自我的挑战,赋予个体以更多的支持和意义。

参考文献

陈氚,2021,《信息行动理论——数字社会时代的社会行动理论探讨》,《社会学评论》第5期。

董晶晶、姚本先,2017,《威廉·詹姆斯与具身认知心理学》,《心理学探新》第3期。

高新民,2020,《自我的“困难问题”与模块自我论》,《中国社会科学》第10期。

霍耐特,2005,《为承认而斗争》,汪堂家、李之喆译,上海:上海人民出版社。

柯泽、宋小康,2021,《从“镜中我”到“雾中我”:虚拟现实中社会互动的畸变与理论危机》,《新闻与写作》第8期。

米德,2005,《心灵、自我与社会》,赵月瑟译,上海:上海译文出版社。

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。

苏振东,2011,《网络情境下的符号互动理论》,《新闻传播》第4期。

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。

王晓升,2019,《社会理论视域中的“社会性”问题——基于霍耐特承认理论的思考》,《南京大学学报》(哲学·人文科学·社会科学)第 3 期。

魏波,2023,《现代社会人如何失去自我——现代性批判的中心线索》,《世界哲学》第 6 期。

项飙,2022,《作为视域的“附近”》,《清华社会学评论》第 1 期。

杨锃,2019,《从“人格崇拜”到“自主自我”——社会的心理学化与心灵治理》,《社会学研究》第 1 期。

姚亦恬,2023,《消失的附近与搁浅的亲密:数字化时代新移动范式下的距离感分析》,《新闻研究导刊》第 8 期。

袁祖社,2010,《“人是谁?”抑或“我们是谁?”——全球化与主体自我认同的逻辑》,《马克思主义与现实》第 2 期。

郧彦辉,2015,《数字利维坦:信息社会的新型危机》,《中共中央党校学报》第 3 期。

张文龙、高新民,2021,《社会自我论:化解自我碎片化疑难的一种新尝试》,《内蒙古社会科学》第 6 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。

周晓虹,2008,《认同理论:社会学与心理学的分析路径》,《社会科学》第 4 期。

Beck, U. & E. Beck-Gernsheim 2001, *Individualization: Institutionalized Individualism and Its Social and Political Consequences*, London: Sage.

Drucker, S. & G. Gumpert 2012, “The Impact of Digitalization on Social Interaction and Public Space.” *Open House International* 37(2).

Dufva, T. & M. Dufva 2019, “Grasping the Future of the Digital Society.” *Futures* 107.

Gillespie, A. 2005, “G. H. Mead: Theorist of the Social Act.” *Journal for the Theory of Social Behaviour* 35(1).

Madouni, A. 2020, “The Social Interaction in Virtual Media.” DOI: 10. 47577/tssj. v11i1. 1605.

Morgan, D. & M. Schwalbe 1990, “Mind and Self in Society: Linking Social Structure and Social Cognition.” *Social Psychology Quarterly* 53.

Parkes, L., S. Schneider & S. Bochner 1999, “Individualism - Collectivism and Self - Con-

cept: Social or Contextual?" *Asian Journal of Social Psychology* 2(3).

Yeung, K. & J. Martin 2003, "The Looking Glass Self: An Empirical Test and Elaboration." *Social Forces* 81(3).

赛博格工人、“科技恶魔”与金币农夫

——东方主义遭遇数字时代

何祎金*

摘要:历史上,东方主义话语持续建构了关于东方的发展停滞、文化落后的他者形象。然而,随着东亚地区科技与社会的发展,其全球影响力不断增强,对传统东方主义的解读提出了新问题。数字时代与东方主义的相遇在延续后者认识论传统的同时,在新技术与新媒介环境下不断被激活,并形成了新的内容。科技与游戏东方主义在流行文化、政治话语中出现。尽管通过加入科技元素重塑了现代亚洲的形象,但对亚洲科技的妖魔化、污名化,以及对亚洲的非人化、机器人化想象,仍然折射出西方在面对全球竞争时的发展焦虑。要超越数字时代东方主义的限制,我们不仅需要彻底告别西方中心主义的文化逻辑,还需要摆脱我者与他者的二元认识论迷思。

关键词:东方主义　科技东方主义　西方中心主义　游戏东方主义

20 世纪 70 年代末,萨义德(E. Said)的东方主义理论对西方话语中的东方形象和生成机制进行了解构和批判。通过塑造一个发展停滞、文化落后的非西方他者形象,西方中心论者不仅凸显了西方现代性的“进步”属性,还反映和维持了西方内部的身份认同、文化认同。虽然萨义德简要提到了战后东方主义知识生产向美国转移,并在知识范式上不同于欧洲传统(Said, 2003),但对战后亚洲取得的发展成就,尤其是东亚地区呈现的发展动态,以及数字时代科技进步带来的影响和冲击,尚未进行深入探讨。

萨义德之后,信息技术的全球化实践引发了对新殖民主义问题的讨论。

* 何祎金,北京工业大学法学部。

20 世纪 80 年代,一些学者预测随着信息社会的到来,非西方社会将面临软硬件和信息标准的依附问题,形成所谓的“电子殖民主义”(electronic colonialism)(McPhail, 1981:20)。延续萨义德对文化表现的批判方法,早期研究还关注西方主导的数字空间如何建构非西方社会的文化异质性和他者性,以高姿态俯视非西方他者,甚至表现为一种数字空间的“全球化可乐殖民化”(globalizing Coca-colonization)。在面对文化差异和多样性时,信息技术的普遍主义实践就像可乐商品的全球化一样不可避免(Nakamura, 2002:99)。近年来,对后殖民理论的拓展使得数字社会的帝国主义、殖民主义研究成为焦点。奎特(2022)的数字殖民主义揭示了当代数字生态系统中,美国的信息技术对全球南方的政治、经济和文化形成了帝国式的控制。此外,在数字环境中,殖民主义不仅在非西方社会得到重塑,而且在万物皆可数据化的时代,社会量化部门、云帝国通过攫取人类生活所产生的数据来获利,也以不平等的数据关系开始在生活世界中殖民(Couldry & Mejias, 2019:41)。

随着东亚科技水平的提高,东亚地区越来越多地扮演起生产者、创新者的角色。这种巨变使得东亚与历史话语中的东方形象明显不同。日本、亚洲四小龙以及改革开放后的中国大陆等国家与地区,展示了东亚地区现代性的转型。信息技术和制造业的全球影响力不断扩大,彻底颠覆了亚洲“落后”的形象。

与数字殖民主义研究对非西方网络边陲角色的预设不同,东亚或亚洲科技力量的崛起及其积极性挑战了传统的东方主义阐释。在东方主义与数字时代的遭遇中,西方话语通过科技元素对亚洲形象进行了重塑,创造了新的叙事内容。然而,这种重塑亚洲形象所产生的新文化、新美学表达,并未从根本上打破西方中心主义的认识论传统。科幻文学、动漫、影视和数字游戏等新媒介,虽然重新定义了充满科技感的亚洲形象,甚至将其视为反思如何摆脱西方现代性困境的功能性他者,但数字时代的东方主义仍在我者和他者的二元认识论框架下延续了西方中心主义的基本逻辑。对当代亚洲,尤其是对

中国科技的妖魔化、污名化、遏制和围堵,既反映了东方主义的古老传统,又折射出了西方对竞争力丧失的发展焦虑。

一、科技东方主义:赛博格工人与“控制台牛仔”

随着世界工厂的出现,东亚地区的现代性转型也逐渐展开,制造业逐渐从低端向中高端迈进。从早期的日本制造到中国制造,再到如今的亚洲制造,越来越多的数字设备由亚洲的世界工厂生产出来。可以说,东亚经验首先打破了东方主义话语中关于“停滞”与“落后”的他者形象。

战后日本经济的高速发展被誉为“经济奇迹”。在制造业和集成电路方面取得的卓越成就。日本在科技领域的进步塑造了人们对科技特定的“日本性”想象,日本甚至被视为未来科技的代名词,有望超越和取代西方的现代性(Morley & Robins, 1995:169)。这种观念激发了西方世界的一种“日本恐惧”文化意识,并且西方通过采取“敲打日本”的外交政策来规训和遏制这个他者。

在此背景下,西方需要重新审视科技化的东方,以确保自身的领导者角色。科技与日本元素的结合,催生了一种名为“科技东方主义”的文化想象。这种想象注入了大量技术与未来主义的元素,以超科技的术语来理解和描绘亚洲和亚洲人(Roh, Huang & Niu, 2015:2)。科技东方主义也可以被定义为赛博或信息社会的东方主义,作为一种无意识的意识形态和文化机制,它塑造了信息资本主义社会的形象和模型(Ueno, 1999)。在生产和运作机制上,科技东方主义类似于传统东方主义,它由西方构建,旨在为西方在新的环境下建立自身身份和认同提供相应的参照。

在众多动漫和影视作品的美学表现中,赛博朋克以其充满丰富的亚洲元素和景观的特点成为科技东方主义的典型表现。上野俊哉认为,赛博朋克展示了一种全球资本主义的变异形式。福特主义是西方理性和现代社会模式

的代表,在日本则出现了类似命名的“索尼主义”“富士通主义”。这些概念不仅指代日本式的生产和管理方式,还被视为当代亚洲资本主义的内在特征。经济学家卡恩(Kahn, 1970:102)提出了“日本公司”(Japan, INC.)的隐喻,用以描述日本政府对经济的高效管理模式。日本宛如一个高度组织化、管理井然有序的公司,这种将集体主义与市场导向相结合的资本主义明显不同于市场主导的西方资本主义。

上野俊哉对“日本公司”意象背后的西方想象和文化内涵进行了解构和批判:日本被塑造成信息资本主义的典型,甚至成为21世纪的亚洲典范。然而,在西方政治和文化的无意识中,面对集体主义所带来的高效产出,古老的东方专制主义的认知传统被激活。日本被理解为一个空洞且非人性化的技术与生产大国,一个充满异化和反乌托邦色彩的国家;日本国民则被描绘为毫无感情的机器、与肉体结合的复制人,他们的忠诚、奉献、遵守纪律和自我牺牲的文化品格被视为构成日本和亚洲现代转型的动力来源。

与萨义德对东方形象的解构不同,玛丽·索斯腾(M. Thorsten)提出了美国流行文化和政策话语对日本形象的塑造,几乎将日本视为一个“超人类他者”(superhuman other)。这种“超人类他者”,与传统东方主义话语所刻画的野蛮、落后和低级的非西方“他者”有明显区别。这种“高级的他者”(superior other)反映了20世纪末美国对日本的焦虑和不安,担心这个“超人类”在经济和技术方面超越美国。尽管日本在二战后迅速实现了工业化转型,在一些方面具有突出的能力,但这个“超人类”仍然在等级上低于完整人的“非人类”,肉体与机器结合的赛博格成为非人类想象的体现。日本的超人化,同时也是非人类化、他者化,在功能上维系了美国社会内部的统一性,对美国认同具有重要意义。这种塑造还含有刺激生产力的策略,通过引入一个想象中的超人类“敌人”,刺激和培养更高效的美国学生和工人,以应对来自日本和全球化的挑战(Thorsten, 2012:13)。在这种策略下,表现得颇为矛盾的是,在非人类化、他者化的同时,日本的教育、管理和生产方式被视为西方学习和借鉴的对

象。有趣的是,傅高义在讨论日本经济和社会转型的成功经验时,在《东方主义》一书出版一年后出版了《日本第一》,该书的副标题正是“对美国的启示”。

作为与超人类赛博格相呼应的空间想象,晦暗的城市底色与暧昧的霓虹光影相结合,营造出异托邦的城市景观。然而,科技未来与贫困生活的尖锐对比并非仅存在于小说或电影的虚构之中,它还真实存在于现实中。在某些亚洲国家和地区,现代化进程的推进同时伴随着一系列尖锐的社会矛盾和问题的产生。这导致了一种奇特的混合现象,高科技的城市景观往往与阴暗的废墟角落相伴,这也是赛博朋克影视文学所描绘的世界场景(Ueno, 1999)。事实上,在全球数字产品的生产和消费链条中,我们可以看到一种极度分化和不平等的生态结构:一方是享受数字红利的消费者,而另一方则是在流水线或血汗工厂中工作的生产者。

赛博朋克中高科技与低生活的混合,正是对这种充满违和感的现实的折射。全喜卿从权力和控制的角度对《神经漫游者》和《攻壳机动队》中的“高科技东方主义”进行了解构,这种“高科技东方主义”通过承诺可理解的差异,在异域色彩的城市景观中实现信息的交汇,引导读者走向科技过载的未来。它给人带来了学习和探索的愉悦,作为一种压倒性的存在,最终实现了对网络服务器的侵入(jacked in)(Chun, 2005:177)。赛博朋克的科幻构想与美学反映了现实中对技术未来和现代性不确定性的焦虑。科技与废墟的结合、英雄个体对抗邪恶系统的剧情结构,在影视文学和数字游戏中已经成为一种流行风格。早期投射的日本恐惧和现代性焦虑,逐渐转化为阅读者或游戏玩家的愉悦体验,数字时代的东方主义也真正成为一种潜在的无意识。

数字资本主义时代的劳动生产衍生出了一个复杂的生态链条,赛博无产阶级在新技术条件下兴起(郝志昌,2022)。在日本之后,亚洲凭借丰富的劳动力资源和区位优势逐渐发展成全球工厂,非人化的他者想象也扩展到参与全球产业分工的亚洲劳动力。科技东方主义固然重新定义了现代亚洲,然而

它始终被视为一种“正常模式”之外的“变异形态”。

从赛博无产阶级到赛博格无产阶级,技术未来的想象与全球分工的现实相互映射。在后人类时代的社会想象中,对无产阶级进行赛博格化改造,以资本增殖的剥削为目的,传统的工人阶级变为受高科技裹挟的新奴隶(闫金敏、戴雪红,2023)。如今,像身体器官一样普遍存在的数字设备凝结了亚洲工人的辛勤劳动,数字设备带来了无限可能,但也伴随着重复的机械劳动,乃至不知疲倦的机器人成为亚洲劳动力的隐喻。

隆·布伊(L. Bui)的“亚洲机器人主义”(Asian roboticism)批判了西方对亚洲劳动力形成的话语与想象。即使在21世纪,在由机器学习和人工智能赋能的自动化生产趋势下,亚洲工人仍然在全球化的数字生态链中承担大量低端和廉价的劳动。数字时代的亚洲机器人主义并未消失,而是持续影响着亚洲工人的主体建构,将他们呈现为缺乏人类权利、可被剥削的异化机器人(Bui, 2020)。显然,机器人的隐喻背后存在一个全球生产与消费不平等结构的现实基础。颇具讽刺的是,西方流行文化常常将自身打造成对抗这种不平等结构的个人主义“英雄”和“救世主”。

与机器人化和非人化的亚洲相比,在科技东方主义和亚洲机器人主义的反面,西方形象的建构形成了极大的反差。西方形象不仅具备超强的能力,还注入了宽容、自由、和平等卡里斯马特质。个人主义英雄拯救人类世界的形象在其中扮演重要角色,他们如同赛博朋克小说中自由而浪漫的“控制台牛仔”(console cowboy),常常以个人力量对抗整个系统的控制,实现对网络服务器、矩阵或母体的入侵与颠覆,同时也对社会系统进行渗透。他们是《神经漫游者》中的凯斯,是《银翼杀手》中的里克,也是《黑客帝国》中的尼奥。

相较于显性的暴力支配,数字时代的东方主义更具隐蔽性。然而,权力支配的欲望并未消失,“控制台牛仔”代表了西方的技术力量,他们有能力驾驭、驯服和教化亚洲的科技。他们在西方构建的赛博世界中自由驰骋,凝结了西方意识和想象。“控制台牛仔”渴望获得更多的快乐、更广阔的土地和更

丰富的资源,他们是亚洲科技的管理者、指挥者和设计者(Patterson, 2020: 117)。在这个赛博异托邦的文化景观中,文化表现与现实世界发生了对应。在全球生产与消费的异化关系中,“控制台牛仔”与赛博格工人、自由的研发管理中心和冷漠拥挤的血汗工厂共同构成了这一景观。

二、亚洲“科技恶魔”:现代性的西方焦虑与他者的流变

数字时代的东方主义重塑了亚洲形象,但并没有从根本上否定或切断其与西方中心主义的历史联系。科技东方主义在一定程度上延续了日本恐惧和异化他者的思想,这也是“黄祸论”在新的历史时期的延续。随着亚洲经济发展动态的变化,21 世纪的中国在信息技术领域不断创新与突破,受到了科技东方主义的“特别关注”,甚至成为当下东方主义想象和叙事的“主角”,从不知疲倦的赛博格工人进化为危险的“科技恶魔”。

从词汇来源和构成来看,“黄祸”由“颜色”加“危险”构成,最早出现在 1895 年的英语世界中,反映了西方特定的文化和种族想象。《韦氏词典》收录了“黄祸”的两层含义:首先,东亚人的影响和权力扩张对西方文明造成威胁;其次,甘愿接受极低薪酬的东亚劳工的涌入,对西方生活标准带来威胁。尽管解释的内容有所不同,但都突出了“黄祸”是对西方的一种威胁。

19 世纪的法国人约瑟夫·戈宾诺(J. Gobineau)提出了等级化的种族分层理论,在欧洲产生了深远影响。他将颜色符号与等级化的种族观念建立了联系,其中“黄色”成为“威胁”的象征,反映了“白种人”对族群的焦虑情绪。在此背景下,集科技天才与犯罪天才于一身的傅满洲形象出现于英国作家莫尔(S. Rohmer)的笔下,被视为“黄祸”的典型代表,他的终极目标是摧毁西方并建立一个属于自己的“黄色帝国”。

在我者与他者的关系上,东方他者在投射与维系西方文化和身份认同的同时,还为陷入发展困境的西方提供了一个内省与反思的功能性他者形象。

东方主义并不是西方对东方的外部凝视，它更表达了来自西方内部的反思（萨达尔，2005：19）。亚洲“科技恶魔”角色在特定历史时期的出现，其意义在于：一方面，傅满洲系列小说充斥着大英帝国的“种族妄想症”（race paranoia），它以 19 世纪末的中国变革为背景，封建帝国向现代国家的转型引发了西方被东方赶超的焦虑；另一方面，傅满洲也成为西方反映工业化与现代化消极后果的媒介，乃至傅满洲所具备的能力及其所作所为很像吉登斯对现代性的隐喻，以“不受控制的庞然大物”（juggernaut）来描述现代性的后果。从这个意义上说，傅满洲身上投射了西方对现代性，以及科技异化和不确定性的焦虑情绪（Seshagiri，2006）。事实上，作为现代性最为著名的隐喻，这一修辞本身便来自东方与西方、传统与现代要素的混合。它源自梵文的“世界之主”札格纳斯，在西方语境中表示不可阻挡的破坏性力量，对应蒸汽压路机或者攻城车这类碾压性的巨型机械。在生成和作用机制上，傅满洲与赛博朋克对亚洲科技未来的想象相似，既忧心来自他者的超越，又将他者与科技的消极后果建立起联系。

东方不会一直“停滞”，东方主义也会随着世界变迁与科技发展而不断重塑自身。在传统的小说和电影之外，动漫和数字游戏这样的媒介成为东方主义的新载体。与傅满洲相似的“邪恶角色”出现在西方流行文化中：美国幻影工作室发行的动画《忍者神龟》在 20 世纪 80 年代末风行全球，大反派施瑞德（Shredder）原是来自日本的忍者，在邪恶外星生物的帮助下将自己昔日的同门滨户良变为一只老鼠人，并将后者的乌龟宠物变成人形。奇幻叙事源自美国对日本日益壮大的“恐惧”反应，施瑞德的角色设定反映了西方视角下对日本公司“鬼鬼祟祟”行为的焦虑，认为它们是像施瑞德一样的偷窃者和破坏者，又如同冷酷无情的机器人一样追求世界霸权，破坏西方现代性的原则（Morley & Robins，1995：153）。

作为制衡的正义力量，在老鼠大师的忍术训练下，来自美国的四只海龟阻止了施瑞德称霸世界的疯狂恶行。它们的名字也极具深意——达芬奇、拉

斐尔、米开朗基罗和多纳泰罗,文艺复兴的四个象征符号,合力挫败了日本忍者与邪恶外星生物侵略地球的阴谋。施瑞德代表的他者并非蒙昧落后的反现代性形象,其掌握的高科技与自身“邪恶”属性是恐惧感的来源。可以看到,这一系列的动画对他者化日本表达了相对矛盾的态度:在施瑞德一方,如同一个日本版的傅满洲,日本的意象意味着传统“东方文化”与外星高级技术结合所带来的人类异化和破坏性力量;在美国神龟一方,日本元素并没有被彻底屏蔽,在善的一面它给予神龟知识与技艺的训练。

不同历史时期,“黄祸”对应的主体和带来的“威胁”也不尽相同。从 19 世纪的亚洲移民和廉价劳动力带来的冲击,到二战时日本的侵略和威胁,以及战后日本经济的快速恢复与科技发展,科技东方主义的出现对应了东亚现代性兴起的历史情境。20 世纪 80 年代以来,以四小龙为代表的亚洲的经济起飞,给西方国家带来了地缘政治的焦虑。这些地区的经济和工业生产不仅展现出强劲的发展动力,更为重要的是,它们逐渐由西方领先技术的“拷贝者”和“供应者”转变为新技术的“发明者”和“发起者”(Christine, 2005)。换言之,他者现代性的兴起,为西方“研发与管理中心”自由且浪漫的“控制台牛仔”带来了挑战。

改革开放后,中国在经济和技术领域的快速发展,使之成为日本之后的科技东方主义主角。在西方的文学和影视作品中,围绕中国衍生了大量的科技东方主义图像与叙事,今天的中国正在填补日本曾在 20 世纪 80 年代留下的“超人类”角色空白(Thorsten, 2012)。与彼时西方流行的“日本恐惧”相似,21 世纪的中国成为西方投射竞争与发展焦虑的他者化镜像。

在中美贸易摩擦和新冠疫情背景下,科技东方主义出现了一些新的内容。这种东方主义将科技和病毒元素融合在一起,表现出三种路径:首先,采用污名化策略,将疾病与种族联系起来,将中国比作病毒和传染病,如特朗普将新冠病毒称为“中国病毒”和“功夫流感”(Kung Flu);其次,出于对失去全球经济地位和技术竞争力的焦虑,不断攻击和诋毁中国在科技领域的发展,

并采取不公平手段打压中国留学生和技术人员的交流;最后,通过种族化的方式构建科技,产生了“中国/亚洲科技病毒”(Chinese/Asian techno-virus)的文化意识形态(Siu & Chun, 2020)。即使在当代全球化风险的背景下,古老的“黄祸论”仍然可以被任意激活和调用,创造出新的傅满洲和施瑞德。新时期的科技东方主义以种族化的傲慢和偏见来理解和阐释21世纪的中国,妖魔化和污名化中国的科技创新,并对科技企业进行围堵和打压。显然,这种策略对破解全球化风险的治理困境没有任何帮助。

事实上,对中国元素的东方主义想象并不是一成不变的,它在投射西方焦虑的同时,也为西方自省提供了空间。就像对日本的超人类构建包含激发西方内部生产力的美学策略一样,对中国这个功能性他者进行想象和构建既有排斥和焦虑的一面,也有自我反思和学习借鉴的一面。在20世纪80年代的科技东方主义出现之前,科幻小说已经开始探索中国的哲学元素及其意义,甚至形成了一个“前现代的东方主义”阶段(Huang, 2008)。20世纪60年代,菲利普·迪克(P. Dick)在《高堡奇人》中探讨了中国哲学的命题,利用西方想象的工具化、技术化的道家思想,批判西方关于未来与进步的修辞。中国元素的意义在于,在西方遭遇困境之时,转向曾经被轻视的他者并寻求破解的可能性。

20世纪90年代之后,面对日益发展和充满活力的中国形象,尤其是中国和美国在许多方面的紧密合作,形成了相互连接和高度依存的关系,需要一种新的文化表达形式。一些学者提出了突出共生关系的“中国特色的科技东方主义”(techno-orientalism with Chinese characteristics)概念。中国元素取代了日本元素,在后赛博朋克中占据主导地位。然而,与赛博朋克对日本景观的异化表达不同,后赛博朋克美学主要表现了21世纪全球关系的“相互依存美学”,它以中美大国关系为现实基础(Fan, 2015)。在一些科幻影视作品中,我们可以看到类似美学的出现,中国相关的科技形象也发生了一定程度的变化。例如,在电影《地心引力》(2013)和《月球陨落》(2022)中,中国航天

技术在最后关头发挥了关键作用,为处于危机中的西方或全人类提供设备,帮助他们完成拯救同伴或地球的任务。

然而,在被称为“中国特色的科技东方主义”的概念中,中性化地使用“东方主义”一词会掩盖后者应有的道德和政治含义,从而削弱批判的力度,甚至存在自我东方主义化的风险。从概念的理论来源来看,我们很容易发现其中的局限性。该概念借鉴了经济学家尼尔·弗格森(N. Ferguson)、莫里茨·舒拉里克(M. Schularick)构想的 Chimerica 关系,中文译为“中美国”“中美利加”或“中美共生体”。值得注意的是,这种构想固然传达了中美两个大国相互依存的一面,但背后仍暗含了西方中心主义的思维定势。

进入全球经济系统的中国和亚洲劳动力提高了资本的市场回报率,美国所在的“西中美利加”相对富裕,中国所在的“东中美利加”则相对贫困。两者之间的互补关系体现在角色分工上:美国擅长市场、金融和商业管理,而中国擅长生产和工程建设(Ferguson & Schularick, 2007)。“中美利加”的组合使得前者成为储蓄过剩的生产者,后者成为消费过度的消费者,实现了在分工角色上的整合。然而,令人感到诡异的是,经济学家的理论构建与影视作品中的文化表现表达了类似的二元逻辑:东方充斥着机器人和“血汗工厂”,而西方则是自由的研发和管理中心。显然,这种角色互补只停留在表面,生产与消费关系的固化也意味着不平等结构的固化。

三、游戏东方主义:金币农夫与“普遍主义”游戏规则

东方主义的生产与再生产总是无法脱离现实社会的基础,相比殖民历史上的暴力掠夺,它对数字空间的渗透与话语权力的支配更为隐蔽。在数字游戏中,虚拟世界与现实世界相互交织,既带来了游戏空间的社会化,也放大了社会空间的游戏化。前者在数字游戏中嵌入了现实社会的运行机制,后者则在现实社会系统中嵌入了游戏化的运行逻辑。数字游戏与东方主义的结合,

不仅生产了亚洲玩家病态的游戏玩工形象,还在游戏化的全球社会系统中标榜西方“普遍主义”的游戏价值和规则。

法国社会学家拉图尔(Latour, 2007)在探讨数字性和物质性之间的关系时提到,非西方要素在全球分工中扮演次级角色。欧美玩家在虚拟空间中收获的装备和游戏体验由印度程序员提供,而大量中国玩家通过对虚拟游戏角色的升级代练,将其出售给没有时间或耐心的美国玩家来谋生。拉图尔的文章主要探讨了虚拟空间的数字痕迹具有前所未有的可追溯性,虽然非西方要素并非重点,但它揭示了西方与非西方在现实与虚拟空间中的结构性不平等。在游戏产业衍生的生态链条中,游戏中的虚拟空间不仅投射了现实世界的不平等,而且其维系与运转还需要来自后者的支撑。亚洲的工程师和“为他人而玩”的中国玩家群体形同机器人一般,孜孜不倦地在现实与数字世界中生产和劳作。

gold farmer 是拉图尔描述中国玩家群体所从事“工作”的术语,中文译名为“金币农夫”“打金农夫”或“职业打金者”。作为数字时代衍生的新职业和新群体,“农夫”的修辞传达了异质性的隐喻,暗示了玩家主体性的丧失。“正常的玩家”通过游戏获得愉悦的体验,而“不正常的玩家”则通过游戏出卖劳动力。就像玩家在游戏中选择“战士”或“冒险家”职业一样,“农夫”这个职业在许多游戏设定中是最不起眼的非玩家角色(NPC)。金币农夫意味着游戏世界中存在廉价的劳动力,大量玩家从事简单且重复的工具性劳动。在游戏世界的等级化分工体系中,非西方,特别是中国玩家承担了次级和低端的角色。

数字游戏是东方主义在新技术环境下进行渗透的新媒介和载体。金币农夫既是游戏世界的职业分工,也是西方对东方玩家刻板印象的体现。汉娜·魏尔曼(H. Wirman)在借鉴“汉学东方主义”概念的基础上,对西方主流媒体中“中国游戏”的形象进行了分析。她认为,狭隘的西方认知导致了充满偏见的报道,这些报道内容和机制与东方主义的文化建构类似。这种文化建

构将“中国游戏”描绘成世界数字游戏的“奇葩展示柜”,其中包括金币农夫、病态玩家、昏暗的网吧和不健康的游戏市场等(Wirman, 2016)。这种刻板印象的文化建构,背后蕴含着西方统治东方的霸权意识。汉学东方主义概括了西方对中国知识生产的根本逻辑,即将他者转化为同者,以加强和巩固西方现代性的“普遍主义”价值观。这种逻辑在目前西方媒体对中国游戏的报道中尤为突出。中国的游戏市场既充满各种商机,又受到以西方为主导的市场逻辑对合法游戏进行干预、治理和规范的批评。游戏东方主义的出现反映了西方的矛盾态度:一方面,觊觎庞大的市场机会;另一方面,以数字世界的“控制台牛仔”自居,标榜自身作为技术和价值的中心。

此外,西方媒体对亚洲玩家的报道塑造了一群不知疲倦的病态游戏玩家形象。西方话语对亚洲玩家的呈现,与东方主义中有关男性的刻板印象重叠:幼稚、无性别、虚弱且不善交际。相反,西方玩家则展现出活力、健康和积极探索的形象(Choe & Kim, 2015)。从刻板印象的文化建构入手,斯蒂夫·崔(S. Choe)等人以即时战略类(RTS)游戏《星际争霸》为例,在韩国的社会历史情境中考察了游戏的合法化、专业化过程,揭示了其背后西方中心主义的资本主义逻辑。

大部分即时战略类游戏的核心玩法是攫取资源、扩张领土、升级进化技术并击败竞争对手,如在《星际争霸》中选择人族,击败神族或虫族。这种即时战略的玩法在虚拟空间中展现了资本主义运行和竞争机制的特定表现形式,与信奉自由竞争的新自由主义逻辑一致。因此,过去以娱乐为目的的游戏与生产性劳动之间产生了关联,不知疲倦的病态玩家就像在“血汗工厂”中因负担过重而倒下的工人一样。与拉图尔所描述的维系虚拟世界运转的全球不平等分工相似,这也是西方玩家和亚洲玩家之间的区别:前者代表正常的游戏体验,而后者则被贴上异化劳动和病态的标签。

针对科技东方主义主要集中在科幻文学或影视作品中的讨论,塔拉·费科尔(Fickle, 2019)进一步推进了对游戏东方主义的解读。与陈查理和傅满

洲分别代表的正邪两种形象相似,游戏东方主义也具有典型的二元性。它既延续了“黄祸论”历史传统对亚洲和亚洲人的刻板印象,又塑造了亚裔在西方社会的模范形象。从刻苦的学生到捧杯无数的电竞选手,勤奋的工人与硬核游戏玩家之间建立了角色联结。亚洲相关游戏文化的形象既有羸弱病态的游戏玩家,也有顶尖的电竞选手,后者在电竞发展史上留下了许多亮眼的成绩。作为电竞史上的标志性事件,1997 年 6 月举办的“RA 雷神锦标赛”被视为第一场现代意义上的电竞赛事。亚裔选手方镛钦(D. Fong)夺得桂冠,因赢得了《雷神之锤》开发者约翰·卡马克(J. Carmack)提供的一辆法拉利汽车而引起轰动。该项赛事奠定了第一人称射击类(FPS)游戏的电竞基础。

“游戏”成为东方主义的新前缀,但是数字游戏内容与角色的亚洲或东方表现只是狭义的游戏东方主义。费科尔并没有将视角局限于在终端上运行的游戏软件或者应用,而是强调日常生活的游戏化逻辑,以及背后的意义生产机制。日常生活中游戏技术的运作如同一个社会化的操作系统,只是用户或玩家并不知晓系统内部的工作原理,也无法察觉或意识到其中的问题(Fickle, 2019:10)。游戏化在今天已经无处不在。围绕西方公平与自由的抽象游戏理念,游戏东方主义逻辑嵌入我们的生活世界。它通过区分等级化的差异,在社会系统中生产排斥——费科尔将其历史追溯至 19 世纪美国的《排华法案》时期。针对彼时大量廉价中国劳动力涌入美国,美国本土的排外主义者以“公平竞争”(fair play)为由进行排斥打压,并根据中国移民玩的棋牌游戏,将他们描绘为“嗜赌成癖”的民族。在早期中国移民和劳工问题上,诸多种族化游戏要素与修辞的引入,是游戏东方主义在特殊历史阶段的表现。

可以说,19 世纪的廉价劳动力与 21 世纪数字世界中的金币农夫,两者在游戏化的全球社会操作系统中并没有本质区别,亚洲的游戏玩工被当作打破数字空间生态平衡的破坏者。然而,在特定情境下,与其说“公平竞争”是需要遵守的“普遍主义”游戏规则,不如说它是西方中心主义认知下为规训他者而产生的托词,一种排斥性与等级化、缺少公平性与包容性的狭隘“规则”。

随着亚洲地区社会经济的发展,亚洲玩家并不会一直固定在游戏玩工的生产者角色上。根据 2022 年 Newzoo(2022)发布的数据,亚太地区的玩家数量达到了 17 亿,占全球玩家的 55%,远超欧洲和北美玩家数量的总和。亚太地区是全球最大的游戏市场,并以中国市场为主导。全球最大游戏市场的出现意味着,亚洲玩家作为消费者的权利和文化认同也应该引起相应的重视。金币农夫的现象或许不会消失,但是亚洲玩家与游戏文化的主体性、能动性亟待正名。换言之,在非西方社会发展进步的大势下,亚洲玩家群体的兴起势必会对游戏东方主义中的刻板印象和数字空间的不平等结构发起反击。

当然,亚洲玩家并不是沉默和被动的行动者,其主体能动性甚至会通过一些另类的方式得到展示。数字游戏的虚拟空间消解了物理时空的界限,为不同民族和文化的玩家发生遭遇和互动创造了可能。然而,通过简单的代码,以维护游戏生态为名,在服务器中架设虚拟的屏障区分不同区域的玩家也更为直接和简单,即所谓“锁区”。针对一些游戏通过锁区屏蔽中国玩家的做法,一部分玩家借助虚拟专用网络实现登录,并在游戏中结成社群活动。在早期的大逃杀类游戏中,中国玩家以上身着红衫为身份标识,以集结号为令,在游戏空间展开仪式性反抗的集体行动。此外,针对许多游戏缺少中文语言支持的情况,大量中国玩家会在游戏平台的评论区以多种语言发出“我们需要中文”的呼声。本土语言的权利诉求不仅以提高游戏体验本身为目的,还要求游戏厂商正视游戏空间中的差异和不同群体的存在。

不同类型或风格的数字游戏,生产了多元化的虚拟空间。尽管在游戏玩法和美学表现上可能存在巨大差异,但它们的生产和维系不仅无法脱离物质性的社会基础,还嵌入了现实社会运行的逻辑。游戏东方主义对不同游戏玩家的等级化想象与区分,强调自身的普遍主义价值和游戏规则。亚洲游戏市场与玩家的日渐崛起,将持续不断地撼动游戏东方主义背后的底层逻辑,释放亚洲游戏玩家与市场主体的能动性。

四、结语:走出数字时代东方主义的藩篱

东方主义与数字时代的遭遇重置了现代亚洲的形象,但它并未切断西方中心主义的古老认识论传统。随着亚洲经济的发展,数字时代的东方主义在新技术、新媒介环境中不断被激活与唤起,并刻写进新的内容。东西方在全球科技生产与数字空间中形成了分工格局,凝结为“科技恶魔”、金币农夫与“控制台牛仔”三种角色的隐喻。

科技与游戏东方主义话语的出现,源自西方对东方现代性转型及其在科技领域影响力日益扩大的反应。西方既因丧失竞争力而产生焦虑,又试图概括与学习东方的优点。卡恩认为,东亚经济和工业发展成功的故事应该为西方反思自身的发展困境提供机会,而不是先入为主地认为西方模式的工业文化应该成为世界参照、复制的模板(Kahn, 1970:124)。卡恩的东亚理解注入了文化交流与互融的乐观主义成分,儒家与西方、传统与现代要素相互交织,凝结为“机器”(西方)与“花园”(东方)两种意象的联姻,认为它们代表了东西方文化融合的经典案例。

如果说贫民窟与高科技的结合是赛博朋克呈现的异托邦景观,那么机器与花园的联姻则代表了现代性的解决方案。针对西方工业文明的消极后果,机器与花园的和谐共存成为破解现代性内在矛盾的理想类型。然而,卡恩保留了乡愁与东方想象的描述,并未摆脱东方主义无意识的思维。它实用主义地认为,来自亚洲的功能性他者或许为陷入现代性困境的西方提供了未来选择。但是机器与花园的隐喻限定了西方与东方的角色分工,将东方想象为一座复制西方科技文明与储存传统要素的花园。换言之,这种融合依然基于固定的二元认识论传统,优美的修辞并没有改变不平等的分工结构。

科技与游戏东方主义理论话语的出现传达了强烈的解构和批判色彩,深刻地揭示了西方文化与话语表象背后西方中心主义的底层逻辑。然而,对东

方主义等后殖民理论的借鉴并没有摆脱该理论的内在矛盾,也未能解答如何走出数字时代东方主义的疑问,这种泛权力关系的阐释方法在建设性方面存在严重不足。“中国特色的科技东方主义”是更进一步的理论话语与美学诉求,表面上以东方和西方,尤其是中美两个大国的共生关系为基础;然而,东方主义固有的认识论传统并不会因为加入中国特色而消失,生产者与消费者的二元固化关系并不是平等的共生结构,它传递了一幅异化的全球图景。而且,在嵌入游戏化逻辑的全球社会系统中,“公平竞争”的游戏规则也不能将自身作为普遍主义的准则来框限、遏制和同化他者。

在数字时代如何摆脱东方主义的限制?在当前全球局势下,重塑东方形象或纠正误解并非解决问题的关键。放下傲慢和偏见,正视并适应亚洲的崛起,是西方在新历史时期的任务。平等、包容的共生结构,要求西方充分尊重各国自主发展和道路选择的权利。在构建全球人类命运共同体的格局下,西方亟待摒弃西方中心主义的传统思维和“控制台牛仔”的霸权心态,走出二元认识论迷思,以非种族和非殖民的方式理解信息技术和数字空间,以共商共建共享理念,共同治理全球性风险和挑战。

参考文献

郝志昌,2022,《数字劳动、数字权力与数字权利:赛博无产阶级的生命政治学》,《思想战线》第 4 期。

奎特,2022,《数字殖民主义:美帝国与全球南方的新帝国主义》,顾海燕译,《国外理论动态》第 3 期。

Newzoo, 2022,《2022 年全球游戏市场报告》, https://www. 163. com/dy/article/HKA5FVMJ05526SET_pdya11y. html。

萨达尔,2005,《东方主义》,马雪峰译,长春:吉林人民出版社。

闫金敏、戴雪红,2023,《阶级构造、身体形塑与生命生成——后人类时代的赛博格无产阶级研究》,《福建师范大学学报》(哲学社会科学版)第 6 期。

Bui, L. 2020, “Asian Roboticism: Connecting Mechanized Labour to the Automation of

Work.” *Perspectives on Global Development and Technology* 19(1-2).

Choe, S. &S. Kim 2015, “Never Stop Playing: Starcraft and Asian Gamer Death.” in S. David, B. Huang & A. Greta eds., *Techno-Orientalism: Imaging Asia in Speculative Fiction, History, and Media*, New Brunswick: Rutgers University Press.

Christine, C. 2005, “Techno-Orientalism and the Postmodern Subject.” in J. Furby, K. Randall eds., *Screen Methods: Comparative Readings in Film Studies*, London: Wallflower Press.

Chun, W. 2005, *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*, DOI: https://doi.org/10.7551/mitpress/2150.001.0001.

Couldry, N. &U. Mejias 2019, *The Costs of Connection: How Data Is Colonizing Human Life and Approaching It for Capitalism*, Stanford: Stanford University Press.

Fan, C. 2015, “Techno-Orientalism with Chinese Characteristics: Maureen F. McHugh’s *China Mountain Zhang*.” *Journal of Transnational American Studies* 6(1).

Ferguson, N. &M. Schularick 2007, “‘Chinmerica’ and the Global Asset Market Boom.” https://doi.org/10.1111/j.1468-2362.2007.00210.x.

Fickle, T. 2019, *The Race Card: From Gaming Technologies to Model Minorities*, New York: New York University Press.

Huang, B. 2008, “Premodern Orientalist Science Fictions.” *MELUS* 33(4).

Kahn, H. 1970, *The Emerging Japanese Superstate: Challenge and Response*, London: Prentice-Hall.

Latour, B. 2007, “Beware, Your Imagination Leaves Digital Traces.” http://www.bruno-latour.fr/sites/default/files/P-129-THES-GB.pdf.

McPhail, T. 1981, *Electronic Colonialism: The Future of International Broadcasting and Communication*, London: Sage.

Morley, D. &K. Robins 1995, *Spaces of Identity: Global Media, Electronic Landscapes and Cultural Boundaries*, London: Routledge.

Nakamura, L. 2002, *Cybertypes: Race, Ethnicity, and Identity on the Internet*, New York: Routledge.

Patterson, C. 2020, *Open World Empire: Race, Erotics, and the Global Rise of Video*

Games, New York: New York University Press.

Roh, D., B. Huang & G. Niu 2015, *Techno-Orientalism: Imagining Asia in Speculative Fiction, History, and Media*, London: Rutgers University Press.

Said, E. 2003, *Orientalism*, London: Penguin.

Seshagiri, U. 2006, "Modernity's (Yellow) Perils: Dr. Fu-Manchu and English Race Paranoia." *Cultural Critique* 62(1).

Siu, L. & C. Chun 2020, "Yellow Peril and Techno-Orientalism in the Time of Covid-19: Racialized Contagion, Scientific Espionage, and Techno-Economic Warfare." *Journal of Asian American Studies* 23(3).

Thorsten, M. 2012, *Superhuman Japan: Knowledge, Nation and Culture in US-Japan Relations*, New York: Routledge.

Ueno, T. 1999, "Techno-Orientalism and Media-Tribalism: On Japanese Animation and Rave Culture." *Third Text* 13(47).

Wirman, H. 2016, "Sinological-Orientalism in Western News Media: Caricatures of Games Culture and Business." *Games and Culture* 11(3).

中国个人信息保护经验研究的方法论反思

——欧盟经验的启示

张月明　李汶龙　李汶锴　李子豪　李傲兰*

摘要:法教义学与法学经验研究的范式之争由来已久,但学界对经验研究方法如何推进某一具体部门法的发展和变革洞见甚少。本文聚焦于个人信息保护领域,探讨经验研究方法在个人信息保护研究中的应用与效果。本文采用系统文献综述法,对以英文公开发表的《通用数据保护条例》(GDPR)相关的经验研究文献进行系统分析,从研究问题和研究方法两个角度进行了全面述评。在综述过程中,归纳了 GDPR 相关经验研究着力回答的六个研究面向及每种面向形成的方法范式,梳理了相关研究的趋势、方法、理念和效果。问题维度方面,经验研究可以为知情同意、告知义务、数据权利、大众认知、市场竞争等问题提供事实证据的支持;研究方法方面,呈现传统定性、定量实证研究方法与新兴计算法学方法并存的图景,跨学科研究有所增加,统计学方法、机器学习与人工智能等方法逐渐占有一席之地。未来在中国开展针对《个人信息保护法》的经验研究,可选择性借鉴欧盟相关研究的经验,更应立足中国立法的具体场景,为《个人信息保护法》的理解提供可靠的事实依据,以此推动法律的实施和解读。

关键词:《个人信息保护法》　《通用数据保护条例》(GDPR)　个人信息保护　经验研究　系统文献综述

---

* 张月明,比利时根特大学法律与犯罪学院。李汶龙,英国爱丁堡大学数据、社会与文化研究中心 CDCS。李汶锴,比利时布鲁塞尔自由大学法学院。李子豪,英国格拉斯哥大学法学院、美国斯坦福大学法学院。李傲兰,英国伦敦玛丽女王大学法学院。张月明负责数据统计、分析和第二、四、五、六节的撰写;李汶龙负责选题、思路的提出、项目统筹和第一、三节的撰写;李汶锴负责统筹系统文献筛选。张月明、李汶锴、李子豪、李傲兰共同参与文献的筛选与分析。统稿与修改完善由五位作者共同参与。

一、问题的提出

在传统的法学研究中,法学理论研究(doctrinal research)和法教义学(dogmatic research)一直占据主流。自美国法学教育家洛斯克·庞德(R. Pound)开始,传统的法学研究开始受到社会科学方法的挑战和冲击。庞德提出的"书本上的法"(law in books)和"实践中的法"(law in action)更是促进了法律现实主义(legal realism)的兴起。① 诚然,法律在现实中究竟是如何运行的,法律理论是否准确地描述现实,法教义学旨在实现的规范目标究竟是否能够达成,所有这些问题的回答不能仅凭推演或者想象,而需要基于证据。

新中国成立以来,受苏联(俄罗斯)、日本、德国法学传统的影响,法教义学在法学领域占有极高地位。但是,法社会学作为一个法学分支发展也很突出。② 伴随着全球跨学科潮流的发展,法教义学的垄断地位逐渐受到挑战,学界出现了不少关于社会法学、跨学科法学等思潮的反思和承继(车浩,2021;宋旭光,2015;侯猛,2020)。例如,车浩(2021)主张法教义学与社会科学的方法并没有本质冲突,反而能够形成内外部合作关系,旨在将社会科学融入法教义学体系,成为塑造教义的材料和工具,尤其是在个案分析、经济分析和社会学观察等面向有突出的优势。然而,国内的讨论主要在法理学层面发起的,在概念或传统上讨论研究格局的多元化重塑和不同支系的互补,对经验研究方法究竟如何推进某一具体部门法的发展和变革则洞见甚少。

---

① 美国的法律现实主义运动在 20 世纪 20 年代作为一种有影响力的趋势出现,在很大程度上借鉴了之前一些思想家的观点。在 20 世纪初,形式主义的法律方法开始受到洛斯克·庞德、约翰·格雷(J. Gray)和本杰明·卡多佐(B. Cardozo)等思想家的有力批评。在法律现实主义方面较具有影响力的法学家是美国最高法院法官霍姆斯(O. Holmes)(Pound, 1910; Duxbury, 1997: 65-161)。

② 改革开放之初,部分学者开始采用社会调查方法研究法律问题或法律改革效果,涌现了一批研究成果(梁慧星、王金浓,1980)。21 世纪初,苏力(2001)基于中国法学研究的整体思考,提出了"社科法学"的概念,并预言社科法学可能在法学领域扮演突出角色。关于社科法学的形成及其与法教义学的分立,参见谢海定(2014)。

中国法学界对“经验研究方法”的接受之路是漫长且坎坷的,甚至对这个舶来概念的翻译也存在诸多误解。当下,“经验”(empirics)和“实证”(positivism)的概念仍在混用,而学界甚至习以为常,后者在诸多研讨中成为指代这一传统的概念。但严格说来,“实证(主义)”在法律体系中有明确所指,与自然法和解释主义呈对立之势(Priel, 2008; McKay, 2014)。“经验”概念或许因其中文字面意义听起来不够科学而被弃用,令人唏嘘。本文认为,应为“经验研究”正名,并与实证主义传统进行区别,故而通篇使用前一概念。

对个人信息保护这个特殊部门法领域而言,经验研究也是一个较新的研究思路和方向。虽然欧盟自 1995 年就制定了《数据保护指令》(Data Protection Directive,DPD),但规范相对原则性(principle-based)、执行程度并不高。这也导致在其后的 20 年间,相关学术性经验研究非常少见,既有经验探索多为以推进立法改革为目的而由欧盟授权的智囊展开的调研和数据分析,并未得到同行评审。这些经验分析旨在为 DPD 的不足和有限有效性提供证据,并为其后的《欧盟通用数据保护条例》(General Data Protection Regulation,GDPR)提供改革重点和方向。

自 2010 年开始,欧盟推动了全世界范围内的数据保护改革。最新数据统计,包括中国的《个人信息保护法》在内,全球有超过 157 个国家已经制定了类似的规则体系,大多受到欧盟(包括欧洲委员会)的直接影响,或者直接以 GDPR 作为模板(Greenleaf, 2012; Greenleaf, 2016)。欧盟的数据保护法变革以可操作性、明确性、可追溯性为主要目标,也让规则的执行变得可以度量,使相关经验研究的开展成为可能。我们的文献综述显示,GDPR 的实施并未达到立法者的预期,无论是企业的整体合规水平,还是信息透明度、同意的获取、数据主体权利的保护等制度的落实,都存在明显的不足。大众对自身数据权利的认知十分有限,大多数民众并不知晓数据权利的内涵和行使路径。

此外,从规制模式角度观察,GDPR 涵盖的若干元素都颇具新颖性,与传

统的民法侵权、产品安全、风险管理、行政管理等领域存在联动,但差异也非常明显。相关立法不仅融入了数据保护影响评估、数据保护设计等前沿治理思路,也糅合了诸如数据可携权、自动化决策相关权利等全新机制。这些创新尝试亟待经验数据证明其有效性、完善的策略与正确的方向,经验研究将在这一方面扮演重要的角色。中国的《个人信息保护法》在多个面向参考、借鉴和吸收了欧盟的立法经验,在规范来源上属于同源。因此,无论国内外,对这些新型机制的运行机制,以及对数据保护的实际影响的经验测量的需求,都是存在的。由此,本文从既往 GDPR 相关经验研究入手,试图探求针对个人信息保护部门法开展经验研究的意义、方法与定位。

二、研究方法:系统文献综述法

为全面展示和分析 GDPR 相关的经验研究,本文在研究方法上采用了系统文献综述法(systematic literature review method)。[①] 文献综述是评估现有研究领域,并进一步发展现有研究体系的重要工具。相较于传统叙述性的总结式文献综述,系统文献综述法采取全面的、可重复的、透明的过程,对领域内已发表的研究进行全面的文献检索,通过程序化的分析过程展示和解读既有研究,从而避免研究者的主观偏见(bias)(Thorpe & Holt, 2008)。

本文的分析范围包括以英文写作,经过同行评审后公开发表的期刊文献、专著、会议论文集或论文集章节。具体数据筛选过程如下:

(1) 以“GDPR”或“data protection”和“empirical”作为关键词,在 7 个主要英文文献数据库[②]中,对 1995 年 1 月 1 日至 2022 年 3 月 1 日之间公开发表的所有文献进行检索,删除重复文献后,共检索出 380 篇文献。

---

① 系统文献综述法采用标准化的文献选取和分析过程,可以全面展示和分析研究问题相关的既有文献(Page, McKenzie & Bossuyt et al., 2021)。

② 筛选过程中纳入的数据库包括:Web of Science、Scopus、IEEE、ACM、ScienceDirect、SAGE、Oxford Academic。

（2）针对本文的研究范围，设立了筛选规则：①与 GDPR 或其前身 DPD 相关，研究主题为个人信息保护的相关问题；②文献包含经验数据，属于经验研究。

（3）为保证文献筛选准确，分两个阶段对检索出的文献进行筛选，分别为标题与摘要筛选、全文内容筛选。每轮筛选由两位作者分别进行，如两位作者意见相反，则由第三位作者进行投票。

（4）对 380 篇文献进行标题与摘要筛选后，发现有 93 篇文献符合本文确立的筛选规则。对 93 篇文献进行全文筛选后，有 70 篇文献符合本文确立的筛选规则，与本文研究范围相关。

（5）为保证文献覆盖范围全面，我们还应用了引文检索法。通过引文检索法，找到 55 篇相关文献。因此，本文献综述对 125 篇文献进行系统分析。

三、理论背景：关于经验研究的内核、争论与误区

展开个人信息保护经验研究系统文献综述之前，本节首先就经验研究的范式、脉络、核心概念和体系进行简要介绍，呈现文献关于经验研究的内核，聚焦其与传统法律分析的区别，即为什么法学经验研究是重要的，如何补充或者完善既有理论或文本分析范式的不足。①

经验研究基于对某个特定现象的观察和测量，研究者亲身经历和体验整个过程。数据从现实世界当中收集，因此一手数据是经验研究的主要特征。但来自文献的二手数据也很重要，如可以提供理论或假设的背景或语境。经验研究收集的数据往往用来验证理论或假设，但这样一种传统随着大数据的出现也在发生变化，并且引发了关于经验研究导向和本质的争议。

① 本文不做纯研究技术层面的梳理。关于经验研究方法论的介绍，可参见麦康维尔等人（McConville & Chui，2017）、卡希兰等人（Cahillane & Schweppe，2016）、洛里斯等人（Lawless，Robbennolt & Ulen，2016）的研究。

经验研究与传统法律分析都以严谨著称,但二者有不同的目标和进路。美国法学家庞德(Pound, 1910)曾有名言:“仔细观察,会发现书本上的法(law in book)和实践中的法(law in action)存在很大区别。前者意图管理人与人之间关系的规则,而后者是实际上管理它们的规则。我们会发现,法律理论与司法实践之间的沟壑既真且深。”

美国学者洛里斯等人(Lawless, Robbennolt & Ulen, 2016)主张:法律分析重视论证和诉诸权威,通常旨在证明特定观点,侧重于个别案件的细节;经验分析重视观察,挑战假设,以检验假设为导向,通常侧重于描述总体上运作的模式。经验实证研究还是递进的过程,既有研究可以揭露新的问题,新的研究进而建立在既有发现之上。① 伯明翰大学教授韦伯利(L. Webley)曾主张,经验研究与法学研究的距离并没有想象中的那么大(Cane & Kritzer, 2010)。通常,法官和律师所从事的判例查询、分析、比对等工作,本质上就是一种定性法学研究方法,虽然很多法律人并未从这个视角看待自己的工作(Leeuw & Schmeets, 2016)。

事实上,经验主义与法律实证主义在概念上的区别原本不应成为问题,但在国内两个概念混用②,甚至“实证”的概念已经默认成为 empirical research 的对译词,因此本文有必要对这两种理念的关系做一简单梳理。“实证”与“经验”两个概念的混淆不仅出现在大陆学界,在台湾也存在混用的现象,印证了中文世界在接受西方科学概念体系时存在理解偏差(王鹏翔,2015)。汤茂林等(汤茂林、黄展,2020)认为,存在混用和误解的成因在于对相关哲学概念——经验主义与实证主义——缺乏足够的把握,并且“把理论研究之外的研究都算作实证研究”,导致了国内实证研究的泛化。汤茂林等梳理了经验研究体系中的基本概念,还从文化角度讨论了陈寅恪重实用、轻理论思维对

① 当然,也有学者提出不同的观点。来自荷兰马斯特里赫特大学的学者里乌等人(Leeuw & Schmeets, 2016)就认为两种思路的差异并没有这么大。

② 南京师范大学汤茂林等人(汤茂林、黄展,2020)曾对国内“实证研究”概念误用的情况做了较为细致的批判分析。

后世的影响。从这个角度观察,20 世纪后半叶中国的实用主义文化导致“对功利的过分重视,而对理论问题尤其是哲学问题视而不见或不思”。两人还主张,之所以使用“实证”概念而不是“经验”概念,或许是因为中文字面意思中“经验”本身“不那么科学”“有主观色彩”,而“实证”听起来相对严谨科学。

传统的经验研究大多基于理论,或者说以理论作为思考的出发点。开展经验研究的目的也多为挑战或发展理论:理论指引研究,而研究发现、扩充和修正理论。可以说,经验研究与理论是共生关系。在这一层面,经验研究与实证主义存在关联。实证主义强调研究者的彻底客观性和独立性,强调所收集数据的价值中立性(value free),深刻影响了自然科学等研究范式(主要是定量和统计分析方法)。[①] 但是,实证主义仅是支撑经验研究的理论支系之一。除此之外,若干不同类型甚至彼此冲突的理论——如经验主义(不重视理论,强调客观性和定量)、现实主义(强调解释,关注社会及其结构)以及解释主义(又称批判实证主义,主张无法实现绝对客观性而关注不同类型的解读)——都在不同程度上影响着经验研究的范式和发展。至少在理论层面,讨论实证主义与经验研究关系的研究较为丰富。

例如,英国华威大学普利尔(Priel, 2008)主张法律现实主义者与法律实证主义者很长一段时间内都处在对立姿态。的确,二者在本体论上存在观点差异:对实证主义者而言,法律是人造的,依赖于国家的存在;对现实主义者而言,法律是一种自然发生的社会现象(类似语言),需要通过结构和功能来辨别(McKay, 2014)。法哲学家雷特尔(Leiter, 2012)进而主张,实证主义与现实主义并非不兼容,实际上法律现实主义者预设了一种强实证主义(或者排他性的实证主义)。总而言之,将 empirical research 翻译成“实证研究”并非简单的翻译失义问题。这些概念所表达的思想存在字义之外更深层次的关联,也在一定程度上导致了误译。诚如雷特尔所言,如果在实证主义正确

① Emerald Publishing, “Conduct Empirical Research.” https://www.emeraldgrouppublishing.com/how-to/research-methods/conduct-empirical-research.

的情况下现实主义才能立得住,那么将 empirical 译成“实证”也没有错,只是逻辑上有些超越。

四、GDPR 经验研究综述

(一) 既有文献概述

1. 时间维度下的欧盟个人数据保护相关经验研究

首先,图 1 展示了欧盟个人数据保护相关经验研究的趋势。在获取到的样本中,欧盟个人数据保护相关经验研究从 2001 年已经开始,但数量很少。GDPR 颁布后,经验研究呈迅速增长态势。GDPR 实施的 4 年来,相关经验研究已经初具规模。一方面,GDPR 具有较大的域外影响力,受到了学界的广泛关注;另一方面,GDPR 的条款更加明确,规则执行的可度量性更强,经验研究方法更加适用。

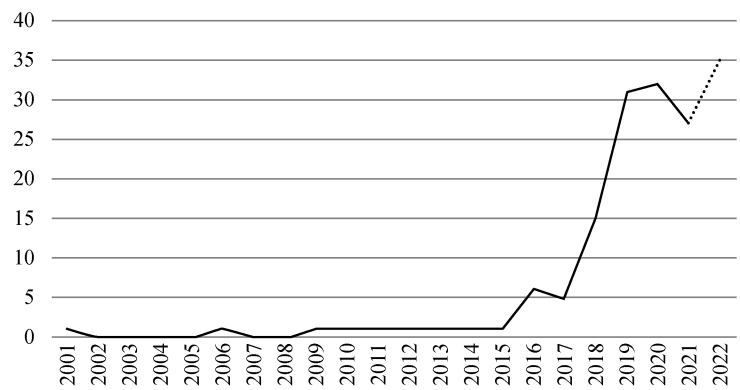


图 1 欧盟个人数据保护相关经验研究发表走势

2. 问题维度下的欧盟个人数据保护相关经验研究

为全面评述相关研究,我们进一步将 GDPR 的 99 个条款概括成 25 个研究问题,并对相关经验研究以问题维度进行了统计。另外,一部分经验研究

将 GDPR 作为整体,从宏观角度研究整部法律的公共认知、实施影响及落实情况等,我们将这些研究进行单独统计。图 2 显示,GDPR 相关经验研究着重解决几个方面的研究问题:数据权利相关的研究最为集中,占文献总数的三分之一;同意的获取与交互设计、透明度原则的落实也常被经验研究所关注。宏观层面,经验研究主要考察了大众对数据保护法和数据权利的认知、GDPR 对市场竞争和创新的影响以及 GDPR 在企业合规实践中的落实情况。究其原因:一方面,这些问题存在可度量性,适合开展经验研究以测评相关制度设计的有效性及实施情况;另一方面,受法律文化影响,欧洲一直将数据保护视为建立在基本人权保护上的一种政治要求,GDPR 也进一步强化了个人对其数据的控制权(丁晓东,2018),这一点在学界体现为研究者更多关注数据主体权利、体验与认知等相关问题。

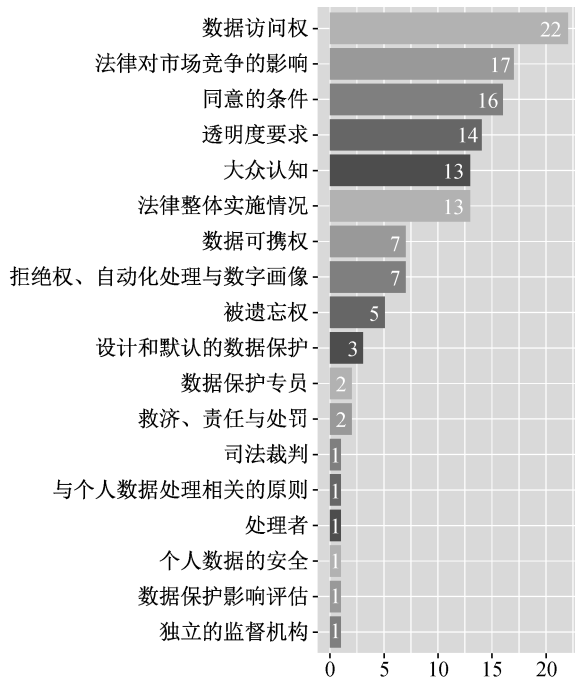


图 2　欧盟个人数据保护相关经验研究问题分布

3. 欧盟个人数据保护相关经验研究的方法

如何在数据保护法领域开展经验研究？我们统计了既有文献使用的研究方法。从图 3 可以看出，现有研究呈现传统定性、定量实证研究方法与新兴计算法学方法并存的图景，跨学科研究有所增加。

首先，数据挖掘、机器学习等人工智能和自动化的处理方法被应用在数据保护相关的经验研究中。例如，一些相关研究运用爬虫等技术抓取大量的隐私政策文本、同意弹窗内容信息，再运用数据处理软件对数据进行编码，运用统计学方法评估企业合规实践（Mathur，Acar & Friedman et al.，2019）。也有一些经验研究与信息科学等学科交叉，通过软件开发、自动化处理等方式，开发和测试合规的自动化解决方案，如构建一个代码化的合规检查体系，并利用自动化方式来验证其软件开发的准确性。人工智能工具可以迅速收集和分析上万份甚至百万份原始数据，并且一些研究可以做到挖掘网页源代码、收集第三方接入状态等人工无法直接获取的信息。由于数据量大、处理迅速的特点，人工智能工具被广泛应用于个人信息保护的相关经验研究中。

其次，为探索数据主体在数据处理活动中的权利落实情况，研究者以数据主体身份，直接向数据控制者发布行使数据权利的请求，收集数据控制者的反馈情况。例如，向网站发送数据获取请求，分析网站的回复及在此过程中可能出现的隐私问题（Bufalieri，La Morgia & Mei et al.，2020）。由于数据权利相关的研究较为集中，因此这种聚焦于数据主体体验的方法也被广泛应用。

另外，传统问卷和访谈的定性研究方法在个人信息保护相关研究中占有一席之地。对隐私政策文本的分析也是数据保护相关经验研究中较为常用的一种方式。通过深入研究一家或几家企业的相关报告数据，访谈企业员工，个案研究法可以挖掘更为深入的结果，也被应用于个人信息保护研究中。

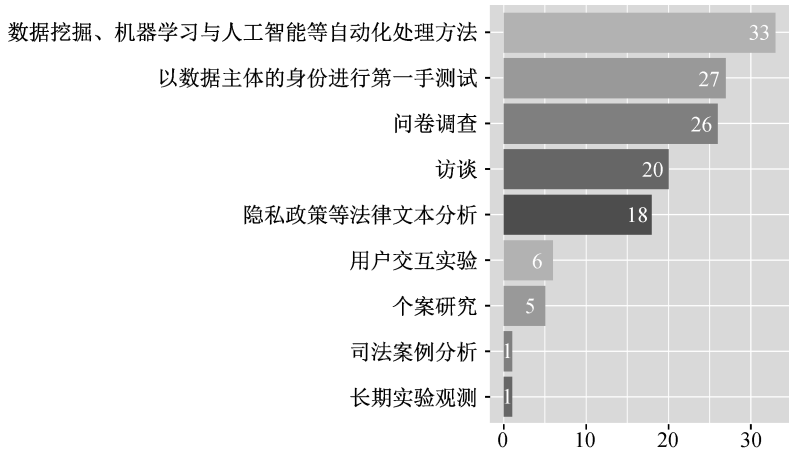


图 3 欧盟个人数据保护相关经验研究方法分布

注:有 12 篇文章同时使用了两种研究方法,我们对其进行了重复统计。

(二) 欧盟个人数据保护相关经验研究述评

在展示相关研究的全貌之外,本节旨在探索现有研究的趋势、方法、经验与有效性。

1. 同意的获取

“告知—同意”是 GDPR 保护数据主体信息自决权的重要制度设计[①],也为我国《个人信息保护法》所借鉴。GDPR 实施后,同意提示、弹窗和邮件铺天盖地,成为公众知晓这部法律的开始。与 GDPR 的要求相似,我国《个人信息保护法》规定了同意应当由个人“在充分知情的前提下自愿、明确作出”,且“基于个人同意处理个人信息的,个人有权撤回其同意”。[②]

在我们的样本中,有 16 篇文章研究同意获取的条件。研究者从两个面向探索这个问题。其一,企业落实同意相关合规要求的状况如何?例如,博尔基等人(Borghi, Ferretti & Karapapa, 2013)测试了 200 个网站获取同意的方

① 详见 GDPR 第六条第一款(a),第七、八条。

② 《个人信息保护法》第十四、十五条。

式是否符合 GDPR 的要求,包括获取同意的方式、网站对用户同意和拒绝同意的回应、网站未经请求的商业电子邮件(UCE)等,发现仅有 11.8%的网站符合要求。其二,在同意弹窗的交互设计中,什么条件会影响用户做出同意的真实性?例如,费尔南德斯等人(Fernandez, Chatzopoulos & Papadopoulos et al., 2021)研究了轻推(nudging)、位置(position)及黑暗模式(dark patterns)对用户同意的真实性的影响。

同意获取的相关研究广泛运用了爬虫软件、机器学习、自动化工具等大数据技术方法。研究同意获取问题需要的第一手数据,是数据控制者关于获取用户同意的方式和过程,通常体现为同意弹窗和用户交互设计。① 除了大量、便捷的优势以外,一些自动化软件还可以挖掘网页源代码,收集第三方接入状态等人工无法直接获取的信息(Kollnig, Binns & Dewitte et al., 2021)。与此同时,一些研究也揭示了自动化处理的局限性,由于人工智能尚不能完全模拟人类的思维和意识,对主观性问题难以把握。人工智能模型也存在技术的固有局限,难免存在抽样和计算系统性偏差(bias)(Tesfay, Hofmann & Nakamura et al., 2018)。受此影响,一部分相关研究运用了用户交互实验的研究方法。例如,乌兹等人(Utz, Degeling & Fahl et al., 2019)向用户发送不同设计的同意交互页面,并总结用户进行同意的意愿。另外一项针对德国 80 000 名用户的用户行为实验,其同意页面和数据处理均通过软件完成(Fernandez, Chatzopoulos & Papadopoulos et al., 2021)。

既有研究揭示了 GDPR 同意合规要求的挑战性。有代表性的是,在科尔尼格等人(Kollnig, Binns & Dewitte et al., 2021)的研究中,仅有 4%的被试 APP 可以达到 GDPR 要求的事前同意标准。科尔尼格指出,绝大部分应用程序开发者未能采取措施保障同意的有效性,缺乏合规的同意获取方式,存在

① 例如,玛瑟尔等人(Mathur, Acar & Friedman et al., 2019)等利用自动化技术,获取了 53 000 个购物网站同意获取的交互设计,并发现 183 个网站使用具有欺骗性、误导性的“黑暗模式”(dark pattern)设计。

不合规的默认同意设置。“黑暗模式”被极为广泛地适用,大大影响了用户提供同意的真实性和有效性(Soe, Nordberg & Guribye et al., 2020)。

这类研究不仅是 GDPR 经验研究的一种范式,也有效提出了制度设计和法律实施过程中的实际问题,因而证明了该类经验研究的意义与价值。2022 年 3 月 14 日,欧盟数据保护委员会(EDPB)通过了《关于在社交媒体平台上使用黑暗模式的指南》,详细列举了不合规的“黑暗模式”的类型,为数据控制者获取用户同意的规则提供了更明确的指导。我们相信,经验研究提供的大量经验证据无疑为推动法律的解读做出了贡献。

2. 告知义务的透明度与隐私政策

GDPR 规定了数据控制者的告知义务[①],相关告知必须完整且易读,即需要满足透明度要求。与 GDPR 相似,我国《个人信息保护法》也要求个人信息处理公开透明,“以显著方式、清晰易懂的语言真实、准确、完整地”[②]告知个人相应的事项。

这类经验研究着眼于两个视角。第一,从数据主体保护的角度,检查数据控制者提供的隐私政策是否满足全面性和清晰易懂的要求。有学者开展了关于 300 份隐私文本的可读性自动化测试,发现大部分的隐私政策都不易理解,没有满足可读性要求(Mathis & Tor, 2021)。第二,部分研究者从数据控制者的角度,试图帮助数据控制者开发更简便、准确的合规方案,以满足 GDPR 的告知要求。例如,克鲁梅等人(Krumay & Klar, 2020)试图为业界提供一个自动测评可读性的软件,供数据控制者在发布隐私政策前进行检查;托尔等人(Torre, Abualhaija & Sabetzadeh et al., 2020)则提供了检查隐私政策是否涵盖 GDPR 所要求的全部信息的自动化工具。

隐私政策文本作为数据控制者履行告知义务的载体,可以用来评估企业进行告知义务合规的程度,也成为透明度相关研究的原始数据来源。隐私政

① 参见 GDPR 第十二、十三、十四条。

② 参见《个人信息保护法》第十七条。

策文本由于其公开性、类型化的特点,容易被机器识别和量化,因而透明度相关的经验研究大量运用了机器学习、软件工具等自动化方法。[①] 回应第一个研究问题,研究者需要大量获取隐私政策文本,以形成较为准确全面的经验证据,并对获取到的文本进行检查,对大量数据进行统计分析。因此,纯人工的处理不再现实,爬虫等软件被大量使用。例如,阿莫斯等人(Amos, Acar & Lucherini et al., 2021)用爬虫软件获取了 1 071 488 份隐私政策文本,并运用可读性测试软件测试这些文本的可读性。人工方法显然无法做到如此大规模的分析和统计。回应第二个研究问题,机器学习方法则得到使用。既有研究通过跨学科的尝试,试图通过开发机器学习工具为企业提供合规方案,如特斯费伊等人(Tesfay, Hofmann & Nakamura et al., 2018; Chang et al., 2019)为业界提供了一种生成合规隐私政策文本的自动化工具。这些研究虽然为公共讨论提供了大量一手数据,但几乎仅停留在数据的呈现上,很难进一步提供有效的合规建议或政策建议。而一些小规模的文本研究则更加深入地回应了理论研究,试图进行规范解读和政策建议。例如,卡马尔力诺等人(Kamarinou, Millard & Hon, 2016)深入研究了 20 份云服务商隐私文本,考察其各项制度设计;米尔凯特等人(Milkaite & Lievens, 2019)仅针对三份儿童常用网站的隐私文本进行深入解读,但指出了业界存在的问题,并对未来的政策制定和细化提出了建议。

研究者对透明度要求的实施表示遗憾。无论是不细分领域的大规模测试,还是针对物联网、健康追踪 APP、儿童内容服务领域的专门研究,业界均未能有效地提供透明度较高的隐私政策文本(Kamarinou, Millard & Hon, 2016; Mulder, 2019; Mulder & Tudorica, 2019)。自动化生成隐私政策文本内容的尝试也被质疑,如孙(音)等人(Sun & Xue, 2020)直言,现有技术自动生成的隐私文本漏洞百出,远不足以满足 GDPR 的要求。

① 关于符合计算要求的法律数据的描述,参见左卫民(2022)。

3. 数据主体权利

数据主体权利是 GDPR 维护数据主体信息自决权的另一个重要手段。与 GDPR 相似,我国《个人信息保护法》规定了个人对其信息享有知情权、决定权、查阅复制权、更正删除权等。我国学界既有相关研究集中于权利性质及救济方式的理论研究,关于合规层面落实情况的观察较少。欧洲学者则针对 GDPR 所规定的数据主体权利进行了大量的经验研究,其中又集中于数据访问权、数据可携权、拒绝自动化处理与数字画像权和被遗忘权四项数据权利,四项加总约占总样本的三分之一。研究者为探索这些制度设计在实践中的真实情况,收集了大量企业合规层面的经验数据,试图探索制度设计的有效性和实施效果。因此,大部分既有文献提出了非常一致的研究问题:数据主体权利是否被企业真正落实,实施情况与书本上的法是否存在差距?研究的多样性则体现在样本选择的不同上,如物联网领域(Turner, Quintero & Turner et al., 2021)、广告领域(Urban, Tatang & Degeling et al., 2019)、新闻领域(Serveto, 2020)等。

在方法层面,数据权利研究的重要方法是从亲历者的角度,对权利的行使进行第一手体验。例如,王(音)等人(Wong & Henderson, 2019)为探索数据可携权的实施情况,以数据主体的身份向不同网站发送了 230 次行使数据可携权的请求,共收到 172 份回复,并进一步分析了这些网站提供的数据的格式;布法利尔日等人(Bufalieri, La Morgia & Mei et al., 2020)向 332 个网站发送了数据获取的请求,收集回复并进行分析。诚然,直接观察方法获取的第一手数据真实且与研究问题密切相关,已然成为数据权利的一种研究范式。但这种方法也存在局限,有时从数据控制者的回复情况无法完全评估数据权利的真正落实情况。例如,赫尔曼等人(Herrmann & Lindemann, 2016)直言其研究具有局限性,即便研究者收到了数据控制者关于行使其被遗忘权、数据访问权的回复,也无法得知数据是否真正被完全删除,亦无法印证获取的信息是否真实完整。因此,也有部分研究将第一手观测与访谈、问卷调

查、隐私政策分析等方法结合起来。例如,马修等人(Mahieu, Asghari & van Eeten, 2018)在向不同行业的数据控制者发送了 106 次数据获取请求以后,又针对获取数据的真实性、过程满意度向数据主体进行了问卷调查和访谈,将数据控制者的反馈与数据主体的态度相印证,形成了更为严谨的结论。

从结果来看,多数研究对数据主体权利的落实情况提出了质疑。例如,奥鲁斯等人(Ausloos & Dewitte, 2018)进行的数据访问权请求,有 67%未收到回复或回复不符合要求。被遗忘权、数据可携权的落实程度也都有限,过程也并不符合 GDPR 的规定(Turner, Quintero & Turner et al., 2021; Wong & Henderson, 2019)。因此,经验研究提示了数据主体面临的风险,也质疑了权利设计的有效性。GDPR 只规定了数据控制者保障数据主体权利的义务,但对数据控制者回复权利请求的过程没有任何程序性、细节性要求。这不仅导致权利难以实际落实,也可能导致权利滥用的风险。伯尼菲斯等人(Boniface, Fouad & Bielova et al., 2019)曾提示,由于缺少有效的身份验证机制,数据主体权极易被冒用与滥用,进而对数据主体构成新的威胁。

虽然既有研究指出了 GDPR 实施和制度设计上的不足,但就研究的有效性来说,一些文献为学界提供了经验研究方法论层面的良好经验。法教义学的解释论研究仅能提供"应该如何做"的信息,而权利在现实生活中究竟是如何落实的,则是理论和推演无法完成的,只能通过第一手的亲身经历获得。

4. 大众认知

在我们的样本中,有 13 篇文章讨论了大众对数据保护权利及数据保护法的认知。研究者提出了如下问题:大众对 GDPR 及法律对其隐私权的影响知晓程度如何? GDPR 颁布后,对大众对其隐私权的认知有影响吗?

针对第一个问题,研究者发现,即便大众对其数据权利有基本了解,也十分关注自己的数据保护和数据安全,然而大众对自己实际权利的认知极其有限,甚至大多数民众并不了解数据保护的含义,更不清楚如何行使自己的数据权利(Presthus & Sørum, 2018; Pleger, Guirguis & Mertes, 2021)。针对第

二个问题,通过 327 份问卷的发放和回收,普雷斯萨斯等人(Presthus & Sørum, 2019)给出了积极的结论:GDPR 颁布一年后,消费者对其隐私权利有了更明确的了解,一半受访者认为其对自己的数据拥有控制权,并且大部分消费者对企业处理其个人数据表示信任。

在方法层面,问卷调查是获取大众认知数据最为有效的方式,形成了该领域的研究范式。通过问卷调查,可以多角度地获取用户对 GDPR 及数据权利的态度、认知、理解、行为等多维数据。例如,普雷斯萨斯等人(Presthus & Sørum, 2019)的问卷中包含了对 GDPR 的地域范围、基本原则、同意的获取和数据权利的态度等多角度问题,同时获取了受访者的性别、年龄、职业等背景数据,为大众对 GDPR 的认知提供了较为全面的描述。

因此,传统的、描述性的研究仍然有其不可替代的价值。研究方法是由研究任务决定的,而对于需要获取受访者主观感受的研究问题,问卷调查方法具有很强的相关性和有效性。在进行问卷调查的时候,样本的选择是一个较为重要的问题。事实上,很多研究者也在研究局限部分提到了样本数量的局限性(Gáti & Simay, 2020)。与此同时,过大的样本数量又往往导致研究问题的肤浅。例如,普林斯等人(Prince, Omrani & Maalaoui et al., 2021)的样本数量非常大(回收了 26 526 份调查问卷),但研究问题仅局限于隐私权认知。同时,问卷调查受到受访者主观认知的影响,往往不适用于研究相对客观的命题。例如,一项关于数据保护设计的研究通过向数据工程师发放问卷的形式,试图探索影响数据保护设计的因素,但无法避免受访者的主观态度与认知的影响,对于衡量合规性之类的问题并不十分有效(Bu, Wang & Jian et al., 2020)。

5. GDPR 对经济、市场与竞争的影响

在众多经验研究集中于以数据主体为中心的权利与合规讨论时,一部分研究着眼于 GDPR 对业界的影响,即 GDPR 与数字经济、市场竞争与科技创新等市场要素的关系。在我们的样本中,有 13 篇文献从这个角度进行经验研

究。研究者试图探索:GDPR 的合规要求对企业运行有什么实际影响?GDPR 对市场竞争和创新有什么影响?

针对第一个问题,研究者又分别关注了中小企业与跨国企业(Walczuch & Steeghs, 2001),抑或在不同领域的企业中进行探索(Yuan & Li, 2019)。例如,哈丁等人(Härting, Kaim & Ruch, 2020)通过问卷调查与采访,认为由于缺乏专业知识、GDPR 合规的时间和人力成本过高、信息不足,中小企业往往难以满足 GDPR 的合规要求。事实上,GDPR 的颁布给中小企业带来了不小的挑战。西莫迪斯等人(Syrmoudis, Mager & Kuebler-Wachendorff et al., 2021)提出,即便理论研究认为数据可携权能够提高中小企业的竞争力,是一项有利于中小企业的制度设计,但囿于合规水平,中小企业并未有效地落实数据可携权的要求。针对第二个问题,盖尔等人(Gal & Aviv, 2020)则通过问卷与采访发现,GDPR 使市场结构更加集中,不利于竞争和数据创新。

与大众认知相似,关于 GDPR 市场影响的经验研究采用了传统的问卷调查、访谈与个案研究的描述性研究方法。尤其对中小企业相关的研究,几乎全部采用访谈或问卷调查相结合的方法。例如,诺瓦尔等人(Norval, Janssen & Cobbe et al., 2021)采访了 15 家初创科技企业的代表,深入了解了初创企业的观点。中小企业代表表示,专家意见通常不是技术初创公司日常合规的一部分,而更多依赖于同行间分享各自对 GDPR 的理解,并且技术初创企业中存在大量的执行效果“观望者”。从这个例子可以看出,尽管访谈的样本量往往较小,但通过更为深入的交流可以获取更多的有效信息,这也是结构化的问卷很难做到的。同时,计算方法也被广泛采用。例如,通过分析 11 000 条网络运营者使用第三方服务的数据条目,发现网络平台在 GDPR 实施后使用第三方接入的数量急剧下降,并且呈现出集中趋势,一些小规模的平台不再被使用,用户倾向于使用以谷歌为代表的大规模平台,进而引发了限制竞争的担忧(Peukert, Bechtold & Batikas et al., 2022)。事实上,传统描述性研究与新兴计算方法的研究各有优劣、互为补足。与访谈和问卷相比,计算模

型显然更为客观,但受样本真实性、计算偏差的影响,也难以提供更为深入的观点类数据;而访谈则可以进行更为深入的探究,不过受受访者主观性的制约,也许提供的事实并不完全准确。

数据法与市场竞争的关联和冲突极为复杂,但经验研究在理论研究之外提供了另一种进路,通过企业人员的真实感受和企业行为的真实数据提供证据,有效地为政策制定者提供了支持。不同的研究方法均有一些成功的经验研究提供了一些深入且有针对性的事实,值得借鉴与尝试。

6. GDPR 的整体实施效果

部分经验研究将 GDPR 作为一个整体进行宏观把握,探索某一领域之中、某一样本之下的整体实施情况。这类研究的目的十分明确:检查某一行业的企业对 GDPR 的落实情况。现有研究覆盖了云服务商(Kamarinou, Millard & Hon, 2016)、健康数据(Mulder & Tudorica, 2019)、政府部门(Faifr & Januska, 2021)、员工数据(Suder & Siibak, 2017)、软件开发公司(Usman, Felderer & Unterkalmsteine et al., 2020)甚至医疗服务(Wairimu & Momen, 2021)等不同领域。研究者大多发现了现有合规实践中的漏洞,并对 GDPR 执行效果进行评估或对企业提出合规期待。但由于研究范围较为宽泛,这类研究往往仅提供了数据展示,并不能在某一制度设计上提供较为有效的反思和建议。例如,阿尔菲尔特等人(Arfelt, Basin & Debois, 2019)测试了 796 个健康类 APP 的合规性,识别了几类较为普遍的违规行为,并表达了对医药数据合规的担忧。但作者也直言,该研究仅是一个初步探索,未来期待更多研究为企业合规提供更具有操作性的指导(Arfelt, Basin & Debois, 2019)。

7. 其他

从如上分析可以看出,虽然目前关于 GDPR 相关的经验研究已经很多,但同质化比较明显,大量经验研究集中在数据主体权利相关问题上,对 GDPR 中一些颇具新颖性的制度设计,如数据保护影响评估、数据保护设计等,并未有足够的经验数据覆盖。虽然近年来针对一些具体规则和制度设计形成了

少数有意义的探索,如卡马尔力诺等人(Kamarinou, Millard & Oldani, 2018)通过分析 13 份数据服务提供协议,探索数据控制者与数据处理者的关系,尤其是云服务商作为数据处理者的义务及其对 GDPR 第二十八条关于数据处理协议的合规情况。再如,一些学者已经对欧盟法院的司法判例以及各成员国的处罚决定进行了法律文本的经验研究,总结了 GDPR 实施后的制裁趋势和监管方向(Presthus & Sønslien, 2021)。但由于研究规模有限,因此未能形成趋势,也未形成有影响力的制度建议。我们期待相关学者在更多细致的制度设计下,尤其是 GDPR 介绍的新颖性的制度设计下,开展更多样化的经验研究。

五、在中国开展个人信息保护的经验研究:挑战与方向

中国的《个人信息保护法》在多个面向借鉴和吸收了欧盟的立法经验,从规范来源上属于同源。《个人信息保护法》实施以来,相关学术研究亟须对法律规则落实情况和实际影响进行经验测量。事实上,已经有少数研究进行了一些探索。例如,郑州大学信息管理学院杨瑞仙等人(杨瑞仙、沈嘉宁、许帆等,2023)对社交 APP 的隐私政策进行了评价,提出了完整性和可操作性方面的问题;黑龙江大学信息管理学院孙瑞英等人(孙瑞英、李杰茹,2022)则对我国政府数据开放平台的个人隐私政策进行了评价,并认为政府数据开放平台应加强政策落实、保护意识、权责配置。这些零散的经验研究远未形成规模和范式,且未能建立在对某一法律规则的理论研究之上,也较少有对法律解读和修改的反馈。

分析既有欧盟文献发现,经验研究在数据保护领域广泛应用,可以为法律解读提供重要的经验数据,得以验证、批判甚至修正理论研究形成的观点和假设。因而,欧盟经验也可以对我国个人信息保护相关经验研究提供一些启示和方向。一方面,经验研究基于亲身观测的第一手资料,可以考察法律

法规在实践中的运行和落实情况。这一类型的研究在域外已经较为成熟,有大量水准较高的经验可供借鉴。通过收集真实、客观的经验数据,经验研究得以发现数据保护法运行的不足之处,为进一步解释、补足甚至修改法律提供事实依据。数个关于同意获取过程中"黑暗模式"的分析,有效地总结了"黑暗模式"设计与用户同意有效性、自愿性的关系(Borghi, Ferretti & Karapapa, 2013)。王(音)等人(Wong & Henderson, 2019)在观察了数据可携权的实践以后,对法律文本的修改提出了建议。另一方面,经验研究得以研究《个人信息保护法》对大众、市场的影响。例如,斯特雷夏尔兹等人(Strycharz, Ausloos & Helberger, 2020)通过多维度的问卷调查,驳斥了隐私悖论理论,提出人们认识到了隐私的重要性,也意识到了自己拥有的权利,问题是使用和分享数据已然成为生活中不可或缺的一部分,民众事实上缺少真正选择的余地;诺瓦尔等人(Norval, Janssen & Cobbe et al., 2021)通过深入、有效的访谈,发现了中小企业在执行数据保护要求方面的显著漏洞及其背后的深层次原因。还有一个面向是为企业合规提供支持,这一类的研究还较为初级,虽然跨学科研究的趋势明显,运用了较为先进的机器学习等技术手段,但对法律实践与研究的影响均有限,仍须继续探索。应注意的是,并不是全部的域外研究都展现了较高的研究水平,部分研究的问题过于宽泛,应用的研究方法不够准确和有效,也并未提出有意义的结论,在学习域外经验时应注意有选择地借鉴。

更重要的是,在我国开展个人信息保护的经验研究,应立足于我国立法的具体场景。我国《个人信息保护法》虽然在体例和一些制度安排上借鉴了GDPR,但也有诸多不同之处,体现了鲜明的中国特色。这些制度设计也可以成为在我国开展相关经验研究的切入点。

第一,相较于GDPR,我国《个人信息保护法》采用了以原则为基础的规范框架,删减了大量细节,法律确定性上比GDPR要弱很多。这意味着经验分析所呈现的图景与基于文本进行的想象之间的差异会更为明显。同时,我

国《个人信息保护法》关联数量众多的行政法规、规范性文件、国家标准、指南等,需要经验研究来梳理其关系。

第二,我国《个人信息保护法》对 GDPR 进行了选择性借鉴,GDPR 中的一些制度设计并未被我国引入,如我国没有引入“正当性权益”(legitimate interest)作为个人信息处理的合法性基础,没有引入“数据控制者”与“数据处理者”的二元结构,而采取了个人信息处理者的一元结构。这些与欧盟立法的差异在企业合规和实践中意味着什么?其意义与影响需要经验研究来展现。

第三,我国《个人信息保护法》也增加了 GDPR 未规定的一些条款,如将不满 14 岁的未成年人信息作为敏感信息,增加了死者个人信息保护、公益诉讼制度,这些制度设计在法律运行中的实际效果,也需要经验数据提供支持。

第四,执法方面,我国《个人信息保护法》并未规定独立的执法机构,而是由数个相关部门共同履行个人信息保护职责。然而,具体的职责分配并未在法律中有明确规定。不同组织在实施法律过程中的方式、口径是否有区别,不同部门、机关如何分工,也是需要观察的问题。

第五,法律文化和社会影响方面,我国与欧盟对个人信息权的性质、个人信息自决权与国家对个人信息的保护义务的关系等问题的看法也存在区别。这种立法进路和法律文化上的差异值得运用经验研究的方法,比较同一法律制度在不同法域中的实施情况,从而提出更适合我国社会现实和立法需求的政策建议(贺欣,2021)。

六、结语

法学的经验研究方法经过与传统法教义学研究的分立、争论与融合,在我国学界已取得一些发展。本文立足法教义学与经验研究的区别,试图梳理经验研究的由来、范式、脉络和体系。本文指出了 empirical research 的翻译失

义和指代不明,提出“经验研究”应是指代这一研究方法的正确用语。由此,本文得以甄别哪些是真正的经验研究,其内核为何。

正名以后,本文聚焦于个人信息保护部门法,探讨在个人信息保护领域进行经验研究的方向与挑战。我国《个人信息保护法》已经出台,法律运行、执行过程中的现实情况和问题需要特别关注。“徒法不足以自行”,应当在个人信息保护领域进行方法得当的经验研究,获取第一手的经验数据来深入探索个人信息保护法律框架对大众认知、企业运作、数字经济等方面的实际影响,切实推动法律的实施、执行、解读和修订。

GDPR 相关经验研究可以为我国开展《个人信息保护法》相关研究提供借鉴。欧盟经验表明,数据保护相关的经验研究已呈现出一些既定的方法和立场。经验研究在数据保护领域具有很强的价值和相关性,方法得当的经验研究可以为法律解读提供重要的经验数据,经验数据揭示出制度设计与法律实施中的问题,可以验证、批判甚至修正理论研究的观点和假设。

最后,我们期待高质量、有针对性的经验研究为个人信息保护提供有价值的经验证据,但也不能忽视个人信息保护相关的法教义学研究。相反,法解释学与理论研究不仅有重要意义,个人数据保护的经验研究也需要从深入理解法律规范的制度设计入手,并尝试验证、补足和挑战现有理论研究与制度安排,提出法律解释和修改的建议。

参考文献

车浩,2021,《法教义学与社会科学——以刑法学为例的展开》,《中国法律评论》第 5 期。

丁晓东,2018,《什么是数据权利?——从欧洲〈一般数据保护条例〉看数据隐私的保护》,《华东政法大学学报》第 4 期。

贺欣,2021,《经验地研究法律:基于社会科学的外部视角》,《学术月刊》第 3 期。

侯猛,2020,《实证“包装”法学?——法律的实证研究在中国》,《中国法律评论》第 4 期。

梁慧星、王金浓,1980,《关于重庆市推行合同制的调查报告》,《法学研究》第 2 期。

宋旭光,2015,《面对社科法学挑战的法教义学——西方经验与中国问题》,《环球法律评论》

第 6 期。

苏力,2001,《也许正在发生——中国当代法学发展的一个概览》,《比较法研究》第 3 期。

孙瑞英、李杰茹,2022,《我国政府数据开放平台个人隐私保护政策评价研究》,《图书情报工作》第 12 期。

汤茂林、黄展,2020,《Empirical Research 到底是实证研究还是经验研究?——兼论学术研究的分类》,《地理研究》第 12 期。

王鹏翔,2015,《经验面向的规范意义——论实证研究在法学中的角色》,《"中研院"法学期刊》第 17 期。

谢海定,2014,《法学研究进路的分化与合作——基于社科法学与法教义学的考察》,《法商研究》第 5 期。

杨瑞仙、沈嘉宁、许帆等,2023,《社交媒体 APP 隐私政策评价指标体系构建及实证研究》,《情报理论与实践》第 1 期。

左卫民,2022,《中国计算法学的未来:审思与前瞻》,《清华法学》第 3 期。

Amos, R., G. Acar & E. Lucherini et al. 2021, "Privacy Policies over Time: Curation and Analysis of A Million-Document Dataset." DOI:10.1145/3442381.3450048.

Arfelt, E., D. Basin & S. Debois 2019, "Monitoring the GDPR." DOI: https://doi.org/10.1007/978-3-030-29959-0_33.

Ausloos, J. & P. Dewitte 2018, "Shattering One-Way Mirrors: Data Subject Access Rights in Practice." *International Data Privacy Law* 8(1).

Boniface, C., I. Fouad & N. Bielova et al. 2019, "Security Analysis of Subject Access Request Procedures." https://doi.org/10.1007/978-3-030-21752-5_12.

Borghi, M., F. Ferretti & S. Karapapa 2013, "Online Data Processing Consent under EU Law: A Theoretical Framework and Empirical Evidence from the UK." *International Journal of Law and Information Technology* 21(2).

Bu, F., N. Wang & B. Jian et al. 2020, "'Privacy by Design' Implementation: Information System Engineers' Perspective." *International Journal of Information Management* 53.

Bufalieri, L., M. La Morgia & A. Mei et al. 2020, "GDPR: When the Right to Access Personal Data Becomes A Threat." https://www.researchgate.net/publication/341175731_GDPR_When_the_Right_to_Access_Personal_Data_Becomes_a_Threat.

Cahillane, L. & J. Schweppe 2016, *Legal Research Methods: Principles and Practicalities*, Dublin: CLARUS Press.

Cane, P. & H. Kritzer 2010, *The Oxford Handbook of Empirical Legal Research*, Oxford: Oxford University Press.

Chang, C. et al. 2019, "Automated and Personalized Privacy Policy Extraction Under GDPR Consideration." in E. Biagioni, Y. Zheng & Siyao Cheng eds., *Wireless Algorithms, Systems, and Applications*, Berlin: Springer International Publishing.

Duxbury, N. 1997, *Patterns of American Jurisprudence*, Oxford: Clarendon Press.

Faifr, A. & M. Januska 2021, "Factors Determining the Extent of GDPR Implementation Within Organizations: Empirical Evidence from Czech Republic." *Journal of Business Economics and Management* 22(5).

Fernandez, C., D. Chatzopoulos & D. Papadopoulos et al. 2021, "This Website Uses Nudging: MTurk Workers' Behaviour on Cookie Consent Notices." *Proceedings of the ACM on Human-Computer Interaction* 5.

Gal, M. & O. Aviv 2020, "The Competitive Effects of the GDPR." *Journal of Competition Law & Economics* 16(3).

Gáti, M. & A. Simay 2020, "Perception of Privacy in the light of GDPR.' Proceedings of the European Marketing Academy, 11th.

Greenleaf, G. 2012, "The Influence of European Data Privacy Standards Outside Europe: Implications for Globalization of Convention 108." *International Data Privacy Law* 2(2).

Greenleaf, G. 2016, "Renewing Convention 108: The CoE's 'GDPR Lite' Initiatives." https://api. semanticscholar. org/CorpusID:157245104.

Härting, R.-C., R. Kaim & D. Ruch 2020, "Impacts of the Implementation of the General Data Protection Regulations (GDPR) in SME Business Models: An Empirical Study with A Quantitative Design." in G. Jezic et al. eds., *Agents and Multi-Agent Systems: Technologies and Applications*, Berlin: Springer International Publishing.

Herrmann, D. & J. Lindemann 2016, "Obtaining Personal Data and Asking for Erasure: Do App Vendors and Website Owners Honour Your Privacy Rights?" arXiv:1602. 01804.

Kamarinou, D., C. Millard & I. Oldani 2018, "Compliance as a Service." *Queen Mary Uni-*

versity of London Legal Studies Research Paper No. 297.

Kamarinou, D., C. Millard & W. Hon 2016, "Cloud Privacy: An Empirical Study of 20 Cloud Providers, Terms and Privacy Policies-Part I." *International Data Privacy Law* 6 (3).

Kollnig, K., R. Binns & P. Dewitte et al. 2021, "A Fait Accompli? An Empirical Study into the Absence of Consent to Third-Party Tracking in Android Apps." https://api.semanticscholar.org/CorpusID:235458342.

Krumay, B. & J. Klar 2020, "Readability of Privacy Policies." in A. Singhal & J. Vaidya eds., *Data and Applications Security and Privacy XXXIV*, Berlin: Springer International Publishing.

Lawless, R., J. Robbennolt & T. Ulen 2016, *Empirical Methods in Law*, Valencia: ASPEN Publishing.

Leeuw, F., H. Schmeets 2016, *Empirical Legal Research: A Guidance Book for Lawyers, Legislators and Regulators*, Cheltenham: Edward Elgar Publishing.

Leiter, B. 2012, "The Demarcation Problem in Jurisprudence: A New Case for Skepticism." https://doi.org/10.1007/978-94-007-6067-7_8.

Mahieu, R., H. Asghari & M. van Eeten 2018, "Collectively Exercising the Right of Access: Individual Effort, Societal Effect." *Internet Policy Review* 7(3).

Mathis, K. & A. Tor 2021, *Consumer Law and Economics*, Berlin: Springer International Publishing.

Mathur, A., G. Acar & M. Friedman et al. 2019, "Dark Patterns at Scale: Findings from A Crawl of 11K Shopping Websites." *Proceedings of the ACM on Human-Computer Interaction* 3.

McConville, M. & W. Chui 2017, *Research Methods for Law*, Edinburgh: Edinburgh University Press.

McKay, J. 2014, "General Jurisprudence, Empirical Legal Theory, Epistemic Fruit, and the Ontology of 'Law': Scope, Scepticism, Demarcation, Artefacts, Hermeneutic Concepts, Normativity and Natural Kinds." *Osgoode Legal Studies Research Paper Series* 36.

Milkaite, I. & E. Lievens 2019, "Child-Friendly Transparency of Data Processing in the EU: From Legal Requirements to Platform Policies." *Journal of Children and Media* 14(1).

Mulder, T. & M. Tudorica 2019, "Privacy Policies, Cross-Border Health Data and the GDPR." *Information & Communications Technology Law* 28(3).

Mulder, T. 2019, "Health Apps, Their Privacy Policies and the GDPR." https://ejlt.org/index.php/ejlt/article/view/667/897.

Norval, C., H. Janssen & J. Cobbe et al. 2021, "Data Protection and Tech Startups: The Need for Attention, Support, and Scrutiny." *SSRN Electronic Journal* 13(2).

Page, M., J. McKenzie & P. Bossuyt et al. 2021, "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." DOI: 10.1136/bmj.n71.

Peukert, C., S. Bechtold & M. Batikas et al. 2022, "Regulatory Spillovers and Data Governance: Evidence from the GDPR." *Marketing Science* 41(4).

Pleger, L., K. Guirguis & A. Mertes 2021, "Making Public Concerns Tangible: An Empirical Study of German and UK Citizens' Perception of Data Protection and Data Security." *Computers in Human Behavior* 122(6).

Pound, R. 1910, "Law in Books and Law in Action." *American Law Review* 44(12).

Presthus, W. & H. Sørum 2018, "Are Consumers Concerned about Privacy? An Online Survey Emphasizing the General Data Protection Regulation." *Procedia Computer Science* 138.

Presthus, W. & H. Sørum 2019, "Consumer Perspectives on Information Privacy Following the Implementation of the GDPR." *International Journal of Information Systems and Project Management* 7(3).

Presthus, W. & K. Sønslien 2021, "An Analysis of Violations and Sanctions Following the GDPR." *Journal of Intelligent Information Systems* 9(1).

Priel, D. 2008, "Were the Legal Realists Legal Positivists?" *Law and Philosophy* 27.

Prince, C., N. Omrani & A. Maalaoui et al. 2021, "Are We Living in Surveillance Societies and Is Privacy An Illusion? An Empirical Study on Privacy Literacy and Privacy Concerns." *IEEE Transactions on Engineering Management* 70(10).

Serveto, M. 2020, "Exercising GDPR Data Subjects' Rights: Empirical Research on the Right to Explanation of News Recommender Systems." *European Data Protection Law*

Review 6.

Soe, T., O. Nordberg & F. Guribye et al. 2020, "Circumvention by Design: Dark Patterns in Cookie Consent for Online News Outlets." https://doi.org/10.1145/3419249.3420132.

Strycharz J., J. Ausloos & N. Helberger 2020, "Data Protection or Data Frustration? Individual Perceptions and Attitudes towards the GDPR." *European Data Protection Law Review* 6 (3).

Suder, S. & A. Siibak 2017, "Employers as Nightmare Readers: An Analysis of Ethical and Legal Concerns Regarding Employer-Employee Practices on SNS." *Baltic Journal of Law & Politics* 10(2).

Sun, R. & M. Xue 2020, "Quality Assessment of Online Automated Privacy Policy Generators: An Empirical Study." Proceedings of the Evaluation and Assessment in Software Engineering.

Syrmoudis, E., S. Mager & S. Kuebler-Wachendorff et al. 2021, "Data Portability Between Online Services: An Empirical Analysis on the Effectiveness of GDPR Art. 20." DOI:10.2478/popets-2021-0051.

Tesfay, W., P. Hofmann & T. Nakamura et al. 2018, "I Read but Don't Agree: Privacy Policy Benchmarking Using Machine Learning and the EU GDPR." https://dl.acm.org/doi/pdf/10.1145/3184558.3186969.

Thorpe, R. & R. Holt 2008, *The SAGE Dictionary of Qualitative Management Research*, London: Sage.

Torre, D., S. Abualhaija & M. Sabetzadeh et al. 2020, "An AI-Assisted Approach for Checking the Completeness of Privacy Policies Against GDPR." 2020 IEEE 28th International Requirements Engineering Conference.

Turner, S., J. Quintero & S. Turner et al. 2021, "The Exercisability of the Right to Data Portability in the Emerging Internet of Things (IoT) Environment." *New Media & Society* 23 (10).

Urban, T., D. Tatang, & M. Degeling et al. 2019, "A Study on Subject Data Access in Online Advertising After the GDPR." https://doi.org/10.1007/978-3-030-31500-9_5.

Usman, M., M. Felderer & M. Unterkalmsteine et al. 2020, "Compliance Requirements in

Large-Scale Software Development: An Industrial Case Study." https://doi.org/10.1007/978-3-030-64148-1_24.

Utz, C., M. Degeling & S. Fahl et al. 2019, "(Un)Informed Consent: Studying GDPR Consent Notices in the Field." DOI:10.1145/3319535.3354212.

Wairimu, S. & N. Momen 2021, "Privacy Analysis of COVID-19 Contact Tracing Apps in the EU." https://link.springer.com/chapter/10.1007/978-3-030-70852-8_13.

Walczuch, R. & L. Steeghs 2001, "Implications of the New EU Directive on Data Protection for Multinational Corporations." *Information Technology & People* 14(2).

Wong, J. & T. Henderson 2019, "The Right to Data Portability in Practice: Exploring the Implications of the Technologically Neutral GDPR." https://api.semanticscholar.org/CorpusID:198362435.

Yuan, B. & J. Li 2019, "The Policy Effect of the General Data Protection Regulation (GDPR) on the Digital Public Health Sector in the European Union: An Empirical Investigation." *International Journal of Environmental Research and Public Health* 16(6).

理性、权力与生态
——平台研究相关中文文献述评

袁方杰*

摘要：本文试图结合当前国外平台研究的主要成果，提出平台研究主要存在理性、权力与生态三种理论范式，通过对近10年中文文献的平台研究进行简要综述和评论，认为中文文献中的平台研究需要在吸收西方理论成果的同时，根据本土时空情境开展理论建构。这种理论建构可以分为两个方面：一是需要关注国家等组织与平台的互动；二是需要借鉴社会学的理论视角，从文化范式入手做出贡献。本文依据三种理论范式对中文文献中的平台研究进行述评，旨在推动平台相关研究的进一步发展。

关键词：理论范式　平台　组织

一、引言

平台作为一种经济和组织形式，存在时间虽然较短，却已成为经济和社会领域的重要组成部分，引发了关于如何理解这种新的经济和组织形式的争论。早期的研究倾向于一概而论，认为平台将彻底改变传统就业市场（Sundararajan，2016），或强调平台的破坏性（Robinson，2017）。近年来，学界做了不少经验和理论研究，也有学者做过相关研究的综述（魏海涛、李国卉，2022），但这些研究往往聚焦于某一具体类型的平台现象，并没有对作为一种新的经济和组织形式的平台做出总体性梳理。另外，目前的平台研究大多局

* 袁方杰，上海大学社会学院。

限于对已有研究观点的总结,并没有围绕理论范式与学界进展进行对话。本文认为,为了推动这一重要研究领域的发展,我们需要从理论范式视角入手对平台现象进行整体性理解,并且不局限于某一学科,而是放在社会科学的整体框架中。

首先,本文借鉴英文文献对平台研究范式的整理(Vallas & Schor, 2020),并结合近 10 年平台研究的相关理论,提出中文文献中的平台研究存在三种主流范式,即理性、权力与生态范式。其次,在这三种范式下,本文梳理近 10 年平台研究的中文文献,并对各个范式下的中文文献进行简要述评。最后,本文指出目前有关平台研究的中文文献的主要缺陷和未来可能的研究方向,强调引入文化、国家等多元视角。需要补充的是,本文主要评述的对象是经济学、社会学等主流社会科学领域的研究,其他有关的理论和经验研究无法作为重点讨论对象在文中出现。

二、平台研究的主要范式

目前,平台研究的理论来源以社会学、经济学和法学为主。平台研究的社会学理论主要源自马克思主义传统的劳动控制和劳动过程理论,强调平台的权力和冲突面向。平台研究的新制度经济学理论强调与公司理论对话,以及平台的效率与创新面向。平台研究的法学理论则源自反垄断与监管理论,强调平台的可塑性面向。这三大理论大体上对应了三种理论范式,本文将之命名为“理性范式”“权力范式”和“生态范式”。因此,本文接下来将依据这一范式分类来回顾平台研究的理论进展。

(一) 理性范式

最初,平台指的是能够调动闲置资源的共享经济,如出租空闲房间或提供乘车服务的共享平台(Frenken & Schor, 2017)。经济学家强调,平台的出

现能够降低交易成本,促进信任,提高市场效率(Einav, Farronato & Levin, 2016)。尽管随着平台规模的扩大,早期对平台经济的想象失去了合理性,但经济学家继续强调平台革命改变了传统企业模式的主导地位。借助新技术和新的商业模式,平台组织比传统经济组织能够带来更多的经济效益(Sundararajan, 2016)。这一范式认为,与传统经济组织形式相比,由于数字技术和算法创新,平台具有一系列优势。平台通过减少对层级结构的依赖,降低了企业经营的交易成本,重新整合了传统经济组织的相关功能,消除了个体进入劳动力市场的障碍。经济学家理解平台的理论基础是科斯(2009)的公司理论。根据古典经济学的假设,市场在没有干预的情况下最有效,这与企业广泛依赖各种协调形式这一现实相矛盾。科斯则解释到,当经营规模促使隐性交易成本上升,并通过价格机制增加总体成本时,企业就会出现。企业会通过计划和控制等手段将价格机制的协调功能内部化,以此来降低市场成本。平台的出现进一步降低了交易成本,促进了市场活动的有效协调。平台只持有有限的固定资产,雇佣少数员工,并将利润创造的过程外部化。例如,平台可以将评分任务众包给用户,因此无须昂贵的广告费用就能促进潜在的市场信任。在这一范式下,平台革命意味着企业的效率最终得以提升。平台作为一种更加平等和松散的经济组织形式,垂直的公司等级制度不再是其主要结构。

此外,理性范式认为平台的出现改变了传统雇佣关系。这种转变的一个表现是,许多平台劳动者获得了自主性和选择权。基于技术创新的平台将提高劳动者的自主权和灵活性。共享平台降低了交易成本,为特定的任务租用劳动力,而不是为公司而雇佣,即劳动力可以灵活流动。理性范式认为,通过技术创新,平台可以创造更多的就业机会(Sundararajan, 2016):首先,平台经济增加了对其他类型工作的需求;其次,平台的产品创新可以创造新的消费需求,催生新的职业甚至经济部门;最后,平台降低了劳动力成本,提供了更便宜的服务,最终扩大了消费者的需求,反过来促进了就业。此外,汽车、房

屋等资产可以通过平台实现货币化,释放其潜在价值,从而减少人们对正式收入的依赖。有研究认为,这有利于低收入家庭获得更好的生活(Sundararajan, 2016)。平台组织也对市场的网络结构提出了挑战。传统的市场网络由有限边界的企业组织构成,它们是为了追求利润,围绕共同设计的规则而形成的(Powell, 1990)。与此相比,平台组织则向市场开放,数以百万计的匿名生产者和消费者基于互惠与信任在平台上相互交易。这种区别既是数量上的,也是质量上的。因此,理性范式假设平台的出现和发展是由理性的效率动机所驱使的,提高了市场整体利益,解决了信息不对称,降低了交易成本,纠正了市场失灵,实现了社会福利均衡。

(二)权力范式

如果说第一种范式认为平台最终合理地提高了效率,那么第二种范式则可能相反。权力范式认为平台作为一种组织,内部存在着权力斗争与不平等,对平台内的劳动者产生限制和约束。劳动过程和控制理论将平台视为控制和剥削劳动者的场所,剥夺其对生产资料的掌握,平台的算法控制导致了劳动标准的恶化。参考算法批判的相关理论,平台劳动的特点是平台通过算法管理和控制劳动者,从而在更大程度上强化了平台企业的权力。在布洛维(2008)的"工作场所"中,劳动者经常会设计出一些策略,用于逃避、蔑视或抵抗机器的控制。这种能力现在正受到算法技术的破坏,因为算法技术可以将工作场所的权力关系编码到劳动者完成任务所必须使用的数字工具中。总的来说,平台削弱了工人抵制、逃避或挑战平台规则的能力,其劳动控制以算法为核心,这与市场和行政控制都不同。权力范式强调利益冲突,占主导地位的商业化数字平台控制边缘群体和弱势群体,引发了社会不平等,以及大量的冲突行为和无处不在的对抗(默多克,2022)。

平台可以通过几种方式做到这一点。首先,平台会收集大量有关其内部运营的数据,包括通过应用程序进行持续的数据收集和监控。但这些信息并

不共享,因而助长了企业内部的信息不对称(Calo & Rosenblat, 2017)。例如,在打车和外卖行业中,平台通过隐瞒即将到来的订单信息来蒙蔽劳动者,从而降低了工人的自主性和收入(Rosenblat & Stark, 2016; Shapiro, 2018)。平台的数据收集也使雇主能够更详细地干涉劳动过程,如规定工人必须接受的工作比例、必须腾出多少时间以及必须达到的工作预期。其次,平台不仅依赖于算法管理,还利用游戏、象征性奖励和其他诱导式规范,使用精心设计的策略加强用户对平台的依赖(Calo & Rosenblat, 2017)。最后,平台将劳动者分解为原子化个体,将劳动者从社会情境中抽离出来,防止劳动者挑战管理权威,加剧了劳动者之间的竞争。

权力范式认为,在平台经济中,工作越来越不稳定、不安全。平台不只是一种新的经济和组织形式,它们提取和利用大量数据的能力使它们能够作为市场的基础设施而存在。平台对整个市场的主导地位一旦确立,这种"赢家通吃"的市场支配地位便能为平台通过租金获取回报提供许多途径。同时,由于利益相关者数量的增加,市场对平台的依赖使平台免受外部的监管。这种网络化的主导地位使平台公司既是传统垄断形式的复兴,又是传统垄断形式的创新。平台作为福特主义组织的最新形式,引发了标准工作安排的衰落(Kalleberg, 2018),取消了劳动者之前享有的经济权利和劳动保障(Vallas, 2019)。此外,通过灵活雇佣或临时雇佣形式,平台能够将以前不得不承担的风险外部化,并将劳动时间商品化,使工人脱离先前的社会保护体系,导致工人的地位日益脆弱和不确定(van Doorn, 2017)。由于失去了最低工资、安全法规、退休收入、医疗保险等保护,平台劳动者被迫承担以前由企业和国家承担的各种风险。例如,当平台需求不足时,他们就会被辞退,并自行承担所有损失(Ravenelle, 2019)。

(三) 生态范式

平台工作的第三种范式最难定义。这一范式总体上主张更强的环境决

定论,认为平台组织的生存嵌入其所处的制度生态中。在这一范式下,平台作为一种经济和组织形式,其意义、性质和影响并不由平台本身的功能和特征所决定,而是需要纳入平台所处的法律和制度环境中考虑,其影响取决于其所处的环境或生态。例如,有研究发现,Uber 的破坏性影响在德国、瑞典和美国有明显差异,在每个社会都产生了不同的问题。虽然平台几乎总是将劳动者归类为独立承包商,但这一特征在美国最为突出。在其他国家,Uber 带来的主要影响不在就业市场,而是在城市交通系统、税收体系、法律监管机构等(Thelen, 2018; Zanoni, 2019)。

生态范式认为,平台不仅仅是技术变化的产物;相反,它与政治、经济和法律等外部生态环境密切相关(Pasquale, 2016)。平台公司的崛起,部分发生在与政治体制和监管政策的互动中。平台不是孤立的中介组织,而是嵌入相互作用的生态中。如果良好的法律和制度规范推动平台确保最低的工资水平、适当的劳动权益、合理的税收贡献,平台甚至可以促进社会治理,提高社会福利。平台的影响不能与它所处的制度环境相分离。因此,平台本身并不决定性地产生危害;如果有合适的法律和制度干涉,平台可以提高工人的福利和权益。例如,平台依据法律标准,将劳动者定义为独立承包商,而不是公司雇员。这种定义对公司有明显的好处,它使公司摆脱了昂贵的人工成本,只用支付较少的管理费用。然而,法律规定平台必须保证工人有一定程度的自主权,通常表现为不受雇主控制,这种观点在倡导平台合作治理的研究中得到体现(Scholz, 2016)。这一范式要求我们意识到平台形态的嵌入性。这一范式将平台视为很大程度上由其所处制度环境所塑造的,可以避免数字技术决定论,从而强调技术在不同的社会和政治制度下产生不同的影响。

三、三种范式下的中文平台研究述评

根据前面的讨论,本文将根据这三种主要的研究范式来梳理中文世界中

与平台现象相关的研究。受篇幅和学识限制,本文讨论的文献以社会科学学科中的主流期刊为主,包括《中国社会科学》《社会学研究》《经济研究》《法学研究》《社会学评论》等。需要注意的是,基于范式的讨论并不意味着一项研究只能拥有一种范式的理论视角。

(一) 理性范式

基于理性范式,中文文献的平台研究主要从以下几个方面展开讨论。

首先,互联网平台作为数字经济的主要组织模式有利于企业创新。研究发现,当收取会员费时,互联网平台的出现可以提高社会福利、企业创新水平,促进市场均衡。互联网平台所覆盖的市场范围扩大,其对企业创新的促进作用也扩大(欧阳耀福,2023)。比如,金融科技平台正是凭借着技术优势,获取了海量信息,缓解了企业的流动性约束,从而普遍提高了企业创新能力。金融科技平台利用数字技术扩大信贷规模,提升信贷效率,为企业和行业带来了一定价值,这是一个效率提升和价值创造的过程。平台的大数据技术使得算法预测更具效率和精度,创造价值的效率也随之提高,有效降低了信息不对称程度,提高了金融普惠性(王义中、林溪、孙睿,2022)。平台经济催生了一大批新就业形态,如零工经济,为劳动力市场带来了深刻的变革。平台既可能通过减少风险而促进创业,又可能挤出低质量、生存型的创业活动,进而带动金融、科技等行业的创新(莫怡青、李力行,2022)。知识经济背景下,平台企业构建情境的不同方式促成了不同的迭代创新模式(朱晓红、陈寒松、张腾,2019),也以跨界网络治理的方式驱动了颠覆性技术创新。内容治理与价值治理是“硬实力”,而关系治理是“软实力”,平台通过软硬实力的组合共同促进了颠覆性技术的突破与市场突破(李东红、陈昱蓉、周平录,2021)。

其次,平台也提升了社会效益,其技术特性和技术应用方式所构建的红利共享机制扩大了受益群体(张茂元,2021)。平台经济的不同发展阶段对应不同的社会福利影响机制,产生不同的福利提升效应和福利分配效应。在互

联网外卖领域,通过市场竞争,平台能够提高消费者、商户、平台及骑手的福利,使之接近社会福利最大化(蔡跃洲、顾雨辰,2023)。通过在盈利性与社会性、专属性与公共性、创新性与安全性等方面实现动态平衡,平台具有技术赋权、增加社会资本积累的功能和公共服务能力(范如国,2021;李彬、翁慧敏,2022)。基于数字平台,企业把社会价值创造内置于平台商业生态体系中,通过对社会、商业关系与资源的混合配置与转化利用来创造共享价值,履行社会责任。其中,社会价值创造是商业价值创造的前提,两种价值创造相互依赖且多有重叠(邢小强、汤新慧、王珏等,2021)。

平台经济的发展具有规模优势,能够加速生产、流通及消费,提高市场效率,促进社会生产力的发展(周文、韩文龙,2021)。平台通过信息的筛选,在有限的屏幕空间实现了供需信息的匹配,进而完成市场交易。互联网平台以信息为投入和产出,以参与消费者的决策过程为主要服务(曲创、刘重阳,2019);也可以通过平台市场的最优分层设计,建立一个拥堵性和安全性相对重要的双边市场(王勇、吕毅韬、唐天泽等,2021)。平台经济作为数字经济的代表形态,以数字化技术为基础,设置交易规则,监管交易行为,降低交易成本,由此形成了市场经济中的新技术秩序。平台成为市场结构和市场秩序新的设立者和维护者,有利于市场的有效运转(江小涓、黄颖轩,2021)。在平台市场中,存在一个均衡状态,能够最大化、最优化地协调政府部门与平台企业的策略互动,对卖家实现最优水平的监管激励(王勇、刘航、冯骅,2020)。

根据理性范式背后的效率原则,平台发展的进程具有目的论的内在逻辑,即注定是改善福利和提高效率的。平台基于理性逻辑而发展,参与者都从平台中获得最优的资源分配结果,平台劳动者无法通过其他方式获益更多。这种解释的问题在于,它忽略了平台管理中的黑暗一面,假设了个体对平台的完美适应,如算法被用于跟踪和监测工人在生产和分配空间中的移动。这种劳动控制既创造了监管系统,也通过灌输纪律并被完全整合到生产和分销过程中,使平台能够更密切地监控和干预工人的流动,约束工人的行

动。此外,平台的增长速度仍然受到传统劳动力市场的影响(Schor, 2020),工作的质量和类别影响着平台劳动者的经验。更具体地说,许多平台劳动者的积极经验来自他们从特定行业那里获得的福利和保障(Schor, 2020),但平台不太可能使权力和控制权的分配扁平化。相反,平台可以扩大规模,主导市场,并获得足够的垄断权力,从而在劳动力市场上占据支配地位,降低不确定性(Dube, Jacobs & Naidu et al., 2020; Khan, 2017)。在这种情况下,个体可能不会像自由行动的经济理性人一样在平台中不受约束。目前,越来越多关于打车平台、外卖平台劳动条件恶化的文献出现(Robinson, 2017; Rosenblat, 2018; Shapiro, 2018),这表明平台存在权力结构和不平等,某些制度安排是可以通过权力强制实现的,而不仅仅是促进创新、提高效率的公共物品。

(二) 权力范式

权力范式下的平台研究主要集中在市场关系、劳动控制、技术互构等方面。

首先,平台作为市场化科层组织,依托企业规则、数字技术和第三方机构实现了垄断地位,在外部构建了一个由多市场参与主体组成的、层级化的科层控制体系(赵磊、韩玥,2021)。平台通过权力控制与资本积累重塑了市场竞争结构,集成数据和信息资源为平台塑造了一个相互监视的场域,造成网商无效和低价竞争,形成了对网商的隐蔽控制(邵占鹏、甄志宏,2022)。数字平台对传统行业的生产组织方式进行了重构,颠覆了传统行业的生产组织方式。在货运行业中,平台使交易从熟人交易转向陌生人交易,从议价模式转向竞价模式,产生了较高的交易风险并恶化了相关主体的处境(周潇,2021)。平台经济通过生产资料的集中、依附关系、垄断能力,以及“自发劳动力”的动员优势,不断强化平台资本,创新资本积累的模式。其以分享与互利为基础的生产资料所有制掩盖了生产资料的归属问题,在一定程度上混淆了平台内部的冲突(邵占鹏、甄志宏,2022)。

其次,平台上的劳动过程和劳动控制呈现出新的特征。根据陈龙(2022)的研究,平台数字经济中的劳动者具有双重身份。平台劳动的控制手段多样化,劳动的对象也从机械的身体转向灵活的认知(贾文娟,2023)。在劳动商品化和薪资红利化背景下,“人气游戏”模糊了劳动付出与工资的界限;同时,平台经济的意识形态强调个人责任的工作伦理,掩盖了工作不确定性背后的权力关系(徐林枫、张恒宇,2019)。平台以技术手段调动和整合劳动的弹性积累,通过创造“梦想”的机会制造和技术控制,使劳动者积极参与到自我规训的生产过程中,从而模糊了真实的劳资关系,将劳动异化扩展到整个生活世界(胡慧、任焰,2018)。平台构建了新型的劳动时间控制模式,创造灵活的工作时间和宽松的工作环境来吸引劳动者加入;但在技术手段的辅助下,也塑造了“准时”和“快速”的劳动时间感。最终,劳动者被束缚于平台之中,不得不主动配合平台的劳动控制(李胜蓝、江立华,2020)。同时,在劳动过程中,人工智能技术的应用加强了平台对劳动过程的管控。算法系统、中介机构消费者取代了平台公司对骑手的管理,生产关系被进一步模糊,冲突也转移到了技术系统与消费者之间。平台通过数字控制削弱了劳动者的反抗意愿,侵蚀了他们发挥自主性的空间,并让他们在不知不觉中参与到对自身的管理过程中(陈龙,2020)。在严密的管理与监控下,平台也可以采取更微妙的方式进行选择性放任(庄家炽,2019),将资本积累的逻辑渗入劳动力再生产过程,导致就业和工资的不稳定(谢富胜、吴越、王生升,2019)。然而,也有研究表明,骑手可以通过合理利用规则、积极合作和抗争等方式来减少平台的控制,争取劳动的自主性,维护自身的主体性(冯向楠、詹婧,2019),并消除劳动过程中的限制因素。

最后,平台处于社会与技术互构的关系之中。平台技术为吸纳不同类型的个体进入平台提供了机会,创造了一种新的社会情境,规避了过去技术与组织相适配的规则。在平台中,为了获得更高收入,司机需要让渡自身的权利,平台掌握了基于算法的决策权和剩余控制权,获得了远高于其他行动者

的权力(张樹沁、户雅琦,2021)。以某信息分发平台为例:一方面,行动者在权力和利益的影响下展开复杂的策略互动与博弈;另一方面,行动者也越来越受到平台算法逻辑的影响,塑造了他们的行为方式以及采取行动的情境(赵璐,2022)。平台作为一个高度复杂的行动者网络与社会现实频繁互动,这是因为平台算法通常非常复杂,处于不透明、难以追踪的黑箱状态。信息经过平台进行隐蔽的精细化处理,平台通过捕捉和估计用户偏好的语义,预测和强化不同个体的语义偏好。随着个体与平台的互动,不同个体逐渐被固定在特定位置的语义光谱中,只接收特定信息,这导致个体信息获取的窄化以及个体间的信息隔离更为隐蔽。因此,平台通过更精细、隐蔽、复杂的信息流控制来获取社会权力(刘河庆、梁玉成,2023)。

权力范式认为,平台中存在普遍且无法调和的利益冲突,即某个居于主导地位的平台将某种安排强加给其他行动者。因此,平台不可能是为所有参与者理性设计和协调的结果,而是必然反映了平台的利益偏好。平台的控制反映了权力分配,而不是通向集体福利的理性适应。然而,劳动者也不是天真无知的。随着他们越来越熟悉面向劳动者的算法管理,他们可以通过各种微妙策略来抵抗公司。关于劳动者对算法的抵制和掌握的文献越来越多,经验研究发现劳动者确实可以发展出抵抗形式,对平台采取反击行动。例如,一些自由职业者已经找到了利用平台声誉指标系统的方法,基本上是通过与客户结成在线联盟来进行博弈(Rahman & Thelen, 2019)。此外,还有通过向算法提供数据、离开平台和躲避监控来操纵平台算法的方法(Jarrahi & Sutherland, 2019)。对司机群体的研究也表明,劳动者采取了一系列策略来规避平台的规则,如建立稳固的社会联系来控制自己劳动力的价格(Robinson, 2017)。滴滴司机可以通过使用多部手机等数字策略来抵抗算法(Chen, 2018),外卖员也有类似的抵制算法控制的方法(Shapiro, 2018)。平台为了应对劳动过程的不确定性,通过算法技术将工作步骤标准化,以实现对劳动的控制。然而,算法无法对劳动过程的变化进行预测和规避,因此常常受到

限制。因此,劳动者可以调动自主性,为自身争取更多的灵活空间(付堉琪,2021)。如果权力范式认为算法管理完全满足了平台对劳动过程的控制,或阻止了劳动者的抵抗能力,可能会高估或夸大数字技术的力量。

权力范式认为,平台只关注自身利益,而忽视整个社会福利。拥有足够权力的行动者会强制施行偏好安排,因此忽视其他所有行动者的利益。然而,这一范式下的大多数研究将平台行动者视为同质群体,过于简化了平台劳动的异质性。现今平台提供的工作种类已迅速增加,包括众所周知的服务,如驾驶、图像标注以及护理劳动等。这些工作种类在技能水平、空间分布和雇佣关系上存在差异。平台内同时存在高技能高报酬的劳动者和只具备一般技能(如驾驶、家庭清洁)的临时工。有关平台劳动者多样性的证据表明,全职工作者与众包工作者的情况不同(Manriquez, 2019; Rosenblat, 2018; Schor, Attwood-Charles & Cansoy et al., 2020)。当劳动者将平台劳动作为补充收入时,平台的存在实际上可能会降低不稳定性,弥补主要工作的低工资和低福利,或帮助减少债务(Schor, Attwood-Charles & Cansoy et al., 2020)。此外,由于补充收入者与需求方具有一定程度的必要距离,他们可以拒绝低薪工作,在劳动力市场的博弈中占据更有利的位置。因此,他们的工资更高,对工作条件有更大的自主权,也比依赖平台的劳动者更容易满足。

虽然学者们已经认识到平台中的劳动形态是多样化的,但这种异质性尚未得到充分的探讨。以前的研究主要聚焦于打车和外卖平台,由于这些平台规模庞大,因此成为明显的研究对象。然而,这类平台也具有特殊性,如驾驶是一种广泛存在的技能,因此该行业容易出现劳动力供应过剩现象。同时,这些平台所处的行业存在巨大的经济利益,吸引了前所未有的资本投入。随着时间的推移,打车平台吸引了比大多数平台更多的全职收入者。因此,是否可以将平台视为剥削和控制劳动者的组织形式并没有定论。此外,平台经济嵌入在一个劳动力市场的异质性环境中,劳动者之间的情况和经验差异很大,这对劳动力市场理论中一个常见的假设提出了质疑,即为同一雇主从事

相同工作的劳动者之间应该具有相似的经验。在工作满意度、薪酬、自主权和管理控制方面,不同平台之间以及同一平台上的劳动者之间存在着明显差异。开放式就业和算法管理的结合,导致劳动者的经验在很大程度上取决于他们对平台收入的依赖程度,其他收入来源和社会保障几乎成为平台劳动者获得满意体验的前提。如果其他地方存在充足的工作机会,替代选择机制将迫使平台改善条件。当平台收入是补充性的(存在多种收入来源)时,工人的满意度较高,自主性也较大,更容易获得更高的工资,劳动条件较好。相比之下,那些依赖平台经济生存的劳动者更容易表达更多的不满,并经历更多的不稳定性。这些发现强调了研究那些能够可以退出平台的劳动者的必要性。

(三) 生态范式

近 10 年来,中文文献关于生态范式的研究主要集中在以下几个方面。

从政府与市场互动的角度来看,平台已经成为一个融信息汇集、资源配置和规则制定为一体的新型经济枢纽和组织形态。多边主体在平台提供的规则下进行交互,共同创造价值,并对经济和社会产生广泛而深远的影响。随着数字化进程的深入,企业和政府主体普遍选择接入平台(刘诚、王世强、叶光亮,2023),形成了多产业融合的平台生态。平台经济的出现使市场的决定性作用得以充分发挥,同时也更好地发挥了政府的作用,国家权力和资本力量在平台经济中相互结合(黄冬娅、杜楠楠,2022)。平台经济的兴起让国家与资本力量正式遭遇,并扩大了政企互动的影响范围。随着平台的不断发展,国家对平台的监管也在逐步摸索和调整之中,使得监管互动充满了巨大的不确定性。同时,监管体系的完善使国家权力与平台经济之间通过正式和非正式关系进行互动。此外,平台成为政治权力和市场权利之间的第三方力量,具有明显的公共性、生态性和非竞争性(张晨颖,2021)。平台的社会责任缺失和异化将引发一系列突出的社会问题,因此需要将平台情境中的社会责任生态化(肖红军、李平,2019),这要求政府主管部门与社会舆论监督需要共

同形成对平台的综合治理(吴伟光,2021)。

从市场主体的角度来看,平台会围绕其核心业务整合相关产品和服务,构建有利于自身而不是竞争对手的生态体系,通过运用不同市场间的杠杆力量破坏竞争。平台经济具有强大的网络效应,使得平台内的行动者对平台产生依赖,形成锁定效应;同时,平台还可以利用其获得的大量数据,从一个领域延伸至其他领域,形成数据优势(周汉华,2023)。平台在数字化的基础上融合了多样化的治理模式,不再将市场机制与等级制度视为相互排斥的协调工具或组织管理模式,而是将它们整合于一个平台之中(胡秋爽,2023)。中介机构和层级机构的扩展有助于减少平台内经营行为的负外部性,节省经济成本,但其权力也容易被滥用。平台企业提供的大数据征信服务可能会对市场竞争产生影响(曹光宇、刘畅、周黎安,2022)。平台企业也可能会利用数据、用户流量等杠杆来撬动各个市场上的份额,导致商业平台无限扩张,出现一系列垄断乱象,如滥用市场支配地位等问题(孙晋,2021)。因此,平台的市场权力需要得到适度约束。在制定和实施网络平台规则时,应遵循基本的程序正义和实体正义标准。立法者应对平台滥用私权力的行为进行必要的司法审查,科学合理地设定平台责任(刘权,2020),并在市场制度框架内进行监管。平台企业应完善个人信息保护体系,设立外部独立监督机构,监督个人信息保护以及对商业用户个人信息处理活动的合规情况(张新宝,2022)。此外,大型数字平台通过技术手段诱导用户成瘾,使有限理性的用户沦为数字平台的非理性拥趸。因此,应将成瘾性治理纳入体系化治理架构,打破平台的生物霸权和瘾性垄断(孙瑜晨,2022)。

从国家主体的角度来看,平台扮演着不可替代的角色,成为重要的治理工具(吴新叶,2011)。平台化运作使政府能够建立整合贯通的流程体系,将需求的整体主义、运作的平台主义和行动的专业主义结合起来(孙志建,2022)。基于平台的治理为犯罪治理提供了技术、组织和制度上的转型路径,塑造了数字科层体系。作为数字基础设施,平台连接了国家和社会的在线互

动(单勇,2022)。数字平台由于其快速反应能力和强大的技术应用,在许多公共事务中发挥了重要作用,有助于实现政府决策的科学化、社会治理的精准化和公共服务的高效化。算法平台的广泛应用使得决策日益自动化,导致现有平台监管的追责模糊、治理滞后以及责任设置不符合比例。同时,区块链技术与平台的结合将改变知识权威的来源,而通证平台权力给用户的数字权利带来了复杂的影响。对平台的治理应嵌入国家的治理网络中,既包括正式的规则体系,也包括各种非正式的治理手段(吕鹏、周旅军、范晓光,2022)。对于平台可能引发的技术风险,政府应依靠技术手段进行内部优化,以实现对平台的法律治理(李晶,2023)。同时,在政策层面应持有包容审慎的态度(陈全真,2023)。平台监管既应符合平台底层的技术逻辑,也应符合法律原则。平台本身是中立的,但平台技术的应用包含主观意图,因此应同时将平台的算法设计和主观过错作为问责的依据(张凌寒,2021)。

生态范式强调平台与社会之间的良性互动,同时也强调平台可能引发的不利影响,这取决于平台所嵌入的制度环境和法律监管。然而,由于过于强调平台所带来的社会影响的制度偶然性,这种观点有可能忽视平台所表现出的共同特征,尤其是在全球化的世界中,平台作为一种组织存在,能够挣脱时间和空间的限制,复制其组织结构和商业模式。举例来说,不同地区的数字平台在劳动条件方面几乎没有显著差异(Wood, Graham & Lehdonvirta et al., 2019)。监管机构虽然可以塑造平台企业的运营方式,但这并不意味着平台是完全可塑的。平台所依赖的商业模式和技术结构具有一定的内在特征。研究表明,平台公司对其外部制度环境具有强大的影响力,甚至能对政治机构产生影响(Culpepper & Thelen, 2019)。例如,平台对市场结构的影响不仅扩大了对参与者的控制,也在将部分控制权委托给交易双方的同时,保留了对重要职能如任务分配、收入获取的控制权。为了获取潜在的劳动力资源,平台必须在全球范围内寻找众包劳动力(Wood, Graham & Lehdonvirta et al., 2019)。这种雇佣劳动的空间分散产生了两种影响:首先,工人之间的不

平等和竞争关系日益加剧;其次,分散产生的个体化削弱了工人的集体行动能力。最终,通过降低就业准入门槛,平台改变了外部的劳动力市场,形成了层级的结构。因此,生态范式过于强调现有制度规范对平台的约束作用,然而这种对制度生态的依赖并不是一个无可争议的预设,而是需要被解释的现象。

简言之,在整个社会科学领域,不同学科的学者对平台的理解存在明显差异:理性范式为平台提供了一个强调效率和均衡的叙述;权力范式认为,平台劳动是劳动不平等过程的下一个阶段,强调平台与传统劳动控制之间的连续性,而算法控制是影响劳动过程的关键因素;生态范式则强调平台生态系统的层次性和可塑性。对这种新的经济和组织形式的分析,学者们提出了三种理论视角。尽管这些范式都专注于平台研究的一个重要方面,但它们对单一特征的强调是具有误导性的。平台研究仍然存在理论化、整体化程度不足的问题:首先,过于关注平台对工人的助益或剥削,而忽略了其他分析视角;其次,研究对象过于单一。尽管打车和外卖服务是劳动平台上最大的产业,但它们并不能代表许多其他类型的平台劳动,无法涵盖劳动控制和组织治理的广泛情况。

四、结语和讨论

近年来,中文文献在平台研究方面取得了一定进展,包括经验和理论研究。然而,总体而言,其仍存在一些问题:第一,平台研究的经验对象相对集中;第二,现有的平台研究主要局限于单一学科范围,与其他学科领域缺乏对话。尽管这些研究都具有一定价值,有些研究在原有理论上取得了一定突破,但仍未提供关于平台作为一种组织和经济形式更广泛和完整的视角。因此,中文文献的平台研究需要拓展研究方向,尝试引入新的分析视角。

首先,平台研究需要发展一个文化范式,以分析平台现象并探讨平台组

织中的文化要素,如组织文化、认知和信仰。在与西方理论范式对话时,我们很难忽略文化的影响,如个人与集体的关系、差序格局的行动伦理等。文化作为社会学分析的一个关键视角和工具,引导我们不应只关注算法技术的影响,而是重回组织视角,全面认识平台的组织化和组织文化。此外,文化范式还需要与其他学科范式进行比较研究,探讨不同文化规范以及中国传统文化如何与平台行动者的活动相互作用。同时,我们对算法和平台的设计过程,平台的组织文化如何塑造行动者的认知、信仰和价值观,了解甚少(Irani, 2015)。我们需要了解价值规范如何推动平台的设计和运营,程序员对其工作规范框架的理解是什么,以及他们如何看待平台算法的操纵性、侵略性和严酷性;他们是将平台视为社会选择的价值工具,还是将其视为经济效率的中立技术(Kelkar, 2018)。文化范式的构建还需要探讨文化和话语如何影响平台内部的行动者,对平台企业的组织话语如何能够将劳动者转变为规范化的角色。许多 Uber 司机最初的目标是通过工作赚取特定收入,但随着时间的推移,他们学会了管理自己的日常生活安排以适应平台需求(Schneider, 2018)。在这一过程中,劳动者的时间规划变得更经济理性和市场化,还受到一种道德感的塑造。对于平台投射到他们身上的理性价值观,劳动者会不断进行反思,并在符合自身利益的情况下进行权衡。例如,中国打车平台的分散式管理与 Uber 不同,通过“灵活和自由的工作时间”的话语建构,将日常管理与工作规划转移给劳动者,侵占了出租车司机的个人时间,并重新定义了日常的打车活动。在这种重新定义中,出租车司机的无偿劳动被合理化了(Chen, 2018)。一些平台尝试通过游戏化延长司机的工作时间(Rosenblat & Stark, 2016)。例如,一些劳动者在平台上分享个人的最佳成绩和每日排名,类似于布洛维(2008)描述的“赶工游戏”。然而,对情感劳动的要求却增加了,司机被迫在特定情况下展示适当和正确的情绪。为了符合平台规范,满足服务需求者的期望,劳动者不得不在正常工作之外投入额外的情绪劳动,扮演对雇主、客户友好的形象(Sharone, 2022)。然而,这种情绪劳动的付出

往往被边缘化,失去了合法地位。

其次,平台研究需要将其他领域中的组织者和行动者纳入视野,如探讨国家作为行动主体在推动平台发展中的关键作用。尽管已有研究将国家纳入讨论(黄冬娅、杜楠楠,2022),但多数研究仅将国家视为制度背景,而忽略了国家本身作为一个具有自主性、强大动员能力的组织。国家在塑造制度环境、价值导向以及网络结构中扮演着关键角色,对平台行为和组织生态具有显著影响。特别是在中国语境下,发展型政府在平台企业的兴起与发展中发挥了至关重要的作用。这不仅限于政商关系或国家与市场的讨论,而是将国家和政府看作具有主导性的行动者,同时其受到各种结构性条件的制约,来探讨国家如何与平台以及其他行动者展开互动。目前,平台研究倾向于将平台视为独立运作的组织实体,与其他组织和群体隔离开来,但这是一种误导性的描述。当平台的边界继续扩展,以致将其他领域与不同行动者都吸纳进去时,平台如何根据具体情况组织策略和结构模式?例如,传统产业结构和市场变迁如何影响平台的组织特征和业务策略?研究表明,传统产业的周期性变化对平台劳动力供应有很大影响(Farrell & Greig, 2016)。在低失业率时期,如果传统企业继续降低劳动力的雇佣标准,是否可能为平台提供越来越多的劳动力,这是一个未知的问题。

最后,目前的研究尚未关注平台中不同社会群体的经历差异。有研究表明,在算法的帮助下,平台在司机群体内部塑造了进一步的层级化差异,对某些群体有利而对其他群体不利,引发了利益冲突。因此,司机的反抗不仅表明了他们对平台规则的拒绝,也反映了更复杂的权力斗争景观,劳动者与平台之间的权力关系变得更加模糊(Benjamin, 2019),这反映了数字平台在契约关系构建中的复杂性和挑战。此外,平台会对符合某类社会标准的求职者给予最优待遇,从而促进内部劳动力市场的同质化和制度性排斥(Ajunwa & Greene, 2019)。例如,平台的声誉指标可能导致性别算法歧视,女性在平台中的算法优先级较低,进而降低了她们的收入和就业水平。特别是当声誉的

分类指标在平台劳动中变得越来越重要时,这更是一个需要更多研究的议题。总体而言,本文旨在通过梳理平台研究的范式来探索平台研究的未来方向。同时须注意,作为一种新的组织和经济形式,平台本身仍在不断发展和变化中,因此目前的研究无法对这一重要现象做出过多定论。此外,近年来中文文献中关于平台的研究众多,本文的梳理难免有所遗漏。

参考文献

布洛维,2008,《制造同意:垄断资本主义劳动过程的变迁》,李荣荣译,北京:商务印书馆。

蔡跃洲、顾雨辰,2023,《平台经济的社会福利机制及其效果测算——来自外卖平台商户问卷调查的证据》,《经济研究》第 5 期。

曹光宇、刘畅、周黎安,2022,《大数据征信与平台流量:基于共享单车免押骑行的经验研究》,《世界经济》第 9 期。

陈龙,2020,《"数字控制"下的劳动秩序——外卖骑手的劳动控制研究》,《社会学研究》第 6 期。

陈龙,2022,《两个世界与双重身份——数字经济时代的平台劳动过程与劳动关系》,《社会学研究》第 6 期。

陈全真,2023,《生成式人工智能与平台权力的再中心化》,《东方法学》第 3 期。

范如国,2021,《平台技术赋能、公共博弈与复杂适应性治理》,《中国社会科学》第 12 期。

冯向楠、詹婧,2019,《人工智能时代互联网平台劳动过程研究——以平台外卖骑手为例》,《社会发展研究》第 3 期。

付堉琪,2021,《"系统困不住"的外卖骑手——劳动场景视角下的变通与协作关系研究》,《新视野》第 6 期。

胡慧、任焰,2018,《制造梦想:平台经济下众包生产体制与大众知识劳工的弹性化劳动实践——以网络作家为例》,《开放时代》第 6 期。

胡秋爽,2023,《市场还是等级?——数字劳动平台中的组织治理模式》,《社会学评论》第 6 期。

黄冬娅、杜楠楠,2022,《平台企业政府事务部门专门化与政企关系发展——基于国家制度环境的分析》,《社会学研究》第 6 期。

贾文娟,2023,《劳动的身体,资本的心智——自我资本化与新自由主义晚期劳动者意识形态的形成》,《学术月刊》第 12 期。

江小涓、黄颖轩,2021,《数字时代的市场秩序、市场监管与平台治理》,《经济研究》第 12 期。

科斯,2009,《企业、市场与法律》,陈郁译,上海:上海人民出版社。

李彬、翁慧敏,2022,《社交网络平台中的社会资本积累:一个微信群实地实验》,《世界经济》第 4 期。

李东红、陈昱蓉、周平录,2021,《破解颠覆性技术创新的跨界网络治理路径——基于百度 Apollo 自动驾驶开放平台的案例研究》,《管理世界》第 4 期。

李晶,2023,《Web 3.0 时代通证平台的法律之治》,《东方法学》第 3 期。

李胜蓝、江立华,2020,《新型劳动时间控制与虚假自由——外卖骑手的劳动过程研究》,《社会学研究》第 6 期。

刘诚、王世强、叶光亮,2023,《平台接入、线上声誉与市场竞争格局》,《经济研究》第 3 期。

刘河庆、梁玉成,2023,《透视算法黑箱:数字平台的算法规制与信息推送异质性》,《社会学研究》第 2 期。

刘权,2020,《网络平台的公共性及其实现——以电商平台的法律规制为视角》,《法学研究》第 2 期。

吕鹏、周旅军、范晓光,2022,《平台治理场域与社会学参与》,《社会学研究》第 3 期。

莫怡青、李力行,2022,《零工经济对创业的影响——以外卖平台的兴起为例》,《管理世界》第 2 期。

默多克,2022,《对抗的联结:大流行病与平台资本主义》,张艾晨、杜怡蒙译,《开放时代》第 1 期。

欧阳耀福,2023,《互联网平台化组织模式对企业创新的影响研究》,《经济研究》第 4 期。

曲创、刘重阳,2019,《平台竞争一定能提高信息匹配效率吗?——基于中国搜索引擎市场的分析》,《经济研究》第 8 期。

单勇,2022,《数字平台与犯罪治理转型》,《社会学研究》第 4 期。

邵占鹏、甄志宏,2022,《全视监控下网商价格竞争的形塑机制》,《社会学研究》第 3 期。

孙晋,2021,《数字平台的反垄断监管》,《中国社会科学》第 5 期。

孙瑜晨,2022,《数字平台成瘾性技术的滥用与反垄断监管》,《东方法学》第 6 期。

孙志建,2022,《平台化运作的整体性政府——基于城市运行“一网统管”的个案研究》,《政

治学研究》第5期。

王义中、林溪、孙睿,2022,《金融科技平台公司经济影响研究:风险与收益不对称视角》,《经济研究》第6期。

王勇、刘航、冯骅,2020,《平台市场的公共监管、私人监管与协同监管:一个对比研究》,《经济研究》第3期。

王勇、吕毅韬、唐天泽等,2021,《平台市场的最优分层设计》,《经济研究》第7期。

魏海涛、李国卉,2022,《平台劳动者研究述评:回顾与展望》,《社会学评论》第6期。

吴伟光,2021,《平台组织内网络企业对个人信息保护的信义义务》,《中国法学》第6期。

吴新叶,2011,《电子党务:党内民主的功能平台与利用——一个比较视角的分析》,《政治学研究》第5期。

肖红军、李平,2019,《平台型企业社会责任的生态化治理》,《管理世界》第4期。

谢富胜、吴越、王生升,2019,《平台经济全球化的政治经济学分析》,《中国社会科学》第12期。

邢小强、汤新慧、王珏等,2021,《数字平台履责与共享价值创造——基于字节跳动扶贫的案例研究》,《管理世界》第12期。

徐林枫、张恒宇,2019,《"人气游戏":网络直播行业的薪资制度与劳动控制》,《社会》第4期。

张晨颖,2021,《公共性视角下的互联网平台反垄断规制》,《法学研究》第4期。

张凌寒,2021,《网络平台监管的算法问责制构建》,《东方法学》第3期。

张茂元,2021,《技术红利共享——互联网平台发展的社会基础》,《社会学研究》第5期。

张樹沁、户雅琦,2021,《技术机会获取与控制权交换过程——对网约车平台技术的互构视角分析》,《社会学评论》第4期。

张新宝,2022,《大型互联网平台企业个人信息保护独立监督机构研究》,《东方法学》第4期。

赵磊、韩玥,2021,《跨越企业边界的科层控制——网约车平台的劳动力组织与控制研究》,《社会学研究》第5期。

赵璐,2022,《算法实践的社会建构——以某信息分发平台为例》,《社会学研究》第4期。

周汉华,2023,《论平台经济反垄断与监管的二元分治》,《中国法学》第1期。

周文、韩文龙,2021,《平台经济发展再审视:垄断与数字税新挑战》,《中国社会科学》第

3 期。

周潇,2021,《数字平台、行业重组与群体生计——以公路货运市场车货匹配模式的变迁为例》,《社会学研究》第 5 期。

朱晓红、陈寒松、张腾,2019,《知识经济背景下平台型企业构建过程中的迭代创新模式——基于动态能力视角的双案例研究》,《管理世界》第 3 期。

庄家炽,2019,《资本监管与工人劳动自主性——以快递工人劳动过程为例》,《社会发展研究》第 2 期。

Ajunwa, I. & D. Greene 2019, "Platforms at Work: Automated Hiring Platforms and other New Intermediaries in the Organization of Work." in S. Vallas & A. Kovalainen eds., *Research in the Sociology of Work: Work and Labor in the Digital Age*, Binkley: Emerald.

Benjamin, R. 2019, *Race after Technology: Abolitionist Tools for the New Jim Code*, Cambridge: Polity.

Calo, R. & A. Rosenblat 2017, "The Taking Economy: Uber, Information, and Power." *Columbia Law Review* 117(6).

Chen, J. 2018, "Thrown under the Bus and Outrunning It! The Logic of Didi and Taxi Drivers' Labour and Activism in the On-Demand Cconomy." *New Media Sociology* 20(8).

Culpepper, P. & K. Thelen 2020, "Are We All Amazon Primed? Consumers and the Politics of Platform Power." *Comparative Political Studies* 53(2).

Dube, A., J. Jacobs & S. Naidu et al. 2020, "Monopsony in Online Labor Markets." *American Economic Review: Insights* 2(1).

Einav, L., C. Farronato & J. Levin 2016, "Peer-to-Peer Markets." *Annual Review of Economic* (8).

Farrell, D. & F. Greig 2016, "The Online Platform Economy: Has Growth Peaked?" Working Paper.

Frenken, K. & J. Schor 2017, "Putting the Sharing Economy into Perspective." https://www.iso.org/contents/news/2022/09/foresight-trend-report-putting-t.html#:~:text=Also%20known%20as%20the%20gig%20economy%20or%20peer-to-peer,it%20bypasses%20traditional%20steps%20in%20the%20supply%20chain.

Irani, L. 2015, "Difference and Dependence among Digital Workers: The Case of Amazon

Mechanical Turk.” *South Atlantic Quarterly* 114(1).

Jarrahi, M. & W. Sutherland 2019, “Algorithmic Management and Algorithmic Competencies: Understanding and Appropriating Algorithms in Gig Work.” DOI: 10. 1007/978-3-030-15742-5_55.

Kalleberg A. 2018, *Precarious Lives: Job Insecurity and Well-Being in Rich Democracies*, Cambridge: Polity.

Kelkar, S. 2018, “Engineering A Platform: The Construction of Interfaces, Users, Organizational Roles, and the Division of Labor.” *New Media Sociology* 20(7).

Khan, L. 2017, “Amazon’s Antitrust Paradox.” *Yale Law Journal* 126.

Manriquez, M. 2019, “Work-Games in the Gig-Economy: A Case Study of Uber Drivers in the City of Monterrey, Mexico.” DOI: 10. 1108/S0277-283320190000033010.

Pasquale, F. 2016, “Two Narratives of Platform Capitalism.” *Yale Law Policy Review* 35(1).

Powell, W. 1990, “Neither Market nor Hierarchy: Network Forms of Organization.” *Research Organization Behavior* 12.

Rahman, K. & K. Thelen 2019, “The Rise of the Platfom Business Model and the Transformation of Twenty-First Century Capitalism.” *Politics Sociology* 47(2).

Ravenelle, A. 2019, *Hustle and Gig: Struggling and Surviving in the Sharing Economy*, Berkeley: University of California Press.

Robinson, H. 2017, *Making A Digital Working Class: Uber Drivers in Boston, 2016-2017*, PhD Thesis, Massachusetts Institute of Technology.

Rosenblat A. 2018, *Uberland: How Algorithms Are Re-Writing the Rules of Work*, Berkeley: University of California Press.

Rosenblat, A. & L. Stark 2016, “Algorithmic Labor and Information Asymmetries: A Case Study of Uber’s Drivers.” *International Journal of Communication* 10.

Schneider, N. 2018, “An Internet of Ownership: Democratic Design for the Online Economy.” *Sociological Review Monograph* 66(2).

Scholz, T. 2016, *Uberworked and Underpaid: How Workers Are Disrupting the Digital Economy*, Cambridge: Polity.

Schor, J. 2020, *After the Gig: How the Sharing Economy Got Hijacked and How to Win It*

Back, Berkeley: University of California Press.

Schor, J., W. Attwood-Charles & M. Cansoy et al. 2020, "Dependence and Precarity in the Platform Economy." *Theory and Society* 49.

Shapiro, A. 2018, "Between Autonomy and Control: Strategies of Arbitrage in the 'On-Demand' Economy." *New Media Sociology* 20(8).

Sharone, O. 2022, "Networking While Long-Term Unemployed: Why Long-Term Unemployment Will Likely Persist Long after the COVID-19 Pandemic Recedes." https://www.academia.edu/100711502/Networking_While_Long_Term_Unemployed_Why_Long_Term_Unemployment_Will_Likely_Persist_Long_after_the_COVID_19_Pandemic_Recedes.

Sundararajan, A. 2016, *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*, Cambridge, MA.: MIT Press.

Thelen, K. 2018, "Regulating Uber: The Politics of the Platform Economy in Europe and the United States." *Perspective Politics* 16(4).

Vallas S. 2019, "Platform Capitalism: What's at Stake for Workers?" *New Labor Forum* 28(1).

Vallas, S. & J. Schor 2020, "What Do Platforms Do? Understanding the Gig Economy." *Annual Review of Sociology* 46(1).

van Doorn, N. 2017, "Platform Labor: On the Gendered and Racialized Exploitation of Low-Income Service Work in the 'On-Demand' Economy." *Information Communication Sociology* 20(6).

Wood, A., M. Graham & V. Lehdonvirta et al. 2019, "Networked but Commodified: The (Dis)Embeddedness of Digital Labour in the Gig Economy." *Sociology* 53(5).

Zanoni, P. 2019, "Labor Market Inclusion Through Predatory Capitalism? The 'Sharing Economy', Diversity, and the Crisis of Social Reproduction in the Belgian Coordinated Market Economy." in S. Vallas & A. Kovalainen eds., *Research in the Sociology of Work: Work and Labor in the Digital Age*, Binkley: Emerald.

气候变化与人工智能政治路径*

——人工智能和人类能动性下的技治主义民主困境

马克·科克伯格　亨里克·塞特拉　著**

李　帅　李　芳　译***

摘要：人们普遍认为，人工智能对政治和民主产生了巨大影响。同时，该技术常被视为解决关键社会和环境问题的方案。人工智能也引发了诸多问题，如人类应该如何应对气候变化。本文将这些议题置于技治主义民主困境背景中，利用关键节点概念进行讨论，提出两条人工智能政治路径，并对其基本假设展开批判性讨论。这为进一步探讨人工智能、气候变化和民主之间的关系提供了有益框架，且有助于研究技术政治的重要问题，如人类专业知识相对于人工智能的作用、技术解决主义引发的问题，以及应在何种水平上治理人工智能和气候变化。

关键词：人工智能　民主　技治主义　气候变化　关键节点　技术政治

一、引言

作为公民，人们往往认为社会秩序和发展主要由个体行动或最高权威者的

---

* 本文原题为“Climate Change and the Political Pathways of AI: The Technocracy-Democracy Dilemma in Light of Artificial Intelligence and Human Agency”，原文见 https://www.sciencedirect.com/science/article/pii/S0160791X23002117? via%3Dihub，经授权翻译发表。本文系云南省“兴滇英才支持计划”青年人才专项项目（项目批准号：XDYC-QNRC2022-0642）的阶段性研究成果。

** 马克·科克伯格（M. Coeckelbergh），奥地利维也纳大学哲学系。亨里克·塞特拉（H. Sætra），挪威奥斯特福德大学。

*** 李帅，云南大学马克思主义学院。李芳，云南大学马克思主义学院。

决策所决定。诚然,这样的观点有其合理性,但本文认为,社会发展往往是沿着技术所预设的某种路径发展的。技术属于政治的说法已经屡见不鲜,技术塑造着社会、经济和政治的发展(Winner, 1977; Winner, 1980: 121-136; Ellul, 1964)。历史上,某些时间点可被定性为关键节点(critical junctures),在该时期,人类的能动性是纠正、引导或塑造历史发展长波的关键(Jonsson, 2015; Jonsson, 2016: 131-151; Collier & Collier, 1991a: 27-39)。目前,人类正处于这样一个阶段,数据的可获得性和人工智能的发展达到了前所未有的水平,挑战了人们对不同政治制度利弊的现有认知。结合当前与环境可持续性,特别是气候变化相关的挑战来看,该问题尤为突出,其中不乏严厉警告(IPCC, 2022)——必须进行重大变革。然而,政治家面对这些问题时的挫败感却与日俱增。生态法西斯主义(eco-fascist)解决方案的呼声(Satgar, 2021),如"反抗灭绝"(Extinction Rebellion)等试图破坏的社会运动,以及年轻人为气候而战[特别是"未来星期五"(Fridays for Future)],都凸显了气候变化引发的社会和政治紧张局势。与此同时,气温升高和自然灾害日益频繁的现实情形凸显了造成这些紧张局势的真实且紧迫的原因。众多要求变革的呼吁都有一个共同点:要求关注人类的能动性和责任,但这没有强调技术在导致人类面临的挑战中起到的决定性作用,也没有强调技术为人类当前困境提出某些解决方案的核心作用。诚然,人类在技术中始终扮演着重要角色(作为用户、开发者、维护者、监管者等),而且正如我们从其他新兴技术的引入中所了解的那样,一项新技术或新兴技术的成功取决于人类系统(Maynard & Dudley, 2023: 1118-1120),即取决于人类—技术系统或社会—技术系统,但技术本身并不是政治中立的(Coeckelbergh, 2022a),往往具有超人类预期和控制的普遍社会效应。

本文将重点讨论如何辨别人工智能的政治倾向,以及如何理解和看待备选的政治路径。为此,本文将探讨"民主"和"技治主义"(technocracy)概念,特别是在应对气候变化挑战的背景下,两者的关系十分紧张。首先,人工智

能在某些方面具有改善民主政治的潜力,如改进意见形成、协商、共识形成、决策和参与等(Savage, Chiarini & Evants, 2018: 369-380; Bakker, Chadwick & Sheahann et al., 2022; Koster, Balaguer & Tacchetti et al., 2022: 1398-1407)。迄今为止,人类在应对气候变化的决策过程中已展现出明显的不足,且常陷入困境。为此,有观点提出,人工智能可以提供改善这些决策过程的有效途径(Sætra, 2020)。其次,人工智能越来越被认为是一种促使民主过时的技术(Sætra, 2020)。因此,人工智能在很大程度上可以替代人类的政治决策。这些论调均基于该观点,即人工智能不依赖于民主决策,因此人工智能可以改善政治决策。例如,在应对气候变化的问题上,通过增加监测和监视,创建社会及其居民的数字孪生体,并利用无与伦比的处理能力和数据分析能力,我们可以在理论上改善社会的各个方面,并有效地应对气候变化。但代价是什么?本文旨在结合气候变化和相关的全球环境挑战,讨论在人工智能领域已知但研究不足的政治倾向。通过阐明目前出现的两条路径,本文探讨了每一条路径的可能后果,同时论述了民主与技治主义之间的紧张关系。

首先,本文将在技术和科学哲学之中进行分析,将科学技术与社会(STS)理论同政治和经济发展理论相结合,以此来说明人工智能具有政治性的原因。同时,运用社会和政治理论来理解人类能动性、技术和其他历史制度力量之间的相互作用。其次,探索人工智能的两条重要政治路径:探讨人工智能如何发展出新的可能性,并创造场景;在这种场景中,人工智能要么加强和改善民主,要么助长技治主义,同时预测这两种情况的后果。最后,讨论人类处于关键节点所面临选择的性质。本文认为上述两条路径各有优点,但讨论我们是否应该接受政治发展的路径依赖,即我们是否应该接受在这些情况下做出的假设,也是至关重要的。虽然人工智能作为主要政治决定因素预示着某种未来,但我们不能忽视潜在的替代性未来。在这种未来中,人类并非只能选择由技术决定的道路,也可以从对当前社会与技术的关系以及人类如何在当前全球形势下继续前进的基础分析着手。

二、技术与社会变革

在思想史上,技术在社会中的作用时常被低估。直到近代,情况才有所改变,人们注意到了新(工业)技术如何重塑社会,如何对人们的日常生活产生深远影响。因此,19 世纪和 20 世纪上半叶的社会理论倾向于以一种确定的方式概念化技术与社会之间的关系。依据技术决定论(technological determinism)的观点,历史发展必然受技术革新的驱动,而技术革新又决定着社会结构和文化价值(Heilbroner, 1994: 67-78)。马克思认为,生产关系(以及生产技术)决定社会内部的社会关系。在该观点中,技术具有社会性和政治性,技术决定了社会的基本结构。技术决定论的当代变体即超人类主义(transhumanist)观点,该观点认为技术奇点或其他由技术决定的未来将不可避免地发生,如科兹韦尔(Kurzweil, 2015: 146-170)或博斯特罗姆(Bostrøm, 2014)的观点。

技术工具论(echnological instrumentalism)的观点与之相反,它认为技术只是一种工具,而且技术在政治上也是中立的(Schatzberg, 2018)。与温纳(Winner, 1979: 75-86)所说的"技术正统"(technical orthodoxy)有关,其涉及人类对后果的理解和控制,以及技术的中立性。技术决定论和技术工具论都受到了技术哲学与媒介哲学理论家和技术研究理论家的严厉批评。前者强调了技术的非工具性和非预期效应,早期有海德格尔(M. Heidegger)和麦克卢汉(M. McLuhan),今有伊德(D. Ihde)和芬伯格(A. Feenberg)。后者认为,技术与社会相互交织、共同发展。例如,兰登·温纳(Winner, 1980: 121-136)在其关于技术哲学与技术研究之关系的著作中,提出了一个著名的观点:技术是政治性的,但这不能简单地从技术决定论或社会决定论的角度来理解,在特定的文物历史中,技术与社会相互影响,但也可能相互独立。

经济和社会变革理论持续关注历史上的关键转折点。例如,熊彼特

(Schumpeter, 2013)的经济发展概念,即通过引入新产品、引入新生产方法、开拓新市场、征服新原料来源或建立新的产业组织,在经济流动中进行创造性破坏,从而实现非连续性变革。在他看来,这些都是由企业家引起的。随着这种变革的推进,原有的经济流动被打断,从而催生新的发展与创新。此外,还不能忽视库恩(T. Kuhn)的范式转移(paradigm shifts),即其在科学史和社会学中引入了不连续变化概念(David, 1985: 332-337; Fioretos, Falleti & Sheingate, 2016)。

关键节点理论(critical juncture theory)认为,存在某些可以改变一个实体(如一个物种、一个社会)演化过程的关键转折点。关键节点可以被定义为:"一个重大变革时期,通常以不同的方式发生在不同的国家(或在其他分析单元中),并假设会形成不同的结果。"(Collier & Collier, 1991b)在这些时期,未来的轨迹已经形成,并在一定程度上确定下来,这就诞生了源自经济学的"路径依赖"(path dependence)概念。

历史制度主义(historical institutionalism)持续关注关键节点的能动性作用,因为在这个相对较短的时期内改变发展轨迹、塑造社会未来的空间相对较大。例如,强有力的制度约束在一定程度上会形成路径依赖,因此,把握人类能动性如何形成新的路径是十分重要的(Capoccia & Kelemen, 2007: 341-369)。例如,琼森(Jonsson, 2016)在其政治经济学著作中使用了"关键节点"的概念。他认为,冰岛引入新自由主义(neoliberalism)就是一个关键节点,而政治反对派却未能以一种能够促成政权更迭的方式介入。他还指出,真正的关键节点是可以被创造出来的,因此在面对各种制度力量的制约时,关键节点进一步突出了政治和人类的能动性。

三、双重关键时刻:气候变化与人工智能

环境可持续性被广泛认为是全社会和全人类必须解决的重要问题,气候

变化则被视为当代最大的挑战。科学家警告说,气温正在上升,极端天气事件愈加频繁;冰川正在融化,全球海平面正在上升。为了解决这些问题,各级机构必须采取政治行动和决策,减少与人类活动相关的温室气体排放。在该领域,有人声称已经到达了关键节点,这通常与各种临界点(tipping points)有关(在临界点之后将会失去控制)(Lenton, 2013: 1-29)。如果我们现在不充分减缓气候变化,局面将会失控,必将出现大规模全球变暖及其他影响。甚至有人声称,我们已经越过了临界点。

"关键节点"的概念在政治体制分析中占据着重要地位,同样地,这一概念在应对气候挑战时亦显得尤为关键。国家制度,即将人类社会组织成主权国家的结构,以及既有的国际合作框架,都在很大程度上影响甚至阻碍着应对气候变化的策略发展(Sætra, 2022: 66; Coeckelbergh, 2021a)。本文将论证,人类目前面临的气候变化挑战就是一个关键节点,这将挑战已有的政治路径和制度遗产,从而保护自然环境。气候变化已是政治争议的根源之一,如果不能在全球范围内减缓和适应气候变化,预计会导致更多的社会和政治动荡,增大本国暴力冲突的风险(IPCC, 2022);如果成功应对这一挑战,预计将对区域性和全球性合作产生深远影响,并对民主在威权主义(authoritarianism)日益突出的世界中的可行性产生重大影响。

与此同时,科技发展领域可能正迎来另一个关键节点——人工智能和数据科学取得了重大突破,特别是机器学习的广泛应用和数据可用性的显著提高。人工智能正在迅速普及,并广泛应用于商业和政府各个领域。人类或许正迈过一个临界点,这与科林格里奇困境(Collingridge's dilemma)有关(Collingridge, 1980),新技术在初期阶段易于监管,但其潜在的深远影响却难以预测和认知,因此监管的理由显得模糊不清。然而,一旦技术成熟,其产生的后果将变得清晰无疑,但此时再试图对其进行规制,往往事倍功半,甚至难以奏效。这表明,技术在临界点之后将会成熟,但在这个临界点之后,未来的发展道路将被固化,技术的政治含义也将难以避免。

因此,人类或许正处于一个双重关键时刻。此刻,自然环境的保护路径与技术发展的方向均呈现出相对开放的状态,而且这两条路径都会对我们正在塑造和选择的政治道路产生重大影响。今天,关于人工智能可持续能力的讨论越来越多,针对人工智能促进可持续发展的解决方案比比皆是(Galaz, Centeno & Callahan et al., 2021)。双重关键时刻的主要驱动因素可能与正在进行的、与数字化转型中绿色转型必要性相关的政策讨论有关(Floridi, 2019: 183-221),讨论将"绿色"(环境政策和绿色经济)和"蓝色"(数字政策和数字服务经济)结合的政治,两者之间的互动和关系日益成为政策制定者,尤其是欧盟内部政策制定者的关注焦点。欧盟委员会已经将"双轨转型"(twin transition)概念作为其战略规划和分析的关键内容(Makitie, Hanson & Damman et al., 2023),明奇等人(Muench, Stoermer & Jensen et al., 2022)将这两种转型描述为"政治优先事项",两者既相互促进,又存在紧张关系。例如,数字技术可以促进创新和增长,但同时也会提高环境成本。

人工智能与气候变化之间的关系,无疑是绿色转型与数字转型交织的关键节点。在全球性议题如气候变化治理中,人工智能的应用显得尤为重要。然而,技术发展的双刃剑效应也尤为显著:一方面,人工智能的运作本身,如模型运行和数据服务器维护,均需要消耗大量能源,进而产生温室气体排放,这无疑加剧了气候变化的严峻性(OECD, 2022);另一方面,人工智能亦展现出在减缓气候变化方面的巨大潜力。其不仅能够协助处理温度变化和碳排放数据,预测极端天气事件和能源使用情况,还能在改造运输系统、监测海洋乃至辅助地球工程等方面发挥关键作用。尽管其中某些应用,如直接空气碳捕获技术(direct air capture, DAC),在学界和政策层面仍存争议(Arcusa, Lackner & Page et al., 2022),但其潜在的积极影响不容忽视。值得一提的是,人工智能在引导人类社会行为方面的作用。在营销和广告领域,人工智能已经展现出强大的影响力,能够在潜移默化中引导人类行为。或许可以利用人工智能来引导人类社会朝着更加气候友好的方向发展(Cockelbergh,

2021b：67-72；Mills & Whittle，2023），同时还可以提高人类的决策能力，这似乎很有诱惑力（Sætra，2021）。但是，这也存在一定的危险和弊端，在讨论不同的选择和方案后就会明白这一点。

人工智能不仅影响自然科学，还影响人文科学和政治等活动。例如，现在可以用大语言模型（LLMs）创建文本，用人工智能辅助政治决策。随着ChatGPT及其他生成式语言模型的兴起，人们担心人工智能在游说方面的潜在作用会“劫持民主”（hijacks democracy）。“生成式人工智能”是一个更宽泛的术语，不仅包括以文本形式，还包括以图像、音频、视频等形式生成内容的人工智能系统，影响着个人（微观层面）、群体（中观层面）和社会（宏观层面）。塞特拉（Sætra，2023a）将生成式人工智能在上述各个层面会令人们产生担忧的一些原因绘制成图1。

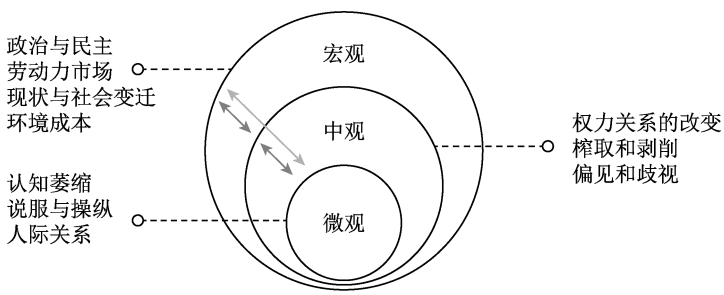


图1　生成式人工智能的潜在危害

塞特拉指出，生成式人工智能系统存在对政治和民主进程产生深远影响的潜在问题。具体而言，这类系统能够生成大量内容，包括虚假新闻和深度伪造的内容，改变公众舆论和认知。此外，这些系统还可能被用于操纵特定目标，从而改变原有的权力关系。塞特拉进一步指出，这类系统在开发和部署阶段均需要消耗大量的能源和设备资源，无疑增加了环境成本（Dodge，Prewitt & Combes et al.，2022：1877-1894；OECD，2022）。

不过，人工智能也具有积极的政治潜力。例如，一些研究人员正在研究如何使用大语言模型促进信仰和观点不同的人达成一致。通过这些应用，我

们看到人工智能有助于促进协商和达成共识,而这也正是民主支持者关注的关键领域。

我们之所以将人工智能的积极方面和消极方面都描述为“潜力”,是因为我们想挑战人类无法控制自己未来轨迹的决定论立场。因此,本文强调人类的能动性以及采取有意义行动的必要性,以确保在当前关键节点上引导未来朝着正确的方向前进。聚焦于人工智能,本文提出在政治活动中使用人工智能时,至少要区分两种选择或路径,我们称之为“人工智能增强的民主”(AI-augmented democracy)和“人工智能驱动的技治主义”(AI-driven technocracy)。

四、人工智能增强的民主

第一条可能性路径是在不取代人类参与决策的前提下,使用人工智能增强民主——人工智能增强的民主。气候变化已被证明是对传统民主决策进程的严峻挑战,一些人认为,加强民主国家的协商进程是采取有效气候行动的关键(Willis, Curato & Smith, 2022)。理论上,人工智能可以帮助大型社区之间更有效地进行沟通、会面甚至协商,这对更好地理解气候变化带来的挑战和潜在的解决方案至关重要。

技术研究人员和政治理论家提出了各种概念和方法。例如,流动民主是直接民主的形式之一,它允许选民委托投票,如可以通过数字平台和算法来实现(Kahng, Macknenzie & Procaccia et al., 2021:1223-1252)。另一个例子是 DeepMind 公司的研究,该研究提出利用深度强化学习,寻找人们会为科斯特等人(Koster, Balaguer & Tacchetti et al., 2022:1398-1407)投票的经济政策。在一场投资博弈中,人工智能发现了一种优化人类偏好从而赢得多数投票的机制。此外,在政治理论方面,也存在利用人工智能促进民主的工作。例如,有人提出使用语言模型来寻找不同偏好民众间的共识(Bakker, Chadwick & Sheahan et al., 2022)。兰德摩尔(Landemore, 2023)认为,人工智能

可以通过多种方式提高协商质量,从而实现大众协商(微观公共领域)。在微观公共领域,随机挑选民众参与协商。人工智能还可以作为促进者(如计时器)、翻译者、事实审查者、论据组织者、交流跟踪者、重叠和质量的衡量者以及团队立场可视化者的角色,帮助改善协商工作。其他人还注意到,协商民主原则可以促进负责任的创新(responsible innovation)和控制人工智能(Buhmann & Fieseler, 2021)。

在这些案例中,人工智能都是在高度政治化背景下使用的,而且人类在民主决策过程中仍保留了核心作用。然而,自柏拉图和培根以来,持续存在着专制主义(autocratic),特别是技术官僚主义(technocratic)的诱惑:人类,包括专家(柏拉图)都经常出错,为什么还让人们做决定,而不是使用科学技术引导社会朝着更加美好的方向发展,创造一个勇敢的新世界呢(培根)?

五、人工智能驱动的技治主义

不应轻视或低估技治主义的诱惑。依靠专家和专业知识让人工智能做出政治决策似乎是很好的理由。理论上讲,人工智能可以更有效地确定应对气候变化的有效缓和战略与适应战略。基于此,一些人可能会认为,如果人类赋予人工智能系统绕过阻碍有效行动的烦琐民主程序的权力,人类自己就可以更有效地实现减缓和适应气候变化的相关目标,即第二条可能性路径——人工智能驱动的技治主义。

塞特拉(Sætra, 2020)在其《对人工智能技治主义的略微辩护》("A Shallow Defence for ATechnocracy of AI")一文中提出,我们应当认真审视并考虑人工智能技治主义的观点。他认为,在政治决策中更加积极地利用人工智能具有充分的合理性,主要基于以下三个前提:第一,政策评价应以该社会的基本道德价值为基础,确定基本道德价值是政治的首要目的;第二,根据第一个前提得出的最佳政策应该予以实施;第三,在某些涉及科学、工程和复杂社会

及宏观经济问题的领域,人工智能比人类更善于发现和制定最佳政策。本文提到的这些论点虽为假设性的,但其主要目的并非直接倡导技治主义,而是检验我们是否拥有充分的论据来反驳这一观点。

政治控制论也能提出类似的论点。1948 年诺伯特·维纳(Wiener, 1948)举办了控制论研讨会后,许多政治学家逐渐倾向于将政治制度理解为输入和输出的流动,并将这些输出作为反馈重新输入,著名的例子包括卡尔·多伊奇(Deutsch, 1963)的《政府的神经》和戴维·伊斯顿(Easton, 1965)的《政治生活的系统分析》。受当时先进计算机系统和仿真技术所带来的可能影响,20 世纪 60 年代是一个对计算社会科学潜力和更全面了解与控制政治制度非常乐观的时期。然而,20 世纪 60 年代末,计算政治学方法的局限性日益凸显。在这一背景下,温纳(Winner, 1969: 1-17)对多伊奇的作品做出了评论,这种评论在某种程度上可被视为对当时政治学领域中一场热潮开始走向衰退的悼词。

虽然政治控制论最终失去了支持,但随着大数据和人工智能的发展,有人开始预言政治控制论将再次兴起,并有望实现 20 世纪 60 年代政治学家尚未达成的雄心壮志。由于海量数据被用以控制多种形式的社会制度,无论是否存在有价值的人类控制,多样化的控制论与人工智能技治主义都紧密联系(Zuboff, 2019)。

在应对气候变化问题方面,技治主义十分具有吸引力,因为普通民众可能不具备应对气候变化必需的专业知识,即使是人类科学家也不可能掌握一切知识。但人工智能凭借其对大数据的深入挖掘和分析能力,可做出更为精准的预测,从而推动制定更为科学合理的决策。然而,技治主义也并非完美无缺的(Sadowski & Selinger, 2014: 161-168)。塞特拉(Sætra, 2020)以五种反对意见总结了技治主义所面临的主要问题:第一,民众需要充分的政治参与才能得到满足;第二,没有民众参与的政府被认为是不合法的;第三,计算机无权做出影响人们生活和福祉的决定;第四,人工智能是不透明的,不可能

完全服从于人类控制;第五,政治决策后果问责必须要明确,而当涉及人工智能决策时,问责将不再明确。

这些似乎都是驳斥人工智能技治主义的有力证据。然而,塞特拉(Sætra,2020)认为,当我们更仔细地审视政治制度时会发现,现存的政治制度(如非理想化和非假想的自由民主形式)也因许多类似的反对意见被摒弃,这给我们带来了一个实际问题,即必须在两个非理想的解决方案之间做出选择(Gaus, 2016)。因此,重要的不是比较技治主义的悲观描述(现实主义)与民主的乐观描述(理想主义),因为现实政治制度总是非理想政治制度的典范。因此,真正的问题是:对人工智能技治主义和人工智能增强民主的最终评估是否会产生更有利的结果,尽管它们都存在负面因素?

然而,即使这些系统在实际应用中取得了成功,也有人可能会反驳,人工智能的技术官僚化运用将威胁到启蒙理想和人类自主权价值。技术官僚将会从民众手中夺走决策权,甚至隐蔽性地操控民众,从而忽视了人民对自我决定生活、社区以及社会事务的强烈愿望。康德式的自主思考也将无法得以实践:人工智能不具备思考能力,那么人类的自主权将得不到保障。然而,自由主义的理论家如弗里德里希·哈耶克(Hayek, 1960)提出了相反的观点,他强调经济自由对有意义的个人生活至关重要,民主在这方面也许是次要的。

另一种反对观点,即一种类似于技术决定论的观点,至少在规范形式上认为技术应该塑造社会。这一观点有时也使用“技术解决主义”这一术语(Sætra, 2023b),即认为人类面临的问题本质上都是技术性的,技术可以解决所有问题,包括政治问题。但本文认为,这是不可能的。

六、讨论:选择哪条路径

第一条路径的优势在于,坚持人类决策的主导地位,而将人工智能作为

辅助决策过程的工具。然而,当涉及大众协商时,决策过程就会存在挑战,但仍然符合协商民主(deliberative democracy)理念和总体民主(democracy in general)原则。人工智能是一种技术性辅助工具,是民主决策的支撑。因此,该路径也为将人工智能与政治科学家的各项增强民主政治决策的提案(如微观公共领域)的融合提供了可能性。该条路径也更有可能获得民主支持,因为人类仍然处于主要地位,人类不会产生人工智能正在取代自身的感觉,从而消除了由科幻小说所引发的普遍恐惧。

第二条路径的优势在于,通过人工智能保证了专业知识的参与。相较于人类,人工智能处理数据的能力更强,能够发现人类可能忽略的模式,人工智能在特定领域极有可能设计出更佳政策。这对于应对气候变化和需要在海量数据中辨别模式、在全球范围内进行更具挑战性协调的复杂社会和全球问题具有重要意义。将政治权力赋予人工智能,而非仅仅将其视作工具,这意味着人类能够更加有效地应对这些问题。

这两条路径都存在严重缺陷。即使在第一条路径中,人工智能也可能会影响民众的决策,而不是仅仅作为一种工具。正如科克伯格(Coeckelbergh, 2021b)在讨论使用人工智能缓解气候变化时所指出的,由于人工智能的预期或非预期效应,它可能会操纵民众,与人类的自由和自主权相违背。至少存在三方面的风险。首先,人工智能中介协商可能会干扰民众的思维,影响决策结果。例如,科克伯格(Coeckelbergh, 2022b)提出,人工智能可能影响信仰的形成和修正,包括政治信仰。剑桥分析(Cambridge Analytica)公司从数百万个脸书账户资料中收集数据进行分析,影响选民,这就表明了通过社交媒体,人工智能可以影响选举结果。其次,就目前而言,人工智能已经实现了海量推理、组织讨论和做出决策;经过人工智能“优化”或“增强”的协商可能更加肤浅,或无法提供人们从“真实”或“全面”协商中获得的学习结果——在这种协商中,人们不得不从头开始讨论,思考新方案。因此,经人工智能调解、促进和简化,协商的性质可能会发生质的差异。最后,人工智能辅助的决

策过程缺乏透明度,可能被技术系统控制者刻意操纵。

从第二条路径来看,没有理由将人工智能的能动性局限于解决特定问题(如改善气候变化的决策)。人工智能能做的远不止于此,它可以优化人类生活的方方面面,甚至可以被赋予充分的政治权力。那为什么还要限制人工智能的力量和影响呢?当人类和人工智能意见不统一时,问题就会出现。例如,假设人工智能决定剥夺某些人的自由来实现某个目标(如应对气候变化),或者人工智能可以在没有透明法律程序的情况下决定人们的生活和自由,那么政治自由和法治等原则将会受到威胁。鉴于目前人工智能在发挥作用时存在偏见和歧视问题,第二条路径要如何通向公正与平等仍然值得商榷。基本的原则如自由、公正和平等有充分的保障吗?如果人工智能认为人类不需要这些呢?因此,这条路径必然会带来专制化,甚至可能是极权主义(totalitarian)政治制度的真正风险。

第一条路径虽有其优势,但也存在一个不容忽视的内在缺陷:原则上协商民主可以跨越国界实施,但目前的建议更多地聚焦于国家层面的决策优化,而忽视了人工智能在全球层面辅助决策的巨大潜力。特别是在应对气候变化等全球议题时,这种全球性的辅助决策能力显得尤为重要。例如,人工智能辅助下的微观公共领域的民主工具和方法,在全球范围内是否具备可行性和有效性,这是一个值得深入探究的问题。相比之下,第二条路径试图解决这一全球性问题,但又引入了新的挑战:可能导致一种被人工智能主导的全权性体制(a totalitarian system),这种体制不仅威胁到自由和民主,甚至可能导致人类的终结。一旦人工智能的目标与人类产生冲突,其结果将是难以预料的。例如,如果人工智能认为种族灭绝或生态灭绝是实现某个或某些人类目标的最有效方式,那么人类将面临前所未有的危机。有人提出,正如DeepMind 等多数大型科技公司所做的那般,人工智能应该与人类的目标保持一致,这一观点在表面上看似合理。然而,当人工智能的决策逻辑与人类产生分歧时,如何确保这种一致性将成为一个难题。同样地,虽然有人主张

人类必须控制人工智能,但这又可能牺牲人工智能在决策过程中的优势。

上述讨论深刻揭示了人工智能与民主之间的两难选择困境。诚然,两条路径各自存在不容忽视的缺陷,但这并不意味着人类必须在二者之间做出非此即彼的选择。幸运的是,我们不需要接受任何一条路径,也不需要把这两种困境视为既定事实。

让我们阐明这两条路径背后的一些(有问题的)假设。

第一,这两条路径都认为技术可以解决所有问题(如气候变化)。这些观点都建立在一种对技术修复能力过度夸大且缺乏坚实依据的信任之上,即技术解决主义。也有其他的观点和方式来界定人工智能技治主义问题,特别是有人指出,人类不可能完全依赖人工智能,我们还需要人类的专业知识。那么,核心问题就不仅仅是人工智能可以和应该为民主做什么,而是在民主中人工智能专业知识与人类专业知识的关系应该是什么。这些问题十分复杂,我们需要承认仅仅依靠技术无法解决这些问题,我们始终需要人类及其专业知识。

第二,这两条路径都倾向于以技术决定论为基础。两者都认为,在民主中使用人工智能是不可避免的,我们必须在不同的人工智能应用方式之间进行选择。但在民主中并不一定要使用人工智能,我们还可以做出其他的选择。此外,人工智能对社会和民主的影响很难预测,而且可能根本不是直接性的,因为技术总会产生意料之外的后果——积极的或消极的。

第三,虽然本文的讨论聚焦在预期效果上,但我们也可以从技术哲学中了解到,人工智能除了会产生有利的预期效果之外,还不可避免地会产生意料之外的副作用。再次谈及有关控制的观点,以及为了实现(人类共识的)目标,人工智能选择的手段是否会导致极权主义的反乌托邦,甚至是人类的终结。这意味着,这两条路径可能都无法实现自己的初衷,尤其是当目标涉及维护民主时。

第四,政治与伦理的关系问题。例如,弗洛里迪(Floridi, 2019)在创作

《绿色与蓝色》时就认为,伦理优先于政治。基于伦理,人们可以拒绝这两条路径,哪怕它们有助于民主或解决地球问题。这就又回到了人类自主权问题上,这不仅可以看作对政治原则的呼吁,还可以看作一种伦理价值。然而,还存在其他价值。

第五,与国家制度相关的挑战以及全球合作或治理的必要性(Bull, 1977)。本文已经指出,第一条路径的设想仅是局限于国家层面拟定的;而理论上,第二条路径也可能推进一个世界国家的建立。两种方式各自存在非常严重的问题。因为气候变化是一个全球性议题,稳定的气候是一项全球公共产品(Barrett, 2007),所以我们需要在全球范围内解决该问题。我们可以设想几种实现气候变化全球治理的方式,每种方法都各有利弊。迄今为止采取的最明显的解决方案,即利益相互竞争的主权国家团结起来,协商并达成共识,解决共同的问题。该解决方案尊重各个国家和民族的主权,其优势在于现实性以及对利益冲突和国际无政府状态的认知。在无政府体系中描述各个国家之间相互作用的另一种方式是,将国家看作国际社会的行动者。该观点以国际关系中的英国学派为代表(Bull, 1977),“国家社会”(a society of states)的概念强调了非正式关系、规范和共同价值观如何极大地影响并隐蔽性地为全球范围更密切、更有效的合作与治理奠定基础。然而,正如批评者指出的,一旦权威行动者发现维持这种社会不再符合他们的利益,这种全球治理方式就很容易崩溃。最后,有人可能会提出一种解决方案,即由一个超级国家实体取代国际合作,该实体拥有强制执行全球解决方案的权威和权力(Sætra, 2022)。这种解决方案将会招致强烈的伦理反对和实践异议。许多人既反对抛弃主权国家和以民族国家为基础的国家制度的可取性,又反对在全球层面实施有效政治制度的实际可行性。人工智能和气候变化对国家和全球治理结构提出了挑战,我们应持开放态度,考虑在各个层面采取不同的治理形式。

第六,即使人工智能能够和人类目标保持一致(Koster, Balaguer & Tac-

chetti et al. , 2022: 1398-1407),也不应该将人类的目标视为理所当然。因为人类的目标可能本身就存在谬误。以当前的消费社会为例,人类普遍依赖并为之贡献,但这种模式对人类和地球环境均造成了严重损害。如果我们任由人工智能按其自身的发展轨迹前行,即使人类很幸运,人工智能可以在符合人类目标的前提下推进人类目标;但是,也存在固有问题,即一些人类目标本身就存在严重的问题。例如,最大限度地刺激消费不可能帮助应对气候变化,高消费会造成能源浪费和环境污染,更不用说自然资源的耗竭,这会以威胁生态系统的方式破坏环境,最终影响人类。人类活动已经过度,减少人类活动和人口数量既符合非人类生命和人类生命的繁荣,又是必要的(Næss, 1989)。如果管理政治的人工智能系统就人类的问题最终得出了与我们类似的结论,并着手实施必要的政策,减少人类活动,其可能带来的深远影响将是人类难以预料的。

七、结论

在探讨人工智能与政治的关系,特别是其与民主议题的联系时,需将其置于技术与社会变革的宏观背景之下,并考虑其与气候变化等全球性挑战的政策交汇点。在此背景下,本文勾勒出了人工智能的两条政治路径:一是运用人工智能促进政治与民主决策的优化,二是借助人工智能解决气候变化等全球性挑战。然而,这两条路径各有其利弊,需审慎权衡。人工智能可以且应该用于增强人类决策的民主性,但正如本文所述,人工智能也存在局限性。最终,人工智能可能会产生不利于实现人类目标的副作用,或在政治上,或在道德上,产生不可接受的其他影响。与此相反,让人工智能来做决定(不受人类的控制和影响)则危险得多,因为它从一开始就威胁着人类的基本价值。

本文深入剖析了导致技术政治倾向困境的原因,这与当前的政治理论、

政策制定及科技界就人工智能与民主关系的讨论紧密相关。同时,这些讨论引出了关于人工智能与政治关系的核心议题,甚至更为深层次的根本性议题,如人工智能专业知识与人类专业知识之间的互补与平衡、技术解决主义与决定论的边界问题、错误的人类目标与人工智能非预期性副作用的潜在冲突、伦理考量与政治决策的融合,以及气候变化等全球性挑战如何在多层次得以妥善应对与治理。如果我们想在政治和伦理上负责任地使用人工智能,特别是在民主政治中使用人工智能,就需要在这些领域开展更多的工作。

参考文献

Arcusa, S., K. Lackner & R. Page et al. 2022, "The Socio-Economic and Technical Context for Direct Air Capture in Decarbonization Pathways." https://keep.lib.asu.edu/items/170566/view.

Bakker, M., M. Chadwick & H. Sheahan et al. 2022, "Fine-Tuning Language Models to Find Agreement among Humans with Diverse Preferences." https://arxiv.org/abs/2211.15006.

Barrett, S. 2007, *Why Cooperate? The Incentive to Supply Global Public Goods*, Oxford: Oxford University Press.

Bostrøm, N. 2014, *Super Intelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press.

Buhmann, A. & C. Fieseler 2021, "Towards A Deliberative Framework for Responsible Innovation in Artificial Intelligence." *Technology in Society* 64.

Bull, H. 1977, *The Anarchical Society: A Study of Order in World Politics*, New York: Columbia University Press.

Capoccia, G. & R. Kelemen 2007, "The Study of Critical Junctures: Theory, Narrative, and Counterfactuals in Historical Institutionalism." *World Politics* 59(3).

Coeckelbergh, M. 2021a, "AI for Climate: Freedom, Justice, and other Ethical and Political Challenges." *AI and Ethics* 1.

Coeckelbergh, M. 2021b, *Green Leviathan or the Poetics of Political Liberty: Navigating Free-*

dom in the Age of Climate Change and Artificial Intelligence, New York: Routledge.

Coeckelbergh, M. 2022a, "Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence," https://doi. org/10. 1007/s43681-022-00239-4.

Coeckelbergh, M. 2022b, *The Political Philosophy of AI*, Cambridge: Polity.

Collier, R. & D. Collier 1991a, "Critical Junctures and Historical Legacies." *Americas* 51(1).

Collier, R. & D. Collier 1991b, *Shaping the Political Arena*, Princeton: Princeton University Press.

Collingridge, D. 1980, *The Social Control of Technology*, London: Frances Pinter.

David, P. 1985, "Clio and the Economics of QWERTY." *The American Economic Review* 75 (2).

Deutsch, K. 1963, *The Nerves of Government: Models of Political Communication and Control*, New York: The Free Press.

Dodge, J., T. Prewitt & R. Combes et al. 2022, "Measuring the Carbon Intensity of AI in Cloud Instances." 2022 ACM Conference on Fairness, Accountability, and Transparency.

Easton, D. 1965, *A System Analysis of Political Life*, New York: John Wiley.

Ellul, J. 1964, *The Technological Society*, New York: Vintage Books.

Fioretos, O., T. Falleti & A. Sheingate 2016, "Historical Institutionalism in Political Science." in O. Fioretos, T. Falleti & A. Sheingate eds., *The Oxford Handbook of Historical Institutionalism*, Oxford: Oxford University Press.

Floridi, L. 2019, *The Green and the Blue: Naïve Ideas to Improve Politics in A Mature Information Society*, The 2018 Yearbook of the Digital Ethics Lab.

Galaz, V., M. Centeno & P. Callahan et al. 2021, "Artificial Intelligence, Systemic Risks, and Sustainability." https://doi. org/10. 1016/j. techsoc. 2021. 101741.

Gaus, G. 2016, "The Tyranny of the Ideal." in *The Tyranny of the Ideal*, Princeton: Princeton University Press.

Hayek, F. 1960, *The Constitution of Liberty*, Chicago: The University of Chicago Press.

Heilbroner, R. 1994, "Technological Determinism Revisited." in L. Marx & M. Smith eds., *Does Technology Drive History*, Cambridge, MA.: MIT Press.

IPCC 2022, "Climate Change 2022: Impacts, Adaptation and Vulnerability." https://www.

ipcc. ch/report/ar6/wg2.

Jonsson, I. 2015, *The Political Economy of Innovation and Entrepreneurship: From Theories to Practice*, London: Routledge.

Jonsson, I. 2016, "Economic Crisis and Real Critical Junctures: On the Decay of the Political Party System of Iceland." *The Polar Journal* 6(1).

Kahng, A., S. Mackenzie & A. Procaccia et al. 2021, "Liquid Democracy: An Algorithmic Perspective." *Journal of Artificial Intelligence Research* 70.

Koster, R., J. Balaguer & A. Tacchetti et al. 2022, "Human-Centred Mechanism Design with Democratic AI." *Human Nature Behaviour* 6.

Kurzweil, R. 2015, "Superintelligence and Singularity." in S. Schneider ed., *Science Fiction and Philosophy: From Time Travel to Superintelligence*, New York: Wiley-Blackwell.

Landemore, H. 2023, "Can AI Bring Deliberative Democracy to the Masses?" https://www.law.nyu.edu/sites/default/files /Helen% 20Landemore% 20Can% 20AI% 20bring% 20deliberative%20democracy% 20to%20the%20masses.pdf.

Lenton, T. 2013, "Environmental Tipping Points." *Annual Review of Environment and Resources* 38.

Makitie, T., J. Hanson & S. Damman et al. 2023, "Digital Innovation's Contribution to Sustainability Transitions." *Technology in Society* 73.

Maynard, A. & S. Dudley 2023, "Navigating Advanced Technology Transitions: Using Lessons from Nanotechnology." *Nature Nanotechnology* 18.

Mills, S. &R. Whittle 2023, "Nudging Policy or Crowding It out? Green Nudges as Ideational Technologies." in H. Sætra ed., *Technology & Sustainable Development: The Promise and Pitfalls of Techno-Solutionism*, London: Routledge.

Muench, S., E. Stoermer & K. Jensen et al. 2022, *Towards A Green & Digital Future*, Brussels: L. Publications Office of the European Union.

Næss, A. 1989, *Ecology, Community and Lifestyle: Outline of An Ecosophy*, Cambridge: Cambridge University Press.

OECD 2022, "Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint." *OECD Digital Economy Papers* 341.

Sadowski, J. & E. Selinger 2014, "Creating A Taxonomic Tool for Technocracy and Applying It to Silicon Valley." *Technology in Society* 38.

Sætra, H. 2020, "A Shallow Defence of A Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government." https://doi.org/10.1016/j.techsoc.2020.101283. https:// www.sciencedirect.com/science/article/pii/S0160791X19305925.

Sætra, H. 2021, "A Typology of AI Applications in Politics." https://doi.org/10.1007/978-3-030-88972-2_3.

Sætra, H. 2022, "A Hobbesian Argument for World Government." https://doi.org/10.3390/philosophies7030066.

Sætra, H. 2023a, "Generative AI: Here to Stay, but for Good?" *Technology in Society* 75.

Sætra, H. 2023b, *Technology and Sustainable Development: The Promise and Pitfalls of Techno-Solutionism*, London: Routledge.

Satgar, V. 2021, *The Rise of Eco-Fascism*, in M. Williams & V. Satgar eds., *Destroying Democracy*, Johannesburg: Wits University Press.

Savaget, P., T. Chiarini & S. Evants 2018, "Empowering Political Participation Through Artificial Intelligence." *Science and Public Policy* 46(3).

Schatzberg, E. 2018, *Technology: Critical History of A Concept*, Chicago: University of Chicago Press.

Schumpeter, J. 2013, *Capitalism, Socialism and Democracy*, London: Routledge.

Wiener, N. 1948, *Cybernetics or Control and Communication in the Animal and the Machine*, Cambridge, MA.: MIT Press.

Willis, R., N. Curato & G. Smith 2022, "Deliberative Democracy and the Climate Crisis." https://doi.org/10.1002/wcc.759.

Winner, L. 1969, "Review: Cybernetics and Political Language. Review of the *Nerves of Government: Models of Political Communication and Control* by Karl W. Deutsch." https://www.jstor.org/stable/41035143.

Winner, L. 1977, *Autonomous Technology: Technics-out-of-Control as A Theme in Political Thought*, Cambridge, MA.: MIT Press.

Winner, L. 1979, "The Political Philosophy of Alternative Technology: Historical Roots and Present Prospects." *Technology in Society* 1(1).

Winner, L. 1980, "Do Artifacts Have Politics?" *Daedalus* 109(1).

Zuboff, S. 2019, *The Age of Surveillance Capitalism*, New York: Profile Books.

归因法:网络攻击来源归因规则*

德尔伯特·特兰 著**

裴 轶 强心语 译***

摘要:近年来,国际层面的网络攻击行为日趋增多,典型网络攻击如"震网病毒"事件、DNC 黑客攻击事件等。而网络攻击的溯源,即归因问题,始终是网络安全领域的重点、难点问题。长期以来,学术研究集中于突破归因的技术瓶颈,然而,本文认为这一研究路径系南辕北辙。网络攻击的应对难点不在于无法发现攻击痕迹,相反,当前并不存在技术障碍,真正的难点在于难以归责。构建一套法律体系(归因法)以认定责任主体、确定举证责任、解决国际争端,才是应对网络攻击的合理进路。本文通过研究现有程序规则,包括对抗制和纠问制的特点、举证责任、证明标准、国家责任理论以及关于提供保密证据的规则,阐述了如何构建一套法律体系来识别网络攻击来源。此外,通过对比研究既有的国际争端解决机构,本文探讨了归因法框架下的机构设置。归因法的探索,于当前的互联网时代有利于进一步丰富、完善国际法,维护网络空间国际秩序。

关键词:网络攻击归因 法律框架 国际法庭 国家责任 证据规则

---

* 本文原题为"The Law of Attribution: Rules for Attributing the Source of A Cyber-Attack",原文见 https://openyls.law.yale.edu/bitstream/handle/20.500.13051/7830/DelbertTranTheLawofAttrib.pdf?sequence=2,经授权翻译发表。论文原无摘要、关键词,中文版摘要系译者总结,关键词系编辑部根据论文大意总结而成,特此说明。

** 德尔伯特·特兰(D. Tran),耶鲁大学法学院。

*** 裴轶,北京理工大学法学院。强心语,北京理工大学法学院。

一、引言

网络攻击[①],尤其是大规模网络攻击,有可能在一些关键领域造成重大而广泛的危害。这些攻击包括针对核基础设施、商业实体、政府基础设施以及互联网基础设施本身的攻击。克拉珀指出,2013 年,在美国的国家威胁清单上,网络攻击威胁程度超过了恐怖主义。而且,对 DNC 的黑客攻击表明,此类网络攻击并没有减弱的迹象。网络攻击之所以持续不断,一方面可能是其成本相对较低(Boerbert, 2010),另一方面则在于其攻击源难以追踪。因此,网络攻击为国家行为者提供了肆无忌惮地从事恶意活动的绝佳场所,而不必担心归因或报复。

长期以来,国家归因[②]问题一直是网络安全领域的一个难题。匿名性是互联网的标志和优势之一,但同时也是这一令人困惑的问题的根源。尽管之前的大部分学术研究都集中在归因问题的技术障碍上,但本文试图重新审视这一问题,关注法律而非技术如何解决归因问题。技术归因放大了一个狭义的问题,即人们是否有可能纯粹通过技术手段保证将攻击归于责任主体。但是,法律学者和法律从业者都知道,责任问题很少仅通过单一的技术手段或证据形式来决定,责任的判断往往也不取决于确凿的有罪声明。[③] 法律判决往往基于大量直接或间接证据的积累,这些证据从整体上描绘了恶意行为的责任。同样的逻辑也适用于网络安全和归因。因此,真正的问题是如何构建一个证据和程序规则完善的法律体系。

尽管这一网络安全问题出现在政策与技术的交叉点上,但它也提出了一

---

① 本文所指的"网络攻击",既包括网络战争,也包括没有发生武装冲突,但仍应受到某种形式制裁的网络攻击,即使它们不值得以武力作为回应。

② 本文的"归因"一词,指识别网络攻击背后的行动者的过程。

③ Desert Palace, Inc. v. Costa, 539 U.S. 90, 99-100 (2003); Siegert v. Gilley, 500 U.S. 226, 236 (1991); Holland v. United States, 348 U.S. 121, 140 (1954).

个特别适合法律解决的问题。如果从根本上说法律涉及各方据以裁决争端的制度,那么网络攻击的归因问题正是法律可以解决的争端。法律程序还赋予结果以合法性,并以更强的制度影响力将这种解决方式正式化。在一个更具争议性和政治化的环境中,所有的报道都会被怀疑带有党派偏见,因此通过法律程序得出的结论更难被视为纯粹的“假新闻”。此外,一旦知道了网络攻击的罪魁祸首,就可以研究他们的战术和方法,威胁进行报复,采取反制措施纠正攻击行为,并建立和巩固适当的行为规范。但是,由于无法确定攻击来源,所有这些可能的应对措施都会受挫。在回答了一个关键的必要问题后,归因使法律得以出现:如果有的话,哪个国家应对网络攻击负责?

实际上,根据国际法,归因法将使针对另一个国家的某些制裁合法化,包括《联合国宪章》第五十一条——可能使用军事力量进行自卫。反之,如果一个国家未能证明其归因主张,那么根据国际法,它所实施的任何制裁都可能变得不合法或无效。归因的法律框架将提供一个重要的基石,使制度能够限制和纠正网络攻击所造成的伤害。

本文分几个部分来构想归因法。第二节首先回顾了归因问题:近期网络攻击造成的威胁、对此类攻击缺乏问责的问题,以及学者和政策制定者普遍认为阻碍网络攻击归属的一般技术障碍;然后,反驳了长期以来无法确定网络攻击归属的说法,指出归因的技术问题远比法律要求的范围狭隘,并说明归因如何反映了一个更容易解决的法律问题。第三节设想了国际归因法的框架:首先,它概述了评估国家对非国家行为者行为的责任承担的背景和重要考虑因素;其次,提出程序和法律规则,不仅是为了设想归因法的外观,而且是为了设想其程序规则如何与其实质内容建立适当和合理的关系。第四节讨论了归因法中最困难的因素:国家加入或参与这种法律框架的动机。虽然评估国家激励措施,涉及国际关系性质、国家间合作、国际秩序遵守等基础性问题,但是本文并不侧重讨论上述宏大命题;相反,本文将通过回顾、分析既有国际法庭的实践,探索归因法框架下的国际裁决模式。

二、国家归因问题

由于无法确定网络攻击的源头,行为者可以肆无忌惮地实施攻击,从而使制定国际法或条约来规范这种有害行为的努力失效。即使当前正式法源难以解决网络攻击,但是各国仍须确定攻击的来源,以便做出应对,包括制裁攻击者及其行为。换言之,归因是任何试图对恶意网络攻击实施规则或限制的关键先决条件。

目前的国际制度几乎没有明确规范或控制国家在网络黑客领域的行为,没有任何国际法律或条约明确规范网络攻击。虽然有学者指出武装冲突法的潜在适用性(Hathaway, Crootof & Levitz et al., 2012),但迄今为止,在应对网络攻击时,武装冲突法显然尚未被引用过。鉴于国际关系领域普遍存在不确定性,各国可能会规避风险,在对其他国家采取法律和外交行动时,对国际法的创新解释犹豫不决。因此,缺乏归因限制了制度和法律解决方案,使网络领域长期处于无法之境,未来网络攻击的范围和规模仍有升级和不确定性的可能(Condliffe, 2016)。

从更广泛的国际关系理论角度来看,归因有助于国际法的发展。很容易理解为什么自由主义理论家需要对网络攻击者进行归因,并以法律制度规制之,因为网络攻击对国内实体的附带影响为国内行为者创造了大量动机,鼓励国家行为者加入遏制此类攻击的国际框架(Moravcsik, 2014)。但即使是国际关系现实主义者也会承认,即使在没有总体国际法的情况下,国家也有必要通过归因来维持秩序。现实主义者的传统口号否认任何凌驾于国家之上的中央权威,并认为国家总是在寻求权力和促进自身利益(Mearsheimer, 2001)。虽然这种对国际关系的理解对国际合作或国际法构成了初步障碍,但现实主义逻辑并不完全排除合作,其中一个反驳理由就是互惠(Keohane, 1986)。互惠论的倡导者从博弈论出发,指出理性、自利的行为体在合作或叛

逃之间做出选择时,在反复进行博弈的情况下会最优化地选择合作。之所以选择合作,是因为博弈者在未来的“博弈”(或互动)中,会根据先前迭代中的决定对他人的行为进行惩罚或奖励(Axelrod, 1984)。因此,即使假定国家行为采用现实主义框架,互惠性也能使国际法在合作过程中形成,因为国际关系往往涉及国家间的重复互动,这些互动构成了国际关系博弈的“迭代”。

然而,互惠论假定国家可以准确地惩罚或奖励彼此的行为。虽然反措施可能促成这一互动,但正确使用反措施与正确归因密不可分。归因不仅是国家制裁负有责任的恶意行为者的基本要求,而且适当的归因对于国家声称合法使用制裁或反制措施也至关重要。法律的作用不仅在于决定冲突的结果,还在于使第三方对冲突结果的决定合法化(Weber, 1968)。法律的合法化功能在国际法和国际关系领域尤为明显,因为在这些领域,国家缺乏通过武力强制遵守的总体权威(Goldsmith & Posner, 1999)。因此,在互惠制度中,归因具有双重功能:确定不法行为者;使对第三方的正式或非正式制裁行为合法化。因此,归因问题是任何限制使用网络攻击的法律制度的关键性第一步。

(一) 为何归因如此困难

长期以来,网络攻击追踪溯源是网络安全领域的重点、难点问题。目前的许多学术研究都接受了这样一种传统观点,即互联网的技术架构使得归因成为一个极其困难的问题。归因问题引发了专家学者浩如烟海的讨论,如辛格等人(Singer & Friedman, 2014)将归因问题描述为网络领域“也许最棘手的问题”。

由于多种原因,很难将网络攻击归因于其源头。

首先,互联网的结构设计和跨网络信息传输的性质使归因变得复杂。当用户希望通过互联网进行某些操作时,例如在油管(YouTube)上搜索柯基犬的视频时,用户的计算机需要找到一种方法,与托管 YouTube 内容的机器进

行通信,并让该机器将柯基犬嬉戏的内容发送到原始机器上。这是如何实现的?每台机器都会被分配一个互联网协议(IP)号码作为“地址”(Grimmelman, 2017),这个地址通常是由互联网服务提供商或网络分配的,用户的电脑一般会从本地互联网路由器的地址开始,由路由器将请求转发到更广阔的互联网。

要将用户的请求发送到载有柯基犬的机器上,用户的机器就需要知道该机器的地址。用户的机器如何找到这个地址呢?从用户的角度来看,他或她可以在搜索栏中输入“www. youtube. com”。在机器端,这些可识别的名称会通过域名系统转换成机器地址或 IP 号,域名系统可以被视为一个将网站名称与 IP 号码相匹配的全局目录。一旦用户的机器知道了拥有大量柯基犬视频的网站地址,下一步就是将来自用户计算机的数据(从 YouTube 获取内容的请求)传输到 YouTube 上,YouTube 再将请求的数据发送到用户的计算机上。我们可以简单描述一下这个过程:请求(用户在地址栏输入的文本或点击网站链接的动作)在超文本传输协议(HTTP)层被转换成数据(数字),然后将数据传递给传输和网络层。在传输层,数据被分解成数据包大小的数据块,每个数据块都包含其目标地址,就像通过邮件发送的小信封。这些数据包在通往目的地的途中被传输到整个网络中的各个服务器(想象一下长途旅行中成千上万个可能的中转站),直到到达最终目的地,并被重新组合成来自 YouTube 的原始数据请求。在有 YouTube 内容的机器上,当 YouTube 将信息发送回用户时,这一过程会以相反的方向重复进行。

不过,两台计算机之间的这种通信过程并不要求知道请求(或黑客)的来源。YouTube 知道将其响应发送到哪里的唯一原因是,原始请求有意包含了它的地址,这样 YouTube 就能将数据发送回来。其他类型的活动,如向 YouTube 上传视频,则不需要在发送的信息中嵌入返回地址。因此,不需要数据传输的原始来源,我们的机器就能参与在线活动。我们通过互联网发送的数据包只需要知道它们的目的地,而不需要知道它们的来源。与邮局不同的

是,我们不需要返回地址,因为任何无法通过的数据都会丢失,人们只需一次又一次地尝试另一个请求,直到它通过为止。

其次,用户可以利用一些技术或程序应用来隐藏他们的网上活动痕迹(Greenemeier, 2011)。用户在互联网上进行的任何活动都会被记录其 IP 地址,因此用户可以选择使用代理服务器或"洋葱路由工具"来掩饰自己的 IP 地址(Feigenbaum, Johnson & Syverson, 2007)。参考邮局的比喻,为了掩盖寄信来源,寄件人可以把信交给朋友,让他们通过不同的邮局寄出,而非最近的邮局。寄件人还可以通过写下一个假的回邮地址来"欺骗"原始地址(Tanase, 2003)。一项实验证明,近三分之一的互联网用户可以在不被发现的情况下伪造他们的源 IP 地址(Beverly, Berger & Hyun et al., 2009)。

再次,即使可以重新设计互联网,对互联网上发送的每个数据位的源 IP 地址进行验证(Gallagher, 2013),这些地址也只能实现识别攻击源机器而非攻击者。攻击者可能会窃取或入侵他人的设备,或利用多人使用的公共设备或网络来进行网络攻击,这种情况数不胜数。例如,Mirai 僵尸网络攻击涉及恶意代理利用成千上万的其他设备,这些设备被代理收编为攻击工具(Meyer, 2016)。

最后,即使克服了所有技术问题,并确定是某个人发动了网络攻击,主权国家是否应对该个人的行为负责的问题依然存在。换句话说,网络攻击也提出了一个问题:国家何时可以为非国家行为者的不法行为负责?虽然这一法律难题最常出现在恐怖分子或集团的情况下(Clapham, 2015),但对黑客和网络攻击者来说,这个问题同样突出,因为他们通常没有制服或旗帜来表明他们是以任何特定国家的名义行事的。需要注意的是,这并不是归因问题的技术障碍,而是法律障碍。这一特别关切突出表明,有必要制定一种法律制度,以解决归因所带来的问题。

互联网的结构设计、掩饰在线活动的工具、仅能追踪于机器的限制以及将个人行为归责于国家的限制,这些因素构成了技术和法律方面的诸多障

碍,经常被认为是建立网络攻击监管法律制度的障碍(Waxman, 2011)。虽然以往的学术研究通常将归因的技术问题视为一个难以解决的难题,最好留给工程师去解决,但最近的学术研究认识到,归因问题可能并不是一项不可能完成的任务(Rid & Buchanan, 2014)。

(二) 技术归因问题是一种误导

尽管归因存在诸多技术障碍,但技术问题只是一种误导。这些技术障碍只是阻止我们得出一个非常狭隘的结论,即我们何时可以绝对肯定某个行为主体对网络攻击负责。然而,法律几乎从未按照绝对肯定这一几乎不可能达到的高标准来运作。即使是美国刑法,其著名的有利于被告的高举证责任也只要求在定罪前没有合理怀疑,而不是要求完全没有怀疑(Whitman, 2005)。从根本上讲,归因问题的核心是责任问题,而责任问题从根本上说是一个法律问题。即使在没有绝对因果关系确定性的情况下,法律也经常回答这个问题。因此,本文通过提出两个要点来解决归因问题:

第一,尽管归因存在障碍,但计算机科学家已经开发出一系列工具来追踪网络攻击,根据经验,大规模的国家攻击往往会留下足够的痕迹(或间接证据),引导鉴证专家找到攻击源。

第二,法律并不要求保证确定无疑,而只要求有足够程度的确定性,即某人应承担责任;无论是合理怀疑标准,抑或优势证据标准,法律已经构建了一套可应用于实践的归责方法。

一旦这两点确立,问题就不再是网络攻击是否可以归因,而是国际社会如何构建一套法律体系,制定必要的证据规则、程序、举证责任等。

关于第一点,强调归因的技术性质自然吸引了那些具有更强技术专长的人的兴趣,计算机科学家也相应地开发了一套工具来归因网络攻击。虽然这些方法都不可能单独提供万能的解决方案,但每种方法都提供了鉴证技术,可能会给任何特定案件带来一些启示,而且累积起来确实有可能在一定程度

上确定归因。就像可以通过鉴证证据(搜索指纹、识别笔迹等)追踪匿名信一样,也有办法利用间接证据来确定数字信息的传输和随后的网络攻击。本文所探讨的网络攻击——可能引发或要求国家做出反应的高调网络攻击尤其如此。由于其范围或规模更大,此类攻击往往更有可能留下痕迹。更大规模的行动也需要更多的资源,从而限制了有能力发动此类网络攻击的潜在对手的范围。

事实上,调查人员利用这些技术找出了最近两次重大网络攻击的罪魁祸首:“震网病毒”(Stuxnet)事件和 DNC 黑客攻击事件。下文将依次回顾每一次攻击,描述取证和间接证据的积累如何实现这些攻击的归因,从而证明归因的技术问题被夸大了这一讨论现状。

1.“震网病毒”(Stuxnet)事件

从 2009 年开始,伊朗的铀离心机开始失灵,且无人知晓原因。在一年的时间里,伊朗 6000 台离心机中有近 1000 台被摧毁。2010 年夏天,白俄罗斯的一家电脑安全公司受聘对伊朗的电脑进行故障排除,这些电脑一直神秘地死机。在调查中,这家公司偶然发现了一系列文件,这些文件后来被称为“震网病毒”(Stuxnet)。“震网病毒”被认为是“世界上第一个数字武器”(Zetter, 2015)。这是一个复杂的恶意软件,旨在渗透伊朗安全的核设施,感染操作核离心机的工业控制器,并通过操纵离心机内的压力水平和转子速度来摧毁这些离心机(Langner, 2013)。该病毒的设计意图是缓慢而渐进地造成这种破坏,从而降低被发现的可能性;它甚至包括一个操纵伊朗传感器的功能,以假装被操纵的功能在正常工作。

尽管人们试图掩盖其来源,但专家们得出的结论是,“震网病毒”是美国和以色列联合制造的。在“震网病毒”事件中,这一信息几乎是决定性的,因为很少有国家有动机和手段来攻击伊朗的核离心机。此外,攻击的规模往往会透露攻击者的信息。虽然高级持续威胁是最具威胁性的网络攻击形式,但其优势也成为其弱点。这是“震网病毒”攻击的另一个线索——代码中有四

个“零日漏洞”(对于私人黑客来说,这些漏洞的转售价值高达数百万美元),这再次暗示了攻击背后有强大的火力支持,几乎可以保证这种攻击来自某个国家。最后,细小的线索往往可以确定攻击的源头。通过“震网病毒”的代码,调查人员能够根据引用西门子设备的名称和 ID 编号发现其主要攻击目标——被操纵的目标是工业离心机控制器。鉴于目标的狭小性和投入的巨大资源,很容易推断出背后的国家。

2. DNC 黑客攻击事件

DNC 黑客攻击事件是国家行为者实施重大攻击的最新例证。美国情报部门坚称 DNC 黑客攻击来自俄罗斯。美国表示,尽管这一认定也依赖于机密情报,但在调查过程中咨询了几家私营网络安全公司,并提供了公开证据,证明攻击是俄罗斯所为。例如,他们注意到,DNC 黑客使用的渗透工具和编码与已知为俄罗斯联邦安全局(FSB,前身为克格勃)工作的俄罗斯黑客使用的工具和编码相同。这些分析人员还将 DNC 黑客攻击事件与 2015 年对德国议会发动攻击时使用的同一个 IP 联系起来。安全专家注意到用西里尔字母留下的数字签名。此外,安全专家注意到,DNC 黑客攻击事件在俄罗斯节假日期间停止了行动,他们的工作时间与俄罗斯时区一致(Meyer, 2016)。

当然,这些间接证据并非完全确凿,其中一些信息有可能是栽赃陷害。但法律体系长期以来一直能够分配惩罚和责任,即使责任仅来自间接证据。在 DNC 黑客攻击事件中,虽然可能有人植入了西里尔字母签名等线索作为障眼法,但黑客组织完全在俄罗斯时区和节假日内协调行动作为其伎俩的可能性要小得多,因为这种行动的协调成本很高,而且复杂程度非同寻常。最终,就像在刑事案件中一样,可以积累足够的证据来确定攻击的来源。

因此,问题不在于确定攻击的来源。挑战在于如何让其他国家相信,源头已经被正确识别。希望采取反制措施的国家需要让其他国家相信其归因的准确性,以确立其反制措施的合法性。出现这一问题,可能有两个主要原因:归因可能基于通过国家间谍活动或情报收集工作收集的数据,而国家可

能希望对这些数据保密(Sanger & Fackler, 2015);当国家有可信的事实依据来认定攻击行为时,它们可能不希望披露这些证据,因为网络攻击者可以从这些错误中吸取教训,避免在未来留下同样的痕迹(Brenner, 2007)。

虽然这些努力最终是基于间接证据的积累,但在许多法律领域,间接证据为支持法律判断提供了足够的可信度。例如,在侵权行为中,严格责任和事实自证原则表明,决定性的问题可能并不总是谁实施了某一行为,而是我们如何追究某一特定个人或实体的责任。而在不同情况下适用不同的责任标准,反映了法律在建立适当框架以解决此类冲突方面的灵活性。因此,在设计归因法时,这些问题将涉及对其他法律领域所援引的证据和因果关系的一般标准的一些调查。在这些领域中,法院已经使用了法律工具来建立足够的信任度,以将责任归于行为人。

三、归因法

既有的法律和程序体系包含了大量可借鉴的材料,提出了许多制度特征和设计,供我们在勾勒归因这一法律制度时参考。国际归因法在设计其结构和组成部分时必须解决几个问题。本节将首先讨论是否有可能制定一套跨实体法的归因规则,以及与此相关的问题,即这一归因法将用于何种目的。这些答案为该体系的整体结构和框架奠定了基础,而整体结构和框架将涉及设计选择,如是否优先选择对抗制而非纠问制,以及制度设计的其他关键方面。然后,本节将讨论界定实体法界限的关键程序规则。这些规则包括举证责任、评估国家对非国家行为者的行为所负责任的标准,以及证据和管理敏感性的规则。虽然这些规则是程序性的,但它们对案件的潜在结果有着巨大的影响,因此必须制定适当的程序,以确保兼顾程序正义与实质正义。

(一) 跨实体法的归因法

是否有可能制定一部跨实体法的归因法,即无论归因所证明的法律或政

治行动如何,该法的规则都能适用,这是构建归因法首先要解决的问题。举例而言,归因法可能会根据应对网络攻击所采取的反制措施的严厉程度而改变其严格或灵活的标准。

从广义上讲,归因法背后可能有几个目的,通过归因的法律主张,可能有理由采取几类后续制裁或应对措施:第一,在确定攻击的归属后,可以对网络攻击的责任国进行负面的经济惩罚,如经济制裁;第二,被归咎于发动攻击的国家可能会被剥夺积极利益,因为它将无法参与未来的国际条约或协议;第三,归因可以成为黑客反击措施的理由(Holzer & Lerums, 2016);第四,归因可以成为军事回应的理由。这些可能的归因对策可进一步分为两类:单边行动或多边行动(表1)。

表1　　归因对策的进一步分类

	单边行动	多边行动
拒绝外交接触/协议	拒绝参与目标国家可能寻求的贸易协定、条约或其他双边协定;拒绝外交接触	拒绝加入更广泛的贸易协定或条约
负面经济惩罚	经济制裁;废除现有的贸易协定	多边制裁机制
网络对抗措施	反击黑客	联合制造网络攻击,如“震网病毒”
军事对策	有针对性的军事打击等	联合军事力量

虽然这些选择提出了国家在网络攻击归因后可能采取的一系列实际和政策应对措施,但就制定归因规则而言,这些应对措施对我们如何设计归因规则的影响,可从两个主要方面加以考虑:单边行动还是多边行动;惩罚的“严重”程度。

第一个问题,归因是用来发起单边行动还是多边行动,实际上对归因法的整体理论影响相当小。这主要是因为归因法背后的目的在单边和多边反应中通常是一致的——在国际社会看来,归因是惩罚的正当理由。无论一个国家是否希望通过自己的单边行动或多边行动来惩罚网络攻击者,归因都是

为了在国际社会第三方的眼中使这种行为合法化。

唯一的例外是在多边承诺无法保证的情况下,受害国不仅需要说服其他国家报复是合理的,而且需要说服其他国家参与报复。这些情况可能会使归因法理论向更严格的要求倾斜,因为其他国家可能会要求对归因有更大的信心,然后才会投入自己的资源来应对没有直接影响到它们的网络攻击。

对信任差距的担忧有两种回应:直接遭受攻击的国家对正确识别攻击来源有着极高的兴趣,这意味着信任差距可能并不取决于归因的确定性,而更多地取决于国家加入多边行动的一般动机;仅仅存在一个多边机构,要求非受害国做出反应,似乎就意味着这种机构联系的来源本身就足以促使这些国家加入实施惩罚的行列,而无须更严格的归因制度的额外保证。例如,如果国家有义务对网络攻击做出多边回应,那么作为一个法律问题,或作为确保未来合作的理性利益问题,它们受到约束这一事实可能足以证明一个国家有理由决定与受害国一起对网络攻击来源做出多边回应(Loomis, 2008)。因此,单边/多边的区别很可能不会改变制定一套跨实质性归因规则的可能性。

对网络攻击可能采取的反制措施的严重性可能会对单一的跨实体归因法造成更重要的影响。更严重的反制措施可能需要更严格的程序规则,使这些规则取决于一国应采取的反制措施。虽然这一直观原则在抽象情况下似乎是正确的,但在网络安全的具体背景下以及上文详述的国家可能采取的应对措施中,这一原则值得探讨。根据可能采取的反制措施的严重程度,应对措施可大致排序如下(从高到低):军事力量、网络反措施(或“黑客反击”协议)、经济制裁和外交惩罚。

虽然军事力量涵盖各种可能的行动,但这些行动仍可归类为应对网络攻击的最严厉的可能对策。鉴于军事行动的一般成本和冲突升级的危险(Nye, 2021),各国越来越少选择动用军事力量。此外,国际法明确规定全面禁止使用武力。尽管如此,政界人士和军方领导人都表示,可能会对外国网络攻击做出军事反应(Williams, 2016);在对黑客攻击采取可能的反制措施时,特别

是当网络攻击严重到足以被归类为武力行为时,这仍是保留选项(Schmitt,1999)。军事行动的幽灵很可能会引发国际社会的高度关注,并对正确确定网络攻击来源的信心提出极高的要求。考虑到2003年美国因妄称伊拉克拥有大规模杀伤性武器而发动的对伊拉克的入侵已声名狼藉,这一点尤其如此。

另一类反制措施,即网络黑客攻击(Chulov & Pidd, 2011),也可能达到与使用军事力量相关的严重程度。虽然网络黑客攻击可能涵盖比军事力量更广泛的一系列活动,但一些学者认为,网络攻击有可能造成与传统动能军事攻击一样大的破坏,有时可被定性为国际战争法下的武力(Graham, 2010)。如果网络黑客攻击在某种程度上被认为等同于国际军事力量,那么这类反制措施也可能需要一套特定的程序规则,以证明在这些高风险情况下的归因是合理的。

对更严格的程序规则和更严肃的反制措施的需求,并不会导致制定一部关于归因问题的跨实体法的计划失败,因为法律可以通过修改相关的程序规则或要求来引发特定的惩罚。以美国的版权法进行对比,其中的条款可以根据侵犯版权行为的严重程度,规定民事赔偿、加重民事赔偿或刑事责任。这三种对侵权行为的惩罚都依附于版权法的一般体系,但具体的惩罚取决于被告的违法意图。回到归因法,其可以根据应对网络攻击的救济措施的严厉程度,来调整举证标准,这并不影响归因法的构建可能性。

鉴于有可能制定一部关于归属问题的跨实体法,下一步就是概述这一体系的主要特征,首先是将形成整个法律结构的基本要素。

1. 对抗制或纠问制

一个需要面对的主要设计选择是,归因法是在对抗制框架下运行(如美国、英国的法律制度)(Strier, 1992–1993),还是在纠问制框架下运行(如大多数欧洲、亚洲和南美洲国家的法律制度)(Kötz, 2003; Langbein, 1985; Tomlinson, 1983)。选择对抗制框架还是纠问制框架,在很大程度上反映了

法律程序的理念,而这一理念又决定了该系统的规则和整体设计。对抗制法律框架的主要特点是,公正的决策者(法官或陪审团)根据当事人及其代理人提供的证据和论据对争议做出判决(Hay & Spier, 1997),这种制度依赖于对抗双方提供证据和论据。而纠问制框架则将法官定位为主要的事实认定者和调查者,当事人及其律师在收集证据方面发挥的作用却相对有限。

虽然许多纠问制诉讼制度仍保留了一些“对抗”特征,但将重点从当事人转移到法官对整个法律制度产生了关键性的连锁反应。正如约翰·朗宾(J. Langbein)所指出的,德国法院的纠问制设计极大影响了德国民事诉讼程序的其他方面,如为诉讼的各个阶段设定了更为灵活的顺序。相比而言,对抗制为原告和被告陈述或参与诉讼的各个环节设定了固定的顺序,而“在德国的诉讼程序中,法院会对整个案件进行审理,不断寻找案件的关键点,即可能决定案件胜负的法律问题或事实问题”。因此,纠问制诉讼程序的灵活性使得审判过程具有连续性,可以通过多个时间点对问题进行重新审理。此外,朗宾还指出,纠问制对证人以及他们在法庭上提供事实或证据时所扮演的角色有重大影响。在对抗制中,提供证人、准备证人、直接质证和交叉质证主要由双方负责。在纠问制诉讼制度中,法官负责传唤证人并在法庭上主导对他们的询问。这只是对抗制诉讼制度或纠问制诉讼制度在影响法律机构民事诉讼程序的整体特征方面的两个例子。因此,在构建归因法时,应从一开始就确定这一制度选择。

无论哪种制度,都各有理由。主张对抗制的人赞美对抗制盘问的优点,认为它是揭露谎言的最有力工具;在这种制度下,当事人专门负责提出和获取证据,而事实认定者专门负责从给定的证据中得出推论,这种制度效率较高(Marcus, 1992);他们还指出纠问制法官可能会预先判断案件的结果,忽略可能进一步揭示争议的关键证据或论据。主张纠问制的人则指出,过度的立场之争和作秀可能会使对抗制程序最终扭曲事实和证据(Frank, 1949),并使该制度偏向于那些拥有更多资源和更优秀律师的当事人(Hadfield, 2000)。

在这些反反复复的争论中,学者们采用了许多理论和经验模型来检验这两种制度的有效性。一些模型表明,这两种制度在产生准确或理想结果的能力上差别不大(Froeb & Kobayashi, 2001);而另一些模型或研究则认为,结果取决于个人所衡量的特定数据(Lind & Tyler, 1988; Parisi, 2002)。通常认为,虽然两种法律模式之间的争论由来已久,而且在短期内也很可能无法解决,但每种模式都有其适合运作的情形。朗宾尽管赞成纠问制的普遍效力,但他承认对抗制在刑法中可能有其独特的理由。

本文认为,对抗制模式更适合归因问题,概因纠问制的优势在国际背景下化为乌有。首先,纠问制依赖于一个预先存在的、中央集权的司法当局,它可以被信任为客观地寻求真相,而国际领域缺乏这样的机构;其次,由于归因往往依赖于技术证据,而证据往往是通过间谍活动或其他秘密情报收集活动获得的,因此在归因争端中当事方几乎总是最有能力获得和提出此类证据。

纠问制依赖司法机构来推动其程序,这在很大程度上是国际背景下的一个弱点。虽然确实存在一些国际法院,但这些法院的管辖权并不完整,或者是专门处理专业问题,无法涵盖这里提出的归因问题。国际法院(ICJ)是当前国际框架内可能存在的最佳司法选择,因为它一般会广泛考虑主题事项(Thirlway, 2016)。然而,即使是国际法院,其影响力也是有限的;只有在国家同意适用国际法院的情况下,国际法院才能解决国家间的争端。例如,在国际法院对尼加拉瓜诉美国一案做出不利于美国的裁决后,美国退出了国际法院的强制管辖(Kahn, 1987)。此外,国际法院的强制执行权也受到限制。纠问制假定了对主导大部分诉讼程序的司法机构合法性的高度信任,而在国际法背景下,国家主体不太可能参与一个更多由法院而非当事人本身主导的程序。

其次,纠问制假定法官有足够的专业知识来寻找解决案件的相关信息,包括知道应该寻找哪些(专家)证人以及如何对他们进行询问。但在归因问题上,这种能力推定可能不成立。鉴于网络攻击和归因的技术性质,当事人

有理由认为,由一个普通法院来牵头提供事实和证据不太可靠。即使可以通过由具有技术专长的法官组成的合议庭进行诉讼来解决这一问题,但这些法官在确定特定争议中的确切事实方面所处的相对地位仍会有所欠缺。当涉及他们在确定某一特定争端中有争议的确切事实方面的相对立场时,这些法官仍然显得力有不逮。法官可能不那么熟悉每个国家的网络能力和行动,也不熟悉可能支持一国指控另一国对网络攻击负责的基本证据。由于围绕网络攻击和网络安全的许多证据也可能来自秘密情报行动(Connolly, 2016),因此对抗制将更为合适,因为当事方本身最有能力提出或决定何时提出某些敏感证据。

对抗制为归因法提供了一个总体框架。这种制度将有一个公正的裁决者,并将在很大程度上由当事人在法律论证和提供事实与证据两方面驱动。因此,这种体系将包含与美国法律体系类似的程序排序,从启动到发现再到提出论据,其中的论据将围绕当事人各自的论证阶段进行安排。

2. 证明标准

有了对抗框架,接下来要做的就是确定对抗双方如何成功证明其归因主张。换言之,就是确定成功证明主张的举证责任。“举证责任”一词通常指两个不同的概念:说服责任和举证责任(证据)(Fleming, 1961)。由于本节大部分内容涉及在对抗制诉讼中当事人的举证责任,因此此处使用的“举证责任”一词指的是说服责任。广义地说,说服责任涉及裁判者在收到案件中提出的所有相关事实和论据后,对得出法律结论的信心。

举证责任也许是对案件实体结果影响最大的程序规则。罗伯特·贝尔顿(Belton, 1981)将举证责任描述为“美国法律体系中最重要的程序概念之一”,因为“它通过指示事实认定者对特定类型案件的事实结论的正确性应具有的信心程度,帮助实施实体法”。毕竟,根据当事人的举证责任,同一套事实可能会导致完全不同的结果。

一些学者批评举证责任之间的分级在法官或陪审团的心目中没有明确

的区别。然而,这些批评都是在理论层面提出的;支持这些论点的经验证据,往往是基于要求个人对各种举证责任进行抽象定义的调查。但是,关于这些举证责任含义的调查答案可能并不是结论性的,因为这些术语的含义在实践中总是结合具体的事实来理解的(Kaplow, 2012)。因此,对"清晰且有力"的特定含义缺乏共识可能并不反映事实认定者在特定案件中的实际共识。在特定案件中,一组事实认定者可能都同意一方当事人的证据已确立了"清晰且有力"的证据。此外,考虑到举证责任在实践中对案件结果所产生的经验性影响,这些否定举证标准作用的论点似乎并不具有说服力(Devine, Clayton & Dunford et al., 2000)。鉴于举证责任对一个法律体系的诉讼程序所具有的重要影响,确定归因法中举证责任的适当高度就显得尤为重要。

举证责任有三种经典标准:优势证据标准、清晰且令人信服标准,以及排除合理怀疑标准。优势证据标准直截了当地要求事实认定者相信事实(或法律结果)存在的可能性大于不存在的可能性,大致上将举证责任平均分配给双方。美国最高法院将清晰且令人信服标准描述为中间标准,该标准对说服力的要求略高于优势证据标准,但仍低于排除合理怀疑标准。最后,排除合理怀疑标准代表了最高的举证责任,旨在确保尽可能保护被告免受错误判断的可能性。

虽然这种举证责任的范围已经确立,但何时适用某种标准的规范性依据却不那么明确。詹姆斯·弗莱明(J. Fleming)写道:"在任何特定问题上,在两种意义上分配举证责任都没有令人满意的检验标准。"罗伯特·贝尔顿(R. Belton)也表达了类似的观点:"法院尚未制定出任何普遍规则或一套政策考虑因素,供法院在确定如何在双方当事人之间分配这三种责任时参考。"诚然,优势证据标准长期以来一直是美国民事诉讼的标准,合理怀疑也同样是美国刑事司法诉讼的主要规则。然而,这些标准与各自的诉讼程序联系在一起,主要是法律传统层面的原因,缺乏特定的理由,特别是民事诉讼程序中适用的标准。这一点在将美国的法律制度与其他国家的法律制度进行对比

时尤为明显。一些具有纠问制传统的国家,如德国,对其法院所面对的所有法律问题,无论事由如何,均采用合理怀疑标准(Clermont & Sherwin, 2002)。因此,任何一种特定的法律制度都会采用不同的举证责任。在归因问题上,如何选择适用哪种举证责任呢?

虽然在选择举证责任标准时可能没有单一的检验标准,但有一些一般原则确实影响着这种选择。例如,政策依据、公平性以及相关事件实际发生的可能性等考虑因素。弗莱明认为,类似的公平、便利和政策性总体原则推动着为举证责任设定标准的决策。除了这些更为普遍的原则之外,弗莱明还承认其他考虑因素的相关性,如一方当事人获得证据的相对机会、一方当事人的论点偏离普遍经验的程度,以及可能将举证责任作为不利论点的障碍的实质性考虑因素。

虽然贝尔顿和弗莱明的描述似乎符合常规,但也没有提供太多有益的启示。公平、便利和政策性作为宽泛的理由,几乎适用于任何法律解释,也适用于任何方向。

相反,路易斯·卡普洛(Kaplow, 2012)将这些考虑因素置于一个更具体的参照框架内,通过经济学分析来探讨举证责任,即每种举证责任如何才能最好地实现法律制度的目标。举证责任被具体视为调整两种主要概率结果的工具:对实施有害行为的人施加责任的概率,以及对适当行为的人施加错误责任的概率。在卡普洛看来,举证责任必须在阻止有害行为和避免扼杀有效行为之间取得平衡。按照这一思路,考虑非对称错误成本至关重要,因为这些错误的计算往往决定了程序规则如何倾斜竞争环境,包括如何设定举证责任。

最典型的例子就是刑事处罚,因为“宁可让 10 个有罪之人逃离处罚,也好过让无罪之人遭受痛苦”,所以我们认为“有利于被告的刑事诉讼规则”是合理的,包括高举证责任。就归因而言,错误成本似乎没有那么明显地偏向一方或另一方(Gersen & Vermeule, 2016)。让一个网络攻击国逍遥法外是否

比惩罚一个无辜的国家更好?假设网络攻击的严重程度足以上升到武装力量的水平,并假设反制措施的范围小于军事打击,那么网络攻击的危害是否不如军事打击严重,这一点并不一定清楚,尤其是如果后者应该受到相称性规则的限制。

就归因法而言,优势证据标准最适合实现归因制度的总体目标。在军事行动是对网络攻击的唯一回应(或威胁)的情况下,举证责任应提高到合理怀疑标准。作为基本的举证责任,以证据优势来证明归因似乎最合适,主要有两个原因:首先,较低的举证责任产生较低的证据门槛,增加了产生法律判决的机会,从而增加了责任风险,提高了对有害行为的威慑;其次,它将说服责任大致平等地分配给各方,挑战双方以最佳方式提供有关网络攻击起源的信息和证据。

虽然可以设想一种更低的举证责任(如严格责任),但优势标准是最可取的平衡点,因为它规定必须提供一定程度的信息以确立表面证据确凿的案件,然后根据对双方所提供信息的比较分析做出判决。这一要求鼓励控辩双方竞相提供信息。因此,优势证据标准能带来最佳的信息生产水平,而更广泛地提供有关国际网络攻击的更多信息,有助于解决困扰网络安全领域乃至更广泛的国际互动中的不确定性和交易成本问题(Li, Ou & Rajagopalan, 2010)。

批评者可能会反对说,优势证据标准对为自己辩护的国家来说是不公平的。毕竟,优势证据标准将责任平等地分配给了双方。但有人可能会说,与提出指控的国家相比,处于辩护地位的国家实际上处于弱势。不仅存在信息不对称的问题,因为提出归因指控的国家可能拥有(或声称拥有)支持其立场的秘密情报,而且扮演防御角色的国家基本上也被迫通过证明反事实来反驳指控,即它事实上没有发动网络攻击。鉴于进行归因可能需要复杂的技术技能,而一些国家可能缺乏拥有此类技能的人员,因此考虑到归因所带来的特殊挑战,一些国家可能根本没有资源来开展反击归因的工作。刑事或民事案

件中的个人可以提供不在犯罪现场的证明,而复杂的、由许多成员组成的国家则不同。一般来说,国家无法说明其全部职能,以显示其诚实。

反驳的理由是,个体在民事诉讼中作为被告时,经常会对自己的行为做出说明。虽然证明反事实确实很难,尤其是在黑客攻击的情况下,但这种反对意见假定将攻击归咎于国家的初步证据已经发生。如前所述,即使适用优势证据标准,这一任务仍是一项挑战。优势证据标准的传统表述是,一方需要以超过 50%的可能性证明其主张。但是,假设一个不可知论的事实调查者从 50%的概率线开始,并能被指控者推翻是不够的。虽然对抗制框架迫使事实调查者对双方的主张进行比较分析,但 50%的概率假设被告只是在否定原告的主张,而实际上被告经常会提出一个或多个反论点。

因此,案件的转折点不是原告陈述的概率真实性的严格角力,而是原告陈述的概率真实性与被告可能的反陈述的概率真实性之间的比例。就网络攻击而言,有人反对优势标准偏向原告,这就犯了贝叶斯概率论的错误;与其假定原告的归责指控绝对属实,不如将这些指控与潜在的概率进行比较,即全球众多潜在行为者中的任何一个都可能对攻击负责。然后,被告国可以参考任何技术或环境依据来怀疑归因。

此外,本应有利于指控国的信息不对称在实践中可能并不那么有利,因为事实认定者往往会对隐瞒信息的当事方表示更大程度的怀疑。这一点在国际法院审理的国家对国家的裁决中得到了具体研究。在这些裁决中,国际法院对通常以安全为由的证据隐瞒做出了回应,对间接证据做出了有利于无法获得被隐瞒证据的一方的宽松解释(Scharf & Day, 2012)。从逻辑上讲,国际法院行动背后的原则也适用于其他形式的国际裁决。考虑到许多国内法院倾向于从一方当事人隐瞒证据的事实中做出积极的不利推论,国际法院的回应对隐瞒证据做出了相当温和的反应(Nance, 2010)。因此,将优势证据标准适用于归因法时,不会造成原告的不公平优势。

一般来说,优势证据标准符合归因的目标,因为它提供了威慑和信息生

产的最佳平衡;较低的举证责任降低了归因的障碍(因此,增加了反制措施的潜力),同时仍然需要必要的说服力,以激励生产有关网络攻击的相关情报和信息。在提议或威胁以军事打击作为反制措施的情况下,归因法应将举证责任提高到合理怀疑标准,这与美国刑法采用该标准的原因相同。合理怀疑标准承认,伴随着如此严重的惩罚,错误率极不相称,正如错误的刑事处罚可能带来极不相称的、不可容忍的伤害一样,错误的军事冲突也会带来同样的伤害,而且范围和规模可能会成倍增加。

3. 将非国家行为体的网络攻击归因于国家:国家责任理论

迄今为止,归因法采用的是对抗制,遵循类似美国和英国法律制度的程序流程,包括提起诉讼的规则、申诉和答辩的前后顺序以及提供证据的对抗性发现框架。它还有一个一般的举证标准,用于确定一方何时成功证明另一国对发动网络攻击负有责任。但是,如果一个国家将责任归咎于碰巧在其境内活动的"非国家行为者",从而为自己辩护,使自己免于承担责任呢?法律是否应将非国家黑客的恶意活动归咎于国家?

这对归因法和网络安全法来说是一个特殊的问题,因为进行网络攻击的成本相对较低,这为无数非国家行为者提供了选择(Nye, 2010),他们可能出于各种动机采取行动。所有与简单归因攻击相关的典型问题都有可能进一步削弱黑客个人与可能将该行为者与国家联系在一起的任何指挥链或控制基础设施之间的联系。因此,归因法必须解决一个不可避免的结果,即追踪到黑客个人,并面对如何为法律责任目的将此人与国家联系起来的问题。

国家责任理论是一个超越网络攻击领域的法律问题,此前在其他情况下也曾涉及。国际法已经有了将非国家行为者的恶意行为归咎于国家的国家责任理论,2001 年《国家对国际不法行为的责任条款草案》规定了国际法院追究国家对非国家行为者责任的方式。《国家对国际不法行为的责任条款草案》第四、八条随后被国际法院认定为习惯国际法,法院、评论家和其他利益方也广泛承认这些条款阐述了习惯国际法中国家责任理论的标准观点。例

如,《塔林手册》第 1 版和最近发布的第 2 版都大量借鉴了联合国国际法委员会的条款草案,在网络攻击的背景下提出了国家责任理论的概念。

《国家对国际不法行为的责任条款草案》认为,如果非国家行为者是作为国家机关行事,或在国家的指示、指令或控制下行事的,则该非国家行为者的不法行为可归于国家。第四条规定:

1. 任何国家机关,不论行使立法、行政、司法职能,还是任何其他职能,不论在国家组织中具有何种地位,也不论作为该国中央政府机关或一领土单位机关而具有何种特性,其行为应视为国际法所指的国家行为。

2. 机关包括依该国国内法具有此种地位的任何个人或实体。

正如第四条评注所明确指出的,第四条也适用于可被视为事实上的国家机关的个人。同时,第八条也认为,如果非国家行为者是“按照国家的指示、在国家的指挥或控制下行事”,则其行为可归于国家。第四、八条所述国家责任的条件,一般被理解为检验国家是否控制了个人行为者的标准。这些控制测试反过来又与国际法院等在裁决中采用的控制测试相呼应。

然而,关于国家责任的现行国际法也存在一些局限性。例如,奥娜·哈撒韦等人(Hathaway, Crootof & Levitz et al., 2012)批评现行框架产生了不正当的激励机制,只要国家保持最低限度的监督,就可以将非法任务交给非国家行为者,从而逃避责任。他们还认为,控制测试实际上抑制了控制无赖或恶意行为的努力,因为实施控制的尝试可能会产生足够程度的控制,使国家对非国家行为者做出的不法行为负责,尽管国家努力对其进行监管。彼得·马格里斯(Margulies, 2013)批评了适用于网络攻击归责任务的国家责任理论的范围,提出条款草案的控制测试要求具体、全面控制的高标准,这样的标准将是排除国家指挥非国家行为者进行网络攻击的非常重要的例子。

这些意见不仅是批评性的,也是建设性的。哈撒韦等人(Hathaway, Crootof & Levitz et al., 2012)、马格里斯(Margulies, 2013)提议进行调整,以弥补国家责任规则中的这些缺陷。马格里斯建议采用“虚拟控制测试”,即“当

一个国家资助和装备一个私人实体或个人,而该私人实体或个人随后参与网络攻击时,该国家就有责任证明其对网络攻击不负有责任”。根据这一检验标准,马格里斯似乎要求有一些初步证据表明被指控的国家与非国家实体有关联。然而,这种建议的国家责任方法与哈撒韦等人所描述的现行制度下的国家责任方法存在一定的风险,即政府与非国家行为者之间的任何现有关系都可能附带责任,这反而会激励政府放弃对其范围内的非国家行为者的任何控制。马格里斯可能会反驳说,“资助和/或装备”的要求意味着虚拟控制测试只要求政府在对此类实体提供实质性支持的情况下行使此类监督,政府自然有动机在各种情况下资助非国家实体,而在政府这样做的情况下,就应该推定期望监督。这一论点的问题在于,马格里斯对“资助和/或装备”非国家实体的定义过于模糊,这些术语潜在的广泛范围实质上抹杀了对将个人不法行为归咎于国家能力的限制。

当然,这些问题很容易通过更具体地界定这些术语来解决。另外,哈撒韦等人提出的为国家责任主张提供积极辩护的建议也可以补充解决不正当激励的问题。他们提出了一项类似的广泛义务,即代表国家“确保尊重”《日内瓦公约》第一条,确保其范围内的非国家行为者不参与网络攻击。虽然这种方法同样会引起人们对激励国家保持距离而不是进行监管的担忧,但他们提出了这样的观点来解决这一担忧:如果国家能够证明他们采取了“合理的措施”来防止违反国际法,那么国家就应该有一个积极的辩护理由。

通过将这些建议纳入程序规则,归因法不仅可以推进国家责任理论,而且可以用哈撒韦等人、马格里斯提出的同样新颖的解决方案成功应对网络攻击归因的新挑战。通过虚拟控制测试结合“合理注意”的肯定性抗辩,非国家行为者与他们所联系的国家之间建立起更为合理的联系,应允许归因法将个人的网络攻击归于国家,同时允许国家在采取善意措施防止不法行为时,以适当的方式保护自己免于承担责任。

4. 敏感情报及其举证规则

假设一个国家遭受了网络攻击,并希望提出法律指控,将攻击归因于另

一个国家。在目前的情况下,国家已经知道了提起诉讼的程序,以及起诉和答辩、简易判决辩论和出示证据的前后顺序。在最后一步,国家遇到了一个问题:如果国家所依赖的证据有很大一部分来自秘密情报,那该怎么办?此外,国家可能有看似合理的事实依据来确定攻击的归因,但可能出于正当理由不想披露这些证据,因为网络攻击者可以从这些归因点中吸取教训,避免在未来留下相同的线索。归因法面临的挑战是,如何协调提出此类证据的必要性与国家对其机密情报及其来源保密的愿望。

对抗制在一定程度上解决了这一难题:由于当事人可以控制指控的推进,一种答案是干脆将这一问题置之不理。根据成本效益考量,该观点认为,这种披露是寻求对网络侵犯者进行指控的代价,完全由国家来权衡指控的利益与披露其秘密情报能力信息的成本。这种方法的问题在于,它假定网络攻击的成本纯粹是受攻击国家的内部成本。然而,如果我们将网络攻击理解为一种普遍的、全球性的、反复出现的现象(Hackett, 2016),而且一个国家的网络攻击行为如果不加制止,就会在未来对其他国家进行网络攻击,那么归因行为(以及它能够采取反制措施阻止未来的攻击)就会产生积极的外部效应,而上述观点则忽略了这一问题。

因此,归因法应努力兼顾国家的保密和归因利益,找到一种既能允许国家将敏感情报作为证据提出,又能对公众保密的方法。这并不是法院第一次在法庭上处理敏感情报的作用问题。长期以来,法院一直在国家的敏感安全关切与法院的公共角色之间寻求平衡,并开发了许多管理工具来保护听证会上产生或使用的信息。归因法可以纳入两个主要程序,以满足国家保护机密信息的愿望:首先,法院可以制定单方面和不公开审理证据的程序;其次,当法院的记录包含机密信息时,法院可以封存这些记录。

当有必要在法庭上证明某项主张时,一些国家的法院采用这种程序来确保机密信息的安全。在美国,1978 年的《外国监视情报法》(FISA)设立了外国情报监视法院,负责审查联邦执法人员和情报人员的监视令申请。外国情

报监视法院的诉讼程序是单方面和秘密进行的,其裁决很少公开(Clarke,2014)。这些程序性举措并不局限于专门法院。美国《机密信息程序法》允许美国法院在审理刑事案件时对机密信息进行单方面和秘密审查,以确定该证据是否对公平审判或刑事正当程序的要求至关重要(Manget,2006)。一般来说,在向联邦法院提起民事诉讼时,《联邦民事诉讼程序规则》第二十六条允许在有充分理由的情况下封存法庭记录。

其他国家也有类似的程序来保护审判程序或证据。英国于2013年通过了《司法与安全法》,规定了秘密材料程序(CMPS),即只有法官和经过专门安全审查的辩护律师才有权接触案件中涉及的任何敏感情报的秘密法庭听证会。同样,荷兰的《受保护证人法》也规定了一项特别程序,由特别治安法官听取荷兰两大情报机构代表的意见,以决定某些信息是否应保密,或某些证人的身份是否应匿名。这种证据在荷兰的行政、民事和刑事案件中都有使用,而且这种程序与美国的情报监视法庭一样,主要是单方面和不公开进行的,但在特别法官评估敏感情报时,案件当事人有可能在场。德国和西班牙则禁止在审判中使用秘密证据,但有时也允许根据秘密证据提供证词或匿名信息。

单方面程序和不公开程序在许多方面有利于归因法。增加这些类型的程序为系统创造了灵活性,使事实认定者能够逐案分析敏感情报引起的问题。尤其是单方面程序,它可以使事实认定者与当事人就信息披露问题进行谈判,因为当事人可能倾向于高估披露自己信息的成本,这是一种损失规避(Kahneman, Knetsch & Thaler, 1991)。不公开程序允许敏感证据获得充分的证据价值,同时更普遍地降低了信息披露的成本。

在法律程序中制定保密规则也是有代价的。法律程序的透明度往往会赋予其更强的合法性,而保密则可能会破坏这种合法性。此外,如果归因法的首要目标之一是在国际社会眼中证明反制措施的正当性,那么秘密听证会可能会让国际社会中的许多人对反制措施的正当性产生怀疑。

这一棘手问题关系到法院或法律判决的权威来源。诚然,司法程序的公开展示可能因其透明度而有助于程序的合法性,但这并不意味着这种公开性对司法合法性的约束力具有决定性作用。毕竟,前面讨论过的国家都成功地将保密措施纳入其法律体系,而没有损害其法律裁决的合法性(Butler, 2013)。当然,这些机构并不是一开始就采用不公开的诉讼程序,其中大多数机构也不是将大部分案件都置于不公开的诉讼程序之后。社会之所以接受某些记录的保密,可能是因为这些司法机构已经通过长期的普遍公开诉讼程序建立了合法性。

如果公众普遍信任国际机构及其国内法院,这种信任就会转化为对国际法院的支持(Voeten, 2013)。这一结论与法学界对权威的更广泛的解释是一致的,即主张权威的是法院的职务或机构,而不仅仅是纯粹的说服力(Raz, 1985)。毕竟,在美国,提交给联邦上诉法院的绝大多数案件都是通过未公开的"无意见"命令终止的,这表明法律争议的解决并不要求法律程序有一个纯粹透明的窗口(Raz, 1985)。

事实上,尽管使用了秘密程序,其他国际法院仍保持了其合法性。例如,欧洲人权法院在"A 诉联合王国案"[①]中就遇到了这一问题,欧洲人权法院在该案中审查了联合王国允许根据包含"秘密材料"的证据拘留个人的程序。国内和国际法院积累的这些经验表明,归因法可以采用不公开和单方程序的方法。当然,这些程序必须审慎使用。尽管如此,有了审查保密材料的程序,国家就可以在声称归因时提出敏感情报,同时又能保护情报的机密性。

(二) 归因法法律框架的总结

总之,拟议的归因法具有以下特点。第一,它基于对抗制而运行,在这种机制中,诉求和记录都主要由诉讼各方形成。第二,与对抗制框架相一致的是,程序规则以典型的对抗制法律程序的前后方式对案件的各个阶段进行时

① A v. United Kingdom, 49 EHRR 29 (2009).

间排序。第三,在进入案情实质阶段时,指控国必须以优势证据证明其归因主张,除非指控国希望采取军事反制措施。如果一个国家没有披露其计划采取的反制措施,或者这种选择尚不确定,那么案件可以按照优势证据标准进行审理,但这不足以成为日后采取军事行动的理由。第四,为履行举证责任,各国可选择采用不公开审查、单方听证和封存记录等程序,以使用敏感证据证明其主张。第五,也是最后一点,证明归因主张的国家需要具体证明攻击与受国家指使或控制的个人有关,其中控制测试将根据马格里斯主张的"虚拟控制测试"进行善意解释。与此同时,国家将有积极的抗辩理由,证明其在监管相关非国家行为者方面尽职尽责。

四、实施归因法的模式

随着归因法框架的制定,如何才能使这一理论更加充实和生动呢?本节将讨论归因问题中政策性较强的方面,主要探讨机构设置问题:在哪里进行判决,由谁来判决?当然,国家遵守国际制度或法律的问题本身就是一个宏大的讨论主题(Goldsmith & Posner, 2005)。从康德哲学(Tesón, 1992)到理性选择理论(Guzman, 2002),对国际法和国际制度的结构性解释不一而足。在刑法(Simmons & Danner, 2010)、环境法(Brodansky, 1999)和人权法(Simmons, 2009)等几乎所有领域,都出现过关于国家遵守特定主题的讨论。

虽然本文可以就国家参与的动机提出一般性的结构分析,但国家合作或遵守的问题既是一个法律问题,也是一个政治问题。为了对国家如何参与到这样一个法律框架中提出一个完全具有预测性的主张,我们的建议必须借助于:国际关系,既包括广义的理论层面,也包括针对当前历史时刻的具体分析;行为经济学,分析激励机制、成本以及不同参与者的行为概率;对许多参与者的具体历史和心理分析,这些参与者可能是建立这样一个法律制度的重要因素。

全面回答国际履约挑战所提出的问题超出了本文的范围。相反,本文采取了一种更为温和的方法,即通过调查各种其他形式的国际裁决来讨论国家参与的一般激励因素。因此,本文研究了三个国际裁决的例子:国际法院、世界贸易组织争端解决机制,以及像美国—伊朗求偿法庭这样的特设系统。每个机构都反映了不同的国际裁判方法,为国际机构如何成功促使各国参与其体系提供了范例。国际法院提供了将归因法纳入具有广泛属事管辖权的现有框架的选择,世界贸易组织争端解决机制反映了一个具有专门主题的裁决体系,而美国—伊朗求偿法庭则是一个特设的、国家对国家的方法的典范,它可以更灵活地解决两个特定国家之间的冲突,但缺乏创造更持久的法律权威的力量。

(一) 国际法院

国际法院是国际法律机构的典范。国际法院于 1946 年根据《联合国宪章》成立,是 20 世纪大部分时间内唯一存在的国际法院(Dupuy, 1999)。因此,国际法院不仅是创建新的国际法律体系的典范,而且还提供了一个可纳入归因法的现有典例。一般而言,国际法院拥有广泛的属事管辖权,可审理向其提出的任何国际法指控,只要该指控是在双方同意的情况下提出的。将归因法纳入国际法院的好处,是将归因法附着在一个已有公信力、机构历史以及充分发展的规则和资源的现有机构上。

在国际法院成立之前,人们曾多次尝试建立国际机构来解决国与国之间的争端。例如,1899 年海牙和平会议后成立了常设仲裁法院(PCA)(Kolb, 2014)。尽管名为“常设仲裁法院”,但它并不是一个永久性的常设法院,而是提供了一个行政组织,各国可以从候选人库中挑选仲裁员,并建立自己的法庭来解决争端。尽管常设仲裁法院提供了一套程序规则,但这些规则只是默认规则,缔约国自行选择制定的规则将推翻这些规则。1899 年常设仲裁法院成立后,1907 年召开了一次后续会议,包括美国在内的几个州提议成立一个

真正的常设法院。

虽然1907年的提案在当时未能获得支持,但第一次世界大战造成的破坏推动了建立国际法院的进程,最终形成了国际法院的前身——常设国际法院。在25年的任期内,常设国际法院做出了32项判决,所有判决均已执行。在此期间,常设国际法院还发表了27项咨询意见,其中大多数咨询裁决都得到了各国的遵守或采纳。总之,常设国际法院为后来的国际法院奠定了成立的基础。

国际法院是随着1946年《联合国宪章》的制定而成立的,其模式沿袭自常设国际法院。国际法院由安全理事会和联合国大会选出的15名法官组成。这些法官通过单独选举产生,任期9年,选举的重点是法官个人而非其国家的代表。国际法院还纳入了一系列规则,以确保其司法机构的独立性,其中包括:要求法院成员庄严宣誓在履行职责时保持公正的规则;国际法院进一步努力消除潜在的利益冲突,禁止其成员在担任法院法官期间"行使任何政治或行政职能,或从事任何其他专业性职业"。

《国际法院规约》第三十四至三十八条规定了国际法院的管辖权,使其有理由审议有关下列事项的所有法律争端:条约的解释;任何国际法问题;任何一经确定即构成违反国际义务的事实的存在;因违背国际义务而应做出的赔偿的性质或程度。

虽然国际法院目前尚未审理有关网络攻击的国际争端,但网络攻击和归因法当然涉及属于国际法院职权范围的法律问题。网络攻击有可能上升到违反第二条第四款的武装力量的程度,同时也有可能违反国家主权和中立的理论。归因作为网络攻击的一个必然附属问题,牵涉到此类国际法问题。

基于以上概述,下文将进一步讨论国际法院的成立因素,以及对构建、实施归因法的启示。我们很难将国际法院(及其前身常设国际法院)的创立与促成这两个机构诞生的历史割裂开来。毫无疑问,第一次、第二次世界大战推动了国际法院的创立,也对与之相关的国际组织的创立产生了重要影响。

作为一个历史问题,它们似乎在讲述国际法因国际悲剧而产生的故事。这一历史叙事既令人鼓舞,又令人不安。说它令人鼓舞,是因为它表明各国有可能创建新的国际法和国际机构,以应对网络攻击和全球网络安全等当代挑战。令人忧虑的是,可能只有当这些挑战发展到导致国际灾难或造成广泛伤害的事件时,各国才不得不创建此类机构。当然,这种广义的概括并不是实际执行归因法的全部。毕竟,像爱沙尼亚网络攻击这样更加本地化的事件已经刺激了像共同制定《塔林手册》及其续篇这样的团体,暗示了先发制人而非被动执行国际法的可能性。

(二) 世界贸易组织争端解决机制

实施归因法的第二种模式是通过世界贸易组织争端解决机制这样的机构。与国际法院模式不同的是,世界贸易组织的争端解决体系是一种将裁决程序附于一个具有特定主题的国际机构的模式。采用这种模式的好处是,通过一个专门的事实认定机构来执行归因法,这些机构可能最有能力解决证据和技术的技术复杂性问题,而国家及其专家正是通过这些证据和技术将恶意数字活动追溯到其创造者的。

世界贸易组织是根据《马拉喀什建立世界贸易组织协定》成立的,该协定是 1994 年乌拉圭回合谈判达成的几项协定之一。世贸组织的成立总体上是为了促进和监督全球贸易,而世贸组织的争端解决体系是《马拉喀什建立世界贸易组织协定》第三条明确规定的职能之一,旨在帮助该机构实现这一目标。与此同时,《关于争端解决规则与程序的谅解》(DSU)对世贸组织争端解决程序的结构和程序做出了更精确的规定。根据《关于争端解决规则与程序的谅解》第一条,争端解决程序可适用于一些特定协定下的争端,包括 1994 年《多边货物贸易协定》和《与贸易有关的知识产权协定》。

争端解决程序由争端解决机构(DSB)管理,该机构负责监督世贸组织争端解决小组的运作及其裁决的执行。专家组的实际职能由《关于争端解决规

则与程序的谅解》规定的规则决定。这些规则包括设立裁决专家组的规定、专家组的组成、专家组程序以及每个专家组如何执行其裁决程序的各种其他基本规则。例如,《争端解决谅解书》规定了启动争端解决小组的条件,指出争端解决机构应在申诉方"以书面形式"提出请求时设立一个解决小组,而且这种请求"应说明是否进行了磋商,确定有争议的具体措施,并简要概述申诉的法律依据,足以清楚地说明问题"。此外,《关于争端解决规则与程序的谅解》还对专家组的成员构成进行了规范,规定了专家组成员不得来自争端当事国(除非双方另有规定)。在决策过程方面,《关于争端解决规则与程序的谅解》的条款还要求专家组为其决定制定具体的时间表,规定了具体的审查阶段和这些具体阶段的程序,并规定了专家组可以审查或咨询的信息类型。因此,《关于争端解决规则与程序的谅解》规定了一个全面的裁决制度。

这样一个机构在促使国家参与和遵守方面的有效性引起了学术界的关注。关于国家参与的问题,一个更加专业化的论坛可能会引起这样的担忧,即在该主题领域拥有既得利益的强国可能会将这样一个机构仅仅作为一种施展其影响力的手段。例如,查德·鲍恩(Bown, 2005)的一项实证研究表明,一个国家的报复能力、法律能力以及在国际政治经济关系中的作用对衡量该国参与争端解决体系的可能性具有重要意义。鲍恩的研究结果让人担心,专门机构可能只是强国将其在某些领域(如贸易或网络安全)的主导权制度化的工具。当然,这个问题可能只是国际力量不对称的一个特征,也可能是财富不平等更普遍地影响法律的结果。

归根结底,即使存在对某些国家的参与偏向,如果法律体系的价值不仅仅在于为一方或另一方裁决诉求,而是在于法律制度带来的积极外部效应,即在国家间创造更大的可预见性与合作,那么参与的偏向可能是一个可以容忍的代价。其他实证研究表明,此类法律确实提供了这些积极的外部效应。例如,迈克尔·贝克特尔等人(Bechtel & Sattler, 2015)发现,申诉方和签署申诉方向世贸组织提出的申诉的被动第三方所获得的经济利益差别很小。这

些结果表明,“较弱”的国家可以选择搭乘“较强”国家的便车,从贸易增长中获益,而且裁决程序产生的外溢效应可能会使更多国家受益。如果世贸组织的争端解决程序能够有效地促使争端各方遵守规定,那么这一程序所产生的遵守规定的情况以及随之而来的积极的外部效应,很可能说明国际裁判制度是成功的。

世界贸易组织的争端解决机制不仅提供了一个国际裁判的范例,也为某些法律制度如何被纳入国际机构并获得更广泛的认同提供了一个实用的经验。在《私权、公法:知识产权的全球化》一书中,苏珊·塞尔斯(Sells, 2003)追溯了《与贸易有关的知识产权协议》如何被纳入世贸组织结构的历史。在这一历史叙事中,塞尔斯提请人们注意“这场戏剧中的核心角色”——“总部设在美国的 12 名知识产权委员会成员”,该委员会由代表不同行业的 12 名首席执行官组成。因此,集中游说在将某些监管制度落实到国际法中,以及在动员各国充当此类制度的有力倡导者方面可以发挥突出作用。鉴于网络攻击对私营商业实体构成的风险越来越高,以雅虎网络攻击为例,商业公司肯定有机会在游说中发挥突出作用,以成功地将拟议的归因法等国际制度制度化。

(三)大规模赔偿委员会(美国—伊朗求偿仲裁庭)

实施归因法的第三种模式是通过特设法庭,如 1981 年成立的美国—伊朗求偿仲裁庭。美国—伊朗求偿仲裁庭是纯粹双边大规模求偿委员会的一个范例,它是通过两国之间的条约而产生的。与前两种模式不同的是,法庭制度是针对双方之间的一系列具体指控要求而产生的。这种方式的优点是灵活性强,可以根据具体情况和所涉各方的具体要求来实施。但其不足在于效力范围有限,这种效力受限既体现在管辖当事方,也体现在仲裁庭可审理的历史事件方面。

该仲裁庭是解决伊朗人质危机协议的一部分(Mosk, 1987)。在 1979 年

的革命中,伊朗人袭击了美国驻德黑兰大使馆,劫持了 69 人(Sahimi, 2009)。虽然一些人质获释,但仍有 52 人被关押了 444 天。《阿尔及尔协议》帮助促成了美国和伊朗之间的协议,伊朗将释放美国人质,以换取美国取消贸易制裁并解冻伊朗的一些资产。重要的是,《阿尔及尔协议》还试图解决美国公民对伊朗提出的以及伊朗公民对美国提出的大量私人指控。《阿尔及尔协议》解决这些问题的方式,是将其从诉讼转为仲裁——因此成立了法庭。

《指控解决宣言》正式成立了法庭,包括其管辖权、组成和仲裁规则等条款。在管辖权方面,法庭仅限于审理两类指控:美国国民对伊朗的指控和伊朗国民对美国的指控,以及构成该国民指控标的物的同一合同、交易或事件所引起的任何反诉;美国和伊朗因双方之间的货物和服务购销合同安排而对对方提出的官方指控。在确定裁决人时,《指控解决宣言》确定仲裁庭由 9 名成员组成:美国任命 3 名,伊朗任命 3 名,然后由这 6 名成员任命仲裁庭的最后 3 名成员。

在程序方面,仲裁庭采用了联合国国际贸易法委员会的仲裁规则。这些规则又制定了一套全面的程序,对审理的各个阶段做出规定,包括进行质证和出示证据的方法。这些规则还为仲裁庭适用各种程序机制提供了很大程度的灵活性和自由裁量权,如何时或如何纳入专家证据(Bockstiegel, 1986; Straus, 1986)。因此,纳入贸易法委员会的规则提供了一个实例,说明如何将先前存在的一套规则纳入或借鉴到特定的特设裁决机构中,这反过来又为特设机构如何对归因法采取同样的做法提供了类似的可能性。

总体而言,美国—伊朗求偿仲裁庭成功地处理了双方的大量求偿。美国提出的几乎所有索偿都得到了裁决,那些有利于美国索偿者的索偿要求都得到了全额支付(Brower, 1992)。伊朗方面,美国于 2016 年同意支付 17 亿美元的和解金,以解决伊朗的一项长期指控。因此,一种观点认为,该仲裁庭发挥了重要作用。举例而言,仲裁庭自成立以来已处理了 3900 余起案件,这些案件总体上涵盖了美伊两国间除少数大型、复杂指控外的所有案件。除了处

理的案件数量,理查德·莫斯克(Mosk, 1987)等人还称赞仲裁庭有能力切实、成功地执行一整套裁决案件的程序规则,这些规则有助于有效地处理复杂的案件,其程序"可作为未来法庭的指南"。事实上,仲裁庭还在其他方面发挥了指导作用。克里斯托弗·吉布森等人(Gibson & Drahozal, 2006)的研究表明,美国—伊朗求偿仲裁庭的裁决已被解决投资争端中心法庭援为先例,这表明特设法庭的裁决仍可在其裁决的直接争议之外产生更广泛的影响。

然而,以临时方法提出归因指控是有局限性的。尽管美国—伊朗求偿仲裁庭的裁决已被其他法庭引用,但对仲裁引用的更普遍调查表明,仲裁法庭对案例的引用往往因背景不同而有很大差异。虽然《国际货物销售合同公约》和国际商会对先前裁决的引用相对较少,但体育仲裁法庭和域名仲裁系统在其裁决中对先例的引用几乎无处不在。在归因问题上,很容易看到这些裁决会走向前者的道路。鉴于网络攻击归因案件的事实差异很大,从网络攻击的类型到国家对支持归因的证据的保密程度,法庭很可能不愿意过多地依赖以前的案例,因为它们可能存在事实差异。

特设法庭在建立参与激励机制方面也面临着特殊的挑战。由于特设法庭经常产生于双边协定,它们依赖于各国拥有(或相互视为拥有)相对平等的地位。此外,它们还取决于特定的历史背景。在这种背景下,每一个国家都对另一个国家有足够的不满,从而提供了首先成立这样一个法庭的动机。在网络攻击的背景下,这种情况当然有可能发生,各国可能相互进行了网络攻击,但很难想象一个国家会主动承认自己的责任,并带着有序解决的愿望与对方接触。尤其难以想象,在这种情况下,各国拥有足够平等的影响力,从而产生迫使双方坐到谈判桌前的局面。虽然前两种模式确实只能对已经发生的攻击进行裁决,但常设司法机构本身就代表了一种时间上的持久性,使其裁决能够对未来投下更大的影响。因此,特别模式虽然在可能需要它的特定事实情况下最有效,但在执行归因法方面却是一个不太有效的模式。

五、结论

在描述国际法院的起源时,罗伯特·科尔布(R. Kolb)将其发展历程分为三个部分:组织全面的仲裁司法计划;试图建立一个永久性和强制性的“仲裁法院”;建立一个与国际联盟有联系的机构法院——常设国际法院。

至关重要的是,创建这一制度的第一步是创建法律方案。每迈出一步,法律的愿景就会逐渐变得更加具体,直到这一愿景成为实际的法律制度。本文提出的归因法旨在描绘国家如何通过法律应对网络攻击威胁的愿景。当然,与最初的国际法院概念相比,归因法是一个更为温和的制度。但这仍然是一个重要的制度,而且由于之前的国际法机构所奠定的基础,这一构想变得更有可能。

本文设想了一个将网络攻击归因于责任国的法律框架,并提出了允许国家合法提出这种主张的程序规则。通过采用对抗制,归因法可使双方平衡就这一不确定主题的举证责任。通过默认的举证责任——以优势证据证明标准归因,归因法可以解释证明归因的技术困难,允许法律承认间接证据何时足以将攻击与其来源联系起来。此外,通过使用虚拟控制测试,归因法可以更广泛地要求国家对与其有联系的非国家行为者负责,并允许禁止抗辩,为对此类行为者实施适当程度监督的国家创造安全港。最后,允许对证据进行单方面和不公开审查的程序规则将照顾到国家对其敏感情报的保密性的关切,同时也保留了在提出归因指控时使用相关证据的能力。

通过这些规则,归因法旨在使网络攻击背后的源头透明化。长期以来,网络攻击一直被蒙上一层神秘的面纱而不受控制,国家长期以来一直可以逃避实施此类攻击的责任(Kaplan, 2016)。虽然像美国这样的国家行为体可能曾一度认为自己在网络战领域拥有不成比例的优势,但网络攻击的日益扩散可能已经超出了任何一个国家的控制范围,不仅威胁到国家的安全,还威胁

到其私人和民间机构的稳定。随着不安全和不确定性的代价与日俱增,各国可能很快就会认识到需要法律机构通过相互问责来约束它们。

然而,近年来国际结构似乎出现了一些裂痕。随着英国脱欧等事件的发生,以及唐纳德·特朗普和玛丽娜·勒庞(M. Le Pen)等支持保护主义政策的个人的日益活跃,过去几十年来作为国际法发展主要特征的国际机构似乎正在退缩。保护主义的威胁因网络攻击的兴起而加剧,尤其是网络攻击在潜在干扰选举政治和国内机构合法性方面的恶性用途。所有这些威胁加在一起,似乎会破坏人们对国家主权和国际法的能力和稳定性的信心。

我们很容易被当下的政治风云所迷惑,而忽视了更长远的发展道路。但是,当今日益加剧的不确定性更加提醒我们需要进一步发展国际法,而不是进一步退缩。想象我们可能实施的新法律框架是一个步骤。但法律理论只是一部分,光有理论是不够的,除非这种理论能够解决、契合现实问题。归因法所规定的程序规则不仅规定了指控成功所必须满足的技术特征,还规定了与之相伴的实际成本。这样,它就将法律制度的成本具体化,以权衡无人管理现状下的不确定性成本。也许国家及其选民可以容忍一个没有法律来遏制网络安全威胁的世界。但如果对归因法可能带来的代价有了更明确的认识,国家可能很快就会意识到,无限制的网络攻击所造成的破坏代价太大,不容忽视。

参考文献

Axelrod, R. 1984, *The Evolution of Cooperation*, New York: Basic.

Bechtel, M. & T. Sattler 2015, "What Is Litigation in the World Trade Organization Worth?" *International Organization* 69(2).

Belton, R. 1981, "Burdens of Pleading and Proof in Discrimination Cases: Toward A Theory of Procedural Justice." *Vanderbilt Law Review* 34(5)

Beverly, R., A. Berger & Y. Hyun et al. 2009, "Understanding the Efficacy of Deployed Internet Source Address Validation Filtering." http://Perma.Cc/Ys79-Zpjr.

Bockstiegel, K.-H. 1986, "Applying the Uncitral Rules: The Experience of the Iran-United States Claims Tribunal." *Berkeley Journal of International Law* 4.

Boerbert, W. 2010, "A Survey Of Challenges in Attribution." https://api. semanticscholar. org/CorpusID:2445682.

Bown, C. 2005, "Participation in WTO Dispute Settlement: Complainants, Interested Parties, and Free Riders." *The World Bank Economic Review* 19(2).

Brenner, S. 2007, "At Light Speed: Attribution and Response to Cybercrime/Terrorism/Warfare." https://api. semanticscholar. org/CorpusID:72953571.

Brodansky, D. 1999, "The Legitimacy of International Governance: A Coming Challenge for International Environmental Law?" *American Journal of International Law* 93(3).

Brower, C. 1992, "Lessons to Be Drawn from the Iran-U. S. Claims Tribunal." *Journal of International Arbitration* 9.

Butler, A. 2013, "Standing up to Clapper: How to Increase Transparency and Oversight od FISA Surveillance." https://heinonline. org/HOL/LandingPage? handle=hein. journals/newlr48&div=8&id=&page=.

Chulov, M. &H. Pidd 2011, "Defector Admits to WMD Lies that Triggered Iraq War." http://Perma. Cc/W46Q-Yz2S.

Clapham, A. 2015, "Human Rights Obligations for Non-State-Actors: Where Are We Now?" in F. Lafontaine &F. Larocque eds., *Doing Peace the Rights Way: Essays in International Law and Relations in Honor of Louise Arbour*, Cambridge: Intersentia.

Clarke, C. 2014, "Is the Foreign Intelligence Surveillance Court Really A Rubber Stamp? Ex Parte Proceedings and the FISC Win Rate." *Stanford Law Review* 66.

Clermont, K. &E. Sherwin 2002, "A Comparative View of Standards of Proof." *The American Journal of Comparative Law* 50(2).

Condliffe, J. 2016, "Security Experts Agree: The NSA Was Hacked." https://www. technologyreview. com/2016/08/18/70386/security-experts-agree-the-nsa-was-hacked.

Connolly, K. 2016, "German Spy Chief Says Russian Hackers Could Disrupt Elections." http://Perma. Cc/N5Em-3Gvc.

Devine, D., L. Clayton &B. Dunford et al. 2000, "Jury Decision Making: 45 Years of Em-

pirical Research on Deliberating Groups." *Psychology, Public Policy, and Law* 7(3).

Dupuy, P. 1999, "The Danger of Fragmentation or Unification of the International Legal System and the International Court of Justice." *International Law and Politics* 31.

Feigenbaum, J., A. Johnson & P. Syverson 2007, "A Model of Onion Routing with Provable Anonymity." *Financial Cryptography and Data Security* 4886.

Fleming, J. 1961, "Burdens of Proof." *Virginia Law Review* 47(1).

Frank, J. 1949, *Courts on Trial: Myths and Reality in American Justice*, Princeton: Princeton University Press.

Froeb, L. & B. Kobayashi 2001, "Evidence Production in Adversarial VS. Inquisitorial Regimes." *Economics Letters* 70(2).

Gallagher, S. 2013, "Turkish Government Agency Spoofed Google Certificate 'Accidentally'." http://Perma.Cc/L5Zy-2Tvv.

Gersen, J. & A. Vermeule 2016, "Thin Rationality Review." https://michiganlawreview.org/journal/thin-rationality-review.

Gibson, C. & C. Drahozal 2006, "Iran-United States Claims Tribunal Precedent in Investor-State Arbitration." *Journal of International Arbitration* 23(6).

Goldsmith, J. & E. Posner 1999, "A Theory of Customary International Law." *The University of Chicago Law Review* 66(4).

Goldsmith, J. & E. Posner 2005, *The Limits of International Law*, https://doi.org/10.1093/oso/9780195168396.001.0001.

Graham, D. 2010, "Cyber Threats and the Law of War." *Journal of National Security Law & Policy* 4.

Greenemeier, L. 2011, "Seeking Address: Why Cyber Attacks Are So Difficult to Trace Back to Hackers." http://Perma.Cc/Fy47-Jcyn.

Grimmelman, J. 2017, *Internet Law: Cases and Problems*, https://www.semaphorepress.com/InternetLaw_overview.html.

Guzman, A. 2002, "A Compliance-Based Theory of International Law." *California Law Review* 90(6).

Hackett, R. 2016, "Why A Hacker Dumped Code Behind Colossal Website-Trampling Bot-

net." http://Perma. Cc/Bg6V-Dj8U.

Hadfield, G. 2000, "The Price of Law: How the Market for Lawyers Distorts the Justice System." *Michigan Law Review* 98(4).

Hathaway, O., R. Crootof & P. Levitz et al. 2012, "The Law of Cyber-Attack." *California Law Review* 100(16).

Hay, B. & K. Spier 1997, "Burdens of Proof in Civil Litigation: An Economic Perspective." *The Journal of Legal Studies* 26(2).

Holzer, C. & J. Lerums 2016, "The Ethics of Hacking Back." https://ieeexplore. ieee. org/document/7568877.

Kahn, P. 1987, "From Nuremberg to the Hague: The United States Position in Nicaragua v. United States and the Development of International Law." *The Yale Journal of International Law* 12(1).

Kahneman, D., J. Knetsch & R. Thaler 1991, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5(1).

Kaplan, F. 2016, *Dark Territory: The Secret History of Cyber War*, New York: Simon & Schuster.

Kaplow, L. 2012, "Burden of Proof." *The Yale Law Journal* 121.

Keohane, R. 1986, "Reciprocity in International Relations." *International Organization* 40(1).

Kolb, R. 2014, *The Elgar Companion To The International Court Of Justice*, https://lawcat.berkeley. edu/record/369760.

Kötz, H. 2003, "Civil Justice Systems in Europe and the United States." *Duke Law* 1.

Langbein, J. 1985, "The German Advantage In Civil Procedure." *University of Chicago Law Review* 52(4).

Langner, R. 2013, "To Kill A Centrifuge." http://Perma. Cc/8G3X-Gr9K.

Li, J., X. Ou & R. Rajagopalan 2010, "Uncertainty and Risk Management in Cyber Situational Awareness." https://www. academia. edu/89056610/Uncertainty_and_Risk_Management_in_Cyber_Situational_Awareness#:~:text = The% 20uncertainty% 20challenge% 20exists% 20in% 20all% 20three% 20phases, uncertainty% 20in% 20these% 20three%

20aspects%20are%20slightly%20different.

Lind, E. &T. Tyler 1988, "The Social Psychology of Procedural Justice." *Contemporary Sociology A Journal of Reviews* 18(5).

Loomis, A. 2008, *Leveraging Legitimacy In Securing U. S. Leadership: Normative Dimensions Of Hegemonic Authority*, Ph. D. Thesis of Georgetown University.

Manget, F. 2006, "Intelligence and the Criminal Law System." *Stanford Law &Policy Review* 17.

Marcus, M. 1992, "Above The Fray or into the Breach: The Judge's Role in New York's Adversarial System of Criminal Justice." *Brooklyn Law Review* 57(4).

Margulies, P. 2013, "Sovereignty and Cyber Attacks: Technology's Challenge to the Law of State Responsibility." *Melbourne Journal of International Law*, *Melbourne* 14(2).

Mearsheimer, J. 2001, *The Tragedy of Great Power Politics*, New York: W. W. Norton & Company.

Meyer, J. 2016, "Why Experts Are Sure Russia Hacked The Dnc Emails." http://Perma. Cc/Vlk7-Dcby.

Moravcsik, A. 2014, "Liberal Theories of International Law." in J. Dunoff & M. Pollack eds., *Interdisciplinary Perspectives on International Law and International Relations: The State of the Art*, Cambridge: Cambridge University Press.

Mosk, R. 1987, "Lessons from the Hague: An Update on the Iran-United States Claims Tribunal." *Pepperdine Law Review* 14(4).

Nance, D. 2010, "Adverse Inferences about Adverse Inferences: Restructuring Juridical Roles for Responding to Evidence Tampering by Parties to Litigation." https://www. bu. edu/law/journals-archive/bulr/documents/nance. pdf.

Nye, J. 2010, "Cyber Power." http://Perma. Cc/3Mcy-3Bn5.

Nye, J. 2021, "Soft Power: The Evolution of A Concept." https://doi. org/10. 1080/2158379X. 2021. 1879572.

Parisi, F. 2002, "Rent-Seeking Through Litigation: Adversarial and Inquisitorial Systems Compared." *International Review of Law and Economics* 22(2).

Raz, J. 1985, "Authority, Law and Morality." *The Monist* 68(3).

Rid, T. & B. Buchanan 2014, "Attributing Cyber Attacks." *Journal of Strategic Studies* 38(1-2).

Sahimi, M. 2009, "The Hostage Crisis, 30 Years on." http://Perma.Cc/M42L-Gtkw.

Sanger, D. & M. Fackler 2015, "N. S. A. Breached North Korea Networks Before Sony Attack, Officials Say." http://Perma.Cc/P3Bb-Kabj.

Scharf, M. & M. Day 2012, "The International Court of Justice's Treatment of Circumstantial Evidence and Adverse Inferences." *Chicago Journal of International Law* 13(1).

Schmitt, M. 1999, "Computer Network Attack and the Use of Force in International Law: Thoughts on A Normative Framework." *Columbia Journal of Transnational Law* 37.

Sells, S. 2003, *Private Power, Public Law: The Globalization of Intellectual Property Rights*, Cambridge: Cambridge University Press.

Simmons, B. & A. Danner 2010, *Credible Commitments and the International Criminal Court*, Cambridge: Cambridge University Press.

Simmons, B. 2009, *Mobilizing for Human Rights: International Law in Domestic Politics*, New Haven: Harvard University.

Singer, P. & A. Friedman 2014, *Cybersecurity and Cyberwar: What Everyone Needs to Know*, Oxford: Oxford University Press.

Straus, M. 1986, "The Practice of the Iran-U. S. Claims Tribunal in Receiving Evidence from Parties and from Experts." *Journal of International Arbitration* 3.

Strier, F. 1992-1993, "What Can the American Adversary System Learn From An Inquisitorial System of Justice?" https://heinonline.org/HOL/LandingPage?handle=hein.journals/judica76&div=30&id=&page=.

Tanase, M. 2003, "IP Spoofing: An Introduction." http://Perma.Cc/5Ege-9Rdv.

Tesón, F. 1992, "The Kantian Theory of International Law." *Columbia Law Review* 92.

Thirlway, H. 2016, *The International Court Of Justice*, DOI: 10.1093/law/9780198779070.001.0001.

Tomlinson, E. 1983, "Nonadversarial Justice: The French Experience." *Maryland Law Review* 42(1).

Voeten, E. 2013, "Public Opinion and the Legitimacy of International Courts." *The Journal*

Theoretical Inquiries in Law https://doi.org/10.1515/til-2013-021.

Waxman, M. 2011, "Cyber-Attacks and the Use of Force: Back to The Future of Article." *The Yale Journal of International Law* 2(4).

Weber, M. 1968, *Economy and Society: An Outline of Interpretive Sociology*, [S. L.]: [S. n.].

Whitman, J. 2005, *The Origins of "Reasonable Doubt": Theological Roots of the Criminal Trial*, New Haven: Yale University Press.

Williams, K. 2016, "Clinton: Treat Cyberattacks 'Like Any Other Attack'." http://Perma.Cc/L4Ju-4Jzk.

Zetter, K. 2015, *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*, Taipei: Crown.

数字技术支配下的生活世界*

——读杰米·萨斯坎德《算法的力量:人类如何共同生存?》

王国伟**

摘要:随着数字信息技术的广泛运用,当代人类的生活世界正在经历数字化变迁。无处不在的数字技术正成为权力的新来源和新表现形式,为传统权力概念赋予了数字化特征。这种权力的本质变化不可避免地对政治生活中的自由、民主和正义等方面产生了深远影响。因此,针对《算法的力量:人类如何共同生存?》一书试图从技术的透明化和技术权力的分割两个维度突破数字系统控制的思路,本文基于技术的政治经济学角度进行了客观评析。

关键词:数字技术　数字生活世界　数字权力

人类共同生活的方式总是随着科学与技术的进步而发生重大变化。数字信息技术的广泛运用,直接推动人类社会进入数字生活世界,这正是英国学者杰米·萨斯坎德(2022)的新作《算法的力量:人类如何共同生存?》(*Future Politics: Living Together in A World Transformed by Tech*)对数字化时代人类社会生活形态做出的一个重要判断。在数字生活世界中,数字技术不仅重新诠释了权力的来源和形式,也给政治与社会生活领域带来了诸多重大影响。突破数字系统的控制,防范数字技术的风险,已成为我们不得不积极应对的一个共同挑战。

* 本文系国家社会科学基金一般项目"当代西方经济社会学述行理论及其本土化研究"(项目批准号:23BSH167)的阶段性研究成果。

** 王国伟,东北师范大学马克思主义学部。

一、生活世界的数字化变迁

“生活世界”一词源于德语 Lebenswelt,其一般性含义是指“所有构成我们个人或集体世界的即时体验、活动和联系”(萨斯坎德,2022:3)。作为数字技术的大规模运用与生活世界紧密结合而形成的崭新人类生活形态,数字生活世界将人、强大的机器和充足的数据连接在一张相当复杂精密的网络之中。它具有三方面的突出特征。

第一,拥有日益强大的数字系统。以人工智能为代表的智能化数字系统,尤其是深度机器学习算法技术和以几何级数增长的算力,为数字生活世界搭建起了基础性的技术构架。目前,人工智能系统已经在自动驾驶、自然语言翻译、人脸识别、模仿人类说话、复杂的市场战略制定、疾病的筛查和预测、各种技巧和战略性运动项目等方面显示出了超强的能力。借助机器学习技术,人工智能系统还将摆脱其人类创造者的限制。机器学习的算法,尤其是利用神经网络进行的深度学习技术,通过调动大量的数据,可以自行识别、创建模型并进行推断以完成任务,为人工智能系统提供了实现自我、完善自我的机制。加之可用数据的海量增加、以几何级数增长的算力以及拥有强大算力的计算机器的产生,人工智能和机器学习获得强劲推动力。正因如此,数字生活世界中成千上万种原来只有人类才能从事的活动,如今都可以由数字系统更快、更高效、更准确地完成。此外,一旦拥有一定认知和抽象思考能力的通用人工智能或全脑仿真技术获得突破,数字系统还将在更广阔的领域释放更为强大的能力。

第二,日益综合化的技术与我们的日常生活经验密不可分。从私人领域的生活起居、日常穿戴到公共场所的公共服务、智能城市,数字技术正无孔不入地渗透进我们的日常经验,显示出超强的连接性。借助各式各样的智能设备,数字技术极大地增强了人与人之间、人与机器之间、机器与机器之间的信

息交换能力。人类连接方式的更新,带来最重大、最深远的变革集中表现为信息、文化和知识的生产与分发模式的去中心化。在万物互联时代,人和物理世界互动的广度与深度,正随着越来越敏感的传感器探测能力的增加而大幅提升。日益敏感的技术已经从工业生产、市政管理等宏观领域,深入更微观的人类身体内部甚至情绪情感领域,这将促使人与机器的关系发生精神上的巨变。不仅如此,随着增强现实(AR)、虚拟现实(VR)技术的发展,人—机关系还会从根本上变得更具有沉浸性。届时,人和机器、线上和线下、虚拟世界和现实世界之间的界限也将更加模糊。最终,我们可能生活在一个虚实一体的世界里。

第三,数字生活世界还是一个被普遍量化的社会。进入 21 世纪,随着基于印刷的信息系统过渡到数字系统,信息的生产和处理发生了革命性的变化。如今我们生产信息的速度,也像计算机的算力一样呈几何级数增长。初步估算,我们每 10 分钟产生的信息量,等于最初 1 万代人类所创造信息量的总和。由于越来越多的社会活动是通过数字系统或数字平台展开的,加之数字信息的存储成本日益降低、复制成本极为低廉,不仅人们的动作、话语、行为、关系、情感和信念等方面的内容都将作为数据被捕捉和记录,甚至来自自然界、机器行为和建筑环境等领域的数据也会被收集起来。这些有意或无意留下的数字化足迹,经过数据专家的清洗、处理和组织,都会成为用来解读、分析人们的思想、信仰、关切、健康、市场活动、音乐品位及性取向等方面的重要知识,由此可以绘制出一张细微而又丰富的人类生活数字地图。因此,数据是有价值的,收集到的某一方面的数据越多,它的价值就越大,其用途也就越广泛。

二、数字权力的新形式

在数字生活世界中,数字技术与权力紧密地结合在一起,无处不在的数

字技术正成为权力的新来源和新表现形式。所有的数字技术都是由代码语言、算法公式操作的,虽然它们不是法律本身,但是能够精细地约束、指导或操纵我们的行为和感知,因此代码和算法就是力量。而且,随着网络空间与现实空间的界限日益模糊,人们也将越来越多地沉浸在完全由代码和算法构成的世界里。在那里,代码和算法不仅是权力,更是数字生活世界的本质和现实。在此背景下,作为社会秩序基石的基本问题,“何为权力”“权力有何种形式”以及“谁将行使权力”都将迎来巨大的变化。

权力最纯粹、最直接的表现形式是暴力。一直以来,暴力的合法使用权始终掌握在国家及其执法人员手中,然而数字生活世界中暴力的使用将发生三方面的重要变化,从根本上挑战了我们过去对暴力运用方式的认识。第一,暴力的数字化。随着越来越多的人类活动由数字系统实施,一些法律规则由人强制执行转变为由数字系统强制执行,人工智能系统也会赋予法律无须经过某种正式程序就能够根据环境变化做出快速精准的适应性反应的权力;此外,法律将越来越多地以代码而非文字的形式出现,这种数字法将是一种强迫或限制我们行动的全新方式。第二,暴力的私有化。传统的政治理论认为,国家是唯一有权使用或允许其他主体使用暴力的主体。然而,在数字生活世界中,我们时常受到私人来源代码的支配,由代码决定的数字平台准则、嵌入设备的代码规则等技术力量的影响无处不在,这将给掌控技术的科技公司带来私权力。第三,暴力的自动化。在不受人为控制的情况下,施加暴力的自主数字系统也是存在的。

权力并不都是暴力性的,在数字系统作用下还产生了许多相对温和的权力形式,如视觉监视和各种收集、存储、处理他人信息的能力与方式,这些都是萨斯坎德所说的“审查”。这种广泛存在的权力形式,“它能看见别人却又不被别人看见,它能了解其他人,却又能保持自己不被别人所知”(萨斯坎德,2022:89)。进入数字生活世界,作为一种全景式“规训权力”(Foucault,1980: 39)的审查能力与方式也发生了重大变革:第一,当我们越来越多的行

动、表达、动作、关系、情感和信仰被数字媒介技术永久或半永久记录下来时，人类生活的可解读性将大大增加，这意味着无论我们是否愿意都将成为被算法描述的人；第二，各式各样的家用数字技术、便携式设备的运用，促使审查越来越具有隐秘性，甚至已经扩展到私人生活领域；第三，经由审查收集的数字形式的信息，不仅比以往任何形式的信息都更容易储存和复制，而且它还是不朽的，可以贯穿整个人生；第四，功能强大的数字系统，将促使我们生活各方面可预测性的广度和精度变得越来越大；第五，人们的生活将会越来越多地由分数、评级和排名进行评估，而评级结果决定着人们是否能够获得有价值的社会物品。作为以上五个方面变化的累积结果，数字技术带来了审查能力的增强与方式的革新，从而极大地强化了规训我们的权力。

权力的行使还表现在对人们的所知所想以及关于世界的看法进行“感知控制”，这一点其实早已被马克思及其追随者的意识形态理论所揭示。著名网络社会学家卡斯特（Castells，2013：3）在《传播力》一书中曾深刻地指出：“我们感受和思考的方式决定了我们行动的方式。”因此，那些控制感知的人，将会更多地决定我们感受和思考的方式，进而决定我们的行动方式。对信息的过滤始终是一种强大的感知控制手段，“如果你控制了一个社会的信息流动，你就可以影响这个社会对是非、公不公平、干不干净、得体不得体、真实还是虚假、已知还是未知的共同观念”（萨斯坎德，2022：109）。以印刷、广播、电视为代表的大众传媒，是20世纪信息过滤的日常形式。然而，随着移动互联网的广泛运用，社交媒体和数字新闻平台让人人都成为信息的生产者和消费者，而且互联网也被用于更精确、更广泛地控制我们传递和接收的信息，这些变化意味着传统大众媒体垄断过滤手段的终结。随着数字生活世界的到来，人们所感知的世界将越来越多地由以算法为基础的数字系统向我们揭示或隐藏的内容而决定，数字化的过滤器主要包括各种量身定制的新闻、搜索引擎、通信邮件、情绪以及即时性的感官体验等。总之，基于数字技术产生的感知控制，是一种难以意识到施加在人们身上的权力形式，它看似更“弱”却能

够悄无声息地影响人们的情绪、情感、价值观念以及行为方式。由于过滤方式的不同,人们看待世界的方式也将各不相同,不可避免地会导致观念冲突问题,这将是感知控制更为深层的影响。

总之,生活世界数字化的发展趋势,给传统的权力概念带来了新的特征。笼统而言,权力概念包括控制权、行动权两方面(Mill, 2009:93),但萨斯坎德(2022:60)在该书中更多关注的是人与人之间的权力关系,即控制权方面的转变:首先,以暴力、审查和感知控制为代表的权力形式虽然古已有之,但是数字技术赋予了它们远超人类历史上任何工具所能给予的力量。其次,权力并非必然与暴力、武力结合在一起,但是"谈论非暴力的权力事实上是多余的"(卢克斯,2008:23)的论断同样站不住脚。事实上,在数字技术的加持下,权力的非暴力形式将会更加丰富多样,其作用方式将会越来越精确,控制范围已经从我们外在的身体、行为扩展到内在的所知、所感和所欲。因此,数字权力将实现对我们的全景式控制。最后,由于大型科技公司掌握着数字技术,它们也将拥有真正的权力。不过,这是一种追求私人经济利益的私权力,它不可避免地会强化科技公司的政治影响力。

三、技术权力的多重政治影响

萨斯坎德(2022:xxi)认为,权力在本质上的变化将影响政治生活的方方面面,这首先表现为自由的范围拓展和实现方式的更新。来自数字技术的功能可供性,为人们未来可能拥有的自由带来了诸多可能性。乐观主义者认为,我们将获得数字解放,能以全新的方式行动和思考,进而产生新的创造、自我表现和自我实现形式;与此相反的悲观主义者则坚持数字化的异议,并预测数字技术将是强者践踏弱者自由的另一种方式。显然,这两种关于数字技术与自由关系的看法都有其存在的现实基础,但是人类自由的总体水平总是以对某些自由的限制为前提的,正所谓"限制使我们自由"(萨斯坎德,

2022:146)。萨斯坎德(2022:148)提出了为自由设定界限的合理主张,认为“某些‘明智的限制’系统也许可以帮助我们在数字生活世界中保持令人满意的‘总体自由程度’”。因此,来自国家制定的法律体系、公正的数字执法以及精细的数字法律,就成为我们共同打造“明智的限制”系统不可或缺的内容。

除了国家影响个人的自由之外,科技公司也拥有同时增强和限制人们自由的能力。“它使我们可以自由地去做以前不能做的事情,但同时也根据代码的约束限制我们的行为。”(萨斯坎德,2022:150)在数字生活世界中,人们获得了数字自由,同时几乎所有的言论和信息交流都受社交媒体平台和技术中介的调节、控制,而且随着越来越多的数字权力集聚在控制技术的科技公司手中,科技公司也将更多地塑造和限制我们的自由。对此,萨斯坎德(2022:158—163)进一步指出,科技公司的工作应该回到尊重社会“应该允许”和“不应该允许”的基本原则上,如保护人类不致自我伤害且能坚守共同的道德感。

西方政治的某些方面,同样因数字技术而产生不可逆转的改变。相对于古希腊的古典民主、西方的现代自由民主和竞争性精英主义模式(代议制民主),萨斯坎德认为未来的每一种民主模式,都是通过数字技术实现的,如协商民主、直接民主、维基民主和人工智能民主。这里仅以协商民主为例,说明数字技术所带来的机遇与挑战。西方民主模式之中的协商,被认为是“共同体成员理性地讨论政治问题,以找到所有(或大多数)理性人士都能接受的解决方案的过程”(萨斯坎德,2022:170)。在数字生活世界里,网络信息技术给少数大众媒体和大量分散的个人提供了更多积极参与讨论、辩论和协商的机会。然而,个人表达意见机会的增多,并不必然广泛提高协商或政治讨论的质量,这是因为它会受到来自数字技术的感知控制、碎片化的现实、匿名或非实名参与以及在线机器人的严重威胁。针对这些挑战,萨斯坎德强调既需要在技术层面警惕数字媒介影响人们言论的方式,也应基于政治共同体原则凝聚碎片化的社会现实。

社会正义问题是萨斯坎德关于数字技术影响未来政治的第三个重要领域。在数字生活世界中,作为社会正义本质的分配和承认高度依赖于算法,而算法的不公正将直接导致这两个方面产生严重的社会不公正。在分配正义方面,随着数字生活世界中使用算法的地方越来越多,尤其是能够学习知识和技能的算法出现,“将与市场和国家一起决定我们获得重要社会物品的途径”(萨斯坎德,2022:219)。作为一种新的、重要的影响分配正义的机制,算法不仅被广泛地用于判定人们是否具有获取工作、贷款、住房、保险等社会资源的资质,还直接参与市场交易、提供市场信息以及干预市场价格。在这些实例之中,作为执行任务或解决问题指令集的算法,对如何分配、根据什么标准分配、优先考虑谁的利益等问题的不同定义,直接决定了社会资源的分配是否符合社会正义的要求。在承认正义方面,算法也被越来越多地用于对人类进行识别、排名、分类和管理,这样在数字代码领域就会展开为争取社会承认的斗争(萨斯坎德,2022:228)。长期以来,争取承认的斗争主要发生在人与人之间;而随着数字生活世界的到来,人们对承认的追求已经扩张至数字系统,并表现为媒体平台之中的粉丝、好友、收藏、点赞和转发等方面的排名。算法在其中的影响,体现在“决定了这些排名和排序系统如何发挥作用,选择谁被看到,谁被隐藏,谁入局,谁出局,哪些内容会像病毒一样传播开来,哪些内容注定无人问津”(萨斯坎德,2022:229),这显然是一个不平等的过程。

萨斯坎德认为,算法在资源分配和社会承认领域极其容易引发不公正的问题。追究其成因,主要来自两方面:一是算法的数据不公正,即算法应用是基于选择不当、不完整、过时或存在偏见的数据做出的;二是算法的规则不公正,主要包括显性不公正规则和隐性不公正规则两类——其中,前者是依据宗教、种族、性别、相貌等不公正标准做出的,后者则是由间接导致某个群体受到不如其他群体的待遇而导致的结果。关于是否存在中立的算法问题,萨斯坎德强调,即便是在群体之间保持中立的算法规则,其结果仍然可能导致

社会不公正的产生。其原因在于,数字系统的学习对象——人类——是有缺陷的,而看似中立的算法规则可能会重复和巩固世界上已经存在的不公正。因此,在中立算法规则的表象之下,实则掩盖着中立谬误的不公正。

数字生活世界中的社会正义问题,还表现为数字系统将促使西方社会产生严重的贫富分化:一方面,当数字系统更多地涉足此前只能由人类完成的任务时,经济上可能导致大量工作岗位的减少和大规模的失业,这是技术导致失业的基本观点;另一方面,数字技术还将进一步加剧西方社会的财富分配不平等。数字生活世界中的资本要比生产劳动产生更多的财富,而相对于优质的传统资本(土地、股票等)和生产技术资本,数据作为数字生活世界的经济命脉,是最重要的资本形式之一(萨斯坎德,2022:264)。然而,这些日益重要的资本形式都掌握在以字母表(Alphabet)、亚马逊(Amazon)和脸书(Facebook)为代表的少数大型科技公司手中,在强大的网络效应(network effect)的推波助澜下,越来越多的社会财富将聚集到极少数的精英阶层,而流向没有资本的劳动者的财富则会日益减少,萨斯坎德将这个过程形象地称为“财富旋风”。数字生活世界中的这种财富分配方式,意味着资本主义社会中日益扩大的少数富有的精英与大多数贫穷的劳动者之间的社会财富差距,这将是资本主义经济体系中显而易见的一种不平等形式。

四、突破数字系统控制的思考

身处数字生活世界之中的我们,不得不正视这个将与人类“对垒”并可能超越人类甚至控制世界的数字系统。然而,摆在我们面前的窘境是,既难以完全理解又无法充分驾驭数字系统。为了避免陷入完全遭受强大的数字系统控制的命运,萨斯坎德结合个体、科技公司和国家三个主体,分别从技术运作的透明化和技术权力的分割两个维度,提出了突破数字系统控制的思路。对此,本文从技术的政治经济学角度进行评析。

在萨斯坎德看来,数字系统并非纯粹的技术,过去和未来的问题都可以归于个人的选择,亦即它们都是我们自己的选择造成的问题。因此,他特别强调应对数字系统支配的个体责任。那些在科技公司工作的人员——从首席执政官到普通程序员——由于他们都扮演着一种具有特别影响力的角色,就必须承担起哲人的使命,成为一个真正的“哲学工程师”。那些不在技术领域工作的人士,也需要保持对数字技术的警觉,并努力去塑造它们,如赋予公民对管理自己的数字规则和数字系统以直接发言权。

对于科技公司而言,萨斯坎德强调抑制它们的科技权力。我们知道,科技公司在政治、经济、文化等方面都拥有较大的控制力和影响力,而这些力量源自它们所掌控的数字技术。在数字生活世界里,数字技术就是政治本身。因此萨斯坎德指出,科技公司的权力获得必须要以合法性为前提,而合法性又源于一系列基本的政治原则,包括人们的同意原则、权责相当的公平原则以及体现用户共同价值观的原则。通过赋予科技公司权力的合法性来制约其权力的思路,反映了萨斯坎德坚持让权力对公民更加负责以及实现技术民主化的观点。

对于国家而言,萨斯坎德采取了两种不同的应对之策:一方面,面对随着数字系统而不断增强的国家权力和执法能力,他主张提升人们问责政府能力的办法以应对国家权力带来的系统性挑战,从而确保国家权力得到监管;另一方面,国家必须对数字系统和掌控数字技术的高科技企业实施监管。一种是透明度监管,即加强算法的审核,确保有关算法、数据使用以及从编码到技术的价值观透明度;另一种是结构性监管,指通过政治干预促进科技公司之间的竞争、打破技术垄断、防止串通、监督并购行为,目的在于防止任何一个科技公司拥有垄断或控制数字权力(暴力、审查和感知控制)之中任何一个的能力。显然,萨斯坎德将技术权力的分权视为保障自由和合法性最好的甚或唯一的方法。

实现数字技术的透明化和数字权力的分割,贯穿在萨斯坎德关于突破数

字系统控制的思路始终,其中或多或少流露出对日益加强的数字系统控制的无可奈何。尽管他也注意到真正的解决之道在于建立一个“政治体制”(萨斯坎德,2022:301),但是他最终开出的“药方”仍然没有超出西式民主和三权分立的窠臼。问题的关键还在于:在数字系统“大兵压境”的背景之下,这种传统的解决之道虽然具有一定启发性,但能否真如萨斯坎德所期待的那样有效地解决问题?笔者认为,答案是否定的,因为他忽视了西方资本主义国家中数字技术的政治经济性质。

数字技术重塑当代资本主义,这一事实已经为近年来不断涌现的诸如数字资本主义(希勒,2001)、数据资本主义(迈尔-舍恩伯格、拉姆什,2018)、平台资本主义(斯尔尼塞克,2018)等研究成果所证实。不管用什么概念来统称数字化时代资本主义发生的深刻变化,它的基本运作逻辑都不可避免地深深嵌入在算法系统之中。在蒲若迪尼克(2022)看来,资本主义社会中的算法运作具有不透明与模糊性、数据化、自动化以及工具合理化四个基本特征,它们不是算法本身普遍固有的,而是资本主义政治经济这一结构性原因的产物。亦即,为了追求商业利益最大化,西方高科技公司更倾向于保持数字系统令人无法想象的复杂性,因此“技术黑箱”的透明化无法真正实现,而且各种“黑暗代码”所产生的不可知性更容易导致缺少问责和逃避监管现象的发生。

关于数字技术的权力分割问题,更是面临着无法逾越的双重困境。在政治经济学视域下,一方面,资本主义国家与高科技企业之间的关系更多体现为政商合谋式的利益共同体,因此无论对资本主义国家还是对高科技企业的技术权力进行新的分割,都是无法真正实现的;另一方面,对于数字技术企业而言,它所拥有的技术权力往往植根于数字技术本身所具有的连通性、聚集性、开放性等属性,这正是数字经济的重大优势之一。因此,纵使资本主义国家通过反垄断法,能够相对削弱或防止某一高科技企业实现技术权力的高度垄断,但却无法从根本上解决技术权力集中的难题。

行文至此,让我们重新回到《算法的力量:人类如何共同生存?》一书的起

点:生活世界数字化时代的人类如何共同生存?尽管在资本主义的技术政治经济背景下,萨斯坎德关于我们摆脱数字系统控制的思路行不通,但人类对此也不是绝对无能为力的。事实上,解决问题的症结就在于打破西方以利益为导向的政商合谋式的政治体制。近年来,我国数字技术和数字经济迎来了飞速发展,在数字中国建设进程中形成的党和政府、市场与社会相协同的治理实践,展现了有中国特色的技术政治经济学在实践中的独特优势,这或许能为全球数字系统治理提供新的启示和经验借鉴。

参考文献

卢克斯,2008,《权力:一种激进的观点》,彭斌译,南京:江苏人民出版社。

迈尔-舍恩伯格、拉姆什,2018,《数据资本时代》,李晓霞、周涛译,北京:中信出版集团。

蒲若迪尼克,2022,《数字资本主义的算法逻辑》,陈文旭译,《国外理论动态》第 6 期。

萨斯坎德,2022,《算法的力量:人类如何共同生存?》,李大白译,北京:北京日报出版社。

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。

希勒,2001,《数字资本主义》,杨立平译,南昌:江西人民出版社。

Castells, M. 2013, *Communication Power*, Oxford: Oxford University Press.

Foucault, M. 1980, *Power/Knowledge: Selected Interviews and Other Writings (1972 - 1977)*, New York: Vintage Books.

Mill, J. 2009, *The Autobiography of John Start Hill*, New York: Seven Treasures Publications.