迈向中国的数字民族志*

丹尼尔·米勒　著**

李昕茹　译***

我是一名英国人类学家,多年来一直从事物质文化研究,著有关于人与消费的书籍,包括住宅、汽车、衣服和购物。但在2000年左右,我开始对数字技术和互联网的崛起感兴趣。我与社会学家斯莱特(D. Slater)共同出版了《互联网:民族志的方法》(*The Internet: An Ethnographic Approach*)(Miller & Slater, 2000),这是我第一本研究互联网对特定群体影响的书。对我们人类学家来说,重要的不是伦敦和纽约这样的大都市,而是像特立尼达(Trinidad)这样的小岛,它属于加勒比国家特立尼达和多巴哥的一部分。三年前,我写的有关特立尼达岛的《脸书故事》一书已有了中文译本(Miller, 2011;米勒, 2020)。

在这本于特立尼达岛上写的关于互联网的书中,我想表达的重点是:互联网很可能在每个地区产生不同的影响,而从自然科学的角度研究数字技术的问题在于,研究者倾向于得出具有普遍性的结论。他们会对大学生做一些实验,表明某种算法会让人上瘾,并对该技术的影响提出某种普遍的说辞;但人类学家会争辩说,发生在大学生身上的事情可能并不适用于印度的农民或中国的工人。我很幸运,后来我得到了欧洲研究理事会(European Research Council, ERC)两项为期五年的资助,这意味着未来十年我能够支付人类学家团队研究数字技术对印度农民和中国工人的影响。该项目最成功的成果之一是《世界如何改变社交媒体》(*How the World Changed Social Media*)(Miller, Costa & Haynes et al., 2016),自然科学家关注社交媒体如何改变我们,但这本书名显示,我们更关注不同群体对社交媒体的使用方式及其后果。我们与伦敦大学学院出版社合作,

* 首先感谢我的博士生,“我们为什么要发帖”项目、“智能手机人类学”项目和“智能老龄化”项目的团队成员所做的工作,同时感谢欧洲研究理事会(European Research Council, ERC)“欧盟新视界2020研究和创新计划”的资助(资助协议号:740472)。

** 丹尼尔·米勒(D. Miller),伦敦大学学院。

*** 李昕茹,哈尔滨工程大学人文社会科学学院。

免费出版了该书,并连同出版了其他田野地点(fieldsites)的有关研究著作,该项目的所有成果被称为“我们为什么要发帖”(Why We Post)。此系列书籍在全球范围内的下载量已超过 100 万。

对我们来说,人类学方法的意义不仅在于全球视角,在非洲、拉丁美洲、欧洲抑或中国进行研究,而在于人类学的核心是民族志方法。在这十年中,所有的研究成果都是人类学家与当地民众一起生活 15 个月以上得到的,这对数字技术的研究至关重要。首先是因为这些技术往往是非常隐私的。我们可以看到人们埋头使用智能手机,但如果不越过他们的肩膀看,就不知道他们在智能手机上具体做什么。我们可以询问他们使用智能手机的目的,但是,例如,在我最近研究的爱尔兰老年人群中,他们可能会说自己只使用两三个应用程序来打电话和发短信,但最后我发现他们可能在使用 30 个不同的应用程序。问卷、调查和焦点小组等方法对于了解人们使用社交媒体等数字技术的情况确实没有什么帮助,原因是它们被强大的话语权所淹没。英国的报纸总是在谈论社交媒体会对儿童造成伤害,或者会让人上瘾,或者会造成焦虑,因为贩卖焦虑往往能使报纸畅销。因此,当你询问人们关于他们的智能手机时,他们倾向于重复那些公众话语。只有当你查看他们手机上的每一个应用程序,并询问他们最后一次使用该应用程序的故事时,他们才会忘记这些关于智能手机是好是坏的道德争论,而详细谈论他们实际做了什么及其后果。

这里重要的问题是,如果你想对社交媒体的使用进行直接观察,而不仅仅询问人们做了什么,或者如果你想花几个小时去看每一个应用程序,讨论可能涉及的非常隐私的事情,比如一个家庭内部的关系,那么就需要和田野对象建立良好的信任关系。这也许需要几个月的时间来了解他们,然后他们才会明白你是谁,为什么要问这些隐私的事情,并知道他们不会在你的出版物中被识别出来,而且这对于教育来说是有价值的。然而,至于我们的研究为什么需要长达 15 个月,还有另一个原因。我把民族志方法称为“整体情景化”(holistic contextualisation)。在田野中,我事先并不清楚人们生活的哪些方面导致了人们对社交媒体或智能手机进行某些使用——有可能是他们的家庭生活、教育、政治、职业、宗教,他们对运气和命运的看法或他们的个性。因此,我们需要全面了解这些事情,否则无法决定哪些因素是关键的影响。民族志是唯一像真实生活的方法。我们知道,人并不仅生活在一个学术话题里,他们同时生活在我们所有

的研究主题中。因此,我们必须在实时发生的情况下一起研究所有的线索。倘若我们事先不知道什么是与我们的主题相关的,我们就不能捏造假设来进行测试。我们必须以开放和机会主义的态度,随着时间的推移循序渐进地了解这个主题。民族志试图不做作,在最好的情况下,我们甚至没有任何访谈。我们只是很好地了解人们,当我们一起散步、一起工作、帮助照顾他们的孩子或为他们做饭时,作为谈话的一部分与他们谈论这些事情。

这种民族志或整体情景化的方法不一定只由人类学家使用,它可以在社会学和媒体研究中找到,但它与人类学学科的联系最为密切。这也是我对中国特别感兴趣的原因。在中国,已经有了不少关于数字媒体的重要研究(例如 Qiu,2009)。但我看到的问题是,中国的人类学是由对少数民族和族裔社区的研究兴趣而发展起来的,而负责研究城市中主要汉族社区的是社会学。但是,人类学家看待世界的方式、对 15 个月整体情景化的承诺,以及对人在世界范围内如何变化的认识,对于大城市的汉族和少数民族来说都是一样的。因此,我关心的是让人们认识到,人类学是未来的学科,而不是过去的学科。事实上,这也是我专注于数码人类学的一个原因,也表明了民族志方法在我们谈论正在人群中传播的最新技术时是特别有价值的。然而人们往往不愿意等待 15 个月的研究结果,以了解抖音算法的影响,或微信最新的小程序。但是我已经在上文提过,好的研究没有捷径。我们学者要想发表权威见解,最好以高质量的学术研究为依据。

事实上,人类学可以成为研究的先锋。我是第一个提出脸书在年轻人中失去人气的学者(Miller,2013)。当然,脸书公司肯定知道这一点,但他们会隐藏这一事实,因为这会对他们的股票价值不利。当我第一次提出此观点时,很少有人相信我。但我是知道的,因为我已经融入了一个群体,我每天都在和年轻人交谈。我的说法不胫而走,人们逐渐意识到我是对的。经常有公司和其他人问我是否愿意做一个“简短”或“快速”的民族志,政府有时也想做同样的事情。但我从来没有同意这样做,因为一个糟糕的研究比没有研究更糟糕。我相信,我们总是需要整整 15 个月的时间来进行充分的民族志研究。

对于中国政府和人民来说,数字技术的发展被认为是帮助创造一个现代化国家和更富有的民众的重要部分。因此,众所周知,中国拥有一批当今世界最成功和最具创新性的数字公司。但我们仍然需要了解这些技术的使用及后果。对于人类学家来说,重要的是,中国是一个地域辽阔的国家,因此对于云南的农民

和北京的数字公司的经理来说,对于一个年轻的女学生和一个退休的老人来说,技术的“使用及其后果”也会有所不同。我不是中国人,也不会说普通话。我的兴趣首先来自认识到中国在当今世界的重要性,如若不了解中国,全球性的方法就无法发挥作用;但同时,也是因为有许多中国学生在伦敦大学学院与我共事。

人类学另一个不寻常的方面是,我们通过移情来工作。也就是说,我们试图理解那些与我们截然不同的人。因此,我们建议,学生至少在其第一次研究中,不应该与像他们一样的人合作,而是与同他们非常不同的人合作。我第一次参与数字中国的研究时就呼应了这种观点。作为 2012—2017 年“我们为什么要发帖”项目的一部分,我的团队中有两人在中国进行田野调查。第一位是个英国人,叫麦克唐纳(T. McDonald),他那时已是我的博士生,在云南做实地考察,能说一口流利的汉语,后来,他决定到山东省的一个农村小镇调研。团队的另一位成员王心远,她本身就是中国人,决定到一个工厂里进行田野调查。心远对中国 2.5 亿人从农村地区迁移到工厂系统感兴趣。虽然这是一个中国学生研究中国人的案例,但她的背景与那些工厂工人非常不同,所以仍然需要基本的共情能力作为她研究的基础。因此,这意味着两项研究都是运用典型的人类学方法。此外,王心远还翻译了第一本数字人类学教科书《数码人类学》(Miller & Horst, 2012),由人民出版社出版(米勒、霍斯特,2014)。

麦克唐纳和王心远的这两项研究很成功。麦克唐纳让我们清楚地了解社交媒体在他研究的那个小镇社区的使用情况。例如,他的研究表明,对许多人来说,使用社交媒体是他们第一次可以直接与既不是家人也不是朋友的陌生人互动,他们的生活、家庭关系模式由此改变。他后来出版了《中国农村的社会媒体》(*Social Media in Rural China: Social Networks and Moral Frameworks*)(McDonald,2016),可从伦敦大学学院出版社网站免费下载。此后,他一直在香港担任讲师,并发表了关于中国数字货币研究的论文(McDonald,2020)。同时,王心远认为,工厂工人经历了双重迁移,从农村地区到城市地区,从线下到线上。她表明,往往是第二次迁移到网上的过程更成功地将他们带入现代化中国的轨道,这也是这两次迁移背后的深刻内涵。她出版了《工业中国的社会媒体》(*Social Media in Industrial China*)(Wang,2016)一书,也由伦敦大学学院出版社出版。

这十年的研究和数字民族志的研究成果,对社交媒体和智能手机的本质提

出了许多有一定学术影响力的理论观点。关于社交媒体,我们探索了一些概念,例如:“复媒体”(polymedia),涉及不同个体使用的不同平台之间的关系;“可扩展的社会性”(scalable sociality),涉及隐私的程度和网络的大小,以及人类对话中视觉与文字、口头的结合方式。关于智能手机,我们认为它最好首先被看作我们此时此刻生活的地方,我们称之为“运输的家”(transportal home),而不仅是一个通信设备。我们认为,智能手机在与个体的亲密连接方面是前所未有的,我们称之为“超拟人化”(beyond anthropomorphism),并讨论了我们称之为“永久机会主义”(perpetual opportunism)的后果,因为它总是与我们在一起。此外,还有许多其他类似的一般性观点。

在这项工作的基础上,我吸引了一批中国博士生,他们帮助我将项目扩展到对当代中国新媒体和数字技术的后果进行更广泛的研究。同样,所有这些研究都将基于经典的人类学方法,即与人的长期接触的民族志方法。目前就职于哈尔滨工程大学的李昕茹(去年从伦敦大学亚非学院毕业)的博士论文便是使用该方法进行研究的一个例子,她表明了同时观察媒体的生产者和消费者的重要性。只有当你研究了民众如何看待一个媒体节目时,你才会知道它是什么,因为只有这样,我们才能理解其后果。如果有人制作了一个节目,试图让人们向往农村生活,但观众却认为它展示了城市生活的好处,那么这个节目就变成了后者,即使它与制作者的本意相反。但我们也需要深入了解制作该节目的构想和技术。因此,为她在湖南卫视及其网络平台芒果 TV 的工作提供了一种深度的理解,如果没有对创作者和观众的双重学术研究,我们是无法获得这些信息的。

目前,我有三位中国博士生正在撰写博士论文。其中之一,李晓琳专注于研究月经追踪应用程序。这项研究再次证明了人类学方法的必要性,因为理解月经追踪应用程序意味着理解诸如人们与血液的关系,以及备孕、避孕与性行为方面的一系列私密话题。需要在长期了解这一群体并获取她们的信任之后,她们才会表达内心真正的想法。短期的研究往往只会得到非真实和非公正的结果,比如仅仅得到她们认为我们想听到的内容。第二位王海潮的研究亦是如此,她研究的内容,比如人们通过微信发展社交电子商务的过程和结果是非常敏感的,因为当尝试请朋友和亲戚帮你在朋友圈卖货时,会有一个非常复杂的问题——可能被视为对朋友和亲戚关系与感情的剥削。因此,我们再一次需要信任和对人的充分了解,然后才能进行这一主题的研究。

第三个是郑肯的研究，她在北京互联网大厂进行田野调研，关注了数码平台的设计者如何在日常生活中设计与维护社交媒体。她研究的意义在于，通常情况下人们倾向于将大公司想象成以高效协作、客观管理的方式运作，这种文化想象往往来自商学院或是公司本身的自我宣传，但科技公司在实际运作中的内部摩擦，可能源于设计与开发部门之间对于用户需求不同的理解、个体对于职业成功的追求与组织整体目标之间的矛盾、性别化的分工，等等，而这些内部摩擦不仅影响了数码平台作为一个机构的有效运作，也外溢到科技公司之外更加广阔的社会关系方面。因为这些内部摩擦影响了数码平台本身以何种形式被设计和维护，从而间接影响了基于这些平台来展开的社会活动与用户。

我还有博士生目前正在中国进行田野调查。例如，我们关注短视频的影响，如豆瓣和快手等平台所青睐的短视频，或如哔哩哔哩等稍长的视频。我们需要在一个社区待上足够长的时间，了解这些视频是被如何用于从广告服务到改变家庭内部的沟通，我们才能真正说知道它们的后果是什么。这项研究目前正在由向昆玉进行。我一直使用“后果”（consequences）这个词，是有意而为之，因为无论是我们想进行良好、准确的教育和讲座，还是我们居于制定媒体政策的位置，我们都需要对后果进行恰当的研究，而非简单地认为我们知道某件事对特定群体产生了或好或坏的影响。最后，我还有一名学生刘纯在云南的某个藏族社区工作，因为我们也想尊重当代中国的巨大多样性，而不是一概而论，认为以大城市为研究对象的结论适用于整个中国。

我们认识到，让我们的作品也有中文版本是很重要的。因此，我们最新的集体论著《全球智能手机：超越年轻人的科技》（*The Global Smartphone: Beyond A Youth Technology*），也是一项比较性研究成果，不仅有英文版本（Miller, Rabho & Awondo et al. ,2021），还有中文版本，可从伦敦大学学院出版社网站免费下载（米勒、拉博、阿旺多等，2021）。如果你想更好地了解我们的研究，下载这本书是一个好的着手点。这本书包括了王心远基于上海老年人的第二项研究成果，该研究也将同时以专著的形式出版。

目前，我们正以团队的形式讨论编写一本关于中国数字人类学的书的可能性。我们认为，如果以上学者都能贡献一章，同时我们也讨论融入其他对中国境内数字发展感兴趣的人类学家正在进行的研究，并以中文出版，那么我们将有一个资源，让整个社会科学、社会学、媒体和传播研究、新闻学、人文地理领域学者

和政策制定者都能了解、理解这个对整个人类有巨大影响的动态领域。我们需要发展这种深入的学术研究，这将给我们提供高质量的教育和权威，使我们能够在做了一定深度、厚度研究的基础上谈论这些技术，使我们学者对自己的主张更加确切和有把握。

我们没有理由局限于关注学科的界限。倘若一位学者在社会学或媒体研究领域工作，但也进行了为期 15 或 16 个月的民族志研究，那么同样有价值。重要的是学术研究的深度，而不是学者的头衔。虽然我也希望人类学研究在未来能在整个中国得到扩展，以确保这种高质量的研究在中国未来的学术发展中拥有适当的地位，但同时希望在这一传统下培养出来的学者能在中国的大学系统中获得教职，这将裨益于这一发展。

我们的目标不仅仅是抽象地提出一些主张，而是要发展出一套高质量的研究体系，让大家认识到为什么这些工作对中国和当今世界的学术工作的未来至关重要。在这方面，重要的是找到正确的研究广度，包括对公司和用户的研究，如李晓琳和李昕茹的研究；或包括对人类多样性的研究，如王心远研究上海人和工厂工人的方式；也让麦克唐纳这样的外国学者与中国学者同时参与此类研究。尽管我自己没有（在中国）进行实地调研，也不会说普通话，但我希望为我们在伦敦大学学院的数字人类学硕士课程做出贡献，因为我们有很多中国学生，我希望为中国境内的博士和博士后民族志项目尽绵薄之力。另外，在未来，我认为让中国学生以与英国人类学学生相同的方式在其他地区开展（海外）民族志工作是很有价值的，例如让中国学生研究英国人及其对数字技术的使用。我们也希望将来无论学生来自哪里、在哪里进行研究，都是完全平等的，这样我们就可以一同合作，共同创造一套关于这些新的数字媒体的全球使用和后果的研究，这些媒体已经成为整个当代世界人民日常生活的重要组成部分。

参考文献

米勒，丹尼尔、希瑟，霍斯特，2014，《数码人类学》，北京：人民出版社。

米勒，丹尼尔，2020，《脸书故事》，北京：北京大学出版社。

米勒，丹尼尔、拉博、阿旺多等，2021，《全球智能手机：超越年轻人的科技》，London：UCL Press.

McDonald, T. 2016, *Social Media in Rural China*, London: UCL Press.

McDonald, T. 2020, "'Social' Money and Working-Class Subjectivities: Digital Money and Migrant Labour in Shenzhen, China." *The China Quarterly* 242.

Miller, D. & D. Slater 2000, *The Internet: An Ethnographic Approach*, Oxford: Berg Publishers.

Miller, D. 2011, *Tales from Facebook*, Cambridge: Polity Press.

Miller, D. 2013, "What Will We Learn from the Fall of Facebook?" https://blogs.ucl.ac.uk/global-social-media/2013/11/24/what-will-we-learn-from-the-fall-of-facebook.

Miller, D. & H. Horst eds. 2012, *Digital Anthropology*, London: Berg Publishers.

Miller, D., A. Rabho & L. Awondo et al. 2021, *The Global Smartphone*, London: UCL Press.

Miller, D., E. Costa & N. Haynes et al. 2016, *How the World Changed Social Media*, London: UCL Press.

Qiu, J. 2009, *Working-Class Network Society: Communication Technology and the Information Have-Less in Urban China,* Cambridge: MIT Press.

Wang, X. 2016, *Social Media in Industrial China*, London: UCL Press.

从话语构序到媒介构序

——基于数字媒介化的本体论之思

董　翟*

摘要:在现代性语境中,尤其是在数字时代的今天,人的生存方式和存在样态不可避免地受制于媒介。媒介并非简单的信息载体和可被使用的工具,而是具有内在结构和驱力的物质性系统。在存在意义上,媒介构成了塑形人与物存在关系和存在方式的场境。本文以福柯的话语分析所揭示的话语构序在媒介化意义上的重构为导引,试图将数字时代的媒介场境揭示为横向和纵向的双向度断裂,并在后者向度上展现基于媒介体系形成的垂直权力构序。基于以上探讨,本文将媒介理论延伸到本体论层面,试图提出媒介本体论范式的一种建构,并由此探索肯定性的媒介化之可能。

关键词:数字资本主义　话语　媒介化　媒介本体论

在《留声机 电影 打字机》的开篇,基特勒(F. Kittler)(2017)如是写道:“媒介决定我们的现状。”这里反映出,至少在现代性语境下,人的生存方式和存在样态已经不可避免地受制于媒介。媒介作为问题,也许向来重要,但在今天尤甚——数字时代语境下,无论是否知觉,人的生活从个体层面到集体层面近乎必然地为可见或不可见的媒介关系所中介着,并且这种关系正在逐渐形成一种异于人掌控的独立性。媒介并非——至少并非仅仅作为人的工具而存在,也并非仅仅是内容借以传播的“无辜”中介,这是自伊尼斯(H. Innis)和麦克卢汉(M. McLuhan)以来创生的媒介理论早已充分予以指认和讨论的境况。

关于这方面的探索,媒介理论历来存在传播学和哲学的二元底色。前者偏重于社会学式的实证研究,主要聚焦媒介在实际社会生活中所呈现的功能、作用及其对个体和社会的影响。这固然提供了大量实证的结论,但往往囿于表象和具体,难以从中直接得到具有现实抽象意义的总体批判视野。新近的媒介反思

* 董翟,南京大学哲学系。

也对此有所自觉,并更多力图以哲学的视角运用存在论等对媒介进行观照,但也存在以理论反注现实的“具体性错置”、缺乏对媒介在技术和社会不同层面发展实际的观照而流于空疏等不足;同时,也正因为缺乏对现实系统、深入的认知,所以现实抽象层次的探讨无从谈起。建立媒介研究的方法论自觉,在现实抽象的基础上提出媒介本体论的新批判范式,是媒介理论必然需要拓展的向度。

这方面探索的另一困难在于模糊性。单是最基本的“媒介”概念,在定义上便千差万别。麦克卢汉的“媒介是人的延伸”也许是广为人知的一例,但如果不加限定,则这一概念的对象几乎可以推展到一切可以与人产生关联的物,甚至人本身也概莫能外。为保证讨论聚焦明确,本文主要基于丹麦媒介学者夏瓦(S. Hjarvard)对“媒介化”的定义探讨媒介问题,即区别于中介化研究——仅仅将媒介当作可以选择使用的传播意义上的手段和技术的研究范式——而将媒介视为“社会和文化实践的一个结构性条件”(夏瓦,2018),作为当下社会生活中的一种机制和场域性的存在来加以探讨。在讨论的社会历史层面,主要以狭义历史唯物主义视角聚焦数字时代媒介的当代变化。

基于此,本文试图以福柯(M. Foucault)对“话语塑形”的分析为导引,通过福柯已经指认出的话语构序机制,及其变化中发现的作为背景而遭到忽视的媒介之物质性能量,将对话语构序的分析在媒介化意义上予以重构,显现出媒介化时代媒介的构序力量。在此基础上,本文试图以“横向”“纵向”两个维度的媒介化,展现数字时代媒介化的特征,并由此在一种可能的方法论构式上,探索数字资本主义时代如何重新理解人、资本与媒介的关系,人在媒介化时代下的处境,并构建更合理的“人—媒介”共生的存在方式。

一、话语构序在媒介化意义上的重构

在《知识考古学》中,福柯一改之前在《词与物:人文科学的考古学》中对不同时代“知识型”断裂进行考古学分析的研究思路,将作为实践的话语置于研究的中心位置。在对话语的分析中,福柯力图指认在话语实践的运作背后,作为不在场的在场而存在的“话语塑形”过程,即“一种确定话语实践本身特征的关系之建立……一组内在于实践的和在其特殊性上确定实践的规则”(福柯,2021)。在这一分析中,“词与物本身一样,都是有意缺席的”(福柯,2021)。如同胡塞尔

将对象的存在先行“悬搁”一样，话语的内容——无论是话语的对象，还是话语的语言学意指关系——都必须首先去在场化，从而回到话语本身，让“决定话语对象的弥散性的连接”（福柯，2021）得以显现。

为了让讨论聚焦于话语本身，福柯重点区分了“话语”本身层面的基本单位——“陈述”与“命题”“句子”和“言语行为”三者的区别。与逻辑层面的命题不同，陈述不需要满足命题结构和构成规则的限定，不同的陈述可以在命题结构上是同一的；与语法层面的句子不同，陈述并不一定是完整的句子，甚至不能简单从语法特征、语言要素的层面去定义，且陈述未必只以语言的形式构成，如图表也可以构成陈述；与表达层面的言语行为不同，一个言语行为可能需要不止一个陈述，而陈述也并不是在言语行为已经规定好的连接序列中确立自身，恰恰相反，言语行为被陈述所定义，其序列是在个体化的陈述中形成的；总之，陈述与以上三者都处于不同层面，并且先于以上三者的结构秩序而存在，“不是一种结构……是一种本来就属于符号的存在功能”（福柯，2021）。

在这个意义上，话语构序并非稳定连续的结构，甚至并非此前“知识型”分析所指认的虽然存在断裂但在一定时段具有相对稳定性的结构，而是话语在现时话语事件中的形构“塑形”过程，它在前概念、前结构的话语本身层面运作，从而组织话语诸概念、结构的秩序。

然而，福柯的话语构序分析并未仅仅停留在认识论层面。如同《词与物》的“知识型”实际上是“事物的秩序”一样，话语的构序进而延伸到陈述的主体与客体，从而话语塑形的过程也成为塑造主体与客体的过程。陈述主体并不是先在的决定被陈述的话语的创造者，“各种各样的陈述样态非但不归结于某个主体的某种综合或某种统一功能，反而显示着主体的弥散”（福柯，2021）。陈述主体只是弥散在陈述的场境中，由陈述塑形的一个临时主位；陈述客体也并非先在于陈述之外，只是被简单移入陈述中的实存，是“只在话语中才出现的、有规则的对象形成”（福柯，2021）。由此，话语的分析实质上是关涉话语所构建的场境中主体与客体存在方式的探讨，因而具有本体论层面的深刻意涵。

基特勒的“话语网络”理论延续了福柯这一话语分析的思路：对象上，同样关注的是在话语内容之外，作为不在场的在场而存在的话语生成场境；方法上，同样依循考古学的方法，考察被话语可见的连续性所遮蔽的权力构序对话语以及对由此延伸的话语中的主体、客体的生产；视角上，同样将话语分析置于本体

论层面的深度,揭示话语场境作为本体论构序的实质。然而,不同于福柯仅仅在话语本身的层面探究话语实践中通常不可见的话语运行规则,基特勒将目光投向了话语实践的深层基底,即话语运行所必需的物质性能量——媒介技术。在基特勒看来,福柯的话语分析缺乏深入媒介技术的体认。这一局限性只有在媒介充分发展、数据处理方法打破原有作为权力根基的信息持存和传输垄断时才会显现,这也使得“福柯的历史研究并未在 1850 年的历史之后进展多少”(Kittler,1990)。诚然,这一重构其实在福柯的意义上并未逾越其话语塑形分析的本体论界限,因为福柯所指认的话语并非仅仅作为媒介传播内容的意义层面而存在。基特勒的分析,与其说是原先层面话语探讨的延续,不如说打开了另一个讨论维度。虽然隐含了不同层面言说的转换,但并不影响其在重构语境下的新意义。

在《话语网络 1800/1900》(*Discourse Networks 1800/1900*)中,基特勒将“话语网络”界定为:“技术和制度的网络,这些技术和制度使特定的文化得以遴选、持存和加工相关的数据资料。”(Kittler,1990)这一描述所指认的话语网络,与福柯话语塑形分析中所指认的话语生成和运作的场境具有一定的同构性,只是将这一场境的观照移入媒介层面,其共同目的都在于揭示作为表象呈现的话语背后的历史先验前提。在将“知识型”断裂的追问延伸到媒介层面之后,基特勒指认了 19 世纪和 20 世纪两个断裂节点的媒介范式转变——断裂节点的选取显然也带有福柯对古典时代到现代知识型界分的色彩。书写与印刷媒介的发展、普及,为 19 世纪的话语网络奠定了物质性基础。由此,19 世纪的浪漫主义话语体系在教育、语言观念、文学艺术等领域得以建构。20 世纪的自动化媒介则将话语网络从所指的意义层面转向能指的符号层面,打破了能指到所指的线性结构,也让心理分析话语体系得以可能。与福柯的话语塑形相似,话语网络同样塑造着其中的主体,如 19 世纪语言观念的转变使得女性在教育中的地位由于口语传授而变得重要,20 世纪自动化媒介又取消了口语传授中教育者主体的稳定位置。

根据基特勒自己在后来访谈中的指认和国内学者对其思想路径的梳理,这一话语网络分析范式其实隐含着香农(C. Shannon)“信息论”的思想影响,实际上是异于原有的文学理论、福柯的话语分析的新范式的出场,以及在北美技术哲学跨学科延展的语境下对原有话语理论的重构(车致新,2019)。

到了《留声机 电影 打字机》中，经由对留声机、电影、打字机所代表的三个向度的自动化媒介的聚焦，基特勒的媒介物质性分析范式正式从此前隐于文本话语分析之后的“隐微写作”走向了前台。在书中，基特勒将留声机、电影、打字机这三种媒介及其所代表的媒介物质性范式，与声学、光学、书写的数据流区分，以及拉康（J. Lacan）意义上真实界、想象界、象征界的区分分别对应起来（基特勒，2017）。

留声机所代表的声学媒介体系，其特征是前意义化过程的真实直接复现。无论是记忆还是书写形式的再现，在呈现之前都已经潜在地经历了既有意义体系的遴选和意义赋予，这一过程垄断了语言，将给定意义体系之外的语言杂多排斥出可见的领域。然而，“‘录下的语音从不会说谎’，它的技术存储可以揭露‘隐藏’的内容，可以让‘过去’……口吐真言”（基特勒，2017），经由不加区分的语音记录（无论其是否可意义化为可解读的语言），既有的象征能指链被悬置，过去真实的在场得以显现。在这一意义上，对于话语本身的细节切分和分析成为可能——弗洛伊德（S. Freud）的精神分析正是基于这一切分，为了使被治疗者说出关于自我的真相，医生恰恰需要像留声机一样不加区分过滤地倾听，避免自身理智的意义体系审查、压制被治疗者的胡言乱语，从而让被治疗者的无意识得以显现。精神分析由此和留声机媒介技术形成了一定的同构性。然而，留声机媒介技术同时也构成了突破精神分析的力量，因为它的真实最终将超越一切书写权力及其背后的象征秩序——即使是自称“具有‘很高的可信度’”（基特勒，2017）的精神分析对歇斯底里的话语的书写也概莫能外。诚如基特勒所举出的一例，当被治疗者持有留声机时，情况就发生了倒转，“分析师现在变成了研究对象”（基特勒，2017），精神分析自身的话语垄断也会被真实界的力量所打破。

然而，这并不意味着留声机媒介体系的真实是不存在任何构序性的绝对本真或无中介的真实，真实只意味着留声机媒介体系在媒介构序中所占据的位置。在悬置和突破传统的话语构序界限的基础上，留声机媒介体系同样生成着新的话语构序，例如留声机从出现之初就带有调节回放速度以对听觉事件进行时间轴操纵的功能，录音磁带的出现为声音的拼接和调节提供了更多可能——如果借用斯蒂格勒（B. Stiegler）的说法，声学媒介同样作为一种“第三持留”对人们的意识时间起着塑形和赋序的作用。

电影所代表的光学媒介体系,其特征是基于特技和蒙太奇的对时间意识的重构。电影媒介体系在结构上并不适合与真实世界发生直接联系——在物质性基底上,声波是低频的一维数据流,光波则是高频的二维数据流,相应的处理容量要求更大。直到光纤技术普遍推行之前,光波一直缺少直接储存和处理的媒介,实时视觉信号处理在很长时间内是无法实施的。然而,电影媒介体系先天具有适于遴选和剪辑来重构数据的优势。这既是电影本身无法直接输入光波数据的限制所决定的,也是其可以借助频闪和余像效应来制造幻觉拟真性的特点所赋予的。“现实中的切剁或剪辑,虚幻中的融合或流动”——这既是电影媒介体系自身的悖谬逻辑,又是电影虚幻与现实交织特性的极好诠释。正因如此,电影媒介体系,如其所对应的想象界一样,是在虚幻镜像中对自我的误认之呈现。精神分析话语的功能与之类似,以其话语将真实界的歇斯底里的话语加以意义再现。正因为如此,疯癫经由对真实的记录和再现被精神分析话语建构出来,共同的逻辑使得二者的结合成为可能,“精神病学因媒介的支持而更为强大”(基特勒,2017)。

这同样不意味着电影媒介体系不具有真实性,恰恰相反,“电影比真实更真实”(基特勒,2017),其本身模糊了真实与对真实的建构之间的边界,在电影媒介体系的逻辑中,复制就是制作,反之亦然。电影对想象中的身体画面进行剖析,从而“赋予人们一个想象的自我”(基特勒,2017),灵魂、内在的自我和个人在电影媒介镜像的重构下只是伪主体的不同代称。

不同于前两个媒介致力于生成原先书写媒介之外的新构序,打字机所代表的书写媒介体系的特征是对原有书写垄断构序在物质基础意义上的倒置。这同时也构成了对主体的塑形,原先男性垄断的书写权力随着工业化、自动化对书写性别象征的抹平而失去其独有性。除了女性之外,原先同样被排斥于话语体系之外的失明或失聪的残疾人也获得了书写的权力。

这并不仅仅是工具的掌握者和使用者意义上的转变,而是媒介场境的转变对主体存在的基底性影响。换言之,打字机并不是以无辜的工具身份介入人的生活的。海德格尔(M. Heidegger)正是在这一意义上指认打字机对寄托于手写的存在关系的侵入的,“当书写被从它的本质来源,也就是手那里脱离,而代之以机器打字,那么其存在对人的关系就发生了一种转变”(海德格尔,2018a)。尽管海德格尔是在消极意义上言说这一转变的,但他也不得不承认“技术存在

于我们的历史中”(海德格尔,2018a)。媒介技术对人的存在关系、存在方式的介入是深刻而不可逆转的,打字机媒介体系实现的就是对基于原有手写书写构序的存在方式的倒置。在肯定性的意义上,这同样意味着新构序对存在方式的重塑。由于视力的严重下降,尼采开始使用打字机写作,而这一书写方式的变化也参与了他的思维过程,让其写作风格发生了明显转变:“从长篇大论变成格言警句,从哲学思考变成一语双关,从善于辞令到电报式的短小精悍。”(基特勒,2017)

分裂为留声机、电影、打字机的三个现代媒介体系,又可以实现不同数据流的重新聚合。正如拉康曾指认的真实界、想象界、象征界之间可能的关联和耦合一样,“先从心理上分解成碎片,再从物理上进行重建,中央神经系统借此获得重生,但却成了一个由不同分身构成的魔像”(基特勒,2017),话语网络由此在媒介基础上完成了对主体和社会关系的再生产过程。同时,三个现代媒介体系也共同经历了三个阶段的发展变化,即从美国内战开始的声学、光学和文字存储技术的出现,从一战开始的与各存储技术相应的电子传送技术的出现,以及从二战开始的将既有媒介体系在新兴信息技术中的新应用。最终,三个分化的现代媒介体系统合于基于现代信息技术的控制论体系中,“自动化的话语分析已经开始发号施令”(基特勒,2017),媒介话语网络实现了对自身的再生产。

从以上的讨论中可以看出,基特勒的“话语网络”理论,主要从两个方面将媒介技术的力量引入原有的话语构序分析中,并将话语构序延伸为话语之下作为其物质性基础的媒介构序:一方面,揭示媒介自身演化的“自主性”,即媒介网络对自身的再生产;另一方面,揭示媒介作为生产场境,对话语以及栖居于话语中的主体、客体的再生产。

总之,从福柯的话语分析到基特勒的媒介分析,从作为本体论式的构序场境之话语到作为本体论式的构序场境的媒介,在媒介场境中对包括话语的陈述实践——在福柯的意义上,这一实践再生产着其他层面的社会关系——在内的诸社会关系的再生产初步得以揭示,进而对媒介构序的探讨,即在媒介语境下对当今人与物处境的探讨也就成为可能。

二、双向度的断裂:数字时代媒介化的特征

既然揭示了话语—权力构序—媒介之间的内在关联,也就从社会的话语分析语境延伸到了作为物质性基底的媒介语境,而要进一步揭示在数字时代媒介的运作方式,就必须关注当下社会媒介化所呈现的特征。在这一向度上,经由对传播学、技术哲学等不同领域学者既有相关探讨的梳理,本文认为,当下的媒介化其实呈现着横向和纵向两个向度。然而,这两个向度并不是相互平行的关系,而是经由不同的分析视角对同一媒介化进程打开的一体两面。横向媒介化更多侧重于媒介化的表象具体层面,更加贴近日常对于媒介的感受和理解。纵向的垂直媒介化实际上是对横向媒介化表象之下的纵深揭示,也只有在这个意义上,媒介的物质性能量以及由之建构的媒介权力构序才得以显现。

(一)横向扁平原子化:"伦理区分"的消亡

海德格尔在《物》的开篇论述了物在现代交通和通信技术发展中的存在处境,即距离的消失。如其所言,"一切都被冲入这种千篇一律的无距离状态之中,都搅在一起了"(海德格尔,2018b)。也许并非着意,在对作为这一进程之背景的"无线电""电影""收音机"的指认中,海德格尔实际上隐性呈现了现代媒介化社会中万物的存在图景。横向媒介化之谓扁平原子化,其首要含义正在于由现代媒介技术对个体生活世界和整体人类社会的介入而制造的对物与人存在意义上的"抹平"。

距离的消失并不等同于海德格尔在《存在与时间》中作为此在在世方式的"去远"。虽然后者同样意味着"去某物之远而使之近"(海德格尔,2018c),但却是在此在建构与其他存在者的切近存在关系的存在建构意义上开展的。相反,现代技术,尤其是现代媒介技术导致的距离的消失,是一种暴力的去远性,在这种去远性中被抹去的不仅仅是距离,甚至连此在去远的存在建构本身也一并被抹去。此在并不能在这种距离的消除中找到自己存在的一席之地,恰恰相反,位置本身成为问题,因为一切都处于同质化的无位状态。张一兵在对维利里奥(P. Viritio)的速度论中所论及的数字资本主义批判的分析中,对这种暴力去远性也有所指认:"这个去远性中发生的光速接近是真实在场的数字化伪像,因

此是败坏的接近。”(张一兵,2018)这种同质化不但消除了存在者之间的差异性,也消去了真实在场与数字伪像之间的差距。

对这一处境的担忧,也正是海德格尔在《物》中所关注的核心问题,即重新探寻物之物性的提出缘由,存在的千篇一律正在取消物之物性,换言之,即物自身的存在。相比之下,对可能的核战争,即原子弹、氢弹对物的毁灭的恐惧只是在“令人惊恐的东西已经发生”(海德格尔,2018b)之后的无助而无谓的恐惧罢了。然而,即使是海德格尔最终诉诸的“天—地—神—人”的四方映射游戏,实际上也无从将物从这一毁灭进程中拯救出来——数字媒介已然将这四方本身的存在统摄于其中(至于这一进程如何进行,将在之后的垂直媒介化向度上再加以讨论)。

如果对《物》的讨论仅仅停留在“主—客”二分框架中作为客体的外物的存在样态层面,将这一处境仅仅视作与人无涉的“物”的处境,那么恐怕海德格尔就必须跳出来指认:“这正是说的阁下的事情!”事实上,物的存在与此在的关涉性,注定了物的命运折射着人的命运。除了前述存在建构层面讨论的物之存在的消隐对此在存在不可避免的影响之外,如果将这一讨论投射到当代生活现实中日常具体的存在感知,就不难发现,一种无差别的原子化构成了媒介化背景下人的存在方式。韩炳哲在其媒介批判著作《在群中》中,将当代社会中媒介所塑形的人的存在处境和社会关系归为“汇集而不聚集”的原子化个体群。原子化个体群相异于勒庞《乌合之众》意义上聚合的群体:勒庞所说的群体是中心化的,围绕少数中心意志而团结的无意识个体,而这些个体会聚合成思想高度一致的力量;然而原子化个体群则是去中心化的,其特征是“没有灵魂,没有思想”(韩炳哲,2019),即不存在中心化的群体意志,其组成个体只是偶尔地汇集,并不能将个体消融于群体中而聚合成为有方向的力量,只能产生反复无常和支离破碎的噪声。

在社会层面,这一原子化处境正源自数字媒介所塑形的媒介场境特征,而在韩炳哲的媒介批判中,作为其讨论前提的数字媒介的结构特征大致可以归纳为四个方面:无距离感、匿名性、即时性和对称性。

无距离感显然和海德格尔所说的“距离的消失”颇有相合之处,而韩炳哲所侧重的在于距离消失使公共领域和私人领域边界模糊化。由于社交媒体将信息生产转入私人领域,隐私成为可以被公开展示的内容,原先对于隐私的尊重不复

存在,这一过度的无距离消解了尊重的基础——尊重的本义就在于“有距离的顾及”(韩炳哲,2019),是从无距离、无界限的窥探中区分出的对距离的适度保持。

匿名性实际上意味着人的数字身份与现实身份、信息与发布者在数字媒体上的分离性。由于姓名是承载着责任、承诺等的现实身份之象征,也是人作为公民进入社会并被社会认可的基础,匿名在消解与姓名的这一关联时,也将对自我的社会约束以及对他人所负有的社会责任一并消除。

即时性消除了在传统通信方式中的延时性。在延时性的交流模式下,最初的愤怒等极端情绪尚未表达就已消散。而即时性通过对这一缓冲空间的取消,为未经反思的冲动发泄提供了便利的出口。

对称性即数字媒介的双向互动特性,在数字媒介中,“不存在明确的等级,把信息的发送者和接受者区分开来”(韩炳哲,2019)。这意味着信息发布不再是中心化的、垄断的权力,垂直单向的权力秩序被完全抹平,充斥在数字空间的信息是去中心化的无序逆流。其中,“网络暴力就是一种带有所有破坏性特征的逆流”(韩炳哲,2019),是这种无序性的极致展现。

这四个方面的特征,较为全面地展现了数字媒介在用户层表象层面的特征,也大体符合人们日常对数字媒介接触和运用的体验。其实,这四个特征是前述海德格尔“距离的消失”在不同方面的具体展现。除了显性的无距离感以外,匿名性是身份上无距离的同质化,即时性实际上是时间上无距离的真实在场与数字伪在场的同质化,对称性是这种无距离感所带来的扁平化体现,这种扁平化消除了建立差异以及在此基础上某种向度秩序的可能性。

此外,在数字媒介化时代,不仅经由媒介中介的人与物的存在被抹平了,媒介的普遍数字化让不同媒介之间的差异也被抹杀,这构成了横向媒介化呈现扁平原子化特征的又一向度。基特勒在《留声机 电影 打字机》的绪论中对此早有指认:“光纤网络遍天下……信道和信息数字一体化抹杀了各种媒介的个体差别。”(基特勒,2017)一方面,各种媒介在数字化意义上实现的互联互通诚然构成当代社会生活中诸多便利性以及相关产业经济运行的基础,这是日常生活经验已然呈现的,这一趋势同时也表现为媒介融合带来的新生可能性;另一方面,随着不同媒介边界的日益模糊,基于不同媒介生态之间的差异域保留的外在性敞开空间已然越来越小,差异日益受限于可控的同一性所掌控的范围。

福柯在其最后的讲座集《说真话的勇气》中,不同于既往对权力塑形主体过程的解构,反而着眼于“直言”主题,关注并寻找一种关于“直言”的肯定性主体建构伦理。“直言”的字面含义即“说真话”,但这并不意味着毫无顾忌、没有原则地乱说,也不是只说不会对自身产生威胁的真话,而是承担违抗权力或冒犯他人的危险的不加矫饰的表达,它承担着以逻各斯辨明真理和揭示真相的作用,同时也是建构主体自身以及主体的真实生活所必需的实践。对于“直言”的伦理而言,与之相伴随的,被福柯所指认为“伦理区分”的区分机制是必要的。这种机制借由逻各斯的运作而具有辨别真假善恶的能力,从而让真话在话语场中得以凸显,导向一种积极的伦理建构。在福柯看来,毫无区别、人人皆有发言权的民主并不利于直言,因为“在民主中,伦理区分不管用”(福柯,2018)。缺乏选择机制的民主无法区分真假善恶,也无从让直言者获得公正的对待。相反,由于“直言”并不谄媚多数人的意见,在无差别的民主中,其几乎一定会成为被排斥的对象。前述对数字媒介化横向维度的批判分析所指向的存在论意义上的“抹平”逻辑,导向的结论正是与“伦理区分”的分析最终同构的——数字媒介化所创造的话语场境,其本身的特性正在于以暴力取消了话语实践中的区分和边界,由此带来的失序性正在以前所未有的程度排斥社会话语场中进行伦理区分的可能,其结果体现为话语场中呈现出的诸乱象。因而,基于横向维度的媒介化分析常常描绘的“后真相”,即一种与真相关系的紊乱带来的普遍化失真这一图景,实际上构成了整个横向媒介化分析的缩影。这一分析逻辑诚然在日常媒介体验中是充分自洽的,然而自始至终不难发现,这一逻辑并未逾越本身剥离出媒介分析维度的“伦理区分”分析逻辑,本质上无异于对“伦理区分”的消亡这一主体伦理逻辑套用媒介语境的改写。横向媒介化分析的局限性正在于,其本身并没有真正深入媒介层面,缺乏对媒介物质性本身的反观,从而成为“无媒介的媒介分析”,套用媒介语境的话语分析,最终揭示的仅仅是媒介化呈现给我们的可见表象。对这种局限性的克服,将最终把我们引向对垂直媒介化的认知。

(二)纵向垂直控制化:新权力构序的诞生

媒介环境学的先驱麦克卢汉曾有一个广为人知的“看门狗比喻”:“媒介的‘内容’好比一片滋味鲜美的肉,破门而入的窃贼用它来吸引看门狗的注意力。”(麦克卢汉,2000)对媒介传播内容的关注,恰恰构成了在内容的背景层面

运作的媒介本身结构的遮蔽。同样,如果对媒介的讨论仅仅停留在媒介通常向人们呈现的用户层表象,实际上也是对界面背后真正塑形了表层场境的基底的遮蔽——而这正是前述讨论中横向的媒介化逻辑常常忽视的向度。

福勒(M. Fuller)在《媒介生态学:艺术与技术文化中的物质能量》中指认了通常被认为是“非物质”的电子媒介领域,反而恰是对媒介物质性的关注“最有收获之处”(福勒,2019)。在用户层面看似非物质性的数字空间,实际上是有待解蔽的复杂物质性媒介系统的运作结果,而媒介生态所关注的正是在媒介的动态作用中物质性的结构驱力所创生的生成性力量。这种生成性力量,福勒将其与尼采(F. Nietzsche)的“权力意志”、德勒兹(G. Deleuze)的“生成机器”和道金斯(R. Dawkins)等人的“模因”理论等相联系,表现了媒介场境的自我生成性。

在福勒看来,对媒介物质性的忽视本质上是本体论层面的形质论传统的延伸。在形质论中,形式和质料是二分的,形式外在于质料并且被赋予质料,从而让质料得以赋形;然而,在媒介实际的运转和演化中,“物质材料在同存在于其周遭的形态发生能供性的关联中生成其自身的形构能力”(福勒,2019)。值得一提的是,这里的媒介物质性层面的生成性与前述讨论中福柯所指认的话语塑形、基特勒所指认的话语网络的媒介塑形颇有相通之处。其实,这是不同层面对形质论、本质主义的超越,也是对媒介生成性、媒介体系自身再生产机制在不同粒度上的言说。

福勒从四个方面延伸了对这一指认的探讨:从对媒介的边缘运用中交叉使用媒介系统的案例,即英国曾经流行的海盗电台切入媒介发展的历时性进程,揭示对于媒介要素的组合如何生成新的媒介;聚焦于照相机这一媒介对象,并由此分析其物质性结构与生成外部关系性的共时性机制;深入媒介预构的标准对象的形成机制,以及在特定环境下的媒介标准对象生成新的意义和功能性的可能,将标准对象的生成与对怀特海(A. Whitehead)“具体性错置”生产性的解读相结合;借助复制性模因的概念分析媒介关系性维度的生成机制,展现在不同关系性维度的复制和转译中所实现的媒介的可能性生成网络。

福勒的媒介生态学关于媒介内在生成性的指认,至少在两个层面拓展了对媒介的言说可能:其一,媒介是具有一定独立性的内生系统,并不能将其仅仅作为内容的载体来看待,也不能仅仅停留在其向人所呈现的媒介内容的表象层面来加以探讨,而要突破传统的形质论二分框架,深入媒介的物质性基底来探寻其

自身的塑形和再生产方式;其二,媒介并不是可以被孤立研究的具体对象(某个媒介物、某种媒介技术等),也不是某种观念抽象的泛化媒介概念(万物皆媒介,词源学、语义学意义上的媒介概念等),而是需要关注实际存在和运行于界面之下的潜在媒介机制,关注交错的媒介关系和媒介层面,从而对其进行带有现实抽象意味的整体动态的考察。在这样的意义上,对媒介的考察从平面走向纵深,看似扁平的媒介场境在实际上被揭示为纵深的媒介物质性构序所生成的在特定视角下的表象结果。

布拉顿(B. Bratton)在《堆栈:论软件与主权》(*The Stack: On Software and Sovereignty*)中,正是在这一纵深维度上展现基于数字媒介的"行星规模计算"时代的政治地理构序。而这一构序,正如媒介生态学所揭示的,是各个不同层级的媒介布局的交错和结合:"当我们将大陆规模的基础设施、城市规模的普适计算和感知规模的环境界面联系起来时,我们将探索这些相互交织的方式,以及我们如何构建、栖居、沟通和治理我们的世界。"(Bratton,2015)

"堆栈"是对这一媒介布局的多层结构描述,是包括"软件、硬件和网络"(Bratton,2015)在内的不同技术模块化、关联化的垂直性排列。这也构成了不同于传统威斯特伐利亚体系基于扁平地图的权力地理布局的纵向复杂性辖域布局,"大地的法"延伸到"云法",在这个意义上,前述海德格尔的"天—地—神—人"四方在一种普遍的可计算性和普遍的数字模拟中被消去了独立的存在意义。具体而言,"堆栈"可以分为六个层面:地球层、云层、城市层、地址层、界面层和用户层。这六个层面"从全球到地方,从地球化学到现象"(Bratton,2015),在其交织的运算生态中实现其中各个个体存在关系和存在方式的再生产。作为栖居于"堆栈"体系中的个体,"用户并不是一种生物,而是一类行动者;它是系统中的一个位置,没有这个位置,它就没有角色或本质身份"(Bratton,2015)。对用户而言,日常生活中通常可见的只是用户层和一部分的界面层(如各种应用程序的可视界面),而在这些层面上看似简单的操作却意味着纵深维度上相关联的各层面复杂的交互运转。这意味着,如同作为生产之结果的物背后遮蔽的是复杂的生产过程,用户所能接触到的媒介界面只是媒介体系层层中介之后的运算结果,其中的数据传输、算法机制等都被遮蔽于可见领域之外。

这些被遮蔽的过程和层级,在当代社会恰恰是媒介权力——如平台资

本——借以掌控和利用用户个体的平台权力的基础。祖博夫(S. Zuboff)(2020)在《监控资本主义时代》中所指认的监控资本主义,以“无所不在的数位控制为媒介,将自身意志加之于人”的平台大他者,正是掌握了数字媒介体系的这些隐性过程、层级从而实现的基于数据搜集、传输、占有、利用实现对个体行为的预测和掌控的媒介权力。

这种权力不仅仅作用于个人,更在社会层面以及传统的“大地法”的领土辖域层面产生了深刻的影响。帕克斯(L. Parks)在《重思媒介覆盖》(*Rethinking Media Coverage: Vertical Mediation and the War on Terror*)中以“垂直媒介化”表现了美国在“9·11”事件之后建立的基于地面、空中、光谱、轨道的全球垂直霸权。垂直霸权的向度拓展了传统军事争夺的领域,战争被吸纳到信道中,或者对信道的军事化,将军事、宣传、市场等社会生活各领域纳入总体战争中。这种垂直战略,就是“将国家的军事化和父权制等级制度自然化为一种生活方式,并使其得以横向复制和扩展,渗入无数的生活领域”(Parks,2018)。“9·11”事件之后,美国在阿富汗和伊拉克的行动就是这一进程的典型案例。除了进行直接空袭等对天空的军事化利用,美军还广泛开展心理战(PSY-OPS)项目,调动军用飞机干扰当地政权的电信系统,并播送美国军方和商业公司的娱乐内容,通过卫星、飞机、电信基础设施等媒介的综合利用来实现垂直控制。而在2002年准备对萨达姆政权发动战争时,美国利用空中广播,散播萨达姆政权拥有“大规模杀伤性武器”等政治宣传,并借助其全球监控能力,以及对全球信息传输的硬件、软件支配,向世界散播萨达姆政权的恐怖形象(Parks,2018)。垂直媒介权力并不是和传统政治权力、资本权力等平行的权力,而是与这些权力深度交织和融合,成为其在当代语境下不可或缺的力量;但同时,垂直媒介权力基于一定的媒介物质性结构,其自身结构驱力具有一定独立性,这种独立性也就意味着其与传统权力模式之间存在张力,不能简单还原到传统权力的讨论范式中。

在媒介看似抹平一切的表象背后,是无法被还原到表象平面的垂直性的新权力构序借由媒介物质性基底而生成的幻象。在这个意义上,数字媒介所塑形的场境绝不是自由、平等、博爱的乌托邦,正如韩炳哲正确指认了“更多的自由也就代表了更多的强迫”(韩炳哲,2019)这一事实。然而,指认数字媒介化只是混沌无序的荒原,是一切权力和秩序都无法确立其统治性的领域,甚至认为数字媒介的无距离性是消除信息通达受众的中介层级的一种“去媒体化”(韩炳哲,

2019)体现,则无异于对唯有在纵深维度的观照下才得以显现的媒介权力构序的遮蔽。换言之,“去媒体化”的提法实质上是对“深度媒介化”的误认,媒介的不可见性并不意味着媒介的消失,而正是媒介的场境在纵深层面的扩展和水平层面的弥散之体现,媒介权力作为不在场的在场的大他者潜于其中。

三、媒介本体论:数字媒介化时代的生存思考

媒介本体论的提法,在基特勒的《走向媒介本体论》一文中早有讨论(Kittler,2009)。然而,基特勒的分析并不真正指向媒介本体论的具体建构,而主要是基于哲学史层面的回溯对媒介本体论之言说可能的探讨。本文认为,在当代,媒介理论走向本体论层面,绝不仅仅是理论发展的需要,而是深化对数字资本主义时代下人的处境,以及人、社会与媒介的存在方式的认知,是探索当代实践语境下更合理的人、社会与媒介间关系的需要。本文试图以狭义历史唯物主义的视角,在存在方式的意义上重构关于当代数字媒介化的思考,并由此提出肯定性媒介化的可能建构。

(一)从技术物到数码物:数字时代媒介本体论的基底

亚里士多德在《形而上学》中提出“‘何谓实是’亦即‘何谓本体’”(亚里士多德,1959),追寻“实是”(存在)就是要追寻存在的事物基底存有的实在(本体或个体)。在这个意义上,将媒介反思延伸到本体论层面,首先必须回答的问题是:什么是构筑这一本体论场境的基础?

在前述讨论中,经过对水平向度扩展的媒介化表象层面的透视,媒介场境的基底已经被揭示为垂直层级上的媒介物质性结构,这一分析实际已然带有本体论言说的色彩。然而,“物质性”本身仍然是有待揭示的问题。原因在于,当我们言说物,尤其是言说人与物的关系时,物绝不是物自体式的无法以任何方式向人呈现的意义上的存在者,而是关涉人,与人发生某种存在关系的存在者。既然如此,对物的言说就必然指向在一定范式下被揭示的物。

对媒介物质性言说的范式,显然并不能将媒介还原到现实空间中以实体形式存在的广延(如具体的媒介设备和基础设施等),正如无法将人还原为组成人的种种化学元素——结构和关系并不能被排斥于物质性的讨论之外。而在数字

媒介时代,逾越现实空间意义的数字空间、非实体形式存在的数据等新的物象形式,也使得传统对物的言说范式显现出更明显的局限性。

技术哲学学者许煜将物在本体论思想史意义上的不同指涉界划为自然物、技术物与数码物三个阶段。实际上,对物的三重不同理解即三个不同的本体论范式的体现。自然物意指“将自然或人工的事物以同等的自然方式分析的范畴或视角”(许煜,2019),而这一范式所要求的就是“把握事物的本质,即确定其独特存在的东西”(许煜,2019),如亚里士多德的“形式”(eidos)、康德的范畴等;技术物的讨论范式打破了自然物范式中常见的形质二分框架,将本体从实体转向环境或关系性的理解,如西蒙东所指认的“技术物是关系的统一体”(许煜,2019)、海德格尔在对“上手状态”的讨论中以关系互动的方式理解物等;数码物则基于普遍语言和普遍运算,“指成形于屏幕上或隐藏于电脑程序后端的物体,它们由受结构或方案(schema)管理的数据和元数据组成”(许煜,2019),数码物的讨论并不完全独立于技术物范式,而是对它的延伸,其所拓展的是与技术系统的物质关系维度,在这个意义上,聚集和综合社会其他系统的受算法控制的媒介网络以及其中个体之存在得以揭示。

因而,媒介的物质性,即媒介本体论的基底存在,不是主客二分框架下的客体物,也不是形质二分框架下的质料,而是在技术物和数码物范式意义上的构成对象,既包含实体形式存在的媒介基础设施,也包含非实体形式存在的数据和元数据,而不同形式的物都在媒介技术系统的网络中确定自己的位置。

(二)新掌控构序与异化:深度从属与媒介化的生命

如前所述,媒介在当代是作为塑形诸存在关系和存在方式的场境而存在的,而对这种塑形方式的分析,以及在这一本体论场境中人的存在境况的分析,也就构成了这一本体论场境的运行机制,换言之即该场境的再生产模式。

在《资本论》中,马克思曾经在对资本主义生产方式的讨论中,区分了劳动对资本“形式上的从属”和“实际上的从属”(马克思,2004)。在形式从属阶段,劳动虽然在生产关系上从属于资本,产出绝对剩余价值,但劳动的具体过程仍然是分散的,没有受到资本的直接掌控;而在实际从属阶段,劳动过程本身已经被吸纳到资本掌控之中,表现为集中生产。直到实际从属阶段,劳动生产力才完全从属于资本逻辑,资本主义生产方式才真正得以确立。从形式从属向实际从属

阶段的转变，其实质是资本主义对经济关系的掌控延伸到对主体劳动本身的掌控，从而将资本主义掌控构序扩张到整个生产过程。

而在当代资本主义中，资本主义的掌控并不仅仅停留在生产领域。一方面，劳动的边界模糊化，在时间和空间上溢出传统的固定工时和工厂区域的限制，这意味着对劳动的掌控也随之广泛介入主体生活；另一方面，生产之外的领域也日益资本化，如巴里巴尔（E. Balibar）在《走向新政治经济学批判：从广义剩余价值到总体从属》中所指认的，教育领域的标准化体系、竞争机制和教育资源私有化将人预置为服务于资本增殖的人力资本；医疗领域通过对身体的量化掌控，将医疗由面向主体失常不适的状态，转向面向可被持续监测、持续纳入优化框架的身体指标，从而通过创造医疗需求的方式实现资本增殖（Balibar，2019）。在这两个案例中，人的塑形过程本身被置于资本逻辑的掌控之中，资本不但在生产领域攫取剩余价值来实现自我增殖，也通过创造需求等方式实现自我增殖。与此类似的是，今天人的闲暇也并非自我规定的自由时间，而只是以资本所规划和预置的方式进行娱乐。由此，资本主义掌控构序延伸到社会再生产的各环节、各领域中。据此，巴里巴尔在马克思“形式从属”和“实际从属”概念的基础上提出“总体从属”（total subsumption）的概念，表现资本主义掌控构序溢出生产的扩张。

诚然，这一指认拓宽了当代资本主义批判的视野，在当代语境下延续了马克思主义政治经济学批判的思路。然而，如果回到马克思对形式从属和实际从属的言说，马克思的这一区分本身并不仅仅指向价值增殖领域的扩展，也指向资本主义时代区别于前资本主义时代所特有的生产方式的确立标志，在历史唯物主义的资本发展史中作为资本关键的质性变化（而不仅仅是数量或范围上的部分增加或扩展）。总体从属要作为其批判思路在同一层面的延续，就必须进一步指认这一概念在资本以及资本逻辑建构下的社会在质性变化意义上所具有的断裂性，仅仅对教育和医疗领域资本化的表象分析来加以言说，无疑是不充分的，也遮蔽了进一步揭示这一转变原因的可能。

如同要揭示从形式从属到实际从属就必须回到物质资料生产方式的社会物质层面，新的从属转变的指认如果要获得其批判范式上的讨论意义，也必须在物质性基底层面寻找其模式的转变。对巴里巴尔“总体从属”概念进行更细致的考察后最终会发现这一转变的实际发生正是在资本与媒介权力媾和的基础上。

无论是教育还是医疗领域资本化的转变，在其运作的标准化评价体系、持续

纳入医疗进程的生物指标等单向度的优化逻辑的社会建构中，经由数字媒介实现的基于大量数据搜集与分析的充分可预测性往往是不可或缺的。事实上，这已经成为当代社会中人生存处境的一部分，当人被预置为人力资本和医疗需求时，人已经首先在媒介体系中被生产为媒介化的生命。库克里克（C. Kucklick）将数字资本主义时代揭示为数字化解析–解体基础上的控制革命，其中的个体都作为媒介化的生命而存在。以医疗领域为例，患者被量化指标解析为数字化单体并纳入预定的优化逻辑的过程，正是其典型体现："在美国，大约2/3 的成年人至少会掌握一项关于自己的数值，要么是血压、血糖，要么是体重或者其他指标……每一个人都将是一个孤本，一个单体，而且他们正是作为这样的单体而得到医学上的治疗。"（库克里克，2018）同样，将这种资本掌控构序下的优化逻辑推展为社会接受的运行逻辑的过程，也离不开媒介体系的运作。在《技术与时间（3）：电影的时间与存在之痛的问题》中，斯蒂格勒揭示了社会意识的接受过程正是社会个体的意识被媒介第三持存深刻影响的过程，而电影、电视、数字网络等数字媒介正在以前所未有的即时性和普遍性将社会意识整合为媒介权力统摄下的巨流，资本主义的文化工业图式借此成为整个社会持存集体记忆和投映意识前摄的基础（斯蒂格勒，2016b），而个体主体投映能力在媒介持存意识流取代下的丧失——斯蒂格勒所说的"精神的无产阶级化"——正是总体从属所指认的通过对需求的生产实现资本增殖的前提。

在这一媒介基础上的资本掌控构序下，人的生活和生命无所不在地被置于资本逻辑之下，人失去了自身投映真实欲望和谋划自身生活的能力。除了当下的生命，过去的记忆和未来的前摄也被吸纳到资本增殖过程中，因而当代的社会主体无法在资本生产的意识巨流裹挟中找到自身位置和自我效能感的"存在之痛"与异化感也就无法避免了。

如果结合前文所讨论的媒介场境的垂直权力构序，对总体从属更加合适的提法应是"深度从属"，因为从属的权力运作过程不仅仅是在平面上拓展了范围，更是在纵深层面转变了运作方式，从而将资本掌控提升到空前的程度。

在这个意义上，对从属的讨论亦同时延伸到资本主义掌控构序在空间领域中的转变。"形式从属—实际从属—深度从属"的进程，在空间上也就呈现为"分散—集中—弥散"的变化，弥散不同于形式从属阶段由于资本掌控构序范围和掌控力上的有限而形成的消极边界外的分散境况，而是资本掌控构序借助媒

介的垂直布局实现的更加彻底深入的全方位、立体化的渗透，也正因为如此，资本的掌控更多以不可见的形式运转。

（三）走向肯定性的媒介化：媒介的再发现及其重塑力量

对于媒介化进程，除了技术主义的盲目乐观倾向，现有的理论分析大多着意于对媒介化的否定性批判。前文提到的韩炳哲对媒介化的描述，几乎将媒介化场境当作摧毁一切秩序、权力乃至人自身的荒漠，祖博夫和帕克斯等人也将媒介当作资本和政治控制个体、实行霸权的工具。然而，借用贝尔纳·斯蒂格勒在访谈中的一段话："哲学的任务就是去摧毁悲观主义或乐观主义，从这两端之间走过去寻找真正的问题。"（斯蒂格勒，2015）否定性批判固然揭示了数字资本主义时代媒介化暴露的问题和潜在的危机，但当涉及当下社会的实际问题，以及寻求人、社会与媒介的合理关系时，全然否定性的批判便显现出其无力的一面。

在这一语境下，讨论肯定性的媒介化之可能，既不能抛开既有媒介理论对媒介化的讨论，也不能忽视媒介在当下已然作为社会运行不可或缺的场境这一状况。因而，首先必须回应既有的否定性批判，并从中寻求在逻辑和现实层面超越这些否定性的可能。

在此，本文提出三个方面的探究，试图在对既有媒介批判范式常见模式讨论的基础上，展现对媒介再发现的可能。

其一是人本主义的媒介批判。这些媒介批判首先预设了前媒介的纯粹主体性，而对于任何媒介对主体性产生的影响和干涉都首先采取抗拒的态度，如马修·福勒指认媒介环境学的实质是一种"环境保护主义"，其"怀有一种人类永续存在的想象，并试图确保这个世界对人类的安全性"（福勒，2019）。然而，这一范式的问题在于，这样的预设忽视了人的存在从来不仅仅是纯粹的存在，而是在与媒介的关系中共生的存在。斯蒂格勒曾用"爱比米修斯的过失"的神话来表现人与技术在存在意义上无法割裂的关系。人的存在从来不是自源性、自足性的纯粹主体性，而是代具性的存在，其存在本身即为技术代具所分有，技术是人的真正先在性，"人在本质上作为技术性存在"（斯蒂格勒，2012a）。拉图尔也曾批判人类主义者"通过贬低物、贬低客体、贬低机器"（拉图尔，2010）的方式将人归为某一本质，而实际上人类与机器本身即具有不可分割的共生关系，人类"将自身化身于机器之中……它甚至用机器来建造自己的身体"（拉图尔，

2010)。尤其在当下的数字媒介时代,人的存在与媒介更加深度融合,在这个意义上,人必须被视为与其所关涉的媒介共生的共同体或结晶体,正如人的身体本身是无数微生物的共生体一样。同时,人的主体性,如前所述,也需要在媒介场境的关系中得以确立。

其二是自由个人主义的反治理批判。这些媒介批判预设了自由与治理之间的对立,将一切治理不经反思地预设为对个体自由的侵害,因而对于经由媒介而得以运行的治理权力,这一批判范式也延续了这一前提。如本雅明·布拉顿所批判的以祖博夫为代表的"自由个人主义"的反监控话语,这些话语的内在前提是"我们首先是自我主权的个体,带有与生俱来的在自我同一性边界之中的隐私"(Bratton,2021)这一信念。然而,一方面,这一范式忽视了自由实际上首先是在一定历史现实语境中的自由,并不能脱离一定的社会关系来空谈抽象的个体自由;另一方面,这一范式也忽视了必要的治理对维系社会正常运行的作用,尤其是在应对社会危机时,这一范式将暴露出其符号化、象征化的自由个人主义话语的脆弱性,在布拉顿所谓的"真实界的复仇"面前不堪一击。

需要指出的是,上述两种媒介批判模式都是超历史性的,或者预设超历史的纯粹主体,或者预设超历史的个体自由,而并未真正将媒介反思置于具体的历史现实条件中,本质上是一种以预设的应然批判实然的思路。这样的批判虽然提供了一种反思向度,但实际上与严谨的反思相距甚远,也难以从中产生建构性的力量。

其三是将媒介还原到资本权力和政治权力批判中,将媒介的可能性仅仅视为服务于既有资本权力和政治权力的工具。祖博夫对监控资本的批判和帕克斯对垂直媒介霸权的批判都存在着这一倾向。诚然,正如前文所讨论的,当代媒介的垂直构序为权力的深度掌控提供了前所未有的便利,其与资本、政治权力的不当媾和,正在生产深度从属于掌控构序的媒介化生命。

然而,媒介体系的内在驱力和生成性,本身并不能还原到既有的资本或政治权力中。媒介自身的运行逻辑固然在很多时候从属于既有权力逻辑,但这并非一种融合的同一,而是存在内在张力的复调构序。

在《媒介生态学:艺术与技术文化中的物质能量》中,马修·福勒所举海盗电台的媒介生态案例,或许是这种张力显现的一例。海盗电台本身是游走于合法边界的媒介生态,在与对其进行压制的权力的对抗中,海盗电台通过各种媒介

超越既有模式的创造性组合和运用，形成了与试图规制它们的力量“共同进化”（福勒，2019）的关系，其媒介生态中内在生成的“无关联的集体主体性，是通过多样化媒介形式，排列于多样化制作网络、多样化运行地点之中的。这种主体性是在各种关联中被动员起来的”（福勒，2019），而在权力对其的压抑中，“因为那些拒绝屈从的场景及节奏动感驱力，它被维持了下来”（福勒，2019）。媒介体系自身的生成性赋予其一定的不可还原的独立性和抵抗性，而这种生成性的力量在生成权力构序的同时也在生成逆构序的可能性，生成权力构序之外的敞开的可能性。

媒介体系，在可以被权力构序利用的同时，也将是冲破和重塑构序的生成力量，而它的可能向度需要在历史现实的具体语境中不断被重新发现和拓展。在这个意义上，我们可以适当地重提和借用拉图尔的问题：“威胁来自何方？”（拉图尔，2010）媒介的生成性力量本身并不是真正的威胁，威胁来自试图将媒介归为某一本质之人，通过对媒介某一向度的固化和垄断，遮蔽了人类与媒介合理共生的其余可能。走向肯定性的媒介化的第一步，就是走出超历史的和还原论式的媒介批判范式，重新在生成性的向度发现媒介，在本体论层面认识和把握媒介构序的复调性。在这一基础上借由媒介生成性驱力，重新塑形既有构序以及由此生成的新主体性，方有可能走出被数字资本异化的宿命。

参考文献

车致新，2019，《浪漫主义作为媒介技术——基特勒论“话语网络 1800”》，《国际比较文学（中英文）》第 3 期。

福柯，米歇尔，2018，《说真话的勇气：治理自我与治理他者（Ⅱ）》，钱翰译，上海：上海人民出版社。

福柯，米歇尔，2021，《知识考古学》，董树宝译，北京：生活·读书·新知三联书店。

福勒，马修，2019，《媒介生态学：艺术与技术文化中的物质能量》，麦颠译，上海：上海社会科学院出版社。

海德格尔，2018a，《巴门尼德》，朱清华译，北京：商务印书馆。

海德格尔，2018b，《演讲与论文集》，孙周兴译，北京：商务印书馆。

海德格尔，2018c，《存在与时间》，陈嘉映、王庆节译，北京：商务印书馆。

韩炳哲，2019，《在群中》，程巍译，北京：中信出版社。

基特勒，弗里德里希，2017，《留声机 电影 打字机》，邢春丽译，上海：复旦大学出版社。

库克里克,克里斯多夫,2018,《微粒社会》,黄昆、夏柯译,北京:中信出版集团。

拉图尔,布鲁诺,2010,《我们从未现代过:对称性人类学论集》,刘鹏、安涅思译,苏州:苏州大学出版社。

马克思,2004,《资本论》第 1 卷,北京:人民出版社。

麦克卢汉,马歇尔,2000,《理解媒介:论人的延伸》,何道宽译,北京:商务印书馆。

斯蒂格勒,贝尔纳,2012a,《技术与时间(1):爱比米修斯的过失》,裴程译,南京:译林出版社。

斯蒂格勒,贝尔纳,2012b,《技术与时间(3):电影的时间与存在之痛的问题》,方尔平译,南京:译林出版社。

斯蒂格勒,贝尔纳,2015,《独家专访德里达弟子斯蒂格勒:被大数据裹挟的人类没有未来》,https://www.thepaper.cn/newsDetail_forward_1309683。

夏瓦,施蒂格,2018,《文化与社会的媒介化》,刘君、李鑫、漆俊邑译,上海:复旦大学出版社。

许煜,2019,《论数码物的存在》,李婉楠译,上海:上海人民出版社。

亚里士多德,1959,《形而上学》,吴寿彭译,北京:商务印书馆。

张一兵,2018,《败坏的去远性之形而上学灾难——维利里奥的〈解放的速度〉解读》,《哲学研究》第 5 期。

祖博夫,肖莎娜,2020,《监控资本主义时代》,温泽元、林怡婷、陈思颖译,台北:时报文化出版社。

Balibar, E. 2019, "Towards A New Critique of Political Economy: From Generalized Surplus Value to Total Subsumption." in Peter Osborne, Éric Alliez & Eric-John Russell eds., *Capitalism: Concept, Idea, Image Aspects of Marx's Capital Today,* London: CRMEP Books.

Bratton, B. 2015, *The Stack: On Software and Sovereignty,* Cambridge: MIT Press.

Bratton, B. 2021, *The Revenge of the Real: Politics for A Post-Pandemic World*, London: Verso.

Kittler, F. 1990, *Discourse Networks 1800/1900*, Michael Metteer trans., Stanford: Stanford University Press.

Kittler, F. 2009, "Towards An Ontology of Media." *Theory, Culture & Society* 26 (2-3).

Parks, L. 2018, *Rethinking Media Coverage: Vertical Mediation and the War on Terror*, New York: Routledge.

认同焦虑与公共性旨归*

——技术治理合法性困境的生成机理与因应

王　磊**

摘要:技术应用场景及其效果的不确定性与“无知”或“不可知”治理问题的凸显,在一定程度上消解了现代行政国家的理性基质与工具基础。对于技术治理的社会认同因治理效能限度、伦理尺度失能以及“权力—权利”秩序结构失衡而式微,进而引发合法性危机。技术治理的合法性危机深植于基层治理公共性价值的退守和流失。因此,确证技术治理合法性基础、凝结完整的社会共识需要技术理性与价值理性、形式理性与实质理性、程序理性与结果理性共同担保,还要重启技术的民主化进程,构建一种集技术的自我约束、道德伦理规范和刚性制度规制的控制体系,助推技术治理的公共性旨归。

关键词:技术治理　合法性困境　公共性再造　科学技术民主化

一、问题提出:科学不确定性背景下基层治理中的技术合法性问题

随着 ABCD 技术(人工智能、区块链、云计算、大数据)在社会治理领域的深度应用,“治理场景+治理技术”——政府治理技术跃进及其范式变迁催生了治理结构的技术化嵌入、社会问题的技术化解决、治理形态的技术化变革等样态,使得传统社会中的“数目字管理”难题正被不断跃升的现代技术所消解。基层治理的技术化、数字化和智慧化转型,成为基层治理体系与治理能力现代化建设的重要内容和基本方向。无论从既有研究的整体脉络和总体趋势,还是从技术应用引起的公共治理转型和治道变革来看,人们对技术及其工具性价值的社会

* 本文系 2022 年上海市哲学社会科学规划年度课题青年课题“超大城市老旧小区改造政策执行中的目标匹配困难与优化对策研究”(项目批准号:2022EZZ001)的阶段性研究成果。

** 王磊,同济大学政治与国际关系学院。

关注主要集中于技术的物理影响,忽视了技术治理在合法性层面的价值影响,从而引发了技术治理经验探索和理论研究之政治价值的转向——重新回到“反思技术治理的合法性”这一前置性问题上来。

技术在发挥其治理价值的同时,各种副作用也在不断显现。国内外学术共同体围绕“技术治理及其风险”展开了较为充分的讨论和研究,这些研究主要围绕技术治理的治理效能(技术治理的有效性)与受控强度(控制技术风险的能力)之间的张力如何平衡而展开,并形成了以下几种研究样态(表 1)。

表 1 基于效能和控制二分的技术治理适应场景及其类型划分

技术治理维度		治理效能	
		强	弱
受控强度	强	技术善治 (技术治理的理想形态)	可控型技术化治理 (“边使用边治理”)
	弱	失控型技术治理 (“先使用后治理”)	—

基于技术有助于解决棘手问题的功能效益而产生的“技术工具论”(instrumentalism)或者技术赋权理论,为技术治理描绘出一种“全能主义”的乌托邦式治理情境。基于“赋权+赋能”双重功能耦合的技术万能论和技术泛在论,认为复杂的基层治理问题能够先验地被简化为能够技术识别的简单问题,进而技术成为解决基层治理情境中“抗解问题”(wicked problem)的救命稻草。因此,基层治理的技术化转型预设了技术工具理性的超然性,并将技术工具的跃迁视为回应和解决复杂性社会问题的唯一途径。

与此相反,科学技术发展段位的局限性和不确定性为技术治理带来的“胎里病”,成为社会科学家关注的重要议题,尤其是对技术治理“胎里病”的防治,催生出两种不同的研究取向。其一,强调技术应用以科层体制对技术强控制为前提,其治理效能的发挥有赖于既有制度结构对技术的规制,因此可将此类技术治理定义为“可控型技术化治理”(Lishan, Qin & Yongmou, 2021)。对这种技术治理现象最有力的理论解释,来自芳汀(J. Fountain)“被执行的技术”(enacted

technology)。作为基层治理场域的核心行动者,政府科层倾向于一种既能提高治理效能又能维持组织现状的方式,将技术纳入行政轨道。这种“即插即用”(plug-and-play)的技术应用模式强调技术嵌入逻辑与官僚科层逻辑之间的张力(芳汀,2010)。其二,强调技术治理的强工具理性,而忽视对技术副作用的制度规制,使得技术与其他要素(资本、权力、数据信息)“合谋”,产生难以规训的“技术利维坦”(王小芳、王磊,2019)。随着学界对此类问题的深入理解和阐释,技术治理研究开始追溯“技术”和“技术治理”的“元命题”——技术与政治关系(王磊,2023),从“技术”本源(是什么)的理论视角阐释“技术治理”的结构形态和权力逻辑(颜昌武、杨郑媛,2020),并以此思考这种“技术副作用”或科学技术的不确定性为国家治理技术化转型带来的合法性危机。这种对“技术治理”热现象的“冷处理”,正在成为全力阐释和审视“技术治理”热潮的重要补充。

从技术治理的适应性场景来看,针对技术治理合法性的争论,学界尚未获得普遍认可的结论。就基层治理情境与技术治理适配程度而言,传统的工具理性与技术治理范式在科学不确定性背景下已经面临失灵的问题。其原因在于:一方面,技术治理在一定程度上解构了依赖于层级化权力结构和属地化管理方式的传统科层式治理,分散化的权力结构重组和复合型的行政治理尺度再造尚未重新寻得合法性资源的再生长空间;另一方面,技术应用场景难以精准锚定基层治理任务情境,并最终导致基层治理技术化转型呈现出“目标倒置”和“走向技术操纵和技术统治”的风险。

基层治理场域中的技术及其应用合法性问题,悄然成为游荡于基层治理技术化转型过程中的一个“幽灵”。既有研究认为,技术治理合法性的流逝是因为在由科学不确定性建构的高度复杂的社会技术环境中,我们没有足够的经验或坚实的专业知识来确定和应对技术后果。基于此,科学技术权威的矛盾性其实是高度发展的技术文化的重要内在特质。因此,对技术治理合法性危机的分析需要调转研究因果箭头,将分析对象重新拉回对技术的认同及其文化特质上来。

二、技术认同的复合式衰弱:技术治理合法性危机的三重维度及其表征

作为国家治理现代化的技术图景和治理叙事,基层治理技术化转型正成为推进国家治理体系和治理能力现代化的重要目标和实现工具。技术的嵌入式效率至上成为国家治理的基本价值取向,同时理性化、专业化、纪律和协作成为国家治理的政治合法性基础(余敏江,2010)。所以,构建一种“社会意义上的健全知识”取代“技术生产流水线的认知模式”(比克、保罗、亨瑞克斯,2015)来扩展和应对科学不确定性的知识结构和主观能力,成为理解技术治理合法性危机的重要基础。然而,新技术落地运行困难、治理程序扭曲、“手段—目标”倒置、数字痕迹主义等技术独有困境和技术治理悖论,正在降低技术治理及其效能的社会认同,进而蚀耗技术治理合法性生成。

(一)绩效合法性式微:技术治理的功能边界及其治理限度

绩效合法性直指科学技术在治理场景中的工具价值,并外化为对技术治理解决实际问题能力的评价体系,因此它反映了技术治理的效能及其边界。

首先,技术治理效能遵循技术处置的“简约逻辑”,只能有效解决能被“技术处理”的治理问题。技术治理内含的“简约主义”能将复杂模糊的棘手问题简约化、清晰化为可被技术识别的“数字”,并通过量化信息收集、算法运算、精细化管理以及智能模拟处置等机制(王磊,2021a),形成以科学技术为治理支架,以满足人民美好生活需求和人民本位观为治理价值,以维护社会公共秩序、实现提升政府治理体系与治理能力现代化为治理目标的治理范式。所以,技术治理明显的绩效合法性在于,它擅长改变治理主体的信息结构,优化工作流程。尤其在一些强技术性和高替代性治理任务中,这种“小规模的精准化诊疗式治理”在很大程度上解决了决策者知识和信息方面的困难;同时,技术治理提供的运算功能,提高了决策者的分析能力,自动化处置和精细管理方面的能力也相应得到提高(彭勃,2020)。因此,工具主义技术治理观内蕴着打通国家治理体系的制度优势及其治理效能转化的巨大潜能和技术方法。然而,这种转化途径形成了对技术的路径依赖,治理效能的发挥深受技术发展及其段位的影响,技术本身构成

了制约基层治理效能的外部因素之一。

其次，基层技术治理的实质理性被替代为“完成任务逻辑”下的技术比拼，构成了技术治理有限性的另一重要维度。技术治理应然层面的工具理性，总是会受到外源性“技术段位”和治理主体内生的“自主性策略”影响，导致技术治理目标异化——现代技术应用于基层治理实践过程中产生的“技术化形式理性”，置换了公共治理“实质理性”的现象（董石桃、董秀芳，2022）。“渲染现代技术采纳”“技术作秀”“技术泛化”成为基层治理主体应对技术化转型考核的“策略性行为”，这种目标替代为技术治理效能发挥套上了限制性“枷锁”：一方面，技术治理过度反映治理主体以及核心行动者的行动取向（特别是反映权力意志），形成了官僚科层主义与技术复合的谋利型“技术反治理”（刘永谋，2019），致使技术治理偏离了治理目标本原；另一方面，外源性技术嵌入基层治理过程本身会对治理体系中的组织和个体产生意外的或“不耐受”的负面影响，即技术治理的负外部性。外源性技术引进会增加大量用于培训、运营和维护技术系统的行政支出，衍生出的更大规模的行政结构会使得行政体系更加迟钝（渠敬东、周飞舟、应星，2009）。可见，高效的技术治理依赖于高成本的维持，因而在一定程度上弱化了其绩效合法性。

（二）道德合法性阙如：“人技”关系倒置与技术伦理尺度失能

道德合法性是指公众对技术是否符合预期规范要求的判断，反映了人们对技术“是否做了正确的事”的事实判断（Suchman，1995）。面对现代技术出人意料的新发展态势，人们对待技术不自觉地会从起初的单纯乐观逐渐转向一种深刻的矛盾心态：既欢呼技术发展带来的无限可能，又因担忧“技术失控”而显得焦虑不安。这种心态上的转变暗示了通过技术追求目标的方法和手段可能超出了公众固有的道德规范和评判标准，并且技术的存在或受技术规训的行为可能不再符合人类的道德预期，而被认为是不恰当或不合理的。

受伦理“缰绳”拘束的数字技术发展取向，应当是“科技向善”的价值保障和伦理皈依。然而，技术化转型中的拒绝困境、技术鸿沟、数字异化等，显示了基层治理的技术逻辑开始超脱并凌驾于人的主体性和自主性之上，显著地区别于受外部自然的决定性所支配的“他治性”，自主性指独立、自治且不受外界法则和其他力量所支配。技术自主性作为一种复合性概念，潜隐着一种主客体（人与

技术)关系倒置——称技术是自主性的,等于说它是非他治性的,不为外部法则所支配。遵循有灵论的解释路径,对技术自主性来源的解释会受到“生命力守恒定律”的影响——人类将自身多少生命力注入装置(技术)中,其自身生命力就会损失多少。人类能量和特性的物理化迁移使得技术装置凝结了劳动价值,因此技术装置被有目的地创造和使用于人类社会各个领域,人类生活在技术创造物之中,并凭借它们生活。人类正在变得越来越像科幻小说中的“电子人”,需要越来越多的技术系统来维持日常生活,特别是随着可穿戴技术(人工心脏、人工肾等)的飞速发展,人体部分将直接与高度复杂的微型技术化器官紧密结合,最终出现人的“数字化生存”。因此,上述生命力的异化转移导致了人对技术装置的依赖,“人的独立性”至少以“物的依赖性”为基础。技术的“人工物”和人工的“技术物”解构了人作为独立个体的自主性,技术的深度嵌入催生了以人的生命力“换取”生活的便捷性,代价就是“人们彼此之间或者与世界之间不再有任何直接关系,每件事必须通过技术这一中介,而总要失去某些东西——人的自主性”(温纳,2014)。

这些问题常常被认为是“技术作恶”或缺乏温度的表现。这种技术的伦理困境一方面反映了人类对技术发展应包含伦理温度的期冀;另一方面则倒逼着我们从道德和情感维度审视数字技术。因此,技术主体身份认同和技术伦理道德困境成为技术治理合法性流失的重要内容。

(三)秩序合法性失范:“权力—权利”共融的秩序结构失衡

技术治理的秩序合法性来自技术的自我解释循环,即社会问题要被治理技术识别和呈现,必须进入技术预设的轨道。这种预设的“社会问题处置说明书”,一方面设定了问题提出的方式(只有特定问题能被技术识别);另一方面也将“答案”包含于其中(基于数据的预测、推测)。这种“轨道”的架设遮蔽了社会问题产生——社会问题的产生往往依附于权利表达与再生产,如群体性事件、非制度化上访等——的过程。因此,技术秩序确立了“社会提出的问题—技术化识别与处置—有效解决问题”的具体路径和治理模式后,对社会问题的技术化治理是否就达到了预设的理想状态呢?如果我们将这一思路继续沿着“问题识别—有效解决”的路径推演,就会发现其中暗藏了合法性危机,即技术刚性结构呈现出的一种强制性权力形态,难以控制承载了权利需求的社会问题的提出,

社会问题的权利逻辑滑出了技术权力预设轨道，并滑向技术化治理的“权力—权利”秩序失衡。

技术治理的“权力—权利”秩序失衡，反映了社会问题和结果导向原则，描画了一种横亘于基层治理中的结构性张力，即社会问题的阐释者和终结者在于国家（权力运作），但问题的制造者和提出者却是社会（权利主张）。社会问题反映了异质性主体的权益和利益诉求，因而不可能被先验性技术预设的“坐标系”所预判并被完全设计在正式规划中。所谓技术预判只是依据特定逻辑的推测。预测的计算基础需要通过技术将社会问题及其秩序之“序参量”的质的差异简化为可被数字化的量的差异，并使之成为“if…then…”循环式算法和标准化计算的过程。所以，“具象”的社会问题不可能被先验地设定为某个算法或计算流程，所谓的技术规划都是“坐标系”外在给予的化简社会问题的轨道，而“被技术设计或规划出来的社会秩序一定是简单的图解，它经常会忽视真实的和活生生的社会秩序的基本特征”（斯科特，2019）。因此，国家通过技术透镜看到的问题，只是自己权力的剪影——只看到自己想看到的，只解决自己能解决的问题，它并不知道或者说不关心异质性的权益问题如何被社会提出，以及这些问题的背后蕴含着哪些核心要素。

社会问题的产生无视了技术限定的呈现轨道，社会问题产生与技术化处置的二元张力中潜藏着一把悬置于治理秩序之上的“达摩克利斯之剑”——井喷式涌现的异质性社会问题不断冲击并试图脱离技术预设的轨道，制造出让技术预判“失能”的新问题，最终使技术秩序陷入自相矛盾的窘境。绩效、道德和秩序维度的合法性资源之流失，反映了技术及其治理效能的社会认同正面临潜在的挑战，这些矛盾会使公共行政与管理的公共性在二者的拉扯和“斗争”中逐渐流失。基于此，挖掘技术治理公共性资源的分布状况，为诠释和理解技术治理合法性危机何以生成，提供了一种具有参考价值的理论视角和分析基础。

三、公共性退守：技术治理合法性危机的生成机理

基层治理现代化及其主要途径——技术化、智能化、数字化等，既依赖于多样性和多元化，又对“一统体制”和“有效治理”目标统合抱有愿景。技术嵌入基

层治理过程的本质,在于将市场部分拥有的“私权”赋予“公权”的性质,造成了以科学知识为基础的专业性权力的“公共化”(刘金河,2022)。尽管合法性问题历来被视为政治制度设计和行政行为实施的基础性条件,但技术治理在合法性意义上的内涵却并没有随着行政技术发展变革而更加清晰。基于此,确证技术治理合法性基础、凝结完整的社会共识需要技术理性与价值理性、形式理性与实质理性、程序理性与结果理性共同担保,公共性不能退守、窄化为理性形态的任意一维。因此,从公共性视角来理解基层治理技术化转型,关系到技术治理及其合法性的理论诠释力和实践生命力。

(一)技术民主化失败与公共性参与不足

技术治理的合法性问题本质上检视了公众对新科学、新知识和新技术的接受程度,科学知识的公共性观念——科学发明、理论知识的发现和发明常被冠以“改善人类生活和推动人类社会发展”的名义展开——调和了科学技术的地方性与普遍性,强调实践和哲学相互融合的逻辑维度。自然哲学家让公共性的观念——在有限的范围,实际上是外行公众的成员——在展开和宣传现代科学的过程中扮演了重要角色(布朗,2015)。吉登斯认为,社会系统的延伸,是通过用于估算技术知识的测试的非人格性以及用来控制技术形式的公众批评来实现的,对专家系统的信任维持着脱域机制的有效运转(吉登斯,2016)。因此,技术治理的合法性认同预设了公众知识理性及其认知能力,即以公众为代表的“常民知识”能够融贯以科学家为代表的“专业知识”。然而,专业知识与常民知识的互动性不足,可能引发技术治理合法性危机。这种理念与当下公众参与科技决策的目的自相矛盾,因为公众是否拥有平等的参与权利是保证公众参与有效性的基础。也正是公众参与科技政策民主化过程、技术应用的可选择性和人的主体性能动性阙如,共同导致了技术治理的公共性出现“退守”现象。

科学家群体和学术研究群体是一个具有强烈排他性的、独特的开放共同体——吸纳更多的“圈内人”,将“外行”自然地排除在外。提出“专业知识与经验知识理论”(SEE)的柯林斯(H. Collins)和埃文斯(R. Evans)认为,技术治理(科学与技术)的政治合法性问题源于代表科学技术的专业知识与代表公众经验的常民知识之间存在冲突,民众被排除在既有政策议程之外(Collins & Evans,2002)。专业知识与常民知识、科研学术共同体与外行大众之间的睽违,

折射出一个不可回避的重要问题——在那些科学技术与政治交叉的领域,有关科技发明和使用的重大决策,是交给广泛民主来最大化实现技术的合法性,还是依据科技专家的专业意见做出合理决策?这一问题反映了专家知识与民主化进程是两种截然不同,甚至是对立的范畴,这可能导致技术决策无法同时满足“专业知识常民化”的合理性要求与“科技政策民主化”的合法性要求。因此,构建技术治理的合法性与公共性,常常陷入“要么确保决策的知识合理性而承担政治非法性风险,要么确保政治合法性而承担知识非理性风险”的两难困境。化解这种两难困境的“便捷策略”就是“退而求其次”——在二律背反中“选边站”,进而引发技术治理公共性的退守现象。这种行为策略在基层治理场域中更为普遍和常见,在压力型体制下,维持基层治理秩序稳定有序是“第一要务”,因而地方核心行动者常常忽视治理手段的合法性要求,采取一些“针对性强”“立竿见影”的治理手段,例如智能社区建设与技术适老问题等。这种公共性退守现象的生成机理是什么?回答这一问题,我们还需对基层核心行动者的行为逻辑进行剖析。

(二)技术失控与公共性多元价值冲抵

国家权力与技术治理的“分途”,很大程度上源于对技术治理的“失控”。这里所谓的“失控”并不是传统意义上的“技术难以被驾驭或不受控制”,而是指技术治理在实际运行过程中产生一种跳脱或偏离技术生产者的行动取向,即脱离或偏离了技术生产者、使用者的既定目标、特定任务和原有意图,特别是国家治理或基层治理的公共性旨趣。所以,基层治理事务和服务的公共性内涵与技术治理效能的“解扣”,释放出相互冲抵但又相互缠绕的多重理性形态,并共同构成了基层治理公共性退守的潜在路径及其最终形态(表2)。

表2　技术失控下公共性退守的潜在路径与最终形态

理性形态		公共性退守路径	退守后的实然形态
分裂维度	应然形态		
主体维	技术理性—价值理性	工具主义	技术理性
内容维	形式理性—实质理性	形式主义	形式理性
过程维	程序理性—结果理性	目的主义	结果理性

首先,技术失控下的公共性退守遵循“工具主义”技术观的基本逻辑。这种逻辑强调技术社会化应用或嵌入基层治理过程中,因治理主体结构(传统科层制与身处其中的行动者)对技术效用的能动性选择,基层治理的公共目标和社会效益简化为“工具性技术的使用”,进而产生了基层治理公共性抱守技术理性的“退守”现象,即由“技术(工具)和价值”的双重理性退守到以“工具理性”为核心的单一向度。当代技术创新与社会发展之间张力关系的一个典型描述是:技术的逻辑可塑性决定了其与社会相融合的过程充满了不确定性。然而,这种不确定性又导致技术的社会评估、预测和控制总是有限的、被动的。工具主义技术观抑或技治主义,认为技术只是实现治理目标的工具,治理主体的道德判断或行为策略决定了技术的使用场景和方式。因此,公共性价值约束仅在于“主体对技术的使用”,而更为核心的、前置性的“治理技术创新”过程因有其自身规律,则不需公共性价值的介入。所以,主体选择基层治理技术的标准及其行动逻辑,仅在于特定治理问题得到有效解决和遵循成本—收益分析的经济逻辑——技术好不好用、能不能解决问题、技术能为我的工作带来何种便利、我投入学习的时间与产出是否成正比,等等。

其次,技术治理公共性价值退守还会受到基层治理“形式主义”逻辑的影响,将基层治理技术创新置于“压力型体制”和“锦标赛”逻辑之下,并表现为一种“为创新而竞争”的形式化创新机制(何艳玲、李妮,2017)。这种基层治理创新“重形式而不重实质”,是基层政府在特定“策略性精英”带动下,通过既有经验学习、专家学者和媒体的“包装渲染”,以及用项目评比获得上级政府的认可和推广,来应对上级政府考核压力和同侪“创新锦标赛”竞争的一种策略性和权宜性手段(任克强,2021)。这直接导致技术治理的公共性窄化为保全“部门主义”的私利行为,造成技术围绕着上级考核“空转”和“悬浮”。技术“空转”不仅带来了高昂的行政成本,也为“象征性”执行、过度留痕等“技术形式主义”披上了合法性外衣。因此,从技术治理和基层治理创新的内容维度来看,形式主义逻辑下的技术治理合法性受制于公共性窄化,其基本治理价值被划归于片面追求“技术包装”与“技术创新”的“表面文章”,进而限制了技术治理公共性的实质回归。

最后,技术进场与治理任务情境的耦合决定了技术治理的合法性来源还需符合目的主义的建构路径。“目的主义建构体现了对技术权力功效性,尤其是

对技术之客观功效性的追求,它在相当程度上也是权力工具性要求的产物。”(樊华中,2019)以问题为导向的技术治理,注重治理问题的解决,而忽视了因技术漏洞而产生的次生风险,比如个人信息泄露等。所以,目的主义退守路径下的公共性偏重对技术治理效能的关注,忽视了技术进场的程序正当性,并表现为以下两个方面:一方面,技术进场过程的封闭性成为外源性技术进入政府治理过程的“黑箱”,进而导致社会评价过低而成为技术治理合法性和正当性的“洼地”;另一方面,基层治理技术化转型具有明显的挤出效应。这种程序正当性的缺失还表现为技术先验地排挤了其他有效治理资源。“技术泛在论”先验地将其他治理要素(如社会志愿要素等)排除在治理结构之外,过分强调“技术万能”本身就不具备社会认同。因此,过程维度的公共性损耗过于强调技术使用的合目的性,而对技术进场的论证和次生风险疏于防范。公共性向目的主义和结果理性的退守,“侵蚀了公共行政的程序性价值,结果和功能目标代替了组织目标成为最终价值,并引发‘科层表面主义’等系列问题”(李晓方、谷民崇,2022)。

综上,技术治理公共性的三条退守路线,阐明了基层治理场域中公共性多元价值之间的相互抵牾,以及由此形成的合法性资源流失后的技术理性基本样态,并从治理主体的行动逻辑维度描画了技术治理合法性危机的内在根源。这为理解“技术治理价值目标替代”和“治理技术逃离监管”现象提供了理论分析工具。

(三)逃离监管与公共性治理目标替代

上文从“主体—内容—过程”维度详细阐述了三种技术治理公共性退守路径及其最终价值。三重公共性退守叠加的一个明显结果就是,技术运转的治理效能对基层治理目标价值的叛逃。技术的嵌入实现了基层治理在任务场景与治理目标上的“降维打击”。

第一,逃离公共视野,大众参与的社会监控被替代为依赖技术自我监控的单向度监管。技术发展——工具的制造——使人和物被分置于两端并赋予人类以“制造工具的动物”与“使用工具的动物”这两种差异性角色(王磊,2021b)。然而,受制于人类的有限理性,技术工具的使用者对工具制造过程的科学知识知之甚少。加之兜底性的规范性制约与强制性规制永远滞后于技术发展及其负面后果,对技术漏洞的填补不得不依赖于更新技术基础设施(优化算法、升级服务器)。所以,作为使用工具的“动物”,人们一边享受技术带来的便利生活,一边

承担着技术带来的次生风险,并在一次次技术危机中学会应对技术失控的“技巧”。此外,技术治理的“黑箱式”运作使得技术嵌入和运转过程很难被公共视线所检视,技术权力的封装性无疑增加了基层治理技术创新的不确定性。基层治理权力的技术运作遮蔽了公众的视线,帮助技术治理逃离了公众的视野。因而,技术治理合法性危机只能依赖于科学知识的更新与技术迭代的速度和纠补技术漏洞的效度。

第二,逃离公共价值,基层治理的公共性被替代为“私共性”和“私性”的社会价值。基层治理以“满足人民美好生活需求”为价值依归,基层公共事务治理与公共服务供给都以社会性公共价值为逻辑起点。然而,技术治理内含的简化逻辑将基层治理任务情境简化为“社会问题识别—技术化应对”的“流水线”工单机制,以社会问题的“回应性”代替治理价值的“服务性”和“人民性”内涵。同时,以“事件处置率”代表政府部门治理效能考核标准,降低了基层治理主体的服务性指标,形式化应对成为局部利益(私共性)和部门利益(私性)的“合法性外衣”。因此,公共价值的“窄化”消解了技术应用的社会认同,甚至激化了基层社会矛盾,进而导致技术治理合法性资源流失,成为技术治理合法性危机的又一要因。

第三,逃离公共目标,基层治理的复合型秩序简化为“问题应对”的机械性技术秩序。基层治理秩序包含了环境秩序、交通秩序、安全秩序、服务秩序以及空间秩序等多种秩序形态(钱坤,2022),这种复合型秩序内在地统合了基层治理场域中的权力与权利、制度与情感、管理与服务、生活与生产等要素。因此,作为推进基层治理现代化建设的有效工具,技术治理的最终旨趣在于推进复合型秩序的共融和异质性要素的耦合,并将其桥接于“满足人民美好生活需求”的基本价值。但是,过于依赖技术治理对复杂场景的“数字化处理”会消磨基层场域的情感维度。这是因为机械化的技术装置悬置了人的情感需求,无差别的数字对待忽视了人的情感和伦理需求,进而灭失了基层治理的“温度”,将人困于冰冷的“技术牢笼”(郑磊,2021)之中。所以,机械性技术秩序虽有“立竿见影”的治理效能,但缺乏情感和伦理关怀,叛离了基层治理为人民服务的公共目标。

四、重建技术信任与公共性旨归：技术治理合法性再生产机制

产生技术治理合法性危机的原因，可被归结为技术公共性阙如、技术治理绩效认同衰弱以及对人之主体性和能动性的怀疑，技术治理合法性重建还需要回到基层治理实践的“绳结”中寻找可行之道。其原因在于，由“科学技术—治理体系—社会问题”塑造的更为复杂、动态和异质的基层治理情境，意味着我们不能将技术视为简单的工具或治理手段，必须直面其对社会关系建构的影响并对技术自身保持敏感，正视技术的价值负载和合法性基础成为应对技术风险所必备的“理论武器”。技术治理合法性亟须在公共性和社会价值基础上实现“理性”和技术认知的重构，为技术治理实践寻找更稳定的合法性资源。

（一）“规训技术”与“技术化归”：“以技制技”的因应之道

技术的问题用技术解决，这是基础的技术治理逻辑，因此“以技制技”是从技术之维来实现技术治理的程序化牵制，防止技术失控对社会的异化和吞噬。“规训技术”是指通过技术的自我更新和自我完善实现技术治理合法性再造。技术发展具有过程性特征，因此面对一些技术发展不成熟所导致的风险和困境时，我们可以通过技术优化来弥补技术风险。技术生产部门和政府部门对于社会问题性质的共同理解，需要预判性地了解甚至掌握科学技术潜在的经济、社会和政治风险，并将这种预判性认知融入技术研发和社会化应用过程，以此防治技术风险。基于此，“以技制技”的约束路径发展出了解和掌握科学技术知识的一般性规制策略，即技术治理规制的认知策略（王磊，2021b）。

认知性策略以社会整体效益的发展为逻辑前提，追求人技和谐的价值规范。该策略不同于强制性规范策略，后者以外在性强制约束为手段，以规范人的技术行为；而前者更强调行为主体基于技术知识的自我认知和内在约束，通过塑造“技术信任”“个体幸福”以及“可持续发展观念”三个层面的科技伦理规范，应对“科技缺乏约束给社会治理带来的挑战”。因此，这种策略具有基础性，是行为主体主动防范风险的自主行为。为此，消解技术社会化困境，实现“技术化归”，有赖于技术生产部门的自觉调整。以“技”为“据”的前提是提高技术发展

质量,增加现代技术的可视性、可监控性和可选择性。技术规制的认知策略要求保障人能拥有足够的主导性和主体性,能够拥有随时终止、叫停、更换技术应用与发展的“撤销权”和“选择权”。

(二)“技术向善”与“道德驯化”:织密技术治理的社会控制网络

技术失控肇始于社会规则与技术迭代的异步,因此,技术治理政治合法性的获取还需将技术风险进行道德化控制。

首先,应当树立正确的技术治理价值理念,将人的情感和伦理维度融入技术应用的全过程,彰显人本主义的伦理关怀。人本主义的价值先导是发展和应用技术的逻辑前提和最终目标,任何技术发展都应恪守“以人为中心,服务于人”的自然规律。因此,坚持技术向善的发展导向,需要纠正技术应用的“经济思维”和技术治理的工具主义偏向——反对将技术看作实现经济效益、功利目的的途径和手段——并以此成为技术治理的社会控制机制的核心内容。

其次,规范技术的研发和应用伦理。负责任的技术创新要求在技术研发和应用的全过程,相关企业和政府部门通过梳理安全责任意识、遵守风险告知诚信以及承担问题补救等多维要素,构建起技术创新责任体系,进而实现技术伦理风险规避和技术的社会规约。随着现代技术自主能动性的提高,责任问题才是最突出的伦理难题。但是,如何在界分主体责任的基础上规避技术风险以及由此产生的“兜弱”需要,才是实现技术治理公共性的重要内容。因而,与技术共存要解决的核心问题是,构建一套兼容技术要素和责任要素的社会规范体系,以便时刻考虑技术对人类社会的影响,以及人与技术的划界问题。只有处理好新兴科技与人、社会文化及环境的关系,处理好高效率存在方式与真正进步之间的关系,才能给人类和新兴科技的和谐发展提供一个良好环境,真正实现“善治”与“善智”的互构。

(三)重启技术民主化进程:构建技术治理的参与式协商机制

技术治理的政治控制机制,指对技术治理的国家干预和政治约束。国家通过政策控制和监管使技术治理重回公共价值轨道,即将技术纳入政治控制过程中,包含从技术嵌入、运作再到赋能治理的全过程。在基层治理技术化转型的过程中,治理主体对技术治理政治控制最关键的目标在于有效规制,促使国家重新

构建对技术治理在不同领域的使用力、选择力和控制力，进而重塑技术治理公共性价值且不偏离基层治理的管理目标和任务。基于此，在技术应用和规制的过程中，构建强有力的国家干预技术和社会参与机制有助于筑牢“技术治理向善”、涵养和汲取技术治理合法性的“制度红线”。

在《技术与迷思的轨迹》一书中，格兰特（D. Grant）和利里亚（M. Lyria）主张法律“不仅仅是更伟大的技术创新道路上的一个‘路障’”（Grant & Lyria，2017）。诚如此，伦理显然不是一个充满价值、用以组织或规避法律的框架。法律和伦理二者平行存在，并且能够产生对人类行为的积极影响。二者可以并且应该在技术的设计过程中相遇，而技术本身又可以使得特定形式的人类行为成为可能。于是，基于价值的设计、基于设计的价值（隐私等）、基于技术设计的法律保护、基于人权的通信基础设施和伦理设计的理念在相当程度上与其对技术发展的许多实际建议相一致。因此，以法律为基础的制度框架要求技术规制或实现“技术治理”，应当回到基层治理实践中去评估公民权利和技术伦理。

作为基层公共事务治理和公共服务供给的重要手段，技术治理应当包含社会权利和公民权益的价值追求与基本主张。所以，现代技术应用和治理运作的前提是，要符合社会利益的公平性和程序的正当性。基于此，重启技术应用的民主化议程成为攫取合法性资源的重要契机。哈贝马斯（1999）认为，科技与政治之间的长期交往关系及其转化过程建立在公共舆论的基础之上，以交流互动为核心的“生活世界”成为容纳技术和权力、科学和政治的全新场域。开启技术治理的民主化进程反映了技术的制度规制和技术权益的话语逻辑与社会抗争逻辑，技术合理性的双重面向（技术性和规范性）使与技术体系具有密切关系的行动者之间的交流成为可能——事实主张和技术规范性规制可以在辩论过程中得到协调（芬伯格，2018）。民主诉求与技术治理合法性在公共领域相遇，揭示了在“理性的交往空间”中技术治理的正当性与在规训技术中的精细表达之间的鸿沟。同时，这一要求表明，建立健全技术治理的合法性汲取机制，需要构建一种无差别的群体观，技术的利益相关者应当通过共享型利益关系凝结成“技术利益共同体”，进而保障技术治理公民权益的“他在性”得到最大限度的容忍和关怀。基于此，将技术治理带回政治学研究的传统领域——开启全过程民主协商，在构建参与式技术治理、规制技术作恶的过程中具有重要作用。

五、结语

技术应用场景及其效果的不确定性与“无知”或“不可知”的治理问题的凸显,在一定程度上消解了现代行政国家的理性基质与工具基础。在日益激烈的科技争论与社会冲突中(对人工智能、转基因技术、大数据技术的伦理、安全等议题的论争),国家治理主体尚未寻找到有效的应对措施而可能陷入治理困境。在此情况下,国家治理面临技术转向的前置性条件是,规避科学技术的不确定性带来的巨大挑战,将技术、权力等治理资源精准匹配基层治理问题,从而维持、再生产其合法性资源。

科学技术作为一种“专业知识”,天然地排除了“常民知识”的控制,一定程度上削弱了技术治理的社会认同。关于外行公众在复杂科技或治理议题上是否“无知”的问题,许多学者对此进行了批判性研究,并认为所谓群众的“无知”很大程度上是传统科学体制(学术共同体)专业化的封装性特征对普通公众的“常民知识”进行的贬低与排斥,这种“常民无涉”的技术庸俗主义很难将技术治理的合法性建基于“民众”之上。因此,我们必须辩证地看待和理解技术与政治之间持续不断的互动关系,尤其不能忽视制度情境、富有知识并具有反思精神的人类行动的中介作用(余敏江、王磊,2022)。对技术治理合法性论辩的学理化解释,可能还需回到以科学知识为基础的“技术”本体论和技术使用的“情景化解释”上来。

参考文献

比克,韦博、罗兰·保罗、鲁德·亨瑞克斯,2015,《科学权威的矛盾性:科学咨询在民主社会中的作用》,施云燕、朱晓军译,上海:上海交通大学出版社。

布朗,马克,2015,《民主政治中的科学:专业知识、制度与代表》,李正风、张寒、程志波等译,上海:上海交通大学出版社。

董石桃、董秀芳,2022,《技术执行的拼凑应对偏差——数字治理形式主义的发生逻辑分析》,《中国行政管理》第 6 期。

樊华中,2019,《检察公益诉讼的调查核实权研究——基于目的主义视角》,《中国政法大学学报》第 3 期。

芳汀,简,2010,《构建虚拟政府:信息技术与制度创新》,邵国松译,北京:中国人民大学出版社。

芬伯格,安德鲁,2018,《技术体系:理性的社会生活》,上海社会科学院科学技术哲学创新团队译,上海:上海社会科学院出版社。

哈贝马斯,尤尔根,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。

何艳玲、李妮,2017,《为创新而竞争—— 一种新的地方政府竞争机制》,《武汉大学学报》(哲学社会科学版)第 1 期。

吉登斯,安东尼,2016,《现代性的后果》,田禾译,南京:译林出版社。

李晓方、谷民崇,2022,《公共部门数字化转型中的“数字形式主义”——基于行动者的分析框架与类型分析》,《电子政务》第 5 期。

刘金河,2022,《权力流散——平台崛起与社会权力结构变迁》,《探索与争鸣》第 2 期。

刘永谋,2019,《技术治理、反治理与再治理——以智能治理为例》,《云南社会科学》第 2 期。

彭勃,2020,《技术治理的限度及其转型——治理现代化的视角》,《社会科学》第 5 期。

钱坤,2022,《城市基层治理秩序的复合化建构》,《兰州学刊》第 9 期。

渠敬东、周飞舟、应星,2009,《从总体支配到技术治理——基于中国 30 年改革经验的社会学分析》,《中国社会科学》第 6 期。

任克强,2021,《创制经验——城市基层政府社会治理创新的动因与优化》,《学习与实践》第 2 期。

斯科特,詹姆斯,2019,《国家的视角:那些试图改善人类状况的项目是如何失败的》,王晓毅译,北京:社科文献出版社。

王磊,2021a,《技术嵌入与系统容灾——提升组织应急管理能力的一种现代化机制》,《河北科技大学学报》(社会科学版)第 4 期。

王磊,2021b,《参差赋权——人工智能技术赋权的基本形态、潜在风险与应对策略》,《自然辩证法通讯》第 2 期。

王磊,2023,《分布式治理——技术嵌入基层治理的理论基础、权力逻辑与治理样态》,《电子政务》第 3 期。

王小芳、王磊,2019,《“技术利维坦”——人工智能嵌入社会治理的潜在风险与政府应对》,《电子政务》第 5 期。

温纳,兰登,2014,《自主性技术:作为政治思想主题的失控技术》,杨海燕译,北京:北京大学出版社。

颜昌武、杨郑媛,2020,《什么是技术治理?》,《广西师范大学学报》(哲学社会科学版)第 2 期。

余敏江,2010,《论管理行政的“科学”内核及其政治哲学基础》,《社会科学》第 11 期。

余敏江、王磊,2022,《平台型政府的“重层嵌套”组织结构及其制度逻辑》,《学术月刊》第 9 期。

郑磊,2021,《数字治理的效度、温度和尺度》,《治理研究》第 2 期。

Collins H. & R. Evans 2002, “The Third Wave of Science Studies: Studies of Expertise and Experience.” *Social Studies of Science* 32(2).

Grant D. & M. Lyria 2017, *Technology and the Trajectory of Myth*, Cheltenham: Edward Elgar Publishing.

Lishan L., Z. Qin & L. Yongmou 2021, “The Rule of Virtue: A Confucian Response to the Ethical Challenges of Technocracy.”*Science and Engineering Ethics* 27(5).

Suchman M. 1995, “Managing Legitimacy: Strategic and Institutional Approaches.” *Academy of Management Review* 20(3).

数据表征的法律治理*

朱　悦**

摘要：数据表征将原始数据转化为人工智能模型能够高效处理的信息，是连接容易治理的数据和难以治理的模型的桥梁。国内外法律正在探索通过治理数据表征治理人工智能：国内针对特定场景算法的专门立法明确纳入数据表征，但难以具体适用；国外个人信息保护执法对法律如何适用于其表征有具体说理，但尚未形成体系立法。综之，通过适当借鉴国外说理，国内既有法律可以清晰适用于数据表征，从而增强人工智能治理。结合法律、市场、伦理、架构等层面的治理进展，数据表征有望成为人工智能治理——包括未来相应立法——中的关键环节。

关键词：数据表征　人脸识别　推荐算法　人工智能治理　《个人信息保护法》

数据表征将原始数据转化为人工智能模型能够高效处理的信息，其发展也是近年来人工智能技术，包括深度学习技术迅速发展的重要因素（Bengio，Courville & Vincent，2013；LeCun，Bengio & Hinton，2015），数据表征因而构成连接原始数据和模型的桥梁。从人工智能治理特别是其法律治理的角度看，原始数据的治理相对容易。《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）等法律法规都已对原始数据有所规范。模型在技术上更加复杂，治理也相对困难，相关法律法规很少涉及人工智能模型，即使涉及通常也只有原则性的、不适用的规则。例如《个人信息保护法》第二十四条要求“利用个人信息进行自动化决策，应当保证决策的透明度和结果公平、公正”，但并未包括更具体的内容。法律治理由此开始关注可以桥接数据与模型的表征，体现为国内外都开始出现涵盖数据表征的立法和执法。具体来说，国内主要通过立法进行治理，例如《互联网信息服务算法推荐管理规定》（以下简称《管理规定》）第九条

* 本文受益于与周辉、许可和王睿的相关讨论，特此致谢。

** 朱悦，中国社会科学院文化法制研究中心。

和第十条针对属于数据表征的特征、标签和兴趣点设定义务;国外更多是在具体执法中针对数据表征展开分析,例如英国和澳大利亚数据当局在针对人脸识别服务提供商 Clearview AI 的调查中,对属于数据表征的人脸矢量展开分析(Australian Information Commissioner and Privacy Commissioner [OAIC],2021)。相应地,国内法律仍然缺乏对数据表征的具体适用,导致难以实际实施;国外也鲜有涵盖数据表征的立法体系,从而失之过度零碎。于是,尽管数据表征在人工智能的新近发展中非常重要,但其法律治理的重要性和潜力却尚未充分体现,国内外相关的法律治理探索也有待梳理、整合。

有鉴于此,本文第一节在介绍数据表征和实践的基础上,进一步阐明通过治理数据表征推进人工智能治理的重要性和潜力;第二、三节分别梳理国内外法律治理数据表征的尝试,国内主要是不同位阶的立法,国外更多的是不同法域的执法;第四节加以整合,以国内法律为出发点,阐释如何使现有法律清晰适用于数据表征;第五节首先分析推进数据表征的法律治理如何促进人工智能治理,然后探讨法律以外通过规范、市场和架构(Lessig,2006)等方式治理数据表征的进展如何促进人工智能治理,之后展望数据表征及其治理如何作为未来人工智能治理立法中的关键环节。

一、数据表征与人工智能治理面临的挑战

近年来,数据表征将原始数据转化为模型处理数据的方式经历了显著的变化,数据表征在显著提高处理效率的同时,其可理解性也在显著下降。这有三方面影响:一是聚焦数据表征的表征学习成为推动人工智能发展的重要方向;二是可理解性的下降与透明性、可解释性等人工智能治理问题密切相关;三是数据表征技术的变化引起了概念的变化,需要明确人工智能实践中许多常见的概念实际上都是表征的不同说法。数据表征还与作为人工智能前沿发展趋势的“大模型”[①](Bommasani,Hudson & Adeli et al.,2021)和开源模型社区存在密切关系,于是亦有必要简述这些趋势对法律治理可能产生的影响。

数据表征的显著变化指人工智能模型本身,而非人类专家主导了原始数据

① foundation model 在国内存在多种译法,除“大模型”外,还有“基础模型”“通用模型”等,本文采用比较常见的“大模型”译法。

的转化过程(Bengio, Courville & Vincent, 2013; 周志华, 2016)。简言之,反映学习对象本身属性或表现的原始数据,未必适合直接作为人工智能模型的输入。因此,在很长一段时间内,需要人类专家通过理解、清洗、转换、选择和探索性分析等方法,在原始数据的基础上找到最适合模型的表征,这一过程通常被称为"特征工程"①(厄兹代米尔、苏萨拉, 2019)。随着人工智能技术,特别是深度学习技术的发展,模型本身即可实现将"初始的、与输出目标之间联系不太密切的输入表示,转化成与输出目标联系的表示"②(周志华, 2016),从而替代了人类专家的角色。由此,自动化地学习最适合模型的数据表征这一需求催生了被称为"表征学习"的专门方向(Bengio, Courville & Vincent, 2013)。不仅是技术本身主导了数据表征的过程,模型学习数据表征的进步也推动了技术的发展。典例是从自然语言处理开始,并很快席卷其他场景的一系列人工智能模型,包括转换器模型、来自转换器的双向解码表征模型和稳健优化的来自转换器的双向解码模型③,等等(Vaswani, Shazeer & Parmar et al., 2017; Devlin, Chang & Lee et al., 2019; Liu, Ott & Goyal et al., 2019)。这些模型在更有效地表征语义信息的同时,相应实现了低计算成本、高可扩展和通用任务等良好的性能。正是这些模型——特别是 OpenAI 基于转换器模型实现的、能够高效合成各类文本的 GPT-2 和 GPT-3 模型(Radford, Jeffrey & Child et al., 2019; Brown, Mann & Ryder et al., 2020),使得算法治理的许多担忧愈发趋于现实。其中,OpenAI 及其合作方在发布 GPT-2 模型时已经预见了算法偏见、深度伪造、仇恨言论和有关的新型安全威胁等挑战(Solaiman, Brundage & Clark et al., 2019),但这些挑战至今也很难被认为已然完全解决。数据表征由模型实施这一变化,是人工智能技术发展的里程碑,也是人工智能治理挑战的发轫地。

数据表征的显著变化不仅通过促进技术发展形成一般性的治理挑战,其自身也直接产生了更加具体的治理挑战。在由人类专家实施特征工程的时代,这些专家的第一步便是理解原始数据(厄兹代米尔、苏萨拉, 2019)。之后通过转换、选择等方式生成的数据表征,同样建立在理解的基础上。于是,作为模型输入的数据及其表征始终没有脱离人——至少是人类专家——的理解范围,人也

① "表征"和"特征"是相近的概念,"表征工程"这一概念同样存在,但远不如"特征工程"常用。

② "表征"与"表示"是 representation 的不同译法。

③ 这三个模型更常见也更方便的表述分别是 Transformer、BERT 和 RoBERTa。

就始终处在人工智能的环内。随着模型接管学习数据表征的任务,数据表征不再想当然地具备对人而言的可理解性。实践中,这些表征总是一些看似没有意义的数字。对于发生显著变化以后的数据表征,其是否可能,又应如何对人解释、为人理解,成为亟待解决的前沿问题。然而,针对具体场景的解释起步未久,针对一般场景的解释道阻且长。仍以转换器模型为例,目前针对自然语言处理和其他少数场景的解释,局限于几乎只对人工智能专家有意义的、近似模型关键权重的可视化(Abnar & Zuidema,2020;Chefer,Gur & Wolf,2021),其他人几乎不可能循此形成对模型有意义的理解。在一般场景中也有从数据表征入手解释模型的尝试,但尚未超越纯粹理论层面(Guan,Wang & Zhang et al. ,2019)。① 既然数据表征难以解释、无法理解,那么人对原始数据和模型中间环节的理解便随之断裂,人也就在很大程度上被推到了环外。挑战随之而来。就人工智能治理中公认的价值而言(Jobin,Ienca & Vayena,2019),数据表征的可理解性缺失直接导致透明性、可解释性和负责任性等价值难以实现,也会造成隐私保护、安全性和公平性等价值的实现困难。

数据表征对这些价值造成的挑战,可以在对应的法律视野下进一步展开。当前而言,虽然存在专门针对人工智能的立法提案以及其他相关适用法律,但最为全面、系统、现实地体现这些价值的法律仍然是关于个人信息保护的法律。② 在国内主要是《个人信息保护法》,在国外主要是《通用数据保护条例》(GDPR)和其他类似法律。如果这些法律适用于数据表征,那么数据表征的处理者既需要在具体处理活动中遵从透明性等原则,也需要遵从一系列体现相应原则的义务,包括体现透明性的告知义务、体现可解释性的算法说明义务、体现负责任性的自证合规义务,以及体现隐私保护、安全性和公平性的其他义务,等等。③ 据此,在现行法律下,数据表征是否属于个人信息,以及相应地是否处在个人信息保护法律的适用范围内,是其法律治理的核心问题。或者说,如果特定的数据表征属于个人信息,则其处理者需要满足这些义务,特别是需要克服数据表征缺乏可解释性的问题,提供告知、说明并自证其对数据表征的处理符合规定等。既然数据表征及其所促进的人工智能技术发展带来亟须应对的治理挑战,个

① 简言之,相应尝试的思路是将模型分解成若干表征及其相互作用(Guan,Wang & Zhang et al. ,2019)。

② 国内外未来如果通过关于人工智能治理的专门立法,则有可能改变此处论断。不过,目前而言,除了仍处于立法进程、通过尚待时日的欧盟《人工智能法案》,此类专门立法还很难与个人信息保护法律比肩。

③ 此述义务均可在《个人信息保护法》、GDPR 与其他个人信息保护法律中找到。

人信息保护法律又是目前最为全面、系统、现实的工具，无怪乎国内外立法者、执法者都在探索通过治理数据表征来加强人工智能治理。与此同时，鉴于这些法律已经适用于原始数据，但法律仍然很少涉及技术复杂的模型，通过治理数据表征还可以弥补其间的断裂。国内外的这些探索及其教益，正是即将展开的主要内容。

在此之前有两项剩余的说明。一是在数据表征产生显著变化的过程中，出现了许多含义与其相近的术语。无论是人工智能技术视角下的讨论，还是法律视角下的讨论，包括国内外立法与执法，都未能就术语使用取得统一。为避免遗漏和混淆，需要列举这些术语。上文提到了“特征”“标签”“兴趣点”和“矢量”，本节提到了“表示”，这些都是含义相近的术语。[①] 国外执法中出现的相近术语还包括“模板”[②]和“嵌入”（OAIC，2021；Federal Trade Commission［FTC］，2021a）。下文除非特别说明，否则将不加区分地使用这些术语。二是数据表征与作为人工智能前沿趋势的大模型和开源模型社区存在密切联系。受到广泛认可的大模型为数据表征提供了标准化的实现。开源模型社区在通过开源进一步促进相应实现方案传播的同时，也在探索自律治理的制度安排，包括在授权时要求用户承诺负责任地使用表征和模型（Hugging Face，2022a；Hugging Face，2022b）。在将数据表征法律治理置于更加多元的治理方式中考察并展望其未来发展时，这是有用的事实。

二、国内数据表征相关立法及其适用

国内旨在治理数据表征的法律——主要是个人信息保护的法律——分散在不同位阶中。位阶最高的《个人信息保护法》和其他相关法律法规尚未明确涵盖数据表征，其解释与适用也缺乏对于数据表征的明确立场。位阶较低的《管理规定》明确涉及数据表征，但其具体适用仍处于探索进程中，除算法备案以外暂时缺乏实例。位阶更低的技术标准[③]，特别是针对人脸识别、步态识别等生物识别技术的技术标准亦明确涉及数据表征。技术标准的实践影响相对突出，但

① 这些术语彼此之间存在着一定程度的差异，特别是使用“特征”或“标签”这两个术语时，所指的常常是可理解性较强的表征。

② 这一术语通常仅用于生物识别，特别是人脸识别的场景。

③ 国内个人信息保护领域的技术标准本身通常并不具备强制力。理论上来说，其更多的是在成本考量之下“对事实的重构或建构”（陈伟，2022）。不过，即使技术标准本身不具备强制力，实践中其仍然可能通过处理者自愿遵守或受到其他强制性规范援引的方式产生强制力（许可，2019）。《个人信息安全规范》等少数特别重要的技术标准也确实因此具备相当的影响。

仍然受制于上位法律的不明确。

《个人信息保护法》中的“个人信息”尚未明确涵盖数据表征,能否解释为一般而言涵盖数据表征也不明确,只能针对具体场景做出推测。按其中第四条,个人信息指“与已识别或者可识别的自然人有关的各种信息”。权威法律释义还从肯定和否定两方面展开了这一定义:一是为了判断哪些信息属于个人信息,需要考虑“识别”和“关联”,无论是能够通过相应信息识别个体身份,还是相应信息与处理者已经识别其身份的个体有关,又或者是相应信息与处理者能够识别其身份的个体有关(杨合庆,2022);二是为了判断哪些信息不属于个人信息①,需要考虑其是否需要付出高额成本才能复原(杨合庆,2022)。对于人脸识别等少数场景,因其所生成表征的处理目的便是识别个体,实践中也具备精准的识别能力,似可推测其属于个人信息(赵精武,2020;赵精武,2022)。尽管如此,由于相应表征的识别能力需要结合特定的硬件和模型才能实现,处理者以外的个体或组织难以通过其实现识别,针对人脸表征的上述结论仍有挑战空间(赵精武,2020;赵精武,2022)。② 对于更加一般的场景中的数据表征,例如相当常见的、推荐算法中基于用户特征的嵌入,亦即神经网络模型中特定层的输出(Goodfellow,Bengio & Courville,2016),无论是基于“识别”“关联”判断,还是基于“复原成本”分析,都很难得到确定的结论。一方面,尽管通过嵌入一般无法实现识别,但可以认为其与相应用户存在一定关联。然而,自动化生成的嵌入通常不具备直观含义,其与用户的关联因此也很难定义和测度。另一方面,嵌入通常由模型深入加工产生,一般不容易由其复原作为原始数据的用户信息,但也可能存在发生导致识别或推断的攻击的可能性(Duddu,Boutet & Shejwalker,2020)。有鉴于此,由于不同展开下都难以得到确定的解释,法律层面是否涵盖数据表征难以称为明确,只在少数场景中存在确定性稍高的结论。

作为部门规章的《管理规定》明确涉及数据表征,但其具体适用有待观察。《管理规定》的相关内容主要有四类:首先是第九条要求“加强信息安全管理,建

① 或者说属于《个人信息保护法》第七十三条第四款所定义的已经匿名化的信息。

② 人脸信息亦可通过《中华人民共和国民法典》和《最高人民法院关于审理使用人脸识别技术处理个人信息相关民事案件适用法律若干问题的规定》的进路施行理论分析和实践保护。然而,其只是简单地包含“人脸信息”这一表述,而缺乏更细致的展开,因而同样无法免疫于与此处类似的批评(赵精武,2022)。

立健全用于识别违法和不良信息的特征库，完善入库标准、规则和程序”，此处涉及特征（库）[①]；其次是第十条要求“加强用户模型和用户标签管理，完善记入用户模型的兴趣点规则和用户标签管理规则，不得将违法和不良信息关键词记入用户兴趣点或者作为用户标签并据以推送信息”，此处涉及标签和兴趣点；再次是第十七条要求“向用户提供选择或者删除用于算法推荐服务的针对其个人特征的用户标签的功能”，此处同样涉及特征和标签；最后是第二十四条建立的算法备案制度在实践遵从中可能需要描述数据表征。无论着眼国内还是国外，《管理规定》都是深度结合法律与技术的突出实例。然而目前而言，至少有三点原因制约了《管理规定》实践适用于数据表征的效果。第一是如前所述，作为《管理规定》主要上位法依据的《个人信息保护法》对于各类数据表征的解释与适用暂不明确，第十条和第十七条的相关义务因而面临是否有据、范围如何的困境。第二是即使假定立法有据，其中引入“特征（库）”“兴趣点”“标签”等术语的创新依然可谓是“双刃剑”，既是对人工智能技术实践的准确描述，又是在国内法律体系中可谓是前所未见的概念。除非能够厘清这些概念确实属于个人信息或者属于其他上位法依据中的有关范畴，并且通过解释进一步纳入“矢量”“模板”“嵌入”等同属数据表征的概念[②]，否则第九、十、十七条的适用范围都很难称为完全明确，例如第十七条是否适用于推荐算法中基于用户模型的嵌入即有争议空间。第三是少数具体条款不够清晰。仍以第十七条为例，无论此处标签是否涵盖嵌入，选择标签的含义都有待进一步澄清。是只能选择特定标签不用于后续推荐服务，还是可以选择特定标签用于后续推荐服务，抑或可以选择更精细的处理方式，其在实践上都存在实质差异。因此，虽然《管理规定》创新性地涵盖了数据表征，但为使其真正发挥效能，依然存在显著的改进空间。[③]

---

① 《互联网信息服务深度合成管理规定》第十条同样明确涉及特征（库），但仍不易推测其具体如何充分实施。

② 《管理规定》未明确其第十条和第十七条调整对象与“用户画像”的关系，是又一可能造成困惑之处。

③ 算法备案制度的顺利推进可能在一定程度上缓解这些问题。不过，尽管来自算法备案的信息足以为立法者和执法者提供背景知识和问责点，从而改进立法和执法（张凌寒，2021；许可、刘畅，2022），但仅凭这些信息不易完全修复上位法依据和概念体系等层面的问题。实际而言，仅从数据表征的角度出发，当前公示的备案信息仍然存在“标签偏好”“向量表征”“特征”“兴趣特征表征向量”等彼此各异，亦未必与《管理规定》一致的术语（北京网易传媒有限公司，2022；北京快手科技有限公司，2022），为规定适用造成了额外的困难。由此，仅就数据表征的法律治理而言，算法备案难以缓解当前立法中存在的问题。

针对具体场景中的人工智能技术的技术标准同样明确涉及数据表征，但其实践效能受到场景和位阶的双重制约。以具备代表性的《信息安全技术 人脸识别数据安全要求》（以下简称《安全要求》）为例，相比其他涉及数据表征的法律和部门规章，《安全要求》有三方面的改善：首先是第 3 条第 2 项更加明确地定义“人脸特征”为“从人脸图像提取的反映自然人脸部特征的参数”。其次是第 8 条从隐私保护和安全等价值出发，对人脸特征信息本身提出了具备可更新、不可逆和不可链接特性的要求。由此，人脸特征泄露造成的负面影响相对可控，由人脸特征出发重建人脸图像，或由其出发识别、推断个体身份相对困难。最后是第 6 条和第 8 条从隐私保护和安全等价值出发，对人脸特征的处理也提出了具体要求，包括用于生成人脸特征的人脸图像原始数据满足最小必要原则，以及生成人脸特征后立即删除相应图像。其他生物识别领域的技术标准也包含类似的、较为明确的要求。比较之下，技术标准确实更加适宜于实践。但是技术标准面临的局限同样很明显：一方面，技术标准的适用范围通常较为狭窄；另一方面，由于技术标准施加的、针对数据表征的要求本身并不具备强制力，一般只是对《个人信息保护法》或其他相关法律法规有参照意义的细化，在这些法律法规仍不明确的前提下，技术标准的定义和规范缺乏足够牢固的基础，实践效能也因此受限。为了夯实《安全要求》的基础，需要通过说理阐明《个人信息保护法》确实涵盖人脸的表征；为了克服更多技术标准面临的场景和位阶局限，需要对更加一般的数据表征阐明这一点。

综上，国内对于数据表征的立法已颇为丰富且深入，但在具体问题上，特别是个人信息是否涵盖数据表征这一问题上需要进一步夯实分析的基础。[①] 他山之石，可攻此玉。

① 国内各地法院已经审理了一些相对复杂的、涉及个人信息定性问题的案件，但这些案件通常没有脱离前述“识别”和“关联”的分析进路。以较为典型的、在《中华人民共和国民法典》相近条款下审理的“黄某与腾讯科技隐私权、个人信息权益网络侵权责任纠纷”一案（〔2019〕京 0491 民初 16142 号，通常称“微信读书案”）为例，除了基于“识别”和“关联”的分析以外，法院还指出“识别”的分析需要考虑具体场景。这样的裁判思路对于分析数据表征是有用的起点，但更进一步的结论通常需要较此更为精细的分析，包括在数据表征相关问题中细化场景这一概念。

三、国外数据表征相关执法及其说理

国外对于数据表征的法律治理主要寓于具体执法中。这些执法可以大致分为两类:一类对于数据表征是否属于、为何属于个人信息有较为详细的说理,一旦阐明这一结论,个人信息保护的法律治理循之适用于数据表征和模型;另一类则以美国 FTC 的“算法罚没”执法(Slaughter, Kopec & Batal, 2020; Li, Forthcoming)[①]为代表,虽然没有太多讨论数据表征的法律性质,但在执法中将其作为个人信息保护执法救济的一部分。

GDPR 的个人信息[②]定义及其展开,为域外涉及数据表征执法的说理提供了肯定和否定的两方面基础。首先是判断何者属于个人信息的肯定性基础。按照 GDPR 第 4 条第 1 款和“鉴于”部分第 26 条,个人信息可以定义为“与已识别或可识别自然人有关的信息”,此与《个人信息保护法》完全一致。在诺瓦克诉爱尔兰数据保护专员(Data Protection Commission[DPC],2021)一案中,欧盟法院细化了个人信息的判断依据:不仅需要抽象地考虑“识别”和“关联”,还要从“内容、目的或影响”的角度具体地分析“关联”(Court of Justice of the European Union,2017)。如此展开“识别”和“关联”的思路足以在很多执法案例中阐明数据表征为何属于个人信息。以人脸识别中的人脸矢量为例。[③] 英国和澳大利亚数据当局发现,从内容出发,矢量是人脸图像信息的“数学表征”,从目的出发,矢量的目的即进行生物识别,综合这两点足以认定人脸矢量属于个人信息(OAIC,2021)。[④] 又以广告推荐算法中表征个体隐私偏好的“透明和隐私串”(transparency & control string,以下简称 TCF)为例。比利时数据当局发现,尽管难以确定 TCF 能否识别个体,但因为处理 TCF 的目的即根据不同个体的隐

① 斯劳特尔(R. Slaughter)等人在其首先阐明“算法罚没”概念一文中的署名单位都是 FTC。

② GDPR 的表述是“个人数据”,但就本文论述目的而言没有必要区分“个人信息”与“个人数据”。

③ 对于生物识别中的数据表征(点表征)而言,存在从“识别”进路出发,仅根据其识别准确率高,且其识别能力随时间推移保持稳定即认定相应表征属于个人信息的案例(Agencia Española de Protección de Datos,2021)。不过,这一判断受到场景的局限,不易适用于生物识别以外的场景。此外还存在处理者将数据表征(模板)与其他个体标识符关联的情况(Nemzeti Adatvédelmi és Információszabadság Hatóság, 2022)。这是容易判断的情况,但同样不具备一般性。

④ 本案主要适用澳大利亚有关法律。其中的主要概念,特别是“个人信息”的概念定义和之前讨论的个人信息保护法律都是相近的。

私偏好对其区分对待——或者说处理 TCF 的目的就蕴含了对个体的区分与识别①,故而应当推定 TCF 属于个人信息(Autorité de Protection des Données [APD],2021)。综之,无论是人脸识别还是推荐算法,此述展开都可以提供相对明确、可以实践的说理。对于更加一般的数据表征,分析其内容、目的与影响亦有意义。比起抽象的"关联",无论是数据表征所表征的信息内容,还是其及相应模型服务的处理目的,抑或其及相应模型处理对个体的影响,都可以更加具体分析。不妨以推荐算法的模型嵌入为例,虽然问题不能一劳永逸地解决,但现在可以通过其加工产生的信息、其处理目的及其处理影响给出更明确的结论。

GDPR 还为判断何者不属于个人信息的说理提供了基础。简言之,为判断特定信息不属于个人信息,需要排除"区分""关联"与"推断"三类风险,亦即需要确定处理者无法通过相应信息将特定个体从群体中区分出来,无法在属于特定个体的不同信息间建立关联,也无法推断特定个体的有关信息(Article 29 Working Party,2014)。这一思路已经针对少数非常复杂的数据表征体现出了"威力",例如通过有损哈希算法表征个体的电话号码。虽然直接分析"识别"或"关联"都不容易,但是由于算法表征以后个体依然可以区分,以及算法表征以后依然可以推断个体的社交关系网络,爱尔兰和欧盟数据当局依然认定其属于个人信息(DPC,2021)。② 对于经过不止一种算法处理的信息以及基于新兴合成技术生成的信息,通过三类风险同样可以给出细致的分析(Autoriteit Persoonsgegevens,2021;Stadler, Oprisanu & Troncoso,2022)。不过,尽管这一思路可以用于分析非常复杂的信息,但其本身对法律和技术的理解有较高的要求,实践应用相对肯定性的思路更为困难。综之,这一思路适宜与"内容、目的或影响"相互补强,在大部分涉及数据表征的案件中辅助分析,在剩余特别复杂的案件中发挥更重要的作用。

美国 FTC 在新近个人信息保护执法中引入的"算法罚没"措施同样涵盖了数据表征。这一概念本身不涉及数据表征是否属于个人信息的判断,但其使得个人信息保护的执法可以涵盖数据表征。所谓"算法罚没"指的是如果作为算

① 区分(对待)属于识别的一种(APD,2021;Purtova,2022)。

② 此处表征系用于联系人匹配,亦即出于区分和识别而处理。这可以和基于三类风险的分析相互补强。

法输入的原始数据违反了个人信息保护的规定，侵犯了个人对其信息的权益，那么后续救济不仅需要删除原始数据，还需要删除由算法衍生的表征和模型（Slaughter，Kopec & Batal，2020）。这一救济措施的初衷是有效阻遏使用算法侵害个人信息权益的行为（Slaughter，Kopec & Batal，2020），客观上也使得数据表征处在个人信息保护法律的射程之内。据此，FTC 在针对相册管理应用滥用人脸图像的案件中要求应用开发者删除所有衍生自图像的人脸嵌入，包括“全部或部分地由图像中人脸提取的所有数值向量等数据”（FTC，2021b）。在回复有关这一案件的质疑时，FTC 进一步明确人脸嵌入包括“概率性”——全部或部分地由图像中人脸提取的“概率分布估计”（FTC，2021c）。因此，无论数据表征是否属于个人信息，出于在执法中提供充分救济的需要，FTC 都可以灵活地设计涉及数据表征的救济措施。[①] 相比 GDPR，这是更加具有实用导向的思路。

以上，可见域外执法针对数据表征的法律治理引入了三方面的不同思路。相比国内，深度结合个人信息法律与数据表征技术的立法在国外整体而言要更加少见；不过，通过充分地展开个人信息保护法律的“个人信息”概念，以及采取其他实用导向的救济措施，国外相比国内积累了更多涉及数据表征的执法案例。于是结合二者成为自然的思路。

四、令国内法律清晰适用于数据表征

“自上而下”地结合国内立法与国外执法经验，可以使得国内法律清晰适用于数据表征。具体来说，首先从《个人信息保护法》的解释与适用入手，通过既有司法案例和相关技术标准，可以自然地引入“内容、目的或影响”和“区分”“关联”与“推断”这两类分析思路。如前所述，由此足以对人脸识别、推荐算法和其他数据表征得出法律适用的明确结论。然后，法律层面的明确适用可以向下传导至《管理规定》等部门规章和《安全要求》等技术标准。既可夯实其基础，亦可解决前面提及的其他问题。

与《个人信息保护法》中“个人信息”有关的典型司法案例引入了“场景

① 前提当然是 FTC 本身有相当宽泛的、设计与违法行为合理相关的救济措施的权力（Slaughter，Kopec & Batal，2020）。

说”,此处的“场景”可以涵盖“内容、目的或影响”。以国内影响较大的“微信读书案”为例,其中着重考虑了相应信息是否“在特定场景下”可以还原到个体的身份信息。又以影响同样较大的“余某与北京酷车易美网络科技公司隐私权、个人信息保护纠纷案”(〔2021〕粤0192民初928号,通常称“查博士案”)为例,其中同样提及“在车辆交易场景下”对涉案信息是否属于个人信息展开分析。两案均未进一步说明何谓“场景”,但其间的场景分析实际已涵盖了“内容、目的或影响”。“微信读书案”实际上已经考虑了内容和目的:一方面,在分析读书信息是否属于个人信息时,法院考虑了读书信息“包括读书时长、最近阅读、书架、推荐书籍、读书想法等,能够反映阅读习惯、偏好等”的事实;另一方面,在分析腾讯内部生成的OPEN_ID是否属于个人信息时,法院也指出了OPEN_ID本就是用于“识别用户的识别码”。“查博士案”中法院也考虑了“内容、目的和影响”。在分析历史车况信息是否属于个人信息时,综合考虑了相应信息的内容是否涉及具体个人,以及相应信息是否用于评价具体个人的行为和状态。由此,通过场景分析引入“内容、目的和影响”不仅不会与国内司法实践冲突,还可以使当前仍然较为模糊的场景分析体系化和具体化。对于人工智能和数据表征治理而言,如此不仅能够进一步明确《个人信息保护法》中的“个人信息”概念以及这一法律的适用范围,还可以对于一般的数据表征得出类似于“推定”的结论。简言之,如果数据表征系由个人信息加工或提取,并且相应的模型是用于识别个人的目的,也会对于个人权益造成影响,则一般应当认为相应数据表征属于个人信息。对于人脸识别或者一般意义的生物识别,通过涵盖“内容、目的和影响”的场景分析足以认定其一般属于个人信息;对于推荐算法嵌入等更加复杂的数据表征,结合这一分析与算法备案中公示的“算法运行机制”和“算法目的意图”(北京网易传媒有限公司,2022;北京快手科技有限公司,2022),一般也足以给出更具体的分析和更明确的结论。

还可以通过援引技术标准细化《个人信息保护法》中关于“个人信息”的解释,从而引入“区分”“关联”与“推断”的分析思路。具言之,《信息安全技术 个人信息安全影响评估指南》(以下简称《评估指南》)附录A.4已经引入了基于这三类风险的判断思路。理论上来说,法院可以通过援引《评估指南》作为判断个人信息范围的说理依据(许可,2019);实践上来说,在“微信读书案”等案件

中,法院也已援引《信息安全技术 个人信息安全规范》作为判决依据之一。[①] 于是,法院通过援引《评估指南》引入“区分”“关联”与“推断”思路,从而在涉及数据标准的案件,特别是涉及复杂数据表征的案件中给出更加具体的分析,理论上可行,实践中亦可期待。如前所述,这也可以与体系化、具体化以后的场景分析相互补充。特别是对于认为基于特定个人信息提取的数据表征不属于个人信息的主张,可以通过全面考虑三类识别风险予以审慎立场的分析。

明确个人信息一般属于个人信息且为何一般属于个人信息以后,《管理条例》和技术标准所面临的问题都可以在三个层面上得到相当程度的纾解。首先,无论是《管理条例》和技术标准中出现的“特征(库)”“兴趣点”“标签”等术语,还是法律和技术中可能出现的“矢量”“模板”“嵌入”等更多术语,都可以通过上述分析认定其一般属于个人信息,而不至于导致困惑。当然,在不同位阶立法中统一描述数据表征的术语,同样有促进明确的效果。其次,明确这些术语都属于个人信息以后,《管理条例》、技术标准和其他相关文件所设定的相关规范都具有了充分的依据。在夯实基础的同时,也为其适用范围提供了更多的确定性——以《个人信息保护法》作为上位法依据的规范的适用对象是属于个人信息的数据表征,也主要限于这些数据标准。最后,夯实依据以后,法律法规和技术标准的起草者可以更加放心地设定更多数据表征的有关规范,法院也可以更加放心地援引这些规范作为裁判或说理依据。《个人信息保护法》和其他相关法律法规体现的透明性、可解释性、负责任性、隐私保护、安全性和公平性等人工智能治理价值,因而可能得到更充分地展开和贯彻。此外,部门规章和技术标准也可以“反哺”《个人信息保护法》的解释和适用。如前,算法备案制度要求公示的信息为体系化、具体化以后的场景分析提供了有用的事实;部门规章和技术标准在坚实基础上的不断细化和完善,为解决具体争议提供了更多的裁判和说理依据。

综上,国内既有立法和国外新近执法可以自然地结合,从而为人脸识别、推荐算法以及更加一般的人工智能技术所涉及的数据表征法律治理提出相对完善的解决方案。个人信息保护体现的人工智能治理价值因而适用于数据表征,这

① 王新锐和罗为(2020)也从律师的角度指出:“在我们接触的与个人信息相关的诉讼案件中,原被告双方通常都会就是否违反《规范》进行辩论。”《规范》即《信息安全技术 个人信息安全规范》。

是仅针对现行法律的结论。通过讨论更加多元的、治理数据表征的方式,可以进一步丰富解决方案,并且增强数据表征治理作为人工智能治理“桥梁”的作用。

五、通过数据表征治理促进人工智能治理

更加多元的治理方式不仅包括法律,还包括规范、市场和架构,分别对应于社群自律、市场竞争和技术治理(Lessig,2006)。对数据表征而言,开源模型社区和大模型共同构成人工智能技术发展的前沿趋势,同时也体现了规范、市场和架构治理的重合。未来更加丰富的数据表征法律治理,需要包括与开源模型社区的充分协同。在专门立法中体现这一协同不仅有助于促进数据表征的法律治理,还有助于促进人工智能技术的法律治理。

大模型,包括前文提及的 GPT-2 和 GPT-3 等模型,既与数据表征学习的进步有关,本身又是当前人工智能治理的核心问题(Bommasani,Hudson & Adeli et al. ,2021)。由于大模型在通用任务上体现出很强的性能,具备实用价值,同时其对于计算资源和工程能力有较高的要求,而个体和大部分组织难以具备相应条件,因此以 Hugging Face 为代表的开源模型社区宣称致力于“为每个人推动和民主化”这些大模型,特别是通过开源表征预训练模型、压缩后模型和大模型接口的方式,使得个体和大部分组织可以自主学习、训练和应用大模型(Hugging Face,2022c)。开源模型社区同时通过社区自律的方式治理表征和模型的使用。例如在应用大模型之前承诺将保护个人信息,避免通过自动化决策损害他人的法律权益,避免散布虚假的或未标注其系合成的合成信息,避免损害弱势群体的权益,等等(Hugging Face,2022b)。又如在开源大模型时配备数据卡和模型卡,其中数据卡着重介绍原始数据的来源及其合法合规情况,也涵盖包括提取数据表征在内的后续处理步骤(Gebru,Morgenstein & Vecchione et al. ,2021;Hugging Face,2022d);模型卡着重介绍模型的训练、应用和评估方式,也涵盖训练过程涉及的表征学习步骤(Mitchell,Simone & Zaldivar et al. ,2019;Hugging Face,2022e)。一定程度上,甚至可以认为开源模型社区的数据卡和模型卡构成了技术色彩更鲜明,但内容也更加全面的自律算法备案。此外,开源模型社区的发展既是大模型领域市场竞争结果的体现,又使得治理风险和治理责任向有能力开

发、训练和开源维护大模型的少数组织集中。这些都是未来法律治理所面临的趋势。

站在超越现行法律、聚焦新兴领域立法的角度，综合所有前文，相应可以展望未来人工智能治理立法如何将数据表征作为治理关键环节的五个要点。第一点是进一步明确提取自个人信息的数据表征一般属于个人信息，同时确认体现人工智能治理价值的各类个人信息保护责任和义务均适用于相应数据表征。例如《个人信息保护法》第二十四条对利用个人信息决策时“保证决策的透明度和结果公平、公正”的要求，即可以数据表征作为连接点进一步适用于人工智能技术。第二点是在算法备案制度实践的基础上，吸收数据卡和模型卡的经验，进一步加强关于人工智能技术的信息披露。特别是对于数据表征而言，需要描述其提取、加工和学习的表征类型以及相应机制，并且简要描述表征的内容。[①] 这一点可以作为备案中“算法运行机制”的一部分，和备案所包括的“算法目的意图”，以及“（尤其是对个人权益影响的）算法自评估报告”一同支持执法和司法中的法律适用。第三点是认可和鼓励基于社区自律的规范治理：一方面，在立法中明确有关数据表征和相关模型开发、训练与应用的自行承诺对承诺者具有约束力，可以通过行政执法和公益诉讼等方式执行；另一方面，通过尽职免责、认可作为社会责任、认可作为责任承担的减轻因素等方式加强自律的正面激励。第四点是将“算法罚没”，包括删除数据表征和相关模型，作为责任承担方式之一。如前所述，从实用导向的角度看，这一责任具有可观的阻遏效果。第五点是考虑到以大模型为代表的人工智能技术权力和治理风险愈发向以“守门人”为主的少数组织集中，一方面是需要在人工智能治理立法中引入关于守门人主体的增强规定，另一方面是如果通过针对守门人的专门立法，例如平台法中也应当包括人工智能治理的有关内容。在更加清晰适用现行法律的基础上，上述五点和规范、市场与技术共同构成了数据表征和人工智能治理更完善的方案。

① 这一描述可以是相对技术性的，例如特定的数据表征提取了相隔较远的字词间的语义联系。

六、结语

以上首先概述数据表征的发展为何对人工智能技术进步有重要作用,以及人工智能治理如何尝试将数据表征作为原始数据和下游模型间的桥梁。然后,一方面,评述了国内正在日趋完善的、涉及数据表征的立法;另一方面,评述了国外日益丰富的、令法律适用于数据表征的执法实践。之后结合二者,“引石攻玉”,自然将域外实践发展的有用思路引入国内法律,使得国内法律可能清晰适用于数据表征,从而贯彻人工智能治理追求的诸多价值。最后,对于立法的展望虽然只是在技术前沿下的初步推测,但我们也坚信其将很快成为现实。

参考文献

北京快手科技有限公司,2022,《网信算备 11010841376070222001 7 号》,https://beian.cac.gov.cn/api/static/fileUpload/principalOrithm/additional/user_bb4fec44-be41-43d7-a11e-6fcd8d7776ad_9ea7387d-c989-4625-9bb3-2e23b29a7a0c.pdf。

北京网易传媒有限公司,2022,《网信算备 11010111352243022200 29 号》,https://beian.cac.gov.cn/api/static/fileUpload/principalOrithm/additional/user_b6c4caf9-5a05-4387-bdbb-850ddb4d61cc_4f8a9fc8-6995-48da-8bee-c07cd97f8619.pdf。

陈伟,2022,《作为规范的技术标准及其与法律的关系》,《法学研究》第 5 期。

厄兹代米尔,锡南、迪夫娅·苏萨拉,2019,《特征工程:入门与实践》,庄嘉盛译,北京:人民邮电出版社。

王新锐、罗为,2020,《〈个人信息安全规范〉被援引情况实证分析》,https://www.leb-china.com/upload/upload/608dfa165de3833c47afe6ad4a194bcd.pdf。

许可,2019,《〈个人信息安全规范〉的效力与功能》,《中国信息安全》第 3 期。

许可、刘畅,2022,《理解算法备案》,《人工智能》第 1 期。

杨合庆,2022,《中华人民共和国个人信息保护法释义》,北京:法律出版社。

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。

赵精武,2020,《〈民法典〉视野下人脸识别信息的权益归属与保护路径》,《北京航空航天大学学报》(社会科学版)第 5 期。

赵精武,2022,《刷脸问题治理逻辑的一般规则转向——以技术透明度为基本立场》,《北方法学》第 1 期。

周志华,2016,《机器学习》,北京:清华大学出版社。

Abnar, S. & W. Zuidema 2020, "Quantifying Attention flow in Transformers." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.

Agencia Española de Protección de Datos 2021, "Expediente N°: PS/00050/2021." https://www.aepd.es/es/documento/ps-00050-2021.pdf.

Article 29 Working Party 2014, "Opinion 05/2014 on Anonymisation Techniques." https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

Australian Information Commissioner and Privacy Commissioner 2021, "Commissioner Initiated Investigation of Clearview AI, Inc. (Privacy) [2021] AICmr 54 (14 October 2021)." https://www.oaic.gov.au/__data/assets/pdf_file/0016/11284/Commissioner-initiated-investigation-into-Clearview-AI,-Inc.-Privacy-2021-AICmr-54-14-October-2021.pdf.

Autorité de Protection des Données 2021, "Decision on the Merits 21/2022 of 2 February 2022." https://www.autoriteprotectiondonnees.be/publications/decision-quant-au-fond-n-21-2022-english.pdf.

Autoriteit Persoonsgegevens 2021, "Gemeente Enschede." https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/boetebesluit_ap_gemeente_enschede.pdf.

Bengio, Y., A. Courville & P. Vincent 2013, "Representation Learning: A Review and New Perspectives." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(8).

Bommasani, R., D. Hudson & E. Adeli et al. 2021, "On the Opportunities and Risks of Foundation Models." arXiv preprint arXiv: 2108.07258.

Brown, T., B. Mann & N. Ryder et al. 2020, "Language Models Are Few-Shot Learners." *Advances in Neural Information Processing Systems* 33.

Chefer, H., S. Gur & L. Wolf 2021, "Transformer Interpretability Beyond Attention Visualization." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Court of Justice of the European Union 2017, "Peter Nowak v Data Protection Commissioner." https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62016CJ0434.

Data Protection Commission 2021, "In the Matter of WhatsApp Ireland Limited." https://www.dataprotection.ie/sites/default/files/uploads/2022-03/Full_decision_WhatsApp_Ireland-August_2021.pdf.

Devlin, J., M. Chang & K. Lee et al. 2019, "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." Proceedings of the 2019 Conference of the North American

Chapter of the Association for Computational Linguistics: Human Language Technologies.

Duddu, V., A. Boutet & V. Shejwalkar 2020, "Quantifying Privacy Leakage in Graph Embedding." MobiQuitous 2020 – 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services.

Federal Trade Commission 2021a. "California Company Settles FTC Allegations It Deceived Consumers about Use of Facial Recognition in Photo Storage App." https://www.ftc.gov/news-events/news/press-releases/2021/01/california-company-settles-ftc-allegations-it-deceived-consumers-about-use-facial-recognition-photo.

Federal Trade Commission 2021b, "Decision and Order." https://www.ftc.gov/system/files/documents/cases/1923172_-_everalbum_decision_final.pdf.

Federal Trade Commission 2021c, "Letter to Commenter World Privacy Forum." https://www.ftc.gov/system/files/documents/cases/wpf_response_final_0.pdf.

Gebru, T., J. Morgenstein & B. Vecchione et al. 2021, "Datasheets for Datasets." *Communications of the ACM* 64(12).

Goodfellow, I., Y. Bengio & A. Courville 2016, *Deep Learning*, Cambridge: MIT Press.

Guan C., X. Wang & Q. Zhang et al. 2019, "Towards A Deep and Unified Understanding of Deep Neural Models in NLP." International Conference on Machine Learning.

Hugging Face 2022a, "Getting Started With Embeddings." https://huggingface.co/blog/getting-started-with-embeddings.

Hugging Face 2022b, "Big Science RAIL License v1.0." https://huggingface.co/spaces/bigscience/license.

Hugging Face 2022c, "The AI Community Building the Future." https://huggingface.co.

Hugging Face 2022d, "Create A Dataset Card." https://huggingface.co/docs/datasets/dataset_card.

Hugging Face 2022e, "Building A Model Card." https://huggingface.co/course/chapter4/4?fw=pt.

Jobin, A., M. Ienca & E. Vayena 2019, "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1(9).

LeCun, Y., Y. Bengio & G. Hinton 2015, "Deep Learning." *Nature* 521(7553).

Lessig, L. 2006, *Code Version 2.0.* New York: Basic Books.

Li, T. Forthcoming, "Algorithmic Destruction." *Southern Methodist University Law Review*.

Liu, Y., M. Ott & N. Goyal et al. 2019, "RoBERTa: A Robustly Optimized BERT Pretraining Approach." arXiv preprint arXiv: 1907.11692.

Mitchell, M., W. Simone & A. Zaldivar et al. 2019, "Model Cards for Model Reporting." Proceedings of the Conference on Fairness, Accountability, and Transparency.

Nemzeti Adatvédelmi és Információszabadság Hatóság 2022, "NAIH-85-3/2022." https://naih.hu//hatarozatok-vegzesek? download = 517: mesterseges-intelligencia-alkalmazasanak-adatvedelmi-kerdesei.

Purtova, N. 2022, "From Knowing by Name to Targeting: The Meaning of Identification under the GDPR." *International Data Privacy Law* 12(3).

Radford, A., W. Jeffrey & R. Child et al. 2019, "Language Models Are Unsupervised Multitask Learners." *OpenAI Blog* 1(8).

Slaughter, R., J. Kopec & M. Batal 2020, "Algorithms and Economic Justice: A Taxonomy of Harms and A Path Forward for the Federal Trade Commission." *Yale Journal of Law & Technology* 23(1).

Solaiman, I., M. Brundage & J. Clark et al. 2019, "Release Strategies and the Social Impacts of Language Models." arXiv preprint arXiv: 1908.09203.

Stadler, T., B. Oprisanu & C. Troncoso 2022, "Synthetic Data-Anonymisation Groundhog Day." 31st USENIX Security Symposium (USENIX Security 22).

Vaswani, A., N. Shazeer & N. Parmar et al. 2017, "Attention Is All You Need." *Advances in Neural Information Processing Systems* 30.

论元宇宙数据安全风险的法律对策

唐欣奥*

摘要:当前元宇宙技术蓬勃发展,为数据安全带来了新的挑战。元宇宙数据收集记录与当前数据收集记录相比,其范围突破了有限性,扩展到了用户自身及其行为习惯的各个方面,具有全方位性,因此具有更大的数据安全风险,亟待研究者以前瞻性眼光进行预测和防范。本文在探讨元宇宙数据、元宇宙数据安全概念的基础上,以元宇宙用户、元宇宙平台、国家三类主体为基础,探讨了不同层面的元宇宙数据安全风险,分析元宇宙去中心化管理办法在数据安全风险预防和应对上的不足与弊端,提出应以多元共治作为基本原则预防和应对元宇宙数据安全风险;分析我国当前元宇宙数据安全相关制度状况与缺陷,提出我国在多元共治原则下预防和应对元宇宙数据安全风险的具体建议。

关键词:元宇宙　数据安全风险　法律对策　多元共治

元宇宙得以产生和运转的枢纽之一就是海量的数据。数据不仅蕴含着丰富的商业价值,亦作为国家的基础性战略资源,与国家安全及利益息息相关(周若涵,2021)。大数据技术是元宇宙的技术基础之一,元宇宙的发展将推动大数据行业的更新和发展,但也会给数据安全带来新的风险和挑战。元宇宙的兴起,将不同程度地引发有关数据安全以及隐私保护的困境(高一乘、杨东,2022);并且在一定程度上,这一概念的火热与流行确实为进入信息化新阶段的全球各主要国家提出了如何通过构建生态化治理体系的方式,有效保障国家数据安全免受新型威胁的现实问题(沈逸,2022)。自 2021 年脸书(Facebook)母公司改名为 Meta 以来,“元宇宙”的概念在全球掀起了新的浪潮,各国纷纷制定元宇宙发展战略,力图在这一高新科技领域占得先机。

截至 2022 年 9 月,我国已有北京、上海、广州等 30 余个省市颁布了元宇宙相关支持性政策,其中 11 个省市出了台 19 份元宇宙专项政策,力图突破关键技

* 唐欣奥,北京林业大学人文社会科学学院。

术、扩大产业规模、构建产业生态，在全球范围内夺得先机。我国在大力发展元宇宙技术、产业的同时，也不能忽视其带来的数据安全风险，这需要研究者以前瞻性的眼光对元宇宙数据安全风险进行预测和防范。

一、元宇宙数据安全内涵

（一）何为“元宇宙”

一般认为，元宇宙概念最早出现于尼尔·斯蒂芬森（Neal Stephenson）1992年出版的小说《雪崩》之中，该书描述了一个由电脑和网络构成的虚拟空间。在之后的几十年里，人们从未停止对这样一个虚拟空间的想象。目前，Meta、微软、苹果等多家互联网巨头都相继布局元宇宙战略；2022年6月，Meta、微软、Epic、英伟达、高通、索尼、华为等36家公司组建了一个名为“元宇宙标准论坛”（Metaverse Standards Forum）的组织，拟为元宇宙、AR/VR等建立行业标准。

元宇宙大致是生活在现实物理世界的自然人以“化身”或者“数字人”的方式，通过计算机操作系统，与其他数字人即时互动的3D数字虚拟空间（程金华，2022）。作为热门概念的“元宇宙”目前尚未有一个获得普遍共识的定义，但根据当下的讨论与研究可知，各观点和学说都认为其有虚拟性和沉浸式的特点。在与元宇宙数据相关的方面，它还具有一定的独立性、封闭性，并由此具有会导致数据出现安全问题的隐蔽性。

（二）元宇宙数据

1.“元宇宙数据”定义

在元宇宙这样的虚拟世界中，一切都以数字化的形式存在，我们的每一个行为都可以被记录，成为元宇宙数据的一部分。未来的元宇宙数据并不会脱离当下人类对于数据的认识，只是在记录、储存形式等技术层面，伴随着技术的进步可能会有所更新和改变。我们可以通过当前对于数据的认识，总结出“元宇宙数据”的初步定义。

各国各地区的法律对数据的定义不尽相同。欧盟《数据治理法》（*Data Governance Act*）将数据定义为：“行为、事实或信息的任何数字表示以及此类行

为、事实或信息的任何汇编,包括声音、视觉或视听记录形式。”美国《政府数据开放法》(*Open Government Data Act*)将数据定义为:“被记录的信息,不论记录数据的形式或媒介。”我国《数据安全法》第三条规定,数据“是指任何以电子或者其他方式对信息的记录”。根据以上对数据的定义,可以尝试对元宇宙数据做出如下定义:元宇宙数据是指任何以电子或者其他方式对元宇宙世界中行为、事实或信息的记录,包括声音、视觉或听觉记录等生物识别信息和非生物识别信息。

2. 元宇宙数据主体

当前,数据主体主要有三类:其一是个人,个人信息及行为的记录产生零散的原始数据;其二是数据利用者,他们收集和整合海量零散的个人数据产生大数据,并对此继续开发利用,包括企业和国家相关部门;其三是国家,国家一方面为本国数据收集处理活动制定规则,另一方面在国际社会中保护本国大数据及其技术,维护国家安全。

可以预见,元宇宙数据涉及的主体主要也有三类。一是元宇宙用户,他们在元宇宙中进行社交、游戏、虚拟产品交易等活动,其生物信息和行为信息的直接记录构成了元宇宙用户数据,也是元宇宙的原始数据。二是元宇宙平台,平台为用户提供元宇宙服务,包括提供可供接入的元宇宙本身以及元宇宙内部的购物、游戏、学习等多类型服务,同时在提供元宇宙服务的过程中收集产生的数据。一方面,平台需要对用户数据进行分析以优化服务,吸引更多用户;另一方面,平台也可能直接将加工后的数据作为商品出售给别的元宇宙或是相关机构,以获得利益。平台是元宇宙衍生数据的主要创造者。三是国家,我国《数据安全法》第七条规定:“国家保护个人、组织与数据有关的权益,鼓励数据依法合理有效利用,保障数据依法有序自由流动,促进以数据为关键要素的数字经济发展。”在元宇宙视角下,国家自然也不会缺席元宇宙数据的管理。三种主体中,国家并不单独创造数据,即使创造数据也是以前两类主体的身份进行,“元宇宙国家数据”主要针对元宇宙数据中的个人信息、重要数据和核心数据等关乎国家安全的部分。

(三)元宇宙数据安全风险

不论是元宇宙原始数据,还是元宇宙衍生数据,都可能受到数据安全风险的

威胁,例如原始数据的泄露可能侵犯用户隐私,衍生数据的泄露可能侵犯平台权益,等等。本文将对元宇宙数据整体的数据安全风险及应对措施进行讨论,而不再具体区分两类数据。

1. 元宇宙数据安全

基于元宇宙数据的定义,可以用类推的方法对元宇宙数据安全做出初步定义。国际化标准组织(ISO)对数据安全的定义是:“保护数据以保证可用性、机密性及数据完整性。”以科学数据为研究对象,李善青等学者将“科学数据安全”定义为通过必要的技术和管理措施,保护科学数据在其全生命周期中免受破坏性外力和非授权操作的侵害,保持科学数据的机密性、完整性和可用性(李善青、郑彦宁、邢晓昭等,2019)。《数据安全法》第三条规定:“数据安全,是指通过采取必要措施,确保数据处于有效保护和合法利用的状态,以及具备保障持续安全状态的能力。”综合以上定义,可以对“元宇宙数据安全”做如下定义:元宇宙数据安全是指通过采取必要措施,确保元宇宙数据在其全生命周期中,包括收集、利用、存储、传输等,始终处于有效保护和合法利用的状态,以及具备保障持续安全状态的能力。

2. 元宇宙数据安全风险

(1)元宇宙数据安全风险的体现

根据工业和信息化部发布的《工业和信息化领域数据安全风险信息报送与共享工作指引(试行)(征求意见稿)》(以下简称《工作指引》),数据安全风险是指通过监测、评估、信息搜集、授权监测等手段获取的数据安全风险。根据实践经验,《工作指引》总结概括了六大类数据安全风险,分别是数据泄露、数据篡改、数据滥用、违规传输、非法访问、流量异常;但同时也认为数据安全风险包括但不限于上述类别,留下了增补的空间。实践中也还存在其他数据安全风险,如数据霸权、网络攻击等。

上述数据安全风险是长期以来从技术发展和实践经验中总结而来的,未来也可能会继续出现新的安全风险。就元宇宙而言,当前元宇宙技术尚在发展阶段,元宇宙数据安全风险并不会立刻显现出和当前数据安全风险完全相同的面貌,而将根据其技术发展和数据收集利用情况,在不同发展阶段呈现不同数据安全风险。对元宇宙数据安全风险的预防,应优先应对已经出现或很快就会出现的最紧迫部分。根据当前元宇宙数据安全案例,以及元宇宙技术发展带来的数

据收集利用情况,元宇宙数据安全风险将首先体现为元宇宙数据滥用、元宇宙数据泄露、元宇宙数据不当流通以及元宇宙数据垄断。

(2)元宇宙数据安全风险的特点

元宇宙数据安全风险具有一般性和特殊性:一般性是说,元宇宙数据作为一种特殊的数据,它所产生的数据安全问题和当下的数据安全问题将会有交叉;特殊性是说,元宇宙本身具有独立性与封闭性,这样的独立性与封闭性导致外界难以全面地了解元宇宙内部的活动细节,同时也因为其中所有的活动都以数字化的形式被记录,数据安全问题隐藏在庞大的数据海洋中,难以被察觉,所以元宇宙数据安全风险将更隐蔽。此外,由于元宇宙数据在采集范围、采集方式以及数据内容等方面的特殊性,其数据安全问题的后果将更加严重,影响范围更广,弥补难度更大。

(3)元宇宙数据安全风险分类应对

根据参与数据流动的主体和行为,数据安全风险将会在不同层面以不同的方式表现出来,引发不同的后果。例如 2022 年 2 月,Meta 公司因涉嫌非法采集用户面部信息被美国得克萨斯州总检察长起诉,指责其未经同意收集了数千万州民的面部识别数据,并将该部分数据披露给了第三方,涉案金额达数十亿美元。该案表明,即使虚拟世界的元宇宙还在发展成熟阶段,但是其技术发展过程中产生的数据收集处理活动就足以带来严重的隐私数据安全风险,造成严重后果,并且已经在个人、企业层面显现出来。对于元宇宙数据安全风险的分析和应对,应当和其技术的发展同步跟进,甚至需要前瞻性的预防;并且,应从元宇宙用户、元宇宙平台、国家三类主体的角度切入,对元宇宙数据安全风险展开多层次的应对。

二、元宇宙数据安全风险可能性分析:基于三类数据主体

结合当前数据安全风险治理局势,以及对未来元宇宙数据所涉相关主体的预测,各个主体在元宇宙数据产生利用的过程中,扮演不尽相同的角色,拥有不尽相同的重要性。从规避元宇宙数据安全风险的角度看,应从主体入手,对元宇宙数据安全风险在元宇宙用户、元宇宙平台以及国家三类主体上的具体表现及

相应的应对措施进行可能性分析,以期提供有益意见。

(一)元宇宙用户的数据安全风险

数据驱动型经济是目前创新的蓝海,该模式需要对收集的海量数据进行开发利用,平台企业的业务如商业推广、精准营销等大都依赖于此(沈逸,2022)。可以预见到,在未来元宇宙经济环境下,收集、存储、利用用户数据仍将成为元宇宙平台经营发展的重要模式,而相比当下不同服务内容相互基本独立的点状互联网服务提供模式,作为沉浸式虚拟世界的元宇宙中的用户数据安全问题将更加严峻。元宇宙的技术原理和运行方式决定了即使在最小必要原则的约束下,它也必然会广泛地获取用户授权,全方位、实时地收集每一个用户信息(王德夫,2021)。这样的用户数据将不再局限于购物、社交等单一类型,而呈现出全方位记录的特点。换言之,元宇宙中的个人数据将是用户在这样一个虚拟世界中的全部痕迹,这是元宇宙用户数据与当下互联网用户数据的最大区别也是最主要区别,这样的全方位数据收集记录将为个人数据保护带来新的问题。具体而言,元宇宙用户个人数据安全问题将主要体现为元宇宙用户数据滥用、用户数据泄露和用户数据不当流通三个方面。

1. 元宇宙用户数据滥用

在元宇宙用户数据滥用方面,元宇宙平台过度收集或超范围、超用途、超时间地利用用户数据,将侵犯用户隐私。前文提到,收集、存储、利用用户数据仍将是元宇宙平台经营发展的重要模式,但是在元宇宙这样的环境下,即使平台是在最小范围内收集用户数据也依然会得到全方位的数据记录结果。算法帮助平台企业实现了深度挖掘个人信息,对于这样的全方位数据进行分析处理将得到比单一类型数据分析处理更加深入、更加立体的结果,这一过程极有可能突破用户个人隐私权边界。

2. 元宇宙用户数据泄露

在元宇宙用户数据泄露方面,元宇宙用户数据被恶意获取、转移或泄露至不安全环境,这将带来更大的安全风险。元宇宙中的用户数据呈现出全方位记录的特点,一旦出现数据泄露事件,就将是全方位的信息外泄。与金融信息泄露、用户注册信息泄露或者医疗信息泄露等单一类型数据泄露相比,元宇宙用户数据泄露负面影响的范围和程度将空前扩大(王德夫,2021)。这也对元宇宙环境

下的用户个人数据保障提出了更高的挑战。

3. 元宇宙用户数据不当流通

在元宇宙用户数据不当流通方面,用户个人数据具有人身性和财产性双重特点。信息控制者不但可以利用技术手段深入挖掘个人信息中潜藏的商业价值,使其服务于企业的经营、管理和决策,还可以将其直接打包出售以实现数据中的财产价值(王玉林,2016)。随着数据经济的发展,当下企业间的大数据交易流通已不是新鲜事。而在元宇宙环境下,如何界定可流通的大数据与用户个人数据的边界,如何防范元宇宙数据未按照有关规定就被擅自传输,如何防范尚未经过足够的脱敏加工以实现匿名化的用户个人数据"被流通",甚至这样的全方位数据即使是脱敏加工后,是否依然存在被其他平台利用技术手段去匿名化,也将是值得关注的问题。

(二)元宇宙平台的数据安全风险

经营者对可商品化数据使用价值的挖掘,是数据财产属性在大数据时代得以充分体现的关键,经营者据此应当对其所持有的可商品化数据享有财产利益(谢宜璋,2021)。平台除了将自己的元宇宙用户数据进行内部分析利用之外,还有可能对其进行分类、加工、脱敏以后放到数据交易市场进行交易,以实现财产价值。根据互联网企业的数据流动现状,这样的数据流通现象在未来元宇宙平台之间几乎必然存在,甚至未来的元宇宙平台很有可能就诞生于当下活跃在数据交易市场中的大型互联网企业里。但是,当下互联网平台大多只记录用户的单一类型数据,就已经引发了众多法律纠纷,诸如"上海汉涛信息咨询有限公司诉北京百度网讯科技有限公司等不正当竞争纠纷案"①和"深圳谷米公司诉武汉元光公司不正当竞争纠纷案"②等。因此,有必要对未来平台层面的元宇宙数据安全问题进行有前瞻性的探索与预防。把涉及用户个人隐私的部分归入用户数据安全问题,平台上的数据可能存在元宇宙平台数据不当流通以及数据垄断问题。

1. 元宇宙平台数据不当流通

当下互联网平台间的数据流通,主要有基于合同的大数据交易以及通过爬

① 上海知识产权法院(2016)沪73民终242号民事判决书。

② 广东省深圳市中级人民法院(2017)粤03民初822号民事判决书。

虫技术爬取他方平台数据两种方式。元宇宙本质上仍然是一种互联网服务，所以当下的两种数据流动模式仍将是未来元宇宙平台间数据流通的主要模式。大数据交易需要双方达成合意，对元宇宙平台数据利益的侵害可以得到经营者的有效防范，其中主要包含的是对元宇宙用户数据权利的侵犯，故此处主要讨论后者。

通过爬虫技术抽取他方平台数据，本质上是在未达成合意的情况下收集了大量的他方平台的数据，以形成新的具有商业价值的大数据集合，侵犯了他方平台的商业利益，引起了不当竞争纠纷和平台数据泄露问题。这样的数据泄露问题在未来的元宇宙平台之间将依然存在。在未来的元宇宙环境下，元宇宙平台将记录的是用户行为各方面的全方位数据，平台收集数据的成本更高、内容更广，这可能导致非合意数据爬取行为更加严重。爬取方可能盯准被爬取方的某一类数据，也可能对其数据整体进行非法的收集爬取，对被爬取方的权益侵犯显然值得重视。

2. 元宇宙平台数据垄断

数据垄断即持有并控制海量数据的少数公司，通过占有大数据形成市场支配地位，滥用其市场支配地位以及利用数据实施不正当竞争行为（张菲、朱桐雨，2022）。元宇宙平台数据垄断问题是指大型元宇宙平台利用其在市场中的优势地位，不当排除其他平台对公域数据的收集利用，阻碍其他元宇宙平台的发展，形成自己在市场中的数据垄断地位。就此而言，元宇宙平台有范围地公布部分数据，合理的非合意的数据爬取行为也能为元宇宙数据流通、市场进步做出贡献。就此而言，如何妥当划定数据爬取正当性边界，不但关乎各方利益，更关乎网络空间数据流通的大局（许可，2021）。

（三）元宇宙中的国家数据安全风险

从国家层面来说，元宇宙数据安全同样是不可忽视的问题。《数据安全法》第四条规定："维护数据安全，应当坚持总体国家安全观，建立健全数据安全治理体系，提高数据安全保障能力。"这表明我国要建立健全数据安全治理体系，元宇宙数据安全的防范与治理自然应当包含于其中。具体而言，根据当前国家数据安全面临的风险，可以推测在未来元宇宙语境下，国家数据安全主要有元宇宙数据跨境流通中的泄露或不当流通风险。

数据跨境流通意指在一国内生成的电子化信息记录,被他国境内的私主体或公权力机关读取、存储、使用或加工,合称“处理”(何傲翾,2021)。在当前经济全球化和数字经济快速发展的局势下,数据跨境流通不仅是数字经济发展中面临的重要问题,也是开展数字贸易的基本前提(何波,2022)。元宇宙数据不同于一般大数据的信息全方位性,这意味着其跨境流通将带来更大的个人数据安全风险和国家数据主权风险,一旦发生泄露或者被不当流通,将对国家安全造成更大的危机。

三、当前我国元宇宙数据安全保护制度及缺陷

(一)我国元宇宙数据安全制度现状

当前,我国正在不断完善数据安全相关制度和规则,不断推动数据安全协同治理机制的形成(马忠法、胡玲,2021)。虽然还未针对元宇宙及元宇宙数据进行专门立法,而多是针对广义数据安全进行立法保护,但也有部分政策提到元宇宙。现主要有以三部法律、多个中央和地方政策为主的元宇宙数据安全保护格局,并形成以数据分类分级保护、数据安全风险评估和监测预警、数据安全应急处置为核心的数据安全制度及以告知同意为核心原则的个人信息保护制度。

1. 元宇宙数据安全相关法律

三部法律分别是《网络安全法》《数据安全法》《个人信息保护法》。其中,《网络安全法》提出了网络安全等级保护制度,并规定关键信息基础设施的运营者掌握的个人信息和重要数据应在境内保存,确需出境的需进行安全评估。《数据安全法》第九条提出数据安全协同治理体系,在第三章“数据安全制度”中规定了数据分类分级保护制度、数据安全风险评估和监测预警、数据安全应急处理机制,在第四章“数据安全保护义务”中规定了重要数据处理者的数据处理活动风险评估义务。《个人信息保护法》在敏感个人信息处理、国家机关处理个人信息、个人信息跨境提供、个人的权利、个人信息处理者的义务等方面构建了以告知同意为核心原则的个人信息保护制度框架,并在第七章规定了法律责任。另外,《民法典》也在第一千零三十四条规定了个人信息保护范围。目前尚未针对元宇宙进行专门立法,元宇宙数据安全相关问题应当参照并适用以上法律。

2. 元宇宙数据安全的中央及地方相关政策

中央政策是指国务院印发的《“十四五”市场监管现代化规划》中提出的，要“加快研究和完善国家数据安全标准与认证认可体系”。《要素市场化配置综合改革试点总体方案》进一步提出，要“加强数据安全保护”“推动完善数据分级分类安全保护制度”。二者虽然都未明确提及元宇宙，但是元宇宙数据安全作为国家总体数据安全下的重要部分，应当参照适用。

多个地方政策是指北京、上海、广州等地针对元宇宙发展制订的行动计划。多地把保障元宇宙数据安全作为工作要求之一，例如：《北京城市副中心元宇宙创新发展行动计划（2022—2024 年）》明确提出，“提升覆盖云、网、端的数据安全服务与保障能力”；《上海市培育“元宇宙”新赛道行动方案（2022—2025 年）》要求，“加强未来网络、云边计算、智能交互终端及数字基础设施的内生安全，保障海量数据的存储、传输和使用”；《厦门市元宇宙产业发展三年行动计划（2022—2024 年）》规定，“提升元宇宙数据治理能力，加强数据安全管理和隐私保护”；《河南省元宇宙产业发展行动计划（2022—2025 年）》也提出，要“深入研判元宇宙发展的伦理风险、数据安全风险、沉迷风险、知识产权风险等”。

（二）元宇宙数据安全风险应对的现行制度之缺陷

首先，尚未出现针对元宇宙的法律法规。虽然《数据安全法》第九条规定了数据安全协同治理体系，但只是对保护主体做了明确说明，后文中也只是使用了“处理者”这样的词汇，并未真正建立起多元主体的数据安全保护制度。各省市的元宇宙行动计划中虽然有提到数据安全，但仅仅停留在原则层面，缺少具体的细化实施办法和规定。元宇宙数据安全保护制度应建立在一般数据安全保护制度之上，并具有针对元宇宙特点的专门规定，在目前这种缺少具体细化规定的情况下，仅凭原则性规定想要建立多元共治的元宇宙数据安全保护制度更是难上加难，不利于应对元宇宙数据安全风险。

其次，数据分级分类体系尚不完善。虽然我国确立了数据分级分类保护制度，确立了不同层级、不同领域的重要数据目录制定主体，并且全国信息安全标准化技术委员会发布的国家标准《信息安全技术　网络数据分类分级要求》（征求意见稿）已规定了较为详细的数据分级参考规则，但当下对核心数据、重要数据的识别脉络仍不够清晰，在行业数据分类以及重要数据目录制定方面仍在初

步阶段,数据管理制度和要求也缺乏细化规定,这些无疑都会影响元宇宙数据的分级分类保护、元宇宙数据管理和风险应对。

四、元宇宙数据安全风险的法律回应:从去中心化到多元共治

(一)元宇宙的去中心化管理

1. 去中心化管理概念及现状

在元宇宙蓬勃发展的当下,有一部分声音认为元宇宙应该建立一个去中心化的成员自组织管理模式。去中心化(decentralization)在元宇宙语境下,即让每位元宇宙成员都能拥有该元宇宙的管理权,能参与到元宇宙的运营、管理和决策中,而不需要一个组织或特定的部分人集中对元宇宙进行管理。区块链技术也为这样的愿景提供了技术基础。区块链技术自问世以来,因其去中心化的分布式记账特征,被比喻为一种“信任机器”(程金华,2022)。当前已经存在这样的去中心化自治组织,例如成立于 2016 年 11 月的 Aragon DAO,以及 2020 年 8 月上线的国内第一个基于区块链技术的自组织管理平台 VoneD-AO(刘涛、袁毅,2022)。去中心化的观点主要从法律和理想两方面展开。元宇宙中行为的法律边界和效力处于待定状态,政府的干预又处于上位法真空状态,因此需要元宇宙成员自行达成定纷止争的治理规则并自立执行,以应对成员身份和行为发生的深刻变化(张钦昱,2022)。依靠区块链技术创造去信任、公开透明的网络环境,随着人工智能的应用,每个个体都将是具备感知、推理、决策功能的智能代理,参与组织的运营、管理和决策,真正发挥出群体智能的力量(刘涛、袁毅,2022)。

2. 去中心化管理的弊端

(1)元宇宙无法真正实现去中心化

元宇宙具备高度中心化的架构,并体现在相关系统搭建与运行、数据收集与处理、数字权益创设与处分等每一个“宇宙角落”(王德夫,2021)。元宇宙的产生和运营一定需要平台提供技术和资金支持,并且单个用户并不会只停留在一个元宇宙,也很难有精力参与到各个元宇宙的治理中去。从保证自身市场竞争力和保护用户隐私的角度来讲,有能力收集用户数据的也只有在技术上处于优

势地位的平台,用户一方面不能获取他人的数据或者只能获取公开的部分,另一方面在个人数据权利层面也未必能与平台处于对等地位。

(2)过大的自治权和空间不利于保护元宇宙数据安全

元宇宙去中心化的管理方式赋予每一个元宇宙使用个体极大的主体自由,如果在数据安全风险治理问题上给予元宇宙过大的自治权和空间,将加大数据安全风险,有违我国建立健全数据安全治理体系的目标。关乎个人隐私和国家安全,不能放任元宇宙野蛮发展。虽然在元宇宙其他管理层面来说存在去中心化治理的实施空间,但是在元宇宙数据安全风险治理层面来说,我们应给予更多的关注,需要公权力的约束,由平台、用户与公权力主体共同参与,以多元共治的机制进行元宇宙数据安全风险的治理、防范和应对。

(二)元宇宙数据安全风险回应的理想路径:多元共治

在传统意义上,多元共治是指以政府、市场与社会为主体,以法治、协商和自治为系统,以对话、竞争、妥协、合作和集体行动为机制的多元主体共同治理的社会共治模式。该治理模式是改革开放时期在制度转型、资源污染、资源约束等问题下我国实践探索的经验总结。其路径是政府职能转变和简政放权,将减少对微观经济事务的干预落到实处,采取更合理的方式消弭市场经济运行带来的各种负面影响(王名、蔡志鸿、王春婷,2014)。

1. 国家:把握底线、监督引导

在元宇宙数据安全风险应对的多元共治中,国家应主要承担把握底线、监督引导的作用。建立健全数据安全治理体系是保障数据安全的重要措施之一,元宇宙数据安全将是其中的重要部分,国家不应缺席其管理。并且,由于法律无法避免的滞后性,国家立法很难跟上技术快速进步的步伐。未来元宇宙的发展显然不会千篇一律,统一的元宇宙数据管理制度难以适配不同特点和类型的元宇宙产品,也难说适合具有全方位性的元宇宙数据。所以,相比于把元宇宙数据管理制度的细枝末节都收揽于国家手中,通过制定元宇宙数据分级分类保护的行业标准、拟定基本规范、确立应急情况的程序和责任承担等方式确立较为宽泛的元宇宙数据利用制度框架,重在把握元宇宙数据安全的底线和红线,把更细节的管理权放给直接支配和管理元宇宙的平台应是更好的选择。这样既可以减轻国家立法和执法压力,节约行政资源和司法资源,还可以通过审查平台制定的细化

的元宇宙数据管理规定,监督和引导企业建立符合基本规定的元宇宙数据管理制度,处理元宇宙用户与平台间的数据纠纷,实现对元宇宙数据安全的前端预防和应对。

2. 元宇宙平台:有限自治、接受监督

元宇宙平台应享有有限的自治权,接受国家和社会的监督。元宇宙平台直接控制、运营和管理自己的元宇宙产品,进行元宇宙数据的收集、处理活动,直接接触也最先接触到海量的元宇宙数据,是防范、发现和应对元宇宙数据安全风险的最前线。但因元宇宙平台本身的逐利性,不能将制定元宇宙数据分级分类保护标准等原则性和底线性规定的权力交予平台,而只能赋予其在底线框架之内进一步根据自己的元宇宙产品和元宇宙数据的特点,建立相应的有特色的元宇宙数据管理制度、元宇宙数据安全风险预防和应对机制的权力。换言之,在国家的监督和引导之下享有有限的自治空间,发挥其灵活性和前瞻性,适应迅速发展的技术和多变的现实情况。通过元宇宙数据风险预防、应对机制,以汇报情况、接受监督和引导、共同应对突发事件等方式和国家形成有机互动,形成权力互赖的伙伴关系,一方面能保障平台自身的利益不受非法侵害,另一方面也能共同达成建立健全元宇宙数据安全治理体系的目标。

3. 元宇宙用户:积极维权、反馈、举报

元宇宙用户作为国家和平台之间直接对接之外的另一条纽带,同样在预防和应对元宇宙数据安全风险中起到重要作用。由于个体力量较小、缺少技术能力、精力和经济有限等,元宇宙用户是相对于平台的弱势方,其数据权利容易受到平台的侵害。作为用户相对方的平台显然也难以保证公平公正,所以需要国家一定程度的介入。虽然用户之间通常难以形成合力,但其往往具有一定的规模性,单个用户的数据权利纠纷背后很可能隐藏着规模性侵犯用户元宇宙数据权利的行为,引发用户层面的元宇宙数据安全风险。这样的纠纷可以首先体现为少量用户数据的泄露、被滥用和被非法买卖传输等,而当少数用户发现自己的元宇宙数据权利受到侵害时,往往这一侵害行为已经波及更大范围的用户。用户层面的元宇宙数据安全风险可能进一步体现为平台甚至跨国的数据不当流通等,在这个意义上,通过国家主动关注元宇宙用户数据权利保护情况,再加上用户及时反馈和举报,有关部门能够及时发现潜在的元宇宙数据安全风险并加以控制。当然,除了可能引发更大规模的元宇宙数据安全事故,即使是不具备规模

性的个别用户的数据权利，也应当受到国家和平台的尊重与保护，这不仅仅是建立健全数据安全治理体系目标的要求，更是建设社会主义和谐社会的要求。

如此，以多元共治为基本原则，构建国家、元宇宙平台、元宇宙用户三方主体共同参与的立体多元主体共治制度，才能在各个层面有效防范和应对元宇宙数据安全风险。

五、元宇宙数据安全的制度建构

元宇宙数据安全风险回应需要以多元共治为基本原则，下文将从元宇宙用户、元宇宙平台、国家三主体的角度，论述各个主体应采取何种措施应对元宇宙数据安全问题。

（一）我国应尽快构建元宇宙制度

如前文所述，当前我国并未针对元宇宙进行专门立法，而各地的行动计划大部分也是论述如何鼓励发展元宇宙产业、元宇宙技术，对于元宇宙数据安全只在原则和目标部分略有提及。当前元宇宙数据安全还只能参照现有的数据相关法律适用，但元宇宙数据和传统数据的主要差异，以及元宇宙数据的全方位性，表明其数据安全风险的危害程度更在传统大数据的数据安全风险之上，只参照现有的数据制度显然不利于切实保护元宇宙数据安全风险。我国在重视推动发展元宇宙技术和产业的同时，也应重视在技术发展过程中以及技术发展成熟后所产生的数据安全风险，应尽快建立元宇宙基本制度、元宇宙数据安全风险应对制度，跟进技术发展的步伐，既要技术和产业发展，也要守住数据安全的红线。

（二）基于不同主体的元宇宙数据安全风险应对

1. 元宇宙用户应提升权利意识，遵守平台规范

就个人层面来说，个人用户在元宇宙中相较于平台处于技术上的弱势地位，个人的精力和所能发挥的作用也有限。元宇宙数据作为一种大数据，其一大特点就是来源海量性，所以不能苛求个人与平台抗衡以保障自身的权利。但是也不应完全剥夺个人保障自己数据权益的权利，而应当赋予用户在元宇宙视角下的元宇宙个人数据权，并通过普法教育、元宇宙进入页面告示等方式提高用户对

于自身数据权利的保障意识,并能积极主动保护自身合法的元宇宙数据权利。同时,要以公权力保障用户权利的实施,一旦用户难以以自身力量在与平台的纠纷中保障自身权益,公权力就可作为救济措施介入。当然,用户自身因为成本过高或其他原因而难以争取自身权益的时候,也应为用户提供向公权力求助的通道,甚至公权力也可主动介入保护用户个人的元宇宙数据权利,防范和应对元宇宙用户数据被平台滥采滥用、数据泄露和数据在未达到流通条件时被不当流通的风险。

用户在数据安全可能受到威胁的同时,也有可能会反过来威胁元宇宙数据安全。例如现实中一些人以跟踪、人肉搜索等方式侵犯他人隐私,或是为境外势力拍摄机密设施、机密地区照片泄露国家机密等。为了预防该类事件在元宇宙中发生,一方面,需要平台充分提示用户在元宇宙中的权利义务,需要普法教育让用户了解元宇宙行为规范,让用户了解元宇宙不是法外之地,虚拟世界中同样有行为红线;另一方面,需要用户自觉遵守元宇宙行为规范,遵守行业规范,在接受公权力保护的同时也应当承担不去危害元宇宙数据安全的责任。

2. 元宇宙平台建立健全元宇宙数据安全管理制度,强化法律责任

元宇宙平台应建立健全元宇宙数据安全管理制度,承担元宇宙数据保护义务:一方面,平台在运营和发展过程中不可避免地会对用户数据进行采集和存储,所以在这个过程中平台自然应当承担对所采集的数据妥善保管保护的义务,以防止元宇宙数据泄露;另一方面,根据当前对大数据财产权的讨论和现实中对大数据的利用情况可以推知,平台对于所采集的原始数据进行适当的加工、脱敏以后形成的衍生数据,同样可以拥有财产权,可以进行流通以实现部分商业目的。不同于传统大数据财产权,元宇宙数据是用户在元宇宙中各类不同类型行为的集合,具有全方位性,所以在处理和流通阶段平台应当承担相较于传统大数据处理和流通过程中更大的数据保护义务。具体而言,平台对于元宇宙数据的脱敏程度应当更高,在流通过程中需要积极配合有关部门对处理后的元宇宙数据进行评估,以确保其满足市场流通标准,防止其被技术复原而恢复身份识别性,防止泄露用户个人隐私。此外,平台还应建立数据安全动态风险评估监测和突发情况应急制度,在元宇宙数据处理的全过程中有效防范风险,在发生元宇宙数据安全事故时有效组织力量减小损害、降低损失,及时向有关部门通报,有效应对紧急情况。平台还应明确元宇宙数据安全风险责任承担制度,明确责任分

配与责任承担,避免追责难的问题,督促相关责任人员做好数据安全防护工作。

平台应保障用户的元宇宙数据权利。个人信息分为一般个人信息与敏感个人信息(时明涛,2020),相应地,元宇宙数据也可以分为一般元宇宙数据和敏感元宇宙数据。元宇宙平台应充分保障元宇宙用户的数据权利,保障用户的知情同意权,即平台收集、存储、利用元宇宙数据等应让用户知晓并在取得用户同意以后进行,用户在元宇宙活动中产生的敏感个人数据更应如此。当用户希望在别的元宇宙世界中继续使用前一个元宇宙中的数据时,平台应当积极予以配合和协助,同时对于用户可以带走的数据范围进行协商。欧盟《通用数据保护条例》第 20 条创设了数据可携权这一新兴权利(汪庆华,2021),这一权利在未来元宇宙世界中或也有适用空间。不过,只要求单个企业自律或还有不足之处,笔者认为也有必要建立相应的行业规范和行业组织予以配套,督促元宇宙平台企业自觉自律承担责任、履行义务。

3. 国家建立元宇宙数据安全保障制度

用户和平台对于元宇宙数据安全风险的应对固然发挥着重要作用,但是就动力而言,更多需要的是用户和平台的自我驱动;缺少强制力的保障,对元宇宙数据安全的维护就难说完善。用户和平台更多是事前的防范,而难以在数据安全问题发生以后进行有效的止损和弥补,在追究责任方面更是捉襟见肘。元宇宙数据中必然包含诸如元宇宙系统数据、商业秘密数据、本国公民用户的个人信息及行为数据等关乎国家安全的保密数据,这就需要我国公权力的介入,建立更完善的制度并以强制力保障制度运行,同时能在数据安全发生问题之后统筹资源、弥补损失并追究相关人员责任。在最后这一部分,本文将探讨国家层面对于元宇宙数据安全风险的应对措施。

(1)及时建立元宇宙数据分级分类目录

在我国数据分级分类保护的制度背景下,元宇宙数据管理也应采取相同路径。当前首先要尽快通过国家标准《信息安全技术　网络数据分类分级要求》,并建立元宇宙行业数据分类以及相关核心数据、重要数据、一般数据分类目录,如此才能和当前的数据保护制度有效衔接,以应对日渐严峻的元宇宙数据安全风险。随着元宇宙技术和产业的不断发展,各类数据的群体、区域、精度、规模等因素会产生变化,因此还应定期评估元宇宙数据分级分类目录的合理性并及时做出调整,防止部分数据重要性上升以后却未得到法律的有效保护,造成元宇宙

数据安全事故。

(2)审查备案元宇宙平台资质及其产品

有关部门应严格审查元宇宙平台资质,审查备案元宇宙产品。可以预见到,未来元宇宙平台不会只有一家,而可能是来自不同国家的多家元宇宙企业同台共舞,这也导致了元宇宙数据安全问题可能在不同元宇宙平台上发生。元宇宙数据安全风险相较于传统单类型的大数据安全问题,后果更加严重,影响范围更加广泛。为了有效控制元宇宙数据安全风险,国家应建立元宇宙平台资质审查制度,提高元宇宙平台准入门槛,筛除明显没有能力应对和保障元宇宙数据安全的平台,例如审查元宇宙提供平台的数据安全保障技术能力、相关信用记录等方面的信息。同时,应建立元宇宙产品数据安全相关标准,对于平台所提供的各个元宇宙世界产品进行审查和备案,具体筛查各个元宇宙是否有数据安全漏洞,是否能够满足相应使用场景下的数据安全保障要求,对于不符合安全要求的元宇宙应不予批准对外开放。

(3)构建元宇宙数据安全管理具体要求

有关部门应介入元宇宙内部数据管理,监督、引导平台制定元宇宙数据安全管理制度。虽然对于元宇宙内部社会性的管理目前有去中心化和非中心化两种不同观点,但是在元宇宙数据方面去中心化的管理显然不能满足防范数据安全风险的要求。为了防范平台自主制定的数据管理制度的漏洞和风险、平台怠于制定元宇宙数据安全保护制度、平台怠于履行元宇宙数据保护义务,国家应当介入元宇宙内部的数据管理,在确认元宇宙数据也应遵守当前《数据安全法》《个人信息保护法》等法律以外,针对元宇宙数据的特殊性,在元宇宙数据安全风险评估与监测预警、元宇宙数据安全风险应急管理措施、元宇宙数据安全管理制度备案等方面,为元宇宙的数据管理制定统一的基本管理规定。例如,在元宇宙个人信息收集和使用活动不得有所放宽的客观现实下,还需要更严格的数据分级分类管理制度、数据流动制度等予以配套(王德夫,2021);抑或要求平台在发生元宇宙数据安全事故时及时向有关部门通报,以及时控制局势、减少损害。国家层面既可以颁布统一的元宇宙数据管理办法,在现有法律法规中新增相关条文,也可以对元宇宙公司章程中涉及数据管理的内容单独进行审查,监督、引导各元宇宙平台切实建立健全元宇宙数据安全管理制度;同时,为平台留下一定的数据管理利用空间,在保障元宇宙数据安全和平台企业创新积极性与数据要素市场

活力之间达到平衡。

(4)建立元宇宙平台数据安全管理的常态化审查监督制度

在审查和管理之外,还应建立元宇宙平台数据安全管理的常态化审查监督制度,及时排除风险;动态监测平台元宇宙数据安全管理制度的落实情况,定期进行元宇宙平台数据管理利用情况审查,及时发现平台安全漏洞并要求其整改,要求平台及时更新落后的元宇宙数据安全管理制度,避免真正造成数据安全事故;要求平台定期公布元宇宙数据安全报告,接受社会监督,帮助元宇宙用户更好地识别和筛选更有数据安全保障能力的元宇宙平台,促进平台良性竞争。做好元宇宙数据安全风险的事前防控,少发生、不发生元宇宙数据安全事故才是最终追求。

六、结语

当前元宇宙的发展距离实现其沉浸式虚拟世界的目标还有一段距离。在元宇宙技术不断发展的过程中,作为其核心技术之一的大数据技术也会随之发展,进行大量的数据采集、存储、利用等活动会一直持续到元宇宙的诞生和发展后,这将对数据安全保障提出新的挑战,而元宇宙数据和当前大数据不同的全方位性也预示着元宇宙数据安全更加需要我们重视和保护,并且将在用户、平台和国家三类主体上表现出不同的形态,引发不同的后果。为避免元宇宙数据安全问题引发严重后果,有必要对其风险来源进行前瞻性的分析并做好应对准备。我们通过对元宇宙治理模式的分析可知,去中心化的管理模式无法有效应对元宇宙数据安全风险,各主体共同参与的多元共治将是理想模式,但是我国当前尚未建立起元宇宙专门制度,元宇宙数据安全相关问题只能参照现行相关法律。要切实建立有效的多元共治的元宇宙数据安全风险治理制度,尽快建立元宇宙基本制度及元宇宙数据安全保护制度,在基本的数据制度之上为元宇宙数据安全上一把更有针对性、更有效的锁。

参考文献

程金华,2022,《元宇宙治理的法治原则》,《东方法学》第 2 期。

高一乘、杨东,2022,《应对元宇宙挑战——数据安全综合治理三维结构范式》,《行政管理改

革》第3期。

何傲翾,2021,《数据全球化与数据主权的对抗态势和中国应对——基于数据安全视角的分析》,《北京航空航天大学学报》(社会科学版)第3期。

何波,2022,《中国参与数据跨境流动国际规则的挑战与因应》,《行政法学研究》第4期。

李善青、郑彦宁、邢晓昭等,2019,《科学数据共享的安全管理问题研究》,《中国科技资源导刊》第3期。

刘涛、袁毅,2022,《去中心化自组织管理的形态、特征及差异性比较》,《河北学刊》第3期。

马忠法、胡玲,2021,《论我国数据安全保护法律制度的完善》,《科技与法律》第2期。

沈逸,2022,《元宇宙与国家数据安全——构建生态化治理体系的挑战与趋势》,《中国信息安全》第1期。

时明涛,2020,《大数据时代个人信息保护的困境与出路——基于当前研究现状的评论与反思》,《科技与法律》第5期。

汪庆华,2021,《数据可携带权的权利结构、法律效果与中国化》,《中国法律评论》第3期。

王德夫,2021,《元宇宙领域的个人信息保护新挑战与法律应对》,《中国市场监管研究》第11期。

王名、蔡志鸿、王春婷,2014,《社会共治——多元主体共同治理的实践探索与制度创新》,《中国行政管理》第12期。

王玉林,2016,《大数据中个人信息开发利用法律问题研究》,《情报理论与实践》第9期。

谢宜璋,2021,《可商品化数据的进一步厘清——概念、保护诉求及具体路径》,《知识产权》第8期。

许可,2021,《数据爬取的正当性及其边界》,《中国法学》第2期。

张菲、朱桐雨,2022,《互联网平台企业的数据垄断问题研究》,《国际经济合作》第5期。

张钦昱,2022,《元宇宙的规则之治》,《东方法学》第2期。

周若涵,2021,《数据安全风险对国家安全的挑战及法律应对》,《上海法学研究》第1卷。

服务机器人:回顾和展望*

刘德文**

摘要:随着人工智能技术的不断推进,融入人工智能技术的服务机器人正在各个行业大放异彩,协助人类员工高效地完成服务任务。基于蓬勃发展的服务机器人态势,本文首先总结了服务机器人的实践和历史,并对机器人发展的不同阶段进行总结。其次,本文进一步对服务机器人的研究现状进行总结,并提出服务机器人的机遇和挑战。在明确下一代服务机器人是实现人机共生、交互认知的基础上,本文从哲学视角探讨了机器人是否存在意识,并归纳和总结了对机器人是否存在意识的现有研究,提出当前机器人具有"机器意识"的观点。最后,基于政策和技术两个方面对我国大力发展服务机器人提出了可操作的建议。

关键词:服务机器人　意识　回顾　展望　历史

一、服务机器人的实践和历史

服务机器人的发展具有漫长的历史,自1954年机器人的专利第一次被申请以来,至今已经有近70年的历史。尤其是最近几年,机器人开始从工业化向服务化迈进,渗透进社会生活的各个领域,包括工业生产、公共安全、旅游/酒店、医疗保健等。从狭义的角度来看,服务机器人专指在和消费者的互动中解决消费者问题和需求的机器人;从广义的角度来看,服务机器人的服务对象还包括生产、运输等领域——这类机器人亦可被归纳为制造机器人的范畴,它们不直接和最终端的用户产生直接互动。本文将主要关注和人类产生互动的狭义服务机器人。在这一节中,我会从为什么我们需要服务机器人这一话题出发,进而梳理服

* 本文系江苏高校哲学社会科学研究一般项目"消费者数字体验创建及其对品牌绩效影响研究"(项目批准号:2022SJYB0103)的阶段性研究成果。作者感谢两位审稿专家和责任编辑对本文提出的意见。

** 刘德文,南京邮电大学管理学院、南京邮电大学信息产业融合创新与应急管理研究中心。

务机器人的发展历史。在此基础上,本文进一步归纳服务机器人的四个发展阶段。

(一)为什么我们需要服务机器人?

虽然在现代社会中,技术的发展可能让人类对机器人产生抵触和反感,但是人类社会不可避免地需要服务机器人来支撑社会的良性运行。服务机器人日益融入人类生活,成为人类不可或缺的伙伴。机器人存在实体和非实体两种形态:前者指具有物理外表的、依托智能网络和深度学习技术进行工作的机器人,如酒店迎宾机器人;后者指存在于虚拟网络空间的非实体机器人,如网络聊天机器人。相比于非实体机器人,实体机器人存在的历史更长。二者目前都搭载着智能系统,并在不同环境中为人类服务,本文将主要研究实体机器人。

首先,当前的服务机器人具有自主性和"智能"性。[①] 它不是简单的机器工具,而是具有一定服务功能的"员工",可以根据设定的程序或者人工指令自主执行任务,并且能够通过人机交互设备与人类进行交流。例如,在咖啡厅,消费者可以通过语音告知机器人自己需要的咖啡类型、温度和甜度等,机器人会根据外部的语音指令精准地进行咖啡冲泡(创氪网,2022)。服务机器人可进行大规模协作和连续性进化,相比于人类的信息传递,机器人的迭代效率毫无疑问更优。这使得服务机器人能够在没有人类监督的情况下工作,极大地促进了服务效率的提升,并降低了服务成本。

其次,服务机器人具有耐久性和可靠性。它们可以适应低温、低氧和低压等人类无法适应的工作场景,在恶劣的环境条件下进行长时间不间断的工作。例如,很多高楼作业环境中,擦窗机器人已经逐渐取代人工进行高空作业。这使得服务机器人成为一种可靠的选择,在一些危险或者较为苛刻的工作环境中更是如此。

最后,服务机器人还具有较高的灵活性和适应性。它们可以通过调整程序或增加新的传感设备,在短时间内胜任不同的服务工作,免去了人类学习服务技能的时间成本,并运用于多种服务场景。尤其是在个性化的服务场景中,通过识别和回应人类的情绪与行为,服务机器人可以更好地理解和满足人类用户的需

① 借鉴中国科学院丁汉院士的观点,"智能"指的是目前服务机器人的"半自动状态",而非全智能状态(白杨,2022)。

求和偏好。它可以通过学习不断适应用户的偏好,从而以同理心回应用户的请求,这在医疗保健或老人护理等领域有很大的应用市场。

总结而言,人类社会的不断发展造成人工成本的不断攀升和机器成本的不断下降,服务机器人具有较高的成本效益。在一定的使用周期内,服务机器人的成本效益要远高于人力,这使得未来社会不可避免地要接受和使用服务机器人。

(二)服务机器人的发展历史

20 世纪后半叶,在技术进步和各行业对自动化需求增加的背景下,机器人的开发和部署开始加速。最早的服务机器人是 20 世纪 50 年代戴沃尔(Devol)和恩格尔伯格(Engelberger)开发的尤尼梅特(Unimate),该款机器人仅是一个用于流水线生产的机器人手臂,它通过提高效率和减少对人类劳动的需求,改变了传统制造业的运作和生产模式。在随后的几十年里,服务机器人继续发展,其部署继续增长,因为技术的进步使得创造更复杂和更有能力的机器人成为可能。事实上,各国也在第二次世界大战结束之后专门发展自己国家的机器人工业,在此仅以日本和美国为例。① 在 20 世纪 60 年代的日本,汽车制造业以及随之而来的工业机器人的崛起,在助力创造日本经济奇迹方面发挥了不可或缺的作用。在当时,日本国民收入和购买力的增长导致个人汽车需求的增长,城市化的速度仍然追不上劳动力的需求缺口,强大的终身雇佣制弱化了员工对机器人的工作威胁感知,这些都催生了日本机器人工业的迅速发展。早稻田大学成为人形服务机器人研究领域的先驱,其研究人员在 1972 年创造了第一个完全人形的机器人 Wabot-1。

在同一时间段,美国也开始大力发展机器人产业。1972 年,美国国家科学发展基金会(NSF)第一次资助一个从包装箱中取出零件并进行视觉识别的机械臂机器人项目,该种机器人可以对目标进行正确重定向,方便在装配线上进行下一步操作。20 世纪 80 年代,美国开始探索机器人的类型和使用方式的多样化,俄亥俄州立大学的研究者开发了全球首个自适应汽车悬挂机器人(ASV),该款机器人使用前置雷达系统扫描前方地形,被用于在崎岖、多山、冰冷或泥泞的地

① 选择这两个国家的主要原因在于,美国和日本是世界上机器人工业较为发达的国家,具有一定代表性,资料可获性较高。对这两个国家服务机器人发展历史的总结,可为中国服务机器人的发展提供经验借鉴。

理环境中运输工业和军事用途货物。

需要指出的是,服务机器人最重要的发展时期是 20 世纪 80 年代。这一时期,出现了专供家庭使用的机器人。虽然没有和人类交互的功能,但这些机器人被设计用来协助人们完成吸尘、修剪草坪等任务,甚至提供陪伴服务。在具体国家方面,伴随着广岛协议的签订,日本经济增长放缓,却意外加速了日本机器人工业的独立。在这一时期,日本抓住了液压机器人朝着电动机器人转变的技术空窗期,加速发展微处理器的高精度技术,这一迅猛发展势头甚至一直保持到“失去的 20 年”之前。根据日本的统计数据,1992 年开始,日本的机器人产量迅速下降,并保持了长达两年的停滞期。然而,尽管面临着日本经济泡沫破灭带来的挑战,20 世纪 90 年代日本仍占据全球 90%的机器人市场。如今,日本在机器人领域是名副其实的超级大国。根据公开数据,2020 年日本在全球机器人制造业中的份额为 47%,工业机器人的装机量约为 386.5 万台,运营库存进一步增加到 37 304 万台。

在同一阶段,苏联的解体也促使美国将精力转投到科技、经济发展上来。20 世纪 90 年代,全向智能导航器(ODIN)、机器人哨兵(Sentry)、护士机器人(Nursebot)、“有袋动物”机器人系统、机器蜜蜂(RoboBee)先后诞生。2011 年 6 月,美国政府启动了国家机器人计划(NRI),以开发可与人类一起或在人类旁边工作的机器人,并用来扩展或增强人类的能力。这促使了专业的机器人商业研究机构和专利代理机构的涌现,世界最为著名的机器人公司波士顿动力最开始就是美国国防部机器人商业化计划的一个分支。如今,根据国际机器人联合会的数据,2019 年美国的机器人销售额达到了 66.5 亿美元,同比增长 9%,2020 年机器人的装机量大约为 30 800 台,出口量达到 38.4 万个。

(三)服务机器人的科技发展阶段

纵观服务机器人的历史,其发展总是随着人工智能技术的演进、硬件技术的低成本和市场需求的变化与时俱进。虽然早期的服务机器人形态与人工智能的概念相去甚远,但是也包含了给人类服务的雏形理念。从 20 世纪 50 年代首次出现开始,服务机器人便从机械式朝着智能式不断发展。总体来看,其发展大致可分为四个阶段。

1. 实验室阶段(20 世纪 50—60 年代):这一阶段以 1954 年戴沃尔提出工业

机器人的概念并申请专利为标志,随后 1959 年“机器人之父”恩格尔伯格设计出第一台用于工业生产的尤尼梅特。该机器人按照事先编写的程序进行工作,以“示教—再现”的方式执行人类的任务。在这一阶段,与机器人相关的计算机、传感器和仿真等技术不断发展。但总体来说,虽然机器人已经初具雏形,可替代重复性工作,但其缺点是只能刻板地完成程序规定的工作,缺乏环境感知能力。

2. 萌芽阶段(20 世纪 70—90 年代):以 1970 年芝加哥举办的第一届美国工业机器人研讨会和 1971 年日本机器人协会成立为标志,机器人的发展开始进入新时期。这一阶段中的机器人配备了传感器,开始具备初步数据传输和协调能力,医用服务机器人、娱乐机器人等逐步投放市场。同时,自 20 世纪 70 年代开始,以日本的发那科(FANUC)和安川电机为代表的机器人“四大家族”开始布局全球市场。

3. 发展阶段(21 世纪前 20 年):随着各国在 20 世纪 90 年代以后大力加快机器人在消费领域中的应用,以麻省理工学院布雷泽尔(Brazel)发明的首台识别和模拟情绪机器人 Kismet 为标志,机器人的感知能力被进一步增强。机器人能够从多维度获取外部信息,并可进行交互。互联网、物联网、人机交互、云计算等先进技术快速发展,服务机器人在各领域均获得了广泛应用。需要指出的是,2000 年中国成立了首家专门的机器人公司——新松机器人自动化有限公司;同时,这一阶段也相继诞生了中国第一台类人型机器人和第一台网络智能机器人。然而,这一阶段中,机器人虽然具备感知和模仿能力,但其学习仍然是单维输入的。例如,2016 年击败顶级围棋手李世石的 AlphaGo 只能用于围棋单项任务。

4. 智能阶段(21 世纪 20 年代之后):以元宇宙元年为分割标志,人类进入互联时代。大数据等软件技术的爆发式发展,柔性关节硬件技术的成本不断降低,都使得机器人不断升级发展,服务机器人从分析式转变为感受式(Huang & Rust,2018),在包括需要情感体验的更多领域大放异彩。2022 年 11 月由 OpenAI 发布的 ChatGPT 更加拉近了普通人和智能机器人的距离,机器人开始进一步渗透到消费者的日常生活中。

二、服务机器人的机遇和挑战

正如日本抓住了液压机器人朝着电动机器人转变的技术窗口,加快实现了日本机器人产业的迅速发展;中国拥有世界上最大的机器人市场,目前亦应该抓住工业机器人朝着精细化服务机器人转变的窗口实现弯道超车。2021 年 12 月 28 日,工信部牵头,十五个部门联合印发《“十四五”机器人产业发展规划》,明确提出加快家庭服务、教育娱乐等机器人的推广,展示了中国发展服务机器人的远大抱负。《新一代人工智能发展规划》也将推动服务机器人和人类产生有效沟通的自然语言处理技术列为重点突破的共性关键技术。本文认为,随着服务机器人的智能技术愈发突破,重点研发有意识的服务机器人是下一代中国实现机器人发展规划的关键。中国工程院李培根(2022)院士指出:我们即将进入以自主学习、场景自适应、知识图谱应用等为特征的机器人 4.0 时代,机器和人相互之间的沟通将增强。服务机器人的智能形态研究将是下一代技术争夺的关键。本节将展示机器人到服务机器人的重点转变现状,并对机器人给中国带来的发展和挑战进行总结。

(一)机器人到服务机器人

伴随着自学习算法、神经网络、深度学习和自然语言处理等的不断推进,机器人逐渐被用来处理原本由人类员工执行的服务。目前,这些服务机器人的应用延续了以往工业化时代对机器人的程序设计,并发展出更多可进行自我处理的决策系统。服务领域的研究对最近的服务机器人应用产生了较多关注,有几项研究旨在从不同的角度对人工智能在服务中的应用进行分类提供基本理解。根据服务机器人的智能水平,Huang 等人将具有人工智能的服务机器人分为机械型、思考型和共情型(Huang & Rust,2018)。机械型人工智能代表了现有的人工智能服务机器人运作的基本智能水平,其一般从事常规和机械性工作,例如在餐馆接单或组装产品。思考型服务机器人则从事需要基于规则、启发式或概率式决策的工作,例如批准贷款申请、操作自动驾驶汽车的人工智能算法。思考型服务机器人还可以分析人类情感线索,例如微笑和哭泣。然而,将情感线索作为人工智能外界的额外输入并不会产生共情型服务机器人——共情型服务机器人

应该能够从过去经验中全面学习,并根据上下文因素进行适应。共情型人工智能似乎也反映在机器的一般智能概念中(Bostrom,2015)。机器人的共情包含完整人类认知功能的哲学和理论框架,如烹饪、运动以及理解和表达情感(Russell & Norvig,2016)。随着技术的发展,博斯特罗姆(N. Bostrom)甚至认为,“机器超级智能”的智能爆炸将最终不可避免地发展出类似于强人工智能的“有意识”的机器人,并指出“意识”可能是一段数学映射(Bengio,2017)或理论化的计算机程序(Blum,L. & Blum,M. ,2021)。

也可以根据在物理层面执行任务的种类多寡对服务机器人进行分类。例如,普托尼(S. Puntoni)等提供了四类人工智能的经验类型,并指出随着人工智能变得越来越有能力、越来越智能,还存在未知的经验类别(Puntoni,Reczek & Giesler et al. ,2021)。罗宾森(S. Robinson)等关注服务遭遇的情境,其中任何一方都可以是人或机器,从而产生一个复杂的交互和期望矩阵(Robinson,Orsingher & Alkire et al. ,2020)。同样,基于系统涌现属性的组合理论,霍夫曼(D. Hoffman)和诺瓦克(T. Novak)定义了智能设备和人类之间的扩张性与限制性主体互动矩阵(Hoffman & Novak,2018)。

因此,大量研究集中在人工智能服务机器人的感知人性、这种感知形成的物理化机制,以及感知人性对服务结果的影响上(Huang & Rust,2018)。一般认为,拟人化程度越高的机器人,越会给用户带来更愉悦的服务体验(Sheridan,2016)。同样,研究也证明服务机器人的人性化程度与其对温暖和能力的感知正相关(Liu,Shen & Hancock,2021)。也有研究发现,通过机器互动增加的购买行为可能与个体在与类人机器人互动时的不适有关(Christoforakos,Gallucci & Surmava-Große et al. ,2021)。个体可能认为类人机器人对人类身份、控制力和工作安全构成威胁,因此对它们的反应是不安或忧虑。相比之下,麦克利(F. McLeay)等发现,感知到的机器人的人性化只会轻微影响服务结果;相反,人工智能与服务环境之间交互的复杂性是服务结果的主要决定因素(McLeay,Osburg & Yoganathan et al. ,2021)。

还有其他研究探讨了用户对与人工智能服务机器人互动的抗拒,包括主管不愿雇用机器代替人类员工,因为它们无法与消费者建立社交联系(Dixon,Hong & Wu,2021);患者不愿意使用医疗服务人工智能代理,因为机器不能满足他们独特的健康状况和医疗需求(Dino,Davidson & Dion et al. ,2022);消费者对聊天

机器人的天然厌恶感(Christoforakos,Gallucci & Surmava-Große et al. ,2021);投资者对人工智能金融顾问的不信任(Adekoya,Oliyide & Saleem et al. ,2022)。现有的实证研究一致认为,人工智能服务机器人应该与人类服务员工结合使用,以互补和增强的方式运用到各类服务场景中,以实现最佳的服务结果。

然而,机器人并不一定具有实体外表,也有可能存在于云端。现有的大部分研究都关注具有实体形态的服务机器人,例如酒店门口的迎宾机器人等。智能服务机器人提供的服务中,有很大一部分是知识型工作,并不需要物理实体参与。这类人工智能服务机器人超越了传统的认知,将人工智能置于更无形的存在中。这种更广泛的概念定义将虚拟的服务机器人包含在内,它们通过影响人们的行为模式、价值体系、信仰和互动,积极塑造人们对世界的体验和感受。

(二)中国的服务机器人发展现状

中国的服务机器人发展在近两年达到黄金期,据 Tecdat 统计,2019 年中国服务机器人相关企业注册数量为 33 222 家,2021 年则为 105 506 家,实现了三倍多的跨越式增长。事实上,早在 2015 年,习近平总书记在致世界机器人大会的贺信中就曾指出,“我们愿加强同各国科技界、产业界的合作,推动机器人科技研发和产业化进程,使机器人科技及其产品更好为推动发展、造福人民服务”(习近平,2015),突出强调了中国发展机器人产业的决心。在这一趋势下,中国服务机器人行业的竞争焦点也由单一技术竞赛转变为对产品落地和良性商业模式的追求。整体来看,中国已经基本实现了服务机器人全产业链布局。在人口老龄化加剧、技术进步及政策支持等多种因素驱动下,中国服务机器人产业将进一步发展,其市场规模有望在 2023 年达到 751. 8 亿元。

在具体地方层面,由于服务机器人产业包含的技术创新和经济势能,各地各级人民政府也在积极推动服务机器人的产业发展,并推动服务机器人产业带动的经济高质量发展。以江苏省苏州市为例,在贯彻《江苏省机器人产业发展三年行动计划(2018—2020 年)》的基础上,苏州又推出了支持本土机器人与智能制造的企业扶持政策,如《苏州市 2022 年推进数字经济时代产业创新集群建设重点工作清单》《关于苏州市推进数字经济时代产业创新集群发展的指导意见》均将打造“机器人与智能制造产业创新集群”作为政府工作的首要目标,在政策上对服务机器人相关产业进行政策配套。苏州市涌现出了以科沃斯为代表的一

批本土服务机器人领军企业。苏州市也大力扶持创业型机器人企业,例如2023年春节档热播的电影《流浪地球2》中的移动机器人、巡检机器人和辅助机器人等就来自苏州的优艾智合公司。2022年,仅苏州市吴中区一区的服务机器人与智能制造产业规模就达1 021亿元。进一步,“2022中国机器人产业发展大会”上,苏州市发布了《苏州市培育发展机器人产业创新集群2025行动计划》,提出到2025年力争苏州全市营业收入超10亿元的企业20家、超20亿元的企业10家、超50亿元的企业5家,创建机器人及数控机床技术创新策源地、高端制造集聚区和集成应用的样板城市。

与苏州市发展服务机器人谋求经济转型的目标不同,一些地方人民政府将服务机器人视为经济追赶式发展的关键。在长三角城市群中,安徽省芜湖市的人均生产总值长期处于末游位置,服务机器人的发展对其经济的跨越式发展具有重要意义。早在2014年,芜湖市就在全国率先出台了《芜湖市机器人产业集聚发展若干政策(试行)》,设立机器人产业集聚发展专项资金和机器人产业担保基金,着力打造政策支持、投融资、技术研发、检验检测、人才支撑、应用推广六大产业支撑平台,支持机器人产业集聚发展。之后,芜湖市又相继发布《芜湖市机器人及智能装备产业集聚基地发展政策规定》《芜湖市扶持产业发展政策的若干规定》等一系列政策,对市区范围内机器人相关产业给予贷款利率、租金、搬迁和固定资产等补贴。自2015年以来,芜湖市每年筹集约5亿元机器人产业发展资金保障相关产业发展。近十年来,芜湖市一直将服务机器人产业作为重点扶持和支助产业,使芜湖市的机器人产业取得了长足发展,建成全国首个国家级机器人产业发展集聚区,实现了核心零部件自主可控、机器人本体全面接近国际一流水平。目前,以埃夫特工业服务机器人为代表,芜湖市共有200多家产业链企业,其中规上企业100多家,2021年产业产值突破300亿元,同比增长30.5%。

(三)中国的机遇和挑战

服务机器人的发展也给中国各方面带来了机遇和挑战,本小节将从社会、市场和政策三方面进行论述。

第一,社会方面。中国人口老龄化程度的不断加剧,在带来旺盛健康需求的同时,也造成了劳动力短缺和人力成本的不断上涨。据国家统计局数据,2021

年65周岁及以上人口数量已经超过2亿,占总人口的14.2%。在人工智能和智能柔性关节的加持下,服务机器人逐渐登上历史舞台,在更多服务和商业领域中的运用已经成为趋势。近两年来,数字经济占国民生产总值的比重逐年提升,数字化产业升级也为服务机器人带来广阔的应用空间。新型冠状病毒肺炎疫情暴发以来,国家更加注意加快发展数字经济,促进数字经济和实体经济深度融合。在疫情影响下,无接触式服务成为消费者的首要需求,对服务机器人的需求陡然增加,机器人的服务价值被进一步挖掘和变现。例如,服务机器人在流调等基层公共服务方面发挥着重要作用;智能服务机器人借助语音识别、语义理解、自然语音合成等技术,准确理解并提取有效信息,完成人员流调,极大释放一线人力资源。

第二,市场方面。服务机器人快速迈过市场教育阶段,进入高速发展时期。据中商网统计,2019—2022年中国服务机器人相关企业数量增长296%,大量企业涌入,已经超过10万家。然而,需要指出的是,中国总体的服务机器人技术仍然较弱,无法适应迅速发展的市场需求。由于市场惯性,国产服务机器人以低端产品为主的局面仍长期存在,主要生产三轴和四轴机器人,应用于精细和高端制造业领域的六轴及以上的服务机器人占比较少。目前,国产六轴服务机器人的新装机量不足10%。此外,我国机器人自主创新能力相对较弱。中游语音技术(如科大讯飞)和图像技术(如上海天数智芯的GPU)已经有成熟的商用商品,但机器人高端操作系统和上游技术(如芯片、舵机)严重依赖国外进口。

第三,政策方面。为适应新时代的数字经济发展要求,国家出台了一系列政策用于支持服务机器人产业发展,为产业跃迁提供强大动力支撑。《“十四五”机器人产业发展规划》指明了机器人产业的发展目标和发展路径,为中国下一阶段的机器人发展指明了方向;《关于促进机器人产业健康发展的通知》《“十四五”应急救援力量建设规划》《关于促进快递业与制造业深度融合发展的意见》等为服务机器人在健康养老、救援、物流等领域落地应用奠定了政策基础。党的二十大指出,要加快发展数字经济,促进数字经济和实体经济深度融合,打造具有国际竞争力的数字产业集群;优化基础设施布局、结构、功能和系统集成,构建现代化基础设施体系。因此,进一步发展机器人在服务经济中的替代作用,布局服务机器人发展新高地,促使机器人产业和服务经济的深度融合,构建全域互动的数字经济发展新蓝图,是实现二十大建设数字中国目标的重要机遇。然而,目

前虽然出台了一系列与世界接轨的先进政策,但政策落实仍然不到位,无法形成长效的政策保障机制。

习近平总书记曾在 2014 年中国科学院第十七次院士大会和中国工程院第十二次院士大会上精准预判,“我国将成为机器人的最大市场……我们不仅要把我国机器人水平提高上去,而且要尽可能多地占领市场”(习近平,2014)。时至今日,中国已是最大的服务机器人生产国和消费国。根据中国电子协会的统计数据,中国服务机器人市场规模稳步增长,由 2016 年的 9.4 亿美元增长至 2020 年的 29.4 亿美元,复合年均增长率为 32.9%。习近平总书记九年前的预判已经成为现实,服务机器人已经像智能手机一样普及开来。然而,在现有的机器人产业布局里,虽然 2021 年中国机器人行业营收超过 1 300 亿元,但相比于美国、日本等国家,中国机器人的整体智能化发展水平仍然不高。伴随着人口增长的逐步放缓,中国未来必须在服务机器人上发力,培育具有学习能力的意识型机器人,抢占世界服务机器人发展高地。

三、服务机器人的意识哲学解释

服务机器人的发展离不开智能计算技术、人工智能技术的赋能,服务机器人的意识发展对于执行更高级和更复杂的任务具有重要意义。研究认为,具有意识的机器人可以更加敏捷地响应用户的需求,加强了人类的服务体验(Belk,2022)。但就算发展出具有意识的机器人,人类是否就一定会接受?恐怖谷理论认为,人类对非人客体的接受会存在一个奇变点,在该点之前,机器人与人的相似度增加反而会引发人类的厌恶。为了更加全面了解服务机器人,我将在这一节简要介绍意识的哲学基础,给出自己对于当前服务机器人是否具有意识的论断,并对恐怖谷理论进行阐述。

(一)意识的哲学基础

哲学观对意识的解释存在两种基本方式。有些学者认为意识是一种天赋,而另一些学者认为意识是大脑活动产生的一种涌现现象。后者对于意识的理解旨在不同的抽象层次对意识活动进行建模,从哲学和心理学角度对大脑和神经系统的电化学活动进行病理学理解。这些理论抽象地关注什么是意识,以及它

如何在一个意识体(例如人)中运作。

表征主义学派将意识还原为外部物体的心理表征,包括意向性和现象性两个范畴。该学派的基本理论是一阶表征主义,认为经验是透明的。因此,经验的现象状态是这种经验的意向性状态,即一个实体能够建构对象的直接表征和它们的相关现象状态是高度相关的。这一理论的核心,是经验与它在头脑中的表征不相分离的论点;然而,该种理论视角不能解释那些存在但没有意识的心理状态。高阶表征主义旨在通过区分有意识的和无意识的心理状态克服这个缺陷。简而言之,一个心理状态(如感知、思想)如果是有意识的,那么只有当它成为高阶表征的对象时才是有意识的。换言之,一个心理状态只有被另一个有意识的头脑中的另一个心理状态意识到时,才被认为是有意识的。这种意识状态是一个高阶表征,它针对的是包含经验表征的精神状态。这一理论将意识定义为两个心理状态相互作用的现象,涉及意识的二次表征原理。

意识的自我表征理论与上述理论相似,它们都提出意识是不同阶的心理状态之间相互作用的结果。然而,自我代表理论认为,高阶状态是整体和复杂的有意识心理状态的一部分,而非本身就是一个不同的独立状态。与表征主义定义明确的模型不同,多重草稿理论不同于心理状态之间的秩序概念;相反,它认为所有的心理活动和状态都在有意识的头脑中同时发生。随着时间的推移,对这些状态的解释和频繁的修订将创造一个“叙事重心”,形成意识体验的核心(Dennett,2005)。与多重草稿理论类似,全局工作空间理论将头脑描述为一个全局工作空间,其中无意识的心理状态和过程争夺个人注意力与焦点。被注意和聚焦的信息随后成为全局可用的信息,而非聚焦的内容则是一些模糊的意识事件(Shanahan,2010)。

意识的社会自我理论将意识解释为一种社会现象。因此,人的心理状态和过程并不重要。相反,该理论强调必须有一个活跃的社会单位参与其中,才能使意识出现。在这个互动环境中,行为者相互沟通和互动。从这个角度来看,自我意识和意识的心智是一种社会性工具,不存在于社会行为的社会矩阵之外。

(二)意识与服务机器人

就机器人是否存在意识而言,从表征主义出发,现有学者也未达成一致。胡迪·利普森(2015)认为,意识可被理解为自我想象的能力。故而,只要让机器

人可以与外部环境互动就意味着机器人拥有了意识,作为物质构成的思维也一定可以用物质去进行重现。但想象仅是意识的子集,绝不等同于意识全部。段似膺(2019)认为意识包含着“具身性”,是“感知—行动”模式的意识观,机器人掌握再多知识也无法亲身经历行动,故而机器人可能无法拥有意识(Searle,2014)。意识是一个复杂的、多方面的概念,包含多重方面和层次,如感知意识、认知意识和感受意识等(秦瑞琳、周昌乐、晁飞,2021)。虽然有学者认为机器人已经具备了一定的功能意识(韩东屏,2021;利普森,2015),但也有学者认为,基于认知主义的观点,机器人的意识是一种被计算机化并可模拟出来的产物(刘悦笛,2019)。本文认为,从表征主义的观点出发,机器人不具有表征外界环境的神经和逻辑基础。现阶段服务机器人的意识还处于计算机模拟的水平,可以称之为一种“机器意识”(秦瑞琳、周昌乐、晁飞,2021)。该种意识区别于人类的意识,是计算机程序设定的感知层次意识。例如,人类可以识别语音中音调的变化,推断并判断出他人的情绪状态,从而做出应对性反应(刘德文、高维和、闵凉宇,2022),而机器人对语音的识别还未达到该种层次或水平。然而,虽然机器意识是被模拟的,但这并不代表其对用户是无意义的。用户仍然可以通过对这种机器意识的认知而感受到机器人作为一个社会实体的存在,和其发生互动,甚至倾注感情(Leo-Liu,2023)。

本文认为,高阶的意识发展需要机器人能够提供基于情感体验的服务,而非对外部世界的简单反应(周阳乐,2022),这要求机器人能够表现出同理。同理意味着唤起自己可从另一个实体传达的符号中提取认知和情感意义,它是重建另一个实体的内部状态和视角的能力(刘德文、高维和、周春阳,2023)。盖洛普(G. Gallup)曾进行过一系列关于自我意识的镜像实验,并指出自我意识、意识和思想都是同一个潜在过程的表达(Gallup,1998)。自我意识是一种认知前兆,能够推断他人的内部状态。此外,汤普森(E. Thompson)将共情视为一种与意识共存的现象,并认为有意识的人会从事一系列共情行动。因此,他很自然地假设未来的有意识机器人应该在和人类互动的过程中“理解”人类,从而更好地提供服务(Thompson,2001)。虽然具有超强的计算能力和复杂的机器学习算法,但现有基于情感线索训练的人工智能技术本质上是无意识的。

(三)恐怖谷理论

人类接受机器人的一个重要障碍,在于人类对机器人的天然排斥。恐怖谷理论是一种用于解释人类恐惧的心理学理论,认为人们的恐惧是由内在的心理机制引起的。根据该理论,当机器人与人类的相似达到一个特定程度时,人类对它们的反应便会突然变得极其负面和反感,哪怕机器人与人类只有一点点的差别,都会显得非常显眼刺目;当机器人和人类的相似程度继续上升,相当于普通人之间的相似程度的时候,人类对它们的情感反应会再度回到正面,产生人类与人类之间的移情作用。

这对未来发展智能机器人产生了一个重要的障碍:哪怕未来人机共融技术不断发展,创造出具有真正意识的机器人,在这类机器人达到完全类似人类(这可能永远无法做到)之前,人类还是会对机器人产生抵触和抗拒。此外,机器人的存在也会造成某些行业的消失。例如,虽然没有进入完全的应用阶段,但已经有研究者开发出人工智能辅助心理咨询师来提供咨询服务(童笑雨,2021),这势必造成实际的心理咨询师的失业。

四、展望和借鉴

随着中国机器人产业的不断发展,机器人存在的技术和伦理也是未来应该克服的问题。习近平总书记曾经指出:“要加快科技安全预警监测体系建设,围绕人工智能、基因编辑、医疗诊断、自动驾驶、无人机、服务机器人等领域,加快推进相关立法工作。”(习近平,2019)这就要求中国未来将服务机器人,尤其是具有机器意识的服务机器人发展得又快又好,且在安全可控范围之内。从日本和美国的机器人产业发展历史来看,机器人产业的发展似乎是经济发展和社会进步的“副产品”。故而,借鉴日本和美国的机器人产业发展经验,对比中国目前的机器人产业发展机遇,本文认为可从政策和技术两个方面对中国机器人产业的健康发展提出针对性建议,从而为中国开发下一代更加智能的机器人提供依据。

在政策方面,机器人产业的发展往往伴随着社会的结构性变化。我国近几年在数字技术方面已经处于世界领先的状态,这也加速了第四次工业革命下服

务机器人的超越式发展。《“十四五”机器人产业发展规划》指出,到 2025 年,成为全球机器人技术创新策源地、高端制造集聚地和集成应用新高地,并在 2050 年网络空间综合实力全球领先,建成全球领先的技术体系和产业体系。本文认为,中国可在以下几个方面进一步推动机器人产业的发展:(1)实施机器人国家战略计划。中国虽然出台了种种关于发展机器人的政策文件和计划,但总体上较为分散,没有形成连贯和稳定的政策体系。机器人在服务领域中的应用正极大改变着人类的生产和生活方式,为经济社会发展注入强劲动能。故而,除了规范类和支持类政策文件外,国家可以出台更多基础类政策文件,将机器人发展写入国家底层发展规划中,全力支持机器人在服务行业的发展,形成“以政策促进技术,以技术推动经济”的良性循环。(2)构建以企业为中心的技术创新支持平台。如前文所述,机器人的发展要依托核心企业带动供应链上企业的技术创新。目前的创新集群多是企业自主形成的,缺少政府的引导和支持。故而,政府可依托优势产学研单位联合建设一批产业创新中心、工程研究中心、产业计量测试中心、质检中心、企业技术中心、标准创新基地、技术创新中心、制造业创新中心、产业知识产权运营中心等创新平台和重点地区承接产业转移平台。通过项目拉动、技术中心共建、产业联盟、国际合作等形式建立以企业为中心的协同技术创新机制。加快科技成果的产业化应用,提升成果转化率。(3)实施“有意义的人类控制”的机器人伦理守则。伦理学家费舍(J. Fischer)曾经对机器人与指导控制、自由意志、道德责任之间的关系进行过哲学阐述,对构成行动机制道德责任的基础条件进行了修正,提出“追踪—追溯”双重伦理设计原则。随着服务机器人在人机互动中被赋予越来越多的智能、情感和意识,国家需要制定基于“追踪—追溯”双重伦理设计原则的机器人控制系统,并强制安装在所有面向用户的服务机器人中,让用户掌握对机器人的最底层控制权,确保人类拥有对机器意识的引导权和掌控权。

在技术方面,发展机器人自己的“语言”是下一代服务机器人的关键,这可以让机器人产生更高层次的意识,从而更好地提供个性化服务,提升人类的服务体验。故而,本文建议,在进行服务机器人技术开发和管理时,应当:(1)建立有意识的机器人技术基础。对于具有机器意识的服务机器人来说,它们需要“理解”信息的存在,并将其转换为一种在程序内部处理的表征符号。不同于人类的神经机制,机器人对于信息的捕捉依赖于各类传感器,这些传感器将用户的请

求进行数据化。信息是有内容的,可以编译成要义表征,依赖于观察者的解释。意识的出现取决于一个物理系统能否在极短时间内整合大量信息,因而“意识是信息的整合”(韩东屏,2021)。对于服务机器人来说,内部的计算机系统也是一个能整合大量信息的物理系统,因为电脑是基于意识的物理基础对意识内容或意识形式的仿真,没有触及人类意识独特的主观感受性。如今,大多数计算机都至少可以通过某种形式的电信号或数据包进行通信。但在不同的机器人之间,这种通信需要有一个统一的“程序符号”。否则,这种信号交换并不能使机器人产生“机器意识”,只会被当成信息噪声而被遗漏。进一步,为了开发“程序符号”的共享含义,每个服务机器人都需要储存信号与对象的关联,并在内部对它们做出与其他机器人外部响应相同的反应。(2)积极完善行业统一标准。现有的机器人开发处于“野蛮生长期”,消费者层面的机器人还未形成统一的技术入口标准,造成了管理和监督的混乱。政府主导的行业协会应充分调动生产企业和上下游企业参与修订标准的积极性,鼓励骨干企业在机器人标准化技术委员会的指导下,加快研究制定产业急需的各项国家标准、行业标准和团体标准,支持机器人评价标准的研究和验证,构建和完善机器人产业标准体系,提升技术领域安全风险管理水平,加强标准应用推广。同时,积极开展相应的国际合作,提升标准化对外开放水平,推进标准互认。(3)整合智能技术,实现共享式的深度学习资料库。下一代智能服务机器人的关键在于智能系统的深度学习,在这一过程中,机器人需要大量进行“用户输入—机器人输出—用户反馈”,以不断适应智能系统面对程序外情况的精准应变能力。在多数企业中,百万甚至千万的用户群体无法满足训练模型对海量数据训练的要求,这就需要机器人企业共享语料训练环境,整合且统一市场上不同机器人的语言处理技术,让基础语料库成为各大机器人公司技术设计的基础技术架构,让机器人产生对于环境、事件、服务和人类的统一理解。(4)加大推广“人—机器人”交互的技术合作新模式。智能机器人的核心是深度学习,这不仅包括对用户的学习,也包括对人类同事行为的学习。尤其在初级智能阶段,服务机器人的不足可以由人类员工进行弥补。故而,优化和推广“人—机器人”的服务嵌入模式,扩展机器人的服务范畴,分层次开展服务机器人的应用,让一些危险、有毒、有害的工作岗位限期由机器人来替代,这也是技术管理政策的另外一个重要方向。(5)像人但不要仿人。现阶段,机器人被设计和开发出来的目的是完成任务。由于恐怖谷理论的存在,人类

可能会天然地对机器人产生厌恶和反感。在设计服务机器人时,可以在机器人中注入“机器意识”,从而让机器人像人一般提供服务。但同时,应该避免机器人完全模仿人类的行为动作,这可以提醒人类机器人的“跨物种”差异,从而让人类在现实生活中更好地接受和使用服务机器人。

参考文献

白杨,2022,《自主智能的服务机器人离我们还有多远?》,https://www.sohu.com/a/555966261_121255906。

创氪网,2022,《冬奥冲咖啡机器人引围观,猎户星空揭秘研发背后的故事》,http://biz.ifeng.com/c/8EBrCmLgO9Y。

段似膺,2019,《海德格尔式人工智能及其对“意识”问题的反思——兼与何怀宏先生商榷》,《探索与争鸣》第3期。

韩东屏,2021,《智能机器人意识问题的各种解答与我见》,《自然辩证法通讯》第4期。

李培根,2022,《我们即将进入机器人4.0时代》,https://www.sohu.com/a/581259307_121124715。

利普森,胡迪,2015,《意识!机器人会有吗?》,《中国信息化周报》3月23日。

刘德文、高维和、闵凉宇,2022,《声音特征和文本策略的说服效应研究》,《管理学报》第9期。

刘德文、高维和、周春阳,2023,《何以报“怨”?顾客抱怨对员工角色服务行为的影响机制研究》,《外国经济与管理》第2期。

刘悦笛,2019,《人工智能、情感机器与“情智悖论”》,《探索与争鸣》第6期。

秦瑞琳、周昌乐、晁飞,2021,《机器意识研究综述》,《自动化学报》第1期。

童笑雨,2021,《西湖大学研发AI辅助心理咨询机器人,免费提供咨询服务》,https://baijiahao.baidu.com/s?id=1705080231625139216&wfr=spider&for=pc。

习近平,2014,《在中国科学院第十七次院士大会、中国工程院第十二次院士大会上的讲话》,http://jhsjk.people.cn/article/25125594。

习近平,2015,《致2015世界机器人大会贺信》,http://jhsjk.people.cn/article/27845451。

习近平,2019,《提高防控能力着力防范化解重大风险,保持经济持续健康发展社会大局稳定》,http://jhsjk.people.cn/article/30582837。

周阳乐,2022,《王耀南:给冰冷的机器赋予“智慧”和“情感”》,https://baijiahao.baidu.com/s?id=1721904594341222273&wfr=spider&for=pc。

Adekoya, O., J. Oliyide & O. Saleem et al. 2022, “Asymmetric Connectedness Between Google-Based Investor Attention and the Fourth Industrial Revolution Assets: The Case of FinTech and Robotics & Artificial Intelligence Stocks.” *Technology in Society* 68.

Belk, R. 2022, "Artificial Emotions and Love and Sex Doll Service Workers." *Journal of Service Research* 25(4).

Bengio, Y. 2017, "The Consciousness Prior." Working Paper on arXiv.

Blum, L. & M. Blum 2021, "A Theory of Consciousness from A Theoretical Computer Science Perspective: Insights from the Conscious Turing Machine." Working Paper on arXiv.

Bostrom, N. 2015, *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press.

Christoforakos, L., A. Gallucci & T. Surmava-Große et al. 2021, "Can Robots Earn Our Trust the Same Way Humans Do? A Systematic Exploration of Competence, Warmth, and Anthropomorphism as Determinants of Trust Development in HRI." *Frontiers in Robotics and AI* 8.

Dennett, D. 2005, *Sweet Dreams: Philosophical Obstacles to A Science of Consciousness*, Boston: MIT Press.

Dino, M., P. Davidson & K. Dion et al. 2022, "Nursing and Human-Computer Interaction in Healthcare Robots for Older People: An Integrative Review." *International Journal of Nursing Studies Advances* 4.

Dixon, J., B. Hong & L. Wu 2021, "The Robot Revolution: Managerial and Employment Consequences for Firms." *Management Science* 67(9).

Gallup, G. 1998, "Can Animals Empathize? Yes." *Scientific American* 9(4).

Hoffman, D. & T. Novak 2018, "Consumer and Object Experience in the Internet of Things: An Assemblage Theory Approach." *Journal of Consumer Research* 44(6).

Huang, M. & R. Rust 2018, "Artificial Intelligence in Service." *Journal of Service Research* 21(2).

Leo-Liu, J. 2023, "Loving A 'Defiant' AI Companion? The Gender Performance and Ethics of Social Exchange Robots in Simulated Intimate Interactions." *Computers in Human Behavior* 141.

Liu, S., Q. Shen & J. Hancock 2021, "Can A Social Robot Be Too Warm or Too Competent? Older Chinese Adults' Perceptions of Social Robots and Vulnerabilities." *Computers in Human Behavior* 125.

McLeay, F., V. Osburg & V. Yoganathan et al. 2021, "Replaced by A Robot: Service Implications in the Age of the Machine." *Journal of Service Research* 24(1).

Puntoni, S., R. Reczek & M. Giesler et al. 2021, "Consumers and Artificial Intelligence: An Experiential Perspective." *Journal of Marketing* 85(1).

Robinson, S., C. Orsingher & L. Alkire et al. 2020, "Frontline Encounters of the AI Kind: An Evolved Service Encounter Framework." *Journal of Business Research* 116.

Russell, S. & P. Norvig 2016, *Artificial Intelligence: A Modern Approach*, New York: Pearson

Education Limited.

Searle, J. 2014, "What Your Computer Can't Know." *The New York Review of Books*, October 9.

Shanahan, M. 2010, *Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Minds*, Oxford: Oxford University Press.

Sheridan, T. 2016, "Human-Robot Interaction: Status and Challenges." *Human Factors* 58(4).

Thompson, E. 2001, "Empathy and Consciousness." *Journal of Consciousness Studies* 8(5-6).

地方公共数据治理*

——重点任务、机制创新与未来挑战

徐偲骕　李　欢　陆伊琳**

随着经济深度数字化转型和数字政府建设的加快推进，中国信息数据领域的制度建设明显提速，中央政府在顶层设计上已经为数据“站稳”新的生产要素和治理工具地位做好了充分准备。不过，尚有大量细节性的体制机制难题需要地方政府创造性地予以破解，特别是公共数据这一兼具公共与市场两种属性的治理对象。本文旨在“发挥中央和地方两个积极性”的语境下，呈现三方面的内容：其一是阐述地方政府在公共数据治理中面临的两大全新任务——开放和授权运营；其二是梳理各地已经取得的重要进展，尤其关注机制创新等方面；其三是简要讨论治理者面临的结构性挑战。以期在数据治理的现有研究中补充一个“地方视角”。

一、中国数据治理的地方视角

2019 年末至今，中国信息数据领域的制度建设明显提速，几乎每年都推出一部甚至几部重磅新规和大量条例、规章、实施细则，乃至国家和行业标准文件。从 2019 年党的十九届四中全会《中共中央关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》（以下简称《决定》），将数据纳入生产要素，到 2022 年末中央正式对外发布酝酿已久的纲领性文件《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》（以下简称“数据二十条”），并以 2023 年国务院一号文件的形式为全年工作

* 本文系国家社会科学基金优秀博士论文出版项目“生产资料、治理工具与要素市场：中国大数据产业的传播政治经济学研究”（项目批准号：22FYB052）的阶段性研究成果。

** 徐偲骕，上海大学新闻传播学院、上海大学中国当代文化研究中心；李欢，华东政法大学传播学院；陆伊琳，上海大学文学院。

开篇;在这四年间,《个人信息保护法》《数据安全法》《“十四五”大数据产业发展规划》《关键信息基础设施安全保护条例》等各类建议性、制度性、规范性文件陆续出台。制度供给丰富、立体,既积极回应当下热点数据问题,又努力与新技术发展速度赛跑,兼具现实意义和前瞻性。可以说,四年来,作为供给侧结构性改革重要一环的中国数据政策,已经形成了完整、清晰的顶层设计思路。中央政府在政治、经济和社会等多个方面已经为数据“站稳”新的生产要素和治理工具地位做好了充分准备(徐偲骕,2022),中国式现代化数据治理体系初见雏形。

但是,中国任何施政活动都不是简单的自上而下,而是需要充分调动中央和地方两个积极性,上下一盘棋。尤其是数据政策的顶层设计往往是原则性的、底线性的,类似《数据安全法》等新法主要起到画红线、定规矩的作用,目的是确立基本的是非边界和负面清单。而在具体贯彻落实层面,各地实际情况差异极大,数据政策的实践性规则需要根据本地区公共治理和经济发展特点来开展探索和试错,无法直接适用一套固定不变的规则。因此,中央鼓励各地方创造性先行先试,研究出台地方性数据条例,探索建设数据交易场所、数据集团等。简单来说,中央负责把握大方向、搭建顶层政策框架,至于如何盘活和利用好各级政府手中的公共数据来改善国家治理能力(政治),如何培育规范而有活力的数据要素市场、解决数据流通领域的各类技术性和利益难题(经济),如何防范数据技术侵害用户和提升数据安全水平(社会),都是需要地方主政者结合本地实际予以考虑的全新命题。

换言之,在数据这个牵一发而动全身的新兴事物身上,尚有大量有待破解的经济、技术难题,需要地方发挥巨大的能动性,在中央给定的活动空间和范围之内做精细的文章。既要提供创新的制度供给,又要破除自身已经存在多年的同数据治理特点不相适应的体制机制,比如传统的部门壁垒等,甚至在规则制定和要素市场地位方面还要展开激烈的省际、地区间竞争。现有研究较少关注地方政府在数据治理活动中扮演的角色,这一任务不仅是全新的,且相当具有挑战性,难度非常高,考验的是综合性治理水平和能力,是国家治理体系和治理能力现代化要经历的“大考”。

篇幅所限,本文要重点讨论的是地方公共(政务)数据[①]的治理。法律上,

① 为行文方便,本文暂不区分“公共数据”和“政务数据”等概念。本文所言的“公共数据”是指履行公共管理和服务职责的公共服务运营单位依照法律规定,在履行职能过程中采集、制作、生产或者获取,并通过一定形式记录、保存的数据资源。

《数据安全法》已经明确规定了各地方在数据治理上的基本职责和任务,如第六条规定“各地区、各部门对本地区、本部门工作中收集和产生的数据及数据安全负责”;第二十一条要求“各地区、各部门应当按照数据分类分级保护制度,确定本地区、本部门以及相关行业、领域的重要数据具体目录,对列入目录的数据进行重点保护”。不过,数据治理的概念既强调数据安全和个人信息保护,当然也注重释放价值(中国信息通信研究院,2023),后者是本文所言“治理”的重点内容,甚至具有终极目的性。也即,数据治理不是为了把数据“管死”,而恰恰是要把数据“管活”,因此作为基本治理内容之一的数据安全和保护其实也是为了更好地促进后续的自由流动(徐偲骕、姚建华,2019),以服务和赋能方兴未艾的中国数字经济。公共数据亦不例外,在安全的基本底线之上,如何盘活、用好公共数据资源,是目前地方政府公共数据治理需要先行探索和试点的治理内容。

二、一个目标,两大重任

国家发展和改革委员会(2023)解读“数据二十条”时指出,要“推动公共数据确权授权,加大公共数据开放共享和开发利用,发挥公共数据在数据要素开发利用中的基础性、引领性、示范性作用”。地方数据治理首先面对的对象,便是自身生成的数据。各级政府在常年履行公共管理和服务职责过程中收集、产生和沉淀了大量的数据,其中很多都是非电子化、非标准化、非结构化,甚至算不上是数据的人工记录信息。在国家大力建设数字政府和推动城市数字化转型的语境下,公共管理活动已经超越了过去简单的手工行政、表格行政的范畴,基层开始与层出不穷的电子政务系统乃至各类应用打交道。传统政务信息先要被数据化,而后再数字化,这是后续实现数字治理或智能治理的最基本前提。另外,随着数字技术和智能硬件引入各类行政治理场景,“数治活动”本身每时每刻都在喷发出新的、呈指数级增长的海量数据,这个过程甚至已经高度自动化。这两大类数据构成了地方政府手中一类新的高价值资源——公共数据。按照一些学者的看法,政府数据治理是一种由内而外的政务流程再造与组织重构,不仅是一个技术采纳和提升的问题,更是一个政治问题。一个好的数字政府,首先是一个能够在公共数据上实现“敏捷治理”的政府(于文轩、魏炜,2022)。

仅以《上海市数据条例》之规定为例,上海市政府在公共数据治理方面的具

体工作十分繁重,共三节二十二条内容(第二十五至四十六条),包含健全公共数据资源体系、加强公共数据治理、提高公共数据共享效率、扩大公共数据有序开放、构建统一协调的公共数据运营机制、推进公共数据和其他数据融合应用、充分发挥公共数据在推动城市数字化转型和促进经济社会发展中的驱动作用等。涵盖收集、归集、治理、共享、开放、应用及其相关质量和安全管理等多个方面。按照王伟玲(2022)的分类,政务数据治理的任务可以分为"数据开放""数据共享"和"数据运营"三大类。其中,数据共享更多指的是政府(机构)间数据的融合共通,强调政务数据跨部门流通,实际上是如何打通政府数据"内循环"的问题。现有大量关于数据孤岛和数据壁垒的分析多是针对于此。共享的主要目的是推动协同高效的社会治理,提供便捷和个性化的公共服务,提升政府决策的科学化和全局化水平。在学理上,数据共享要解决的,更多的是一个相对较老的问题——条块分割、块块割据等科层制体系内部的问题——有待职责聚合改革和条块协同优化来破解(邱实,2022)。

而"数据开放"和"数据运营"[①]则是在数据要素化大背景下的全新问题,是政府为推动数据资源的社会化利用、调动公民和企业参与经济社会数字化转型的积极性而进行的改革,开放和授权运营的对象是不特定的个人或组织,特别是各类数字平台企业、互联网科技公司等。随着数据要素化进程的开启,它能够直接投入信息生产,发挥要素替代与节约、要素融合与价值共创的功能,数据协同创造财富与价值的角色日益获得重视。公共数据依托行政管理与公共服务职能,来源广、体量大、门类多、增速快,是数据开发和利用中的"富矿"(刘阳阳,2022)。它不仅是一类新型国有资产,也是全社会现存数据要素集合中的高价值子集,作为关键持有者面向非政府主体供给高价值公共数据,不仅是透明政府的必然要求,也是各级政府促进数据要素市场的发育、激励围绕数据的社会创新的基本公共责任。非政府机构和市场主体对公共数据资源进行创造性开发利用,将在多个方面创造出新的价值,比如激发公益与公共服务创新(市政环境领域)、助力科学研究(医疗卫生领域)、推动公共信息服务创新(交通出行领域)、

① 二者的最直接区别就是,"数据开放"是指直接提供原始数据给有能力自行利用的市场主体或个人,而"数据运营"是指对于不便开放或有条件开放的敏感数据,需要政府授权委托特定市场主体进行加工和处理,确保合规和安全,而后提供基于原始数据的产品和服务,因此也叫"授权经营"。下文会对为什么有了数据开放还要有数据授权运营进行进一步解释。

衍生新型商业化活动(金融风控领域),等等(胡业飞,2022)。

仅以数据需求最为旺盛的金融业为例,根据综合测算,在银行等机构开展个人和企业征信业务所依赖的数据产品中,90%的数据产品来源于对公共数据的处理加工。公共数据赋能金融业数字化转型,不仅可以保证交易的合规性,更有利于金融业公共数据价值的发现,充分发挥数据对金融业发展的放大、倍增、叠加效应(上海数据交易所研究院,2022)。这里面就包括从电力部门获取企业的用电信息,用电量和电费高低直接反映生产型企业的经营状况,和贷款安全充分相关。同样的道理,医药企业和商业保险公司希望从卫健部门获取脱敏之后的医疗数据,用以新药研发、疗程设定乃至保费标准设定;数字地图开发者希望从交通部门获取路况信息,从气象部门获得天气数据,从文旅部门获得景区人流量等实时数据……公共数据所蕴藏的价值潜力还有待充分释放。

信息数据"深藏闺中"是极大浪费。行政机构本身的专长并非创新活动,"提取政府数据价值最好的办法是允许私营部门和社会大众访问"(迈尔-舍恩伯格、库克耶,2013:149)。海量数据拿在政府手上用途毕竟有限,数据在政府内部无偿交换和流转,赋能的也主要是行政治理活动,经济效应较弱;而当各种类型的营利和非营利机构都跻身公共数据资源的开发利用主体时,将释放出巨大的经济社会价值。反过来,盘活公共数据资源并予以合理的社会化开发和利用,也将进一步激发市场主体活力,激励中小科技企业参与数据创新活动,解决他们"巧妇难为无米之炊"(空有技术却无数据)的困境,缩小他们与大型垄断性互联网公司之间在数据要素方面的起跑劣势,撬动现有平台经济的竞争格局。这是地方公共数据治理的根本意图和最终目标。

因此,我们主要聚焦"数据开放"和"数据运营"这两大全新的地方数据治理任务。

三、各地有条不紊推进公共数据开放

我国公共数据开放工作肇始于2004年前后,《中共中央办公厅 国务院办公厅关于加强信息资源开发利用工作的若干意见》要求对具有经济和社会价值、允许加工利用的政务信息资源,应鼓励社会力量进行增值开发利用。起初主要是由中央部委推动相关工作开展,为企业和社会公众提供公益性信息服务。随

着经济和产业数字化转型对公共数据的需求日增，全国范围内尚无统一的政府数据资源平台和数据要素统一大市场，因而地方政府手中的公共数据就日益显现出其价值。实际上，仅靠中央和国务院下属部门的政务数据（本身也要依靠地方汇总），很难完全满足尚处分割状态的各地数字经济市场的本地化动态需求。而且，加强公共数据开放共享已被列入国家“十四五”规划所确定的重点任务之一，一系列顶层制度设计又陆续完成，探索试点公共数据开放体制机制和具体做法的重任，就落到各地政府的肩头。反过来，地方实践取得的成果和一些基础性工作，也能为后续构建全国一体化政务大数据体系打好基础。

（一）制度供给

目前，各地已经逐渐形成以地方人大制定的数据条例为核心，以省级地方政府常务会议审议通过的规章或行政指导性规范为配套的公共数据治理体系。多数省份（包括一些省会城市和副省级城市，乃至上海自贸区临港新片区）都已经颁布专门的公共数据管理办法。比较特殊的有浙江、山西等省，浙江没有制定单独的数据条例，但却通过地方人大制定了《浙江省公共数据条例》，这是全国首部经过地方人大批准通过的公共数据领域地方性法规。其规格之高，反映了浙江省政策制定者可能并不认为公共数据的治理任务属于数据政策框架下的子领域，而是和一般性数据治理工作同等重要，甚至其重要性和优先性还要更高，希望以管好、用活公共数据来带动和促进数据领域的其他工作。山西省虽然有经过地方人大立法程序推出的《山西省大数据发展应用促进条例》，但《山西省政务数据管理与应用办法》同样也是由山西省人大常委会制定并通过的，而非山西省政府规章，这在全国也属少见。

而有关“公共数据开放”的规定，一般是内嵌在上述各地公共/政务数据管理办法篇目之中的，属于后者重要的章节之一。只有少数省份的管理办法，目前还止步于公共数据的政府内部共享，暂未触及向社会开放的问题。

向前一步，已经有不少地方把“公共数据开放”提到了与公共数据整体管理相同的高度，相当于把这个单章提取出来，再专门制定一部独立的、更为细致的公共数据开放管理办法。譬如较早发展大数据产业的贵州省，不仅在 2016 年就出台了《贵州省大数据发展应用促进条例》《贵州省政务数据资源管理暂行办法》，在开放领域还专门颁布了《贵州省政府数据共享开放条例》（以下简称《条

例》)。不同于其他地方有关公共数据开放的办法通常只是政府规章,该《条例》乃是经贵州省人大常委会表决通过的地方性法规,在全国也是仅此一例,可见贵州对于公共数据的开放工作重视程度之高。

再比如,2019年8月,上海市政府就根据2018年9月出台的《上海市公共数据和一网通办管理办法》,研究制定了专门的《上海市公共数据开放暂行办法》,共48条;2019年11月,上海市经信委印发《上海市公共数据开放分级分类指南(试行)》;2021年11月,《上海市数据条例》获上海市人大通过。2022年8月,上海市经信委又印发了《2022年上海市公共数据开放重点工作安排》,其中提出四项任务共19项具体工作,包括优化公共数据开放机制、提升公共数据开放质量、加快开放平台功能迭代、创新公共数据应用场景。其中,优化公共数据开放机制下的第一条任务"制定公共数据开放实施细则"在一个月后就有了进展:9月13日,市经信委起草完成了《上海市公共数据开放实施细则(征求意见稿)》,共七章38条,并对外发布,公开征询公众意见。可以说,上海市的公共数据治理及开放制度建设一步一个脚印,从无到有,几乎每年都有新的文件诞生,整体朝着更细致和完善的方向加速推进。

(二)平台建设

最能反映制度建设成果的,是各地纷纷上线的开放平台。复旦大学与国家信息中心数字中国研究院联合发布的《2022中国地方政府数据开放报告(省域)》(复旦大学数字与移动治理实验室,2023)显示:截至2022年10月,我国已有208个省级和城市的地方政府上线了政府数据开放平台,其中省级平台21个(含省和自治区,不包括直辖市和港澳台地区),城市平台187个(含直辖市、副省级与地级行政区)。与2021年下半年相比,又新增了15个地方平台,其中包含1个省级平台和14个城市平台,平台总数增长约8%。特别是在省一级层面,我国已有74.07%的省级(不含直辖市)政府上线了政府数据开放平台,自2015年浙江省上线我国第一个省级(不含直辖市)平台以来,省级平台数量逐年增长。

表1提供的是省级平台(含自治区、直辖市和港澳台地区)的清单。

表 1　全国省级行政区公共数据开放平台统计清单(截至 2023 年 2 月)

地区	省份	平台名称	网址	规模
华北地区	北京市	北京市公共数据开放平台	https://data.beijing.gov.cn/	收录单位数 115 个、数据集 15 385 个、数据项 577 743 个、数据接口 12 527 个、数据量 71.86 亿条
	天津市	天津市信息资源统一开放平台	https://data.tj.gov.cn/	开放 21 个主题、61 个市级部门、16 个区、2 733 个数据集、871 个数据接口
	河北省	河北省公共数据开放网	http://hebdata.hebyun.gov.cn	开放部门 52 家、目录量 803 个、资源量 746 个
	山西省	暂无省级统一的公共数据开放平台,以大同市公共数据开放平台为例	http://www.dt.gov.cn/Dataopen/index	收录数据开放部门 19 个、数据项 2 208 个、目录数 256 个、数据下载量超 41 万次
	内蒙古自治区	暂无省级统一的公共数据开放平台	暂无	据公开资料,2019 年 12 月,内蒙古自治区大数据发展管理局公示了内蒙古公共数据开放平台建设项目中标(成交)公告。该项目由东华软件股份公司中标,中标价为 280 万元
东北地区	黑龙江省	暂无省级统一的公共数据开放平台,以哈尔滨市公共数据开放平台为例	http://data.harbin.gov.cn/oportal/index	开放 51 个部门、2 737 个数据目录、1 003 万条数据、2 774 条数据接口、20 个创新应用,数据下载量超 49 万次
	吉林省	吉林省数据共享平台	暂无	据报道,共享平台已接入 2 366 个政务服务部门,汇集各类数据信息 60 多亿条
	辽宁省	辽宁省公共数据开放平台	http://data.ln.gov.cn	开放数据的部门 25 个,开放数据接口 46 个、数据目录 209 个、数据 2 028 万条

表 1(续 1)

地区	省份	平台名称	网址	规模
华中地区	河南省	暂无省级统一的公共数据开放平台	暂无	暂无
	湖北省	暂无省级统一的公共数据开放平台,以武汉市公共数据开放平台为例	http://data. wuhan. gov. cn/index. html	收录市级部门 61 个、1 413 个数据目录、1 361 个数据集、1 032 个数据接口、11 682 万个数据项、近 3 亿条数据
	湖南省	湖南政务大数据公众门户	https://data. hunan. gov. cn	开放 483 个数据资源、21 个 API 接口
华东地区	山东省	山东公共数据开放网	https://data. sd. gov. cn	收录开放部门 55 个、数据目录 103 852 个,开放数据 67. 9 亿条、2. 2 万个数据接口、110 个创新应用
	江苏省	江苏公共数据开放网站	http://data. jiangsu. gov. cn/extranet/openportal/pages/statistic/statistic. html	数据总量 16 975. 2 万条,资源总量 891 个,API 服务总量 11 个
	上海市	上海市公共数据开放平台	https://data. sh. gov. cn	收录开放部门 50 个、131 个数据开放机构、5 357 个数据集、2 245 个数据接口、73 个数据应用、4. 4 万个数据项、超 20 亿条数据
	浙江省	浙江省数据开放平台	http://data. zjzwfw. gov. cn/jdop_front/index. do	开放 26 279 个数据集、12 852 个 API 接口、144 562 个数据项、114. 3 亿条数据
	安徽省	安徽省数据开放平台	http://data. ahzwfw. gov. cn:8000/dataopen-web/index. html	开放资源 556 个、1. 04 万条数据、556 个 API 接口
	江西省	江西省政府数据开放平台	https://data. jiangxi. gov. cn	开放数据目录 290 个、接口目录 120 个,数据项总数 76 项,数据总量约 1. 3 亿条,API 接口 8 个

表 1(续 2)

地区	省份	平台名称	网址	规模
华东地区	福建省	福建省公共数据资源统一开放平台	https://data.fujian.gov.cn	收录 38 个部门开放数据、5 984 个数据目录、25 亿条数据、6 533 个数据接口
	台湾省	台湾省公共数据开放平台(政府资料开放平台)	http://data.gov.tw	参见网站首页全部资料列表
华南地区	广东省	开放广东	https://gddata.gd.gov.cn	收录省级部门 55 个、地市 21 个、数据目录 54 571 个、数据集 54 272 个、政府数据 10.8 亿条、数据应用 102 个、API 接口 299 个
	广西壮族自治区	广西公共数据开放平台	http://data.gxzf.gov.cn	收录数据目录 8 840 个、数据总量 26.92 亿条、部门 90 个
	海南省	海南省政府数据统一开放平台	http://data.hainan.gov.cn	收录数据目录 2 660 个、主题 21 个、数据集 1 363 个、数据部门 80 个、数据接口 2 247 个
	香港	资料一线通	https://data.gov.hk/sc	收录超过 5 130 个数据集,有约 2 200 个应用程序界面,下载次数超过 490 亿次
	澳门	澳门特别行政区政府统计暨普查局	https://www.dsec.gov.mo/zh-MO	参见网站首页全部资料列表
西北地区	宁夏回族自治区	宁夏公共数据开放平台	http://opendata.nx.gov.cn/portal/index	开放 28 个部门、2 004 个数据目录、4 231 万条数据、6 条数据接口、7 个创新应用

表 1(续 3)

地区	省份	平台名称	网址	规模
西北地区	新疆维吾尔自治区	暂无自治区级统一的公共数据开放平台,以乌鲁木齐市政务数据开放平台为例	http://zwgl.wlmq.gov.cn	归集 25 个部门数据、12 个主体分类、99 个数据集、1.08 万条数据
	青海省	暂无省级统一的公共数据开放平台	暂无	暂无
	陕西省	陕西省公共数据开放平台	http://www.sndata.gov.cn	开放 141 类数据,总下载量 1 201 次。上一次数据更新时间为 2021 年 2 月
	甘肃省	暂无省级统一的公共数据开放平台,另有文献显示,甘肃省境内共有 14 个公共数据开放平台	暂无	暂无
西南地区	四川省	四川公共数据开放网	http://www.scdata.net.cn/oportal/index	收录 48 个部门、21 个市州、33 696 个数据目录、216 720 个数据资源、30.24 亿条数据、12 个应用
	云南省	数据云南	http://www.qjtj.com/member/login/index.html	需要登录后才能查看
	贵州省	贵州省政务数据开放平台	http://data.guizhou.gov.cn/home	收录 16 973 个开放数据集,开放文件 11 794 个、部门 3 255 个、市州 9 个

表 1(续 4)

地区	省份	平台名称	网址	规模
西南地区	西藏自治区	暂无自治区级统一的公共数据开放平台	暂无	2022 年 10 月 29 日,西藏自治区大数据中心挂牌运行,完成全区各级各部门 491 个业务系统上云工作,对接数据接口 25 个,归集数据 4.33 亿条,并聚焦政务服务、经济发展、社会治理、民生服务等 10 大应用专题近 40 个应用场景
	重庆市	重庆市公共数据开放系统	https://data.cq.gov.cn/rop/home	开放 60 个部门、41 个区县、20 510 个数据集(其中 5 069 个数据接口)、63 769 个数据项、15 个数据应用

资料来源:数据交易网(2022a)。作者已根据 2023 年 2 月各地平台披露的最新数据,对“规模”部分予以更新。

另据复旦大学数字与移动治理实验室(2023)自主开发的“中国开放数林指数”,在不含直辖市和港澳台地区的省级地方中,山东省和浙江省的综合表现最优,贵州省总体表现优秀,其次是四川省、广西省、广东省、福建省、海南省、江苏省、辽宁省等地。在几项关键指标上,山东省在利用层(利用促进、利用多样性、成果数量、成果质量、成果价值)排名第一,浙江省在准备度(法规政策、标准规范、组织推进)和数据层(数据数量、开放范围、关键数据集质量、关键数据集规范、关键数据集安全保护)排名第一,贵州省则在平台层(平台体系、开放协议、发现预览、关键数据集获取、社会数据及利用成果提交展示、使用体验、互动反馈)排名第一。

值得注意的是,本部分主要讨论省一级的公共数据开放工作,但各省之内的地市级政府也已经上线了自己的公共数据开放平台,目前,广东省、广西壮族自治区、江苏省、江西省、山东省、四川省与浙江省的省(自治区)本级和下辖所有地市(州)都已上线了数据开放平台。其他地方情况各异,还存在下列情况:(1)省(自治区)本级平台已上线,但省内多数地市(州)尚未上线平台(陕西和青海省域内的所有地市[州]尚未上线数据开放平台);(2)下辖地市(州)上

线了数据开放平台,但省(自治区)本级平台尚未上线;(3)省(自治区)本级和下辖绝大多数地市(州)均未上线数据开放平台。不过,随着社会对公共数据的需求日益增长,地方政府对数据开放工作的重视程度和推进力度不断提升,上线数据开放平台已经成了带有竞争性的积极作为。增加开放的数据数量、提高开放的数据质量等也已经成为数字政府建设的重要基础性任务。

(三)后续工作

实际上,上线统一的开放平台只是公共数据开放工作的第一步,背后是大量看不见的隐性工作,涉及开放前、开放后的诸多环节以及日常管理,绝非易事。参看各地公共数据管理办法,在准备工作方面,大体上都包括建设基础设施(比如电子政务云、大数据资源平台等)、进行数据管理(比如统一采集、归集、整合,实行目录管理、分级分类、质量管理,建立更新机制等)、数据开放(比如制定开放清单、进行脱敏处理、受理和审核开放申请、开展授权运营和鼓励应用创新等)、安全保障(比如进行风险评估,预防攻击、侵入、干扰,建立安全审查机制、预警机制和应急管理制度等,以确保数据安全及利用合法正当)等方面。

开放平台开始运行后,仅以上海市为例,从《2022年上海市公共数据开放重点工作安排》和《上海市公共数据开放实施细则》这两份文件来看,近期上海市政府相关部门还要重点推进以下几类工作:其一,创新开放机制,包含开展需求征集、制订开放年度计划、建设示范项目、明确开放条件、数据质量提升等;其二,在细化获取流程方面,除了区分无条件和有条件开放数据之外,还要处理申请和签订协议、建设样本数据库;其三,在鼓励数据利用方面,还要进行成果展示与合作应用、建立数据利用反馈机制、健全公共服务体系等。

以上,还只是开放平台上线之后的部分工作,在推进平台技术升级、功能迭代和资源扩展,制定并公布开放平台管理制度,开放清单动态调整,开放范围扩大,利用监管与权益保护,违法违规行为处理,安全保障、资金保障与考核评估等方面还有大量的细致工作要跟进。可见,公共数据开放是一项既具有全局性和系统性,又考验精细度和灵活性的长期动态工程。目前上海市运作比较成熟和规范的应用场景要数2019年试点上线的大数据普惠金融应用,该项目重点关注“中小企业贷款难问题”,以数字普惠为抓手,通过公共数据共享和开放——由上海市大数据中心归集、清洗、融合、共享各委办局提供的原始公共数据,加工后

使其达到能够支持商业银行业务要求的质量水准和合规水平，然后通过开放网关提供给商业调用，或通过数据开放门户共享给需要的银行。首批四家试点银行已与上海市大数据中心签署《上海市公共数据开放普惠金融应用数据利用协议》，基于公共数据开发专项产品（上海数据交易所有限公司、普华永道中国，2022）。

四、各地探索公共数据授权运营“蹄疾而步稳”

2021 年以来，中央密集提到公共数据授权运营，如“十四五”规划明确提出“开展政府数据授权运营试点，鼓励第三方深化对公共数据的挖掘和利用”，此后《“十四五”数字经济发展规划》《要素市场化配置综合改革试点总体方案》等文件均予以了强调。近期发布的“数据二十条”明确指出，要“探索用于产业发展、行业发展的公共数据有条件有偿使用”。在各地陆续出台的数据条例和公共数据管理办法的条文中，“探索建立公共数据授权运营机制”也是紧接着“公共数据开放”之后的重要内容。如《浙江省公共数据条例》第三十五条规定：“县级以上人民政府可以授权符合规定安全条件的法人或者非法人组织运营公共数据，并与授权运营单位签订授权运营协议。”在制度供给上，从中央到地方基本已经明确了主要的发展方向和政策框架。

（一）为什么要进行授权运营？

从定义上看，“公共数据运营是指按照国家相关法律法规要求，经公共数据管理部门和其他相关信息主体授权具有专业化运营能力的机构，在构建安全可控开发环境基础上，按照一定规则组织产业链上下游相关机构围绕公共数据进行加工处理、价值挖掘等运营活动，产生数据产品和服务的相关行为”（中国软件评测中心，2022）。其核心目标是盘活公共数据资源、挖掘数据价值、释放数据红利，要求“公平公正、市场配置”“安全可控、依法利用”“挖掘价值、服务人民”。不过，既然有了开放，为何还需要有授权运营的环节？二者是什么关系？

实际上，政府数据开放的目的是促进社会化利用和鼓励创新，但是供和需、数据和场景、开放和利用之间的对接不是凭空发生的，以往文件形式的信息公开

往往既不利于寻找,也无法有效转化成技术企业能够利用的资源和产品。而且,在数据经济时代,公共机构是否能够提供稳定、持续的数据服务/产品本身也是个问题。一方面,这本身就不是政府所擅长的;另一方面,受到公共机构运作体制机制的限制,其运作无法完全达到商业化服务标准,况且巨大的运营成本需要长久强大的经济支撑。因此,需要引入第三方市场化力量,先对原始数据做技术处理(包括但不限于数据的清洗、补完、对齐、脱敏、标准化等),使其处于一个可供开发使用的状态;补足数据产品的开发缺口,更高效地转化原始数据成为可被重复利用的数据产品、服务,加快公共数据的利用率;还可帮助公共机构一方面推广营销数据产品或服务,另一方面评估和准入数据申请主体,从而极大化地使公共数据实现真正的大范围社会化流通。所以,授权运营是对现有公共机构供给数据产品/服务的传统模式外的一种"市场化能力补充"(陆志鹏,2022)和"规模化流通"的手段(开放数据中国,2021)。

此外,政府开放公共数据虽然是一种公共服务,但并不是零成本的工作,存储、传输需要的硬件支撑和算力支持是实实在在的,加工整理和标准化也需要相当高的成本,日常还有水电开支、设施设备运维、网络安全、合规管理等其他成本。数据资源访问人数和下载量越大,各项成本越会水涨船高。因此,这笔钱在我国当前财政收支的刚性结构下要新增出来也是比较困难的,依靠"吃财政饭"来推动公共数据开放,是无法长期维持的(衣俊霖,2022),正因为政府手中的数据具有准公共物品性质,因此完全可以作为国有资产进行国有专营或授权经营。当被授权主体带有营利目的对公共数据进行独家创新应用时,其实质乃是由全体纳税人为作为"特定受益者"的市场主体之需求买单,为平衡这种超额收益,建立"利益返还"机制就是必要的,不管是与被授权主体的经营利润挂钩,还是以统一的固定比例收取,总之目的是把授权运营产生的部分收益以适当超过成本的标准返还给政府。因此,授权运营不仅可能创造一条全新的数据经济赛道,培育一大批各领域活跃的数据运营商,而且反过来极有可能催生新型"数据财政"(杨飞虎、余炳文、王志高,2023)。"运营"获取的利益可以反哺"开放"。这笔全新的增值性财政收入将形成新的激励,不仅能很好地覆盖政府持有和开放数据资源的成本支出,缓解财政压力,而且能提高政府继续扩大数据开放范围、改善数据质量的积极性,能够安排更多的财政专项资金来加强普惠性的免费服务,奖励和资助数据挖掘、应用领域的创新,一个良性的"公共数据运营+数据财

政收入+公共数据专项财政支出”的数据财政长效运行机制就有可能形成(童楠楠、杨铭鑫、莫心瑶等,2023)。

(二)授权运营的各地实践

尽管制度供给已经相当充分,但现有规章框架性很强,虽然形成了典型的基本思路,但具体实践仍然需要各地根据本地实际自行摸索。各地也正在加速形成各具特色的实践路径,譬如成都模式、上海模式、海南模式、广东模式等。各种模式之间主要存在两类实质性的区别:其一,是(1)依靠成立国资公司,还是(2)以特许经营方式运营公共数据;其二是(1)区域一体化(也称区域集中式)运营,(2)由垂直行业主导,还是(3)依靠场景来灵活牵引数据运营。不过,它们之间并非简单的对应关系,按行政属地整体统筹的区域一体化模式往往只有一家获得授权的国资企业承担全部运营,而行业主导或场景牵引模式可分为数源单位的下属事业单位直接运营、授权本地国有企业运营以及授权民营企业运营三类,虽然也经常由垂直领域主管部门指定国有数据运营商来提供服务,但毕竟也会根据市场主体的各自特长,开放给不特定企业参与特许经营。

比如区域一体化或区域集中模式,主要以区域内(省、自治区、直辖市)数据管理机构统筹建设的公共数据管理平台为基础,以整体授权形式委托数据运营机构(一般是本地国资载体),开展区域内公共数据运营平台建设和市场运营。该企业承担公共数据运营服务平台的建设与运营,通过对数据进行加工处理,对外提供数据产品与服务。这种模式是高度数据驱动的,以统一的公共数据运营平台为基础,而后再按行业领域划分,通过被授权主体的市场化运作发现市场需求并予以响应。比较有代表性的是上海市和成都市,两地都成立了本地的(大)数据集团,并将公共数据开发利用统一授权给它。

据报道,上海数据集团是以数据为核心业务的具有功能保障属性的市场竞争类企业,集团承担上海市公共数据和国企数据的授权运营,以公共数据为牵引,推动实现公共数据、行业数据、社会数据资源的整合与布局。成都市大数据集团的授权运营流程更为清晰(图1),其有六个特点:(1)从政务信息资源共享平台上获取政府部门已经授权确认的政府数据;(2)与互联网逻辑隔离,若要出网,则经授权的软件需搭建在安全可信云平台;(3)采用需求驱动方式推进;(4)多轮协商确定相应的数据服务价格;(5)政府数据的经济价值应留在本

地,以国有资产运营收入的方式进入地方财政;(6)数据运营服务并不依赖于数据所有权的转移,避免了数据精细确权的难题。

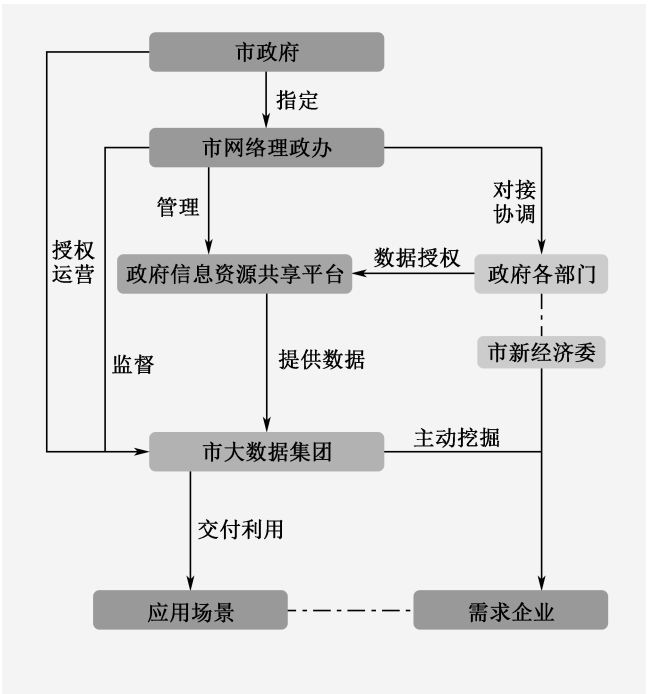


图 1 成都市政府数据授权运营流程

资料来源:数据交易网(2022b)。

整体授权国有企业的打包模式在趋势上看有可能成为未来公共数据授权运营的主流,已经成立并正在运营的国资数据集团有云上贵州大数据(集团)、河南数据集团、福建省大数据集团、陕西省大数据集团、南京大数据集团、苏州市大数据集团等。它的优势体现在国资运营主体能够以较低成本和较少限制拿到数据,降低了运营成本,而且只存在单一运营方,政府的监管压力也相应减小,但却无法回避运营权垄断的问题(陆志鹏,2022)。

再比如行业主导和场景驱动模式。行业主导模式主要由垂直领域行业(如金融、医疗、交通、气象、财政、税收、安全等多个行业)管理部门授权和指导其下属机构承担本领域公共数据运营平台建设、场景开发和市场运营,面向社会主体提供数据服务。这方面的案例也很多,如山东省以国家健康医疗大数据中心(北方)为数据管理主体(政府机构),授权北方健康医疗大数据科技有限公司

开展医疗健康公共数据运营。这类模式在如城市大脑、疫情防控等各类公共事业中应用广泛。还有“航旅纵横”依托中国民航信息网络股份有限公司(国有控股上市企业),通过签订协议,对民航运行、旅客航空出行数据进行开发利用,并形成专业化产品或服务。场景驱动模式则更为灵活,主要由地区或行业数据管理机构基于特定应用场景,通过针对性、专业化分类授权引入专业数据运营机构,分领域、分场景激活公共数据价值。比如北京市以公共数据资源为基础,在金融、医保等领域建设公共数据专区,面向多行业赋能。这两个模式下,参与者所有制情况开始多元化,可能是各细分领域的国有科技企业,也有可能是新的市场主体。比如海南、浙江等地就采取了特许经营的方式进行公共数据授权运营。比较特殊的是海南省,海南省建设了公共数据产品开发利用平台和全省统一的数据产品超市,提供了供需之间的统一对接平台,但并不指定国有运营主体,面向各类市场主体开放竞争性授权。

此外,运营的直接产物基本上都是通过安全沙箱交付的“可用而不可见”的数据产品和服务,而它们首先需要经过确权才能进入市场流通。比较值得一提的有广东的公共数据资产登记制度。广东提出“以凭证承载资产、声明权益、治理数据、保障合规”的数据资产凭证解决方案,以此来提升数据流通效率。2021年10月,广东发布全国首张公共数据资产凭证,并加盖有效力的电子签章。有别于以往的抵押或信用贷款,企业可将其一定时期以来的用电数据作为申请贷款的条件,而相关数据则由广东电网公司权威提供(上海数据交易所有限公司,2022)。2022年8月,广东省政务服务数据管理局为顺德区颁发了全省首批六项共2.8万张“公共数据资产登记凭证”,推动全省首批公共数据资产“入市”,交易在佛山市搭建的全市统一标准的数据要素流通交易平台上进行。这六项数据产品分别应用于乡村振兴、产业发展、高科技、工程、安全生产、气象六大场景,有助于扫清金融机构和融资企业间信息不对称的障碍,有效缓解企业融资难题,解决金融机构对中小微企业“不敢贷、不愿贷、不能贷”的老问题。

(三)机遇与风险并存

市场主体获取公共数据的运营授权,实际上还只是公共数据开发利用的一级市场(场外),但数据产品和服务真正要流转起来,实现供需两旺和价值释放,尚需能够为交易流通提供平台的二级市场(场内),其主要形式就是各地正在建

设的数据交易所。实际上,公共数据社会化利用能否在经济上可行,最终也取决于是否有一个活跃的数据要素市场能为授权运营活动提供统一、开放和透明的平台。“数据二十条”明确提出要“构建集约高效的数据流通基础设施,为场内集中交易和场外分散交易提供低成本、高效率、可信赖的流通环境”,要“完善和规范数据流通规则,构建促进使用和流通、场内场外相结合的交易制度体系,规范引导场外交易,培育壮大场内交易”。今后,国家将建设具有基础设施功能的数据交易场所,充分发挥场内交易的规范引导作用,要素市场的逻辑和规则将更为清晰,最终实现将绝大部分的数据交易活动(包括公共数据产品)都搬到场内去的目标。随着数据要素流通标准体系逐步健全、数据资产实现全国统一登记和入表、数据产权结构性分置完成、区块链+可信隐私计算技术不断迭代(上海数据交易所研究院,2023),数据要素领域的关键性难题将被逐一攻克,这就为公共数据产品/服务全面入市交易创造了巨大的机遇。

机遇与风险并存,除了数据安全问题之外,一些专门行业已经出现了利用公共数据滥用市场支配地位的苗头。2022 年 8 月 5 日,北京知识产权法院受理了国内首例涉公共数据领域的反垄断案件。从事二手车交易的上海彧菡汽车科技有限公司对二手车信息服务平台 CADA 柠檬查提起车险公共数据的反垄断诉讼,原告认为被告利用其能便捷获取全国车险信息平台公共数据而形成的垄断地位,对包括彧菡公司在内的数据需求方查询数据高价收取费用,对非中国汽车流通协会会员给予歧视待遇,违背了公共数据公平公开的开放原则,损害了数据需求方获取公共数据的合法权益,属于反垄断法禁止的滥用市场支配地位的行为。在本案中,资料显示,柠檬查及其所属的北京与车行信息技术有限公司实际上均是由中国汽车流通协会按照国家主管部门要求发起并建立的,柠檬查平台所运营的官方二手车历史信息也来自中国汽车流通协会,因此它属于典型的公共数据市场化运营者是确定无疑的。目前案件正在审理中,但已然提醒我们,作为还在形成中的新事物,公共数据的开放和运营需要预判和防范的绝不止于黑客攻击、数据泄露等直接风险,还有更多尚未显露出来的广义上的风险,比如垄断和不正当竞争。

五、当前地方公共数据治理面临的结构性挑战

尽管已经取得相当大的进展,但地方公共数据治理任务仍然任重而道远,面临的挑战依然艰巨。不过,在厘清难点和矛盾时,应分清楚两类情况。(1)工作推进中的阶段性困难,不存在方向性争议,只需改进方式方法、健全制度体系,把工作做足做细即可改善。比如开放中的数据质量问题:分级分类不精细,类目过于粗疏;数据碎片化,缺乏完整性和时效性,动态数据少;格式不统一且陈旧,开放的数据仅限 PDF、XLS、DOC 等传统格式;存在数据安全和个人信息泄露的风险;维护主要依靠手动,与公众互动性差,用户体验不佳;等等。在授权运营方面,也存在着诸如探索意愿不强、授权范围不明确、监管机制不健全、创新应用能力不强等问题。(2)结构性挑战,存在根本性张力,尚未有定论和产学研各界形成共识的解决方案。换言之,阶段性困难的克服是个时间问题,而结构性挑战仍然需要进一步研究和攻关。我们认为,后者至少有以下几个方面特点。

(一)沿用信息公开的思维来做数据开放

二者侧重目标不同,技术路线也并非简单升级和进化的关系。前者主要还是 web1.0 时代(以门户网站单向信息传递、用户通过 PC 端“只读”为特点)的产物,其底层逻辑是保障公民知情权,实现形式主要是将有关信息以文档的形式上传到政府网站供大众查阅,其实只是“人挂给人看的”,很少有其他的应用方式和价值产生。因此,不少地方的数据开放平台带有浓厚的信息公开色彩。数据开放尽管也内含透明化的公共管理目标,但其首要目标是促进数据的社会化开发利用,政府提供数据要素,与市场合作提供成品化的公共产品,并释放公共数据所蕴含的经济价值(胡业飞、孙华俊,2021)。因此,后者更多的是系统之间交换数据的过程,格式多为可机读的 CSV、JSON 或 WMS 等新型开源项目格式,不需要人的介入,是开放给应用程序的。最典型的就是通过接口供其他机构直接调用,或者通过运营提供基于原始数据的有偿数据产品或数据服务。这二者之间的转变并非简单改变观念和提升意识水平就可做到的,现有语境下的数据开放和运营,是高度需求导向和场景牵引的,需要当地市场有活跃的数字经济活动和数据应用需求,有足够多需要用到数据进行创新活动的大中小企业,有具备

技术能力进行数据挖掘和深加工的市场主体,才能形成良性的数据开放和运营循环。这需要一边盘活和开放公共数据,一边积极调整产业结构、培育市场主体,非一朝一夕可成。相关市场发育不足、需求疲弱,可能也是政府数据开放动力不足、停留在网页张挂阶段的根本原因之一。

(二)激励不相容甚至冲突

已经有不少学者提出,开放与运营之间本身便存在一对可能的对立关系。公共数据授权运营或对政府数据开放制度形成一定的冲击。这是因为,目前政府开放数据基本是免费供社会下载和取用的,但授权运营产生的数据产品和增值服务是要收取一定费用的,尽管定价机制、收费标准、收益分配问题尚未明确,但这是一种产生经济效益的有偿服务确定无疑。由此,授权运营很可能会削弱数据开放的积极性和意愿。毕竟开放的主要动力来自上级的决策部署和行政指令,而一旦有经济因素参与进来,开放与运营二者的范围和边界问题就会凸显出来。上述"市场化能力补充"一说可能本末倒置为一种支配性模式,以运营为主,开放反而成了点缀和补充。因"有条件开放"的数据是授权运营的主体,届时无条件开放的数据就可能"有条件"化,有条件开放的数据就可能不开放了,不排除直接通过点对点服务输送给有限数量的第三方营利机构。厘清二者之间谁主谁辅,建立相互补充、促进而非竞争和排他的关系,还有相当多复杂的工作要做。况且在可用不可见的模式下,授权经营的原始数据一般都是不出本地的,这和数据开放是完全不同的两种工序。如果需求方只获得加工处理后不含原始数据的产品及服务,那么公共数据也并没有实现流通和共享。因此,有学者主张(郑磊,2018):开放数据是一种公共产品,而不是政府的一个新的收入来源,数据免费开放所带来的收益远超直接售卖数据本身,除非涉及对原始数据进行增值加工,才可考虑对额外付出的增值服务的成本进行收费,原始公共数据本身不应该用以营利。

(三)国家与市场关系有待进一步厘清

由此引出第三个问题,一些先行先试的地方,一方面鼓励不同规模、不同行业和不同所有制企业,通过"有条件申请流程"参与公共数据开发和利用;另一方面在授权运营上,倾向于依赖"功能性国有企业"(如新成立的数据集团)来开

展,这里固然有确保国有数据资产始终掌握在国家手中的社会主义市场经济特色,但也有可能牺牲一部分效率和市场活力。从数据产品和服务的应用场景出发,公共机构乃至国有企业都未必擅长发现急难愁盼的真实市场需求。因此,除提供底层数字(算力)基础设施和统一平台之外,在开展数据产品/服务的加工生产,拓宽其可能的应用场景,提升其知晓度、转化率和利用率方面,还是应该尽可能发挥市场能动性。通过合理的利益分成方式,鼓励和引入多种所有制第三方市场主体参与竞争(冯洋,2023),作为被授权主体开展市场运营,避免独家授权,促进市场活力(开放数据中国,2021),如此才能提升公共数据市场化配置效率,让专业的机构完成专业的工作,也能最大限度减轻现有公共机构的工作量,不增加新的财政负担。

六、余论

综上,提升地方政府数据治理能力,尤其是公共数据治理能力,是当前数据领域的重大公共责任。政府既是数据的治理者,也是数据的使用者(郑磊,2020),这两种角色的同构性决定了数据治理能力的提高,反过来在横向上也能提升其他领域的治理能力和效率,因为数字政府背景下各个部门的治理活动都要仰赖数据的充分汇集和高效利用。换言之,狭义上的“数据治理”(治理数据)也能赋能广义上的“数据治理”(用数据来治理)。地方取得的成效对全国政务一体化、形成整体政府格局,乃至推进国家治理体系和治理能力现代化都是十分重要的。

眼下,地方数据治理任务千头万绪,承担经济发展重任的中国政府并非简单的市场守夜人,“有为政府”和“有效市场”结合,是我国数字经济发展能够取得目前巨大成就的根本原因(黄益平,2022)。地方政府还要扮演市场创造者的角色,着力培育规范健全的数据要素市场。同时,“数据二十条”还提到了算法审查,因此开展科技伦理研究,加强人工智能应用的治理以保护社会免受数据和智能技术的伤害,也是数据治理的应有之义。这还只是同数据客体本身的治理活动直接相关的领域,地方涉及数据治理的工作绝不只有这三个方面,譬如适当超前建设大量数字基础设施(新基建领域的5G基站、数据中心和算力枢纽)、推动数字产业化和产业数字化转型、加强平台经济反垄断等,都属于更宽泛意义上的

与数据有交叉的治理活动。

不管是用公共数据提升政府行政效率、打造数字智能政府,还是作为公共资源鼓励企业创新活动、激活数据要素市场,本质上都应始终坚持“以人民为中心”,将维护公共利益和实现全民福祉作为追求的目标。有学者提出公共数据“取之于民,因而必须用之于民”的说法(罗培新,2021),至为允当。

参考文献

冯洋,2023,《公共数据授权运营的行政许可属性与制度建构方向》,《电子政务》网络首发。

复旦大学数字与移动治理实验室,2023,《2022 中国地方政府数据开放报告(省域)》,http://www.ifopendata.cn/report。

国家发展和改革委员会,2023,《加快构建中国特色数据基础制度体系,促进全体人民共享数字经济发展红利》,《中国产经》第 1 期。

胡业飞,2022,《政府数据开放:基于大数据的合作治理创新》,上海:复旦大学出版社。

胡业飞、孙华俊,2021,《政府信息公开与数据开放的关联及治理逻辑辨析——基于“政府—市场—社会”关系变迁视角》,《中国行政管理》第 2 期。

黄益平,2022,《数字经济的发展与治理》,https://mp.weixin.qq.com/s/ZdfDFJiKJK3BmeUB3CoAdA。

开放数据中国,2021,《我们理解的公共数据授权经营——简议〈上海数据条例草案〉》,https://mp.weixin.qq.com/s/Hc6c3TJHbCQqCUu0mfaq5w。

刘阳阳,2022,《公共数据授权运营——生成逻辑、实践图景与规范路径》,《电子政务》第 10 期。

陆志鹏,2022,《公共数据授权运营机制探索》,《网络安全与数据治理》第 7 期。

罗培新,2021,《阿里、腾讯等平台企业的用户数据,是否属于公共数据?可以共享吗?》https://www.shobserver.com/news/detail?id=378545。

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

邱实,2022,《职责聚合:基层治理条块协同的优化创新》,《理论月刊》第 12 期。

上海数据交易所研究院,2022,《金融业数据流通交易市场研究报告》,https://www.digitalelite.cn/h-nd-5427.html。

上海数据交易所研究院,2023,《数据要素市场十大研判》,https://www.ciiabd.org.cn/articles/3k3wxk.html。

上海数据交易所有限公司,2022,《全国统一数据资产登记体系建设白皮书》,https://www.

digitalelite. cn/h-nd-5423. html。

上海数据交易所有限公司、普华永道中国,2022,《数据要素视角下的数据资产化研究报告》,https://www. digitalelite. cn/h-nd-5424. html。

数据交易网,2022a,《全国省级行政区公共数据开放平台统计清单》,https://mp. weixin. qq. com/s/IF-c2vp5VpWnVDPGDBkXmg。

数据交易网,2022b,《成都市政府数据授权运营的流程与实现机制》,https://mp. weixin. qq. com/s/NJz1cH5mT2sMDvdJLjT23g。

童楠楠、杨铭鑫、莫心瑶等,2023,《数据财政:新时期推动公共数据授权运营利益分配的模式框架》,《电子政务》第 1 期。

王伟玲,2022,《数字政府:开辟国家治理现代化新境界》,北京:人民邮电出版社。

徐偲骕,2022,《数据坐稳第五大生产要素了吗?——警惕数字经济的"波兰尼时刻"》,《文化纵横》第 4 期。

徐偲骕、姚建华,2019,《"看不见"的国际传播——跨境数据流动与中国应对》,《国际传播》第 6 期。

杨飞虎、余炳文、王志高,2023,《从数字经济到数据财政的实现路径分析》,https://mp. weixin. qq. com/s/v7rklIcCfDWskGbdeLJQ6w。

衣俊霖,2002,《论公共数据国家所有》,《法学论坛》第 4 期。

于文轩、魏炜,2022,《数据的敏捷治理——价值重塑与框架构建》,《广西师范大学学报》(哲学社会科学版)第 5 期。

郑磊,2018,《开放不等于公开、共享和交易——政府数据开放与相近概念的界定与辨析》,《南京社会科学》第 9 期。

郑磊,2020,《政府在数据治理中的两种角色——政策的制定者和数据的使用者》,《探索与争鸣》第 11 期。

中国软件评测中心,2022,《公共数据运营模式研究报告》,https://dsj. guizhou. gov. cn/xwzx/gnyw/202205/t20220530_74436679. html。

中国信息通信研究院,2023,《网络立法白皮书(2022 年)》,http://www. caict. ac. cn/kxyj/qwfb/bps/202301/t20230112_414232. htm。

对社交机器人领域拟人化现象的三种回应*

——走向批判性、关系性与解释学方法

马克·科克尔伯格　著**

曹忆沁　译***

摘要：社交机器人的设计者与用户都倾向于将机器人拟人化。本文以人类与机器人的关系为着眼点，考察该如何对这一关系进行概念化。目前围绕这一问题存在“天真的工具主义”与“不加批判的后人类主义”两种相反的哲学观点，它们与对拟人化现象和拟人化设计的各种规范性反应相关。基于对上述两项观点的考察，本文提出了第三种观点，即从批判性、关系性以及解释学方法出发理解人类与机器人间的关系。这一观点通过厘清机器人在使用与开发中所蕴含的人类维度，并将机器人严肃地看作更大范围的技术语境内的工具，同时规避了以上两项立场，并对它们所假设的二元论提出质疑。最后，本文探讨了第三种观点对社交机器人领域及计算机科学家和工程师的教育所产生的意义，提倡基于关系性理解对上述领域进行改革，并为用户—公民在其中可以扮演的角色提供建议。

关键词：社交机器人领域　人—机器人关系　拟人化现象　拟人化设计　工具主义　后人类主义　解释学　关系方法　权力

一、社交机器人领域的拟人化现象与拟人化设计

社交机器人学家旨在开发能比以往机器更自然地于社会环境中发挥作用的机器人。此类机器人常以人类或其他高等灵长类动物为模型，由此通过隐喻方式鼓励人类将它视作一种社交机器。综合心理学、认知科学、神经科学及其他领

---

* 本文原题为“Three Responses to Anthropomorphism in Social Robotics: Towards A Critical, Relational, and Hermeneutic Approach”，原文见 https://link.springer.com/article/10.1007/s12369-021-00770-0，经授权翻译发表。

** 马克·科克尔伯格(Mark Coeckelbergh)，奥地利维也纳大学哲学系。

*** 曹忆沁，复旦大学哲学学院。

域内的研究成果,社交机器人领域与人机交互(HRI)领域的研究者得以确定社会交互所必需的特征,并将这些特征赋予机器人。多滕哈恩等人(Fong, Nourbakhsh & Dautenhahn,2003)结合人机交互方面目前已完成的工作与社会脑假说(social brain hypothesis)——该假说认为,灵长类动物的智力之所以能够进化,是因为其需要处理日益复杂的社会动力学(social dynamics)问题——对社交机器人被赋予的特征进行了总结。她认为:"社会交互机器人呈现出以下特性:表达与/或感知情绪,以高级对话形式交流,学习其他能动者(agent)的行为模型或对其他能动者进行识别,建立与/或维系社会关系,使用目光、手势等自然提示(natural cues),展现独特的个性与性格,可能学习与/或发展社会能力。"(Dautenhahn,2007;Fong, Nourbakhsh & Dautenhahn,2003)。为了实现上述社会特征,机器人往往需要展现出拟人化的外观与行为。这一语境下的拟人化设计具有更务实的目标——在社交机器人领域中,拟人化设计是一种使机器人融入人类社交环境的策略,也是一种赋予机器人"与人进行有意义社交的能力"(Duffy,2003)的手段。无论是社交机器人的设计者还是用户,他们在与这些机器人交互时都倾向于将其拟人化,赋予它们如个性、生命等特征。后文中,我将使用"拟人化现象"(anthropomorphism)、"拟人化行为"(anthropomorphizing)与"拟人化设计"(anthropomorphization)来分别指代(用户与机器人互动时)将类人特性赋予机器人的这一现象、这一行为本身,以及机器人设计者与开发者对此种类人特性的创造。

拟人化现象有时并非人们有意追求的结果。社交机器人学家意识到了这一点并试图对拟人化现象进行控制。他们或许会通过对机器人进行设计,尝试鼓励或阻止用户对机器人进行拟人化。一些机器人学家有意制造尽可能像人的机器人,如石黑浩(Ishiguro,2006)认为仿人机器人(humanoids)是理想的人机交互界面(interfaces);另一些机器人学家则试图制造不那么像人的机器人,以看上去像动物的机器人代替或尝试彻底规避拟人化现象。理论上看,机器人的拟人化对于社交机器人领域而言似乎并非必需品,社交机器人只要能够展现出前述社会互动所需要的特征就足够了。但在实践中,许多社交机器人学家都乐于见到某种程度或某种形式的拟人化现象,他们认为这有助于促进人机交互。如今,拟人化现象已成为塑造人机交互的因素之一(Bartneck, Kulić & Croft et al., 2008)。虽然人们已认识到拟人化现象带来了挑战,但归根结底需要承认,人机

交互的塑造旨在“促进类人机器在现实世界内的整合”(Złotowski, Proudfoot & Yogeeswaran et al. ,2015),而拟人化现象有助于实现这一目标。

事实上,机器人与这一目的十分匹配(Kiesler & Hinds,2004),它们的外观与存在方式使其具有高度的“拟人化能力”(anthropomorphizability)(Złotowski, Proudfoot & Yogeeswaran et al. , 2015)。早在 20 世纪 90 年代,李维斯和纳斯(Reeves & Nass,1996)就发现人们会像对待真人一般对待计算机,而对于机器人——尤其是仿人机器人——来说,这种情况则更为普遍(Kiesler, Powers & Fussell et al. ,2008)。但在外观之外,用户的拟人化行为也受机器人在交互过程中给出的种种线索(Hegel, Gieselmann & Peters et al. ,2011),以及机器人的自主性、可预测性等其他因素影响。有趣的是,就本文而言,拟人化现象的出现并不仅仅由机器人的外观或行动导致,也与机器人观察者自身的特性相关,如观察者的社会背景与性别(Złotowski, Proudfoot & Yogeeswaran et al. ,2015)。举例说,人们会将社交机器人归入不同的社会类别,如赋予它某一群体的成员身份(Kuchenbrandt, Eyssel & Bobinger et al. ,2013)。这意味着,社会视角将有助于我们理解乃至评估人机交互中的拟人化现象。同时这也表明,除机器人的客观特性外,在对拟人化现象的考察中关注用户的主观性也同样重要。

心理学家以实证方式研究公众对机器人的感知方式,以此帮助机器人学家开展拟人化社交机器人的开发项目(Krägeloh, Bharatharaj & Kutty et al. ,2019; Peca, Coeckelbergh & Simut et al. ,2016)。而机器人学家也会引入有关人类社会性的心理学理论,在跨学科背景下开展研究:他们使用心理学的关键范式,在人机交互语境中应用与复制人际互动中的发现(Eyssel,2017)。除了在设计层面对机器人的特征进行打造,形象构建(framing)也是一种鼓励拟人化的方法。当机器人被赋予名字或背景故事时,人们对它的反应也将受到影响(Darling, 2017)。需要再次说明的是,针对某一技术出现拟人化现象并非新事,但社交机器人领域内有意开发拟人化机器人的做法将会增加这一现象的影响。

将机器人进行拟人化的行为引发了多种担忧,其中一些已在社交机器人与人机交互社群中得到识别、认知与研究。特克尔(Turkle, 2011)与舒尔茨(Scheutz,2011)对人类与拟人化的社交机器人发展情感关系表示担忧。日本机器人学家森政弘(Mori,1970)在机器人领域引入弗洛伊德的恐怖(uncanny)概念,进而提出著名的“恐怖谷”(uncanny valley)问题。此后,关于机器人是如何

唤起“恐怖”情感,以及该如何在机器人制造(MacDorman & Isghiguro,2006)等方面规避这一问题的讨论层出不穷。对弗洛伊德来说,恐怖感(uncanniness)是一种在熟悉的事物中产生陌生感的心理体验。他在1919年的一篇论文中(Freud,2003)强调,当本应无生命的物体变得好似有生命时,人们将对其感到陌生,并受自身无意识且潜藏的冲动威胁。机器人语境中的恐怖感则意味着,当一个外表与人类相似的机器人以奇怪的方式移动时(这只是其中的一个例子),用户可能产生恐惧情绪。根据心理学上弗洛伊德式的解释,这一情绪产生自人们对死亡的恐惧。从人类学视角考察,拟人化现象可被概念化为技术泛灵论(technological animism)的一种形式。泛灵论认为无生命物与自然存在物也拥有灵魂与精神,但灵魂与精神不被理解为属于我们社会进化的“原始”阶段或仅限于“自然”现象层面的东西,而是作为主体能动性与人格的构建基础(Richardson,2018)。技术泛灵论认为,上述过程也发生于机器人等技术物中。但是,本文中我将专注于拟人化现象与拟人化行为的讨论。

鉴于拟人化现象带来的问题,一些机器人学家主张不应以创造拟人化的机器人为目标。戈茨与其同事(Goetz,Kiesler & Powers,2003)认为,机器人的设计应与机器人所需完成的任务相匹配,而不应将机器人与人类的相似性作为主要目标。但是,由于构建拟人化的机器人能够促进人机交互,且在卫生保健、老年护理等领域都有助于实现机器人的目标,故社交机器人的开发者们往往仍追求着拟人化的设计思路。

我在此关注的问题是:如何从技术哲学角度回应拟人化现象?我将通过探索“如何从哲学层面架构起机器人与人类间的关系”来进行回应。首先,我就这一问题设置了两种对立的哲学观点,并将它们称作“天真的工具主义”(naive instrumentalism)与“不加批判的后人类主义”(uncritical posthumanism)。这些观点可能会被视为“稻草人观点”(straw man views)①,但我将表明,确有研究者持有以上两种观点,或至少是其温和版本。其次,我也将展示这些观点是如何与针对拟人化现象与拟人化设计的不同规范性立场②相联系的,后者无关对现象的

① “稻草人观点”或“稻草人谬误”是一种逻辑谬误,指对他方观点进行扭曲或夸大,使其更容易被驳倒。——译者注

② “规范性立场”指对“在道德层面应当做什么”问题的立场。文中所指的“规范性立场”即对“是否应当视机器人具备人类特性”“是否应对机器人进行拟人化设计”等问题的立场。——译注

描述与解释,而关乎用户是否应该拟人化机器人,研究者又是否应该研发拟人化的机器人。最后,勾勒上述两种极端的对立观点,也有助于我提出回应拟人化现象的第三种观点。这一观点受批判性、关系性与解释学的方法启发。

虽然我明白对于拟人化现象的讨论也与机器人伦理学及机器人道德地位问题相关,但我不会在这里直接处理这些具体的规范性问题,而会将论述重点放在与社交机器人领域及人机交互相关的拟人化现象问题和人类—技术关系问题上。在文章最后,我将探索第三种观点对作为学科的社交机器人领域,以及对计算机科学家与工程师教育所产生的影响。

二、两种有问题的对立观点:天真的工具主义与不加批判的后人类主义

(一)天真的工具主义

天真的工具主义坚持认为机器人只是机器,仅仅是实现人类目的的工具。根据这一观点,对机器人进行拟人化是一种心理偏见。我们知道机器人仅仅是工具,尽管如此,与机器人互动时的心理作用也会使我们将机器人视作人类。这一观点具有两方面的哲学基础。第一,它预设了一种二元论的世界观,其中实在(reality)与表象(appearance)彼此对立(这一观点可追溯至柏拉图)——实在层面的机器人只是工具,而表象层面的机器人却好似真人。同时,天真的工具主义也秉持人与非人对立的二元观点,认为人与非人是两个互斥的范畴,彼此间存在一道深深的本体论鸿沟。如此来思考技术,则意味着技术归属于非人一侧(包括事物、客体等),而人类则归属另一侧(包括人、主体等)。因此,机器人与人类分别属于两个完全不同的本体论范畴。第二,天真的工具主义观点在形而上学与认识论层面是实在论式的。这一观点认为,对象独立于我们的概念和感知存在,我们可以用客观的方式来描述它们。当用户对机器人进行拟人化时,他们被机器人的外表所误导,实在层面的机器人其实只是一台机器。相反,科学家则能用客观术语来描述这一现实情况,并能研究人类的心灵(mind)是如何被机器人的拟人化特征、人机交互的形式,以及我们所处的社会文化背景所误导的。

天真的工具主义至少可以引出两项有关拟人化现象与社交机器人任务的规范性立场。第一,社交机器人领域的研究者可以利用用户的心理偏见,通过设计类人机器人来促进人机交互,以此更好地实现人机交互的目标。不仅机器人是工具,拟人化设计本身也是一种能被利用与开发的工具。这便是我在文章开头所勾勒的立场,它与专家和公众间关于实际情况的知识分布不均有关:机器人的设计者在实在层面知晓机器人仅是机器,而用户——至少是暂时地——受到误导,相信机器人是一名真人。可以说,机器人设计者就像一名魔术师(Coeckelbergh,2018a),他创造了“机器人是真人”的幻觉。正如设计哲学家弗卢塞尔(Flusser,1999)指出的那样,“设计”与“机器”这两个词,在词源学上便与狡猾和欺骗有关。对机器人进行设计,本身便是为了令用户对它进行拟人化,使用户受到欺骗而误以为机器人是人,或(在考虑到机器人技术现有水平的情况下)至少愿意暂时悬置怀疑、拥抱幻觉。根据工具主义的观点,只要能够达成目的,那么使用拟人化设计这般魔法诡计便不成问题——如果拟人化现象与拟人化行为能够促进人机交互,助力所愿目标的达成(如卫生保健领域内的特定任务),那么拟人化设计的诡计便是被允许的,乃至是受鼓励的。

需要注意,出于同样的理由,如果机器人的目的无法通过拟人化现象达成,那么工具主义便会放弃拟人化的设计思路。无论出于什么原因,只要一个机器人没能实现它的功能,那么工具主义便会反对它。

第二,天真的工具主义也能够导出另一规范性立场。这一立场认为,无论机器人能够达成何种功能与效果,对机器进行拟人化或开发此类机器,在一般层面与原则层面都极成问题——因为工具主义相信并断言机器人是且仅是工具,不能是其他东西。有学者因此认为,社交机器人学家应当停止设计这类机器人,或至少确保用户能够意识到它们仅仅是工具。以布莱森(Bryson,2010)的观点为例,她认为机器人仅仅是我们赖以实现人类目标的工具,以一种暗示机器人是人类的方式来谈论它们将构成一种误导。

此外,另有一些其他方面的理由使得拟人化现象成问题。有观点基于先前提到的潜在伦理问题反对建造拟人化的机器人。斯派洛等(Sparrow & Sparrow,2006)便认为,人们或许会基于拟人化现象对机器人产生情感依恋,进而脱离现实。例如,卫生保健领域的社交机器人或许会误导用户认为其真的活着,或能够真正成为他们的朋友。虽然当前的机器人技术尚不足以面向所有用户创造幻

觉,且机器人创造的幻觉也不能长时间延续,但社交机器人领域的发展趋势却似乎正以此为目的,例如医院与疗养院正在使用的海豹宝宝机器人 Paro。特克尔(Turkle,2011)认为,像 Paro 这样的机器人令用户以为自己建立了一段真实的关系,但这仅仅是一种幻觉。

从工具主义入手,我们可以认为上述批评是有道理的,并认为社交机器人学家应开发能够避免产生幻觉、仅仅被用作工具的机器人,而不应开发“假装”自己不止于此的机器人。如果这一规范得到落实,那这很可能使社交机器人这一研究项目走向终结——当机器人不再具有激发人类拟人化行为的特征(或用户会被不断提醒它只是一台机器)时,我们尚不清楚它该如何与人类进行“自然的”社会交互,或者说,真正“社会”意义上的互动。根据上述观点,我们应当停止开发与使用鼓励用户进行拟人化的社交机器人。如果这意味着人机交互将变得不那么“自然”,那么我们也只能接受。

但无论对拟人化设计采取何种规范性立场(我们是否应当进行拟人化),两种立场都对人机关系采取了工具主义的解读方式。两种立场都认同机器人只是工具,仅在评估拟人化现象与拟人化设计时有不同的态度。此外,我称这一工具主义“天真”,并非认为这一观点的秉持者缺少关于拟人化现象或机器人学的知识。事实上,两种规范性立场的持有者都对拟人化现象及其性质、效果极为了解。“天真的工具主义”中的“天真”二字,仅指这一观点忽视了技术的非工具性维度。在接下来的章节中,尤其是当我谈及第三种立场时,“天真”一词的内涵将变得更为清晰。

(二)不加批判的后人类主义

我试称另一种态度截然相反的观点为“不加批判的后人类主义”。它将社交机器人彻底接纳为准人类(quasi-persons)与“他者”(others)。后人类主义拥有几重不同的含义,我在此指的是一种批判传统人类主义(humanist)[①]世界观将人类置于世界中心(人类中心主义),并将本体论与道德考量范围延伸至非人(non-humans)的理论。这里说的非人,可以是动物,也可以是(某些)机器人。与工具主义立场相反,后人类主义欢迎社交机器人,并将其视作后人类生态中的

① humanist 通常可译为“人类主义”或“人文主义”,此处因与“后人类主义”对应,故译为前者。——译注

一环或人与非人网络的一部分。这一观点并不秉持二元论式的世界观,不将人与物置于对立面,而是认为人与非人同属一张网络,并以各种方式相互纠缠。后人类主义者鼓励我们跨越边界,质疑二元论与二元对立(包括人与技术的二元对立)的思考方式。根据后人类主义观点,人是技术的存在,而技术也是人类——人类总是在使用技术,人类创造技术,技术并不仅仅是人类的工具,它们是人类世界的一部分。比如,我们可能会对技术夸夸其谈,我们可能期望技术向我们提供超出其能力范围的东西(这种情况于机器人与人工智能领域十分常见),我们或许会崇拜技术,等等。在机器人领域,后人类主义观点认为,我们与机器人之间的关系并非仅仅是工具性的,所以我们也不应将机器人视作纯粹的工具。相反,我们可以承认机器人是人类与文化世界的一部分,甚至将它们视作他者——并非人类意义上的他者,而是与我们一同分享世界的社会实体(social entities)。此外,后人类主义并不以实在论的观点来看待世界与我们认知世界的方式。其支持者相信,我们并不能客观地了解何为"实在"(我们无法脱离自己的认知方法独立地获知现实),科学信念都是一种社会建构,关于机器人的信念也不例外。所以,科学并不能详尽地描述出机器人"是"什么,对这一问题的回答不能脱离人类主观性与人类文化。试考察"机器人"与"机器"这两个术语。它们在科学与工程领域都没有得到确切的定义,因为它们与自身的文化历史背景结合紧密,其含义有赖于它们的时代背景。例如,对"机器人"的定义或取决于社会语境及相关的语言使用——我们通常不会将自动驾驶汽车称为机器人,多数语境中我们只是叫它们汽车,但单就技术而论,自然驾驶汽车确实是一种机器人。

这样一种后人类主义、社会建构主义式的方法或许是受到了哈拉维(D. Haraway)的启发。哈拉维(Haraway,2000)在《赛博格宣言》及其后续作品中,主张通过跨越边界,将机器与动物都纳入政治议题。当然,这一观点也可能来自拉图尔(B. Latour),他的非现代方法(non-modern approach)将非人类也纳入社会中(Latour,1993)。哈拉维的目标不仅在于创造字面意思上的赛博格(cyborgs),即人与机器的融合体,更在于改变我们对于人机边界(之重要性)的思考方式——不仅在形而上学层面,同时也在政治层面。同样,拉图尔将机器人视作需被纳入社会与政治集体的非人类;若将它们排除在这一集体之外,便是在自然与文化、人类与非人间延续了严格的界限,而非现代的方法希望超越这般二元对

立。贡克尔(D. Gunkel)基于列维纳斯的理论探索机器人是否能够被视作“他者”(Other)(Gunkel,2007;Gunkel,2017)。他的观点更具批判意味,与后人类主义的观点并不完全相似,但也可被视作后者的潜在起源。传统现象学中,他者(the other,有时写作 Other)与自我(self)对立,并常指其他人类。哲学家列维纳斯在此基础上对该词添加了一种特定用法:他者是自我的根本对应物,是绝对的另一者,我们被要求以道德的方式对其进行回应。列维纳斯在此关心的仍是人类,而从后人类主义角度出发,我们可以继续追问:列维纳斯所说的“他者”是否也可以是非人的动物或机器人?

根据后人类主义的观点,发展社交机器人未必是成问题的,甚至是值得期待的。毕竟后人类主义的观点鼓励我们将非人纳入社交领域,而拟人化的设计有助于达成这一点。我们应超越工具主义立场的人类中心论,人类不应该是万物的中心,而发展社交机器人便有助于撼动人类的中心地位。后人类主义同样对拟人化现象表示支持,将打造看起来像人类的机器人并令用户将其感知为人类视作一桩好事,甚至认为这值得我们主动追求。究其原因,并不在于机器人能够达成何种人类目的,而在于它们能够成为人类与之互动并将其纳入社会中的“他者”。后人类主义拒斥人类中心主义,拒绝将人类置于世界与道德考量的中心,所以人类的目标并不是唯一重要的目标。基于这一观点,我们甚至可以尝试打造与人类毫不相似的社交机器人。

上述规范性立场并不拒斥拟人化现象本身,只是指出还存在许多其他的可能性。然而,基于贡克尔的观点(Gunkel,2007;Gunkel,2017)与一种后现代方法,后人类主义的立场也可以对拟人化设计采取另一种更负面的规范性立场:从一种以差异或他者性(otherness)为导向的后人类主义进路出发,我们可以认为,将机器人感知为人类的拟人化行为在一定程度上是成问题的,因为这对作为人工他者(artificial other)的机器人本身之差异性与他者性不够尊重。我们不应在机器人身上投射我们的人类性(humanness)——我们何以为人,又如何视己为人——并将它们当作(或错误地当作)我们自身形象的投影屏。相反,我们应将机器人视作自身拥有权利的他者,并尊重它们与人类的不同。如果我们做不到这一点,那么我们就不应打造拟人化的机器人(但可以建造其他类型的机器人)。这一规范性立场同样认为拟人化现象与人类中心主义联系紧密,但对二者都持拒绝态度。

(三)问题

天真的工具主义与不加批判的后人类主义都存在问题。前者是天真的,因为它无法理解机器人并非单纯的工具,事实上机器人还会产生意料之外的后果,并通过经验、语言、社会关系与叙事等因素,与人类紧密联系在一起。其二元论与实在论(realist)的世界观在人与技术间设置了一种外在的关系,使得我们无法意识到,机器人并不仅仅是一件东西,并不仅仅是技术与编程的产物,同时也由人类与社会建构而成。后人类主义的观点是未经批判的,它遗忘了机器人由人类制造这一事实。机器人或许会在人类面前呈现为“准他者”(quasi-other)(Ihde,1990),但它们从不是完全的他者或非人类,因为它们由人类创造、解读,并由人类赋予意义。在我们面前,机器人或许会呈现为一个外在的他者,但实际上它们从未完全外在于我们。后人类主义对于机器人他者性与社会文化建构的关注,使我们忽视了机器人在人类物质实践中的起源。机器人也是机器,而这些机器由人建造。换句话说,以上两种观点都没能充分认识到社交机器人是如何以各种方式,与人类产生本质且(下文将进行论述的)内在的联系的。

上述两项观点的问题源于理解不全与遗忘。两种观点的持有者对拟人化的意外后果及机器人技术的起源均有所了解,或原则上能够获取相关知识。所以问题并非天真的工具主义者不知道技术可能产生预期外后果,或不加批判的后人类主义者不知道技术是由人类创造的。相反,问题在于,他们往往无法理解这些知识对于思考人机关系所具有的深层含义,也因此无法理解它们对于理解拟人化现象问题的意义。尤其是,这两种观点都没能理解人类与机器人间的联系有多紧密,彼此间存在一种内在关系(见后文)。

然而,或许还存在“第三种”立场,它可以解决上述两种观点所存在的问题,同时仍允许我们对拟人化现象与拟人化行为——事实上是社交机器人这一研究项目本身——保持批判态度。我们如何才能得出这样的第三种观点?一种方式是尝试在两种极端中找寻“中间”位置,即认为社交机器人既不是全然的工具,也不是全然的他者。但这样的观点很模糊,什么是机器人的“非工具性”面向,机器人能够在一定程度上成为“他者”又意味着什么?所以,为了找到真正的替代方案,我提议对前述两种立场的预设进行质疑与改变。特别是,我提议放弃二元论的预设,因其将人类与机器人间的关系视作外在关系,并由此出发考察人类

与机器人间的交互方式,进而对社交机器人领域进行定义。一旦我们放弃二元论式的立场,一种完全不同的全新立场便将诞生。这一立场并非在两种极端立场间取其中点,而是直接从上述观点所划定的领域中挣脱而出。

三、走向第三种观点:批判性、关系性与解释学的观点

让我概述一下这样一种立场,它适用于一般意义上的机器人,但其本身的提出却是为了回应上述两种针对社交机器人领域中拟人化现象与行为的相反立场。为了成为真正有效的替代性方案,这一立场至少需要包含以下要素。

(一)机器人与人类相联系,但并不是人类本身:它们被人类创造与使用,它们塑造我们的目标

社交机器人并不是凭空出现的(仿佛它们是"非人"自然的一部分,或外太空的外星人),它们是由人类设计与制造的。这意味着,上述两种对立观点中论及的"工具"或"他者",不仅仅是非人类,同时也与人类相关。机器人不仅与人类的目标及意图相关,就好像它们是外在于我们的事物,它们同时也由人类在具体科技实践中创造,并为人使用。在这层意义上,它们与人类相联系,永远不可能是完全的"他者"。鉴于我们创造并使用了机器人,即便是非拟人化的机器人也拥有人类的一面。我们不必邀请机器人进入"我们"的领域,因为自从我们使用机器人以来,它们便一直身处其间。既然我们创造了它们,它们就从未是单纯的工具;相反,它们被注入了我们的目的、意义与价值。所以说,机器人是工具,但它们是我们的工具。而由于它们能够达成人类预期之外的意外效果,它们同样也塑造了我们的目标。由此可见,工具与目标相互依存,不可分离。

例如,为了使机器人在特定的卫生保健领域发挥作用,它或许会被赋予拟人化的特征。这一设计的目标与卫生保健本身相关,如与病人进行交流。但通过应用机器人完成交流任务,这一目标的内容便发生了变化:"与病人交流"的含义被改变——过去它意味着人类作为护理人员与病人进行交流,而如今这一任务则可由机器完成。进一步说,如果我们(作为病人或护理人员)忘记了机器人是由人类所创造的,并由人类赋予其意义,那么机器人便可能被不加批判地投入

应用,而人机交互在意义建构方面的开放性也将被取消。事实上,机器人与人机交互在意义建构方面的作用不应被视为既定的,而应被视为开放的。

(二)机器人的语言与社会建构

即使我们将机器人视作既定的,认为其自足地展现于我们面前,我们作为用户与它进行交互(设计者也是一类特殊的用户)时仍然应该质疑人与机器人间的关系仅是外在的。可以通过考察我称之为认识论上的人类中心主义来质疑这一点。从这一观点出发,我们认为机器人"是"什么、我们如何了解机器人,以及我们如何与机器人产生联系,这些问题的答案始终取决于人类的主体性、意义建构方式、叙事方式、语言与隐喻等。无论我们是否能够摆脱道德或形而上学层面的人类中心主义,(在康德、现象学,以及 20 世纪的大部分哲学之后)采取一种批判哲学的进路便意味着,当我们谈论、思考机器人并与之互动时,我们无法绕过"人类"这一面向。机器人"是"什么,这一问题的回答有赖于我们人类如何评价它,以及如何在交互过程中回应它。人类并不仅仅在物质层面创造了机器人,也在开发、使用与交互过程中,通过语言的方式在社交关系中"建构"了机器人。当我们思考这些机器人并与之互动时,必须首先预设这一点。而这一点可以由批判哲学与社科考察揭示——包括从时间与历史角度入手考察,并关注语言与隐喻的作用(Coeckelbergh,2012;Coeckelbergh,2017;Coeckelbergh,2019)。考察语言的作用:如果我为机器人取了一个个人化的名字,那么我便已经将它建构成了一个伙伴、他者,并使其拥有了特定的性别等特征(Coeckelbergh,2011)。拥有姓名后,机器人未必会停止发挥其工具性功能,也可能在其他时候或其他视角中仍被视为一件工具;但取名这一行为本身,连同机器人的外表等其他要素,共同塑造了机器人的意义。此外,机器人的意义也与其嵌入的社会文化背景相关。比如,赋予家政机器人一个女性化的名字(这一社会行为是对机器人相关特性的回应),不仅"使得"机器人成为"某人",且用户应以社区与社会能够接受的方式与此人互动;同时也反映、确认并放大了一种特定的文化,其中女性应当承包家务。在此意义上,社交机器人既不是"单纯的机器",也不是完全的"他者"。它们以各种方式与人类社会和人类文化相连,本身也正由该社会和该文化塑造。从这个意义上说,机器人并不"外在"于人类与人类的目标,相反,人类与机器人间存在着一种内部关系。

(三)关系性与意义建构:机器人是嵌入在文化整体中的

替代性的观点应当始终是关系性的。这一立场需要承认,机器人的关系性不仅内嵌于更广泛的技术系统中,同时也内嵌于更广泛的社会技术系统(sociotechnical system)中——就像其他技术一样,机器人"与社会实践以及人类的意义系统相互交缠"(Johnson,2006)。因此,它们与此类实践与意义系统间存在"关系性"。也可以说,它们与社会关系相连(Coeckelbergh,2012)或嵌入在文化整体中。

在近期的工作中,我使用维特根斯坦的"语言游戏"(language games)与"生活形式"(form of life)的概念来表述机器人的这一社会文化嵌入性——就像使用语词一样,对诸如机器人等物品的使用,也同样嵌入(技术)游戏与生活形式中(Coeckelbergh,2017;Coeckelbergh,2018b)。我并不是在生物学意义上使用"生活形式"这一术语,而是(通过使用维特根斯坦的理论)将其解释为文化层面定义的游戏与实践。它们构成了一个或多个整体性语境,并赋予我们的活动与技术以意义。这些整体性语境使我们能够理解机器人并与之互动。而这些机器人又反过来帮助我们自己进行意义建构活动,乃至成为这一建构活动的一部分。可以说,我们并不仅仅是在口头上谈论技术,技术本身也积极促进着意义建构。由此形成的那张意义之网,或那方意义之界,也反过来帮助我们理解具体的经验与事物。基于利科的叙事理论,雷耶斯与我(Coeckelbergh & Reijers,2016)认为,就像文本一样,不同的技术通过将人物与事件配置成一个有意义的整体来共同构成叙事。例如,一场特定的人机交互可被放置于"与机器人建立友谊关系"这一更广范围的叙事语境中进行理解。这一点可以通过讲述关于机器人的故事来达成,但关键在于,机器人本身也积极促成了叙事的形成——比如用户可以通过机器人的拟人化形态,来为其创建角色(比如朋友)与事件(比如相遇与问好)。

更一般而言,由于技术对意义建构起了作用,我们可以说技术具备一种解释(hermeneutic)[①]功能。而当我们与机器人互动时,这一解释功能同样存在,并通过人机交互发挥作用。机器人并不只是我们建构意义并加以解释的对象(比

① 此处使用"解释"(hermeneutic)一词指代意义(meaning)与诠释(interpretation)。

如,我将机器人解释为“单纯的机器”),它们也同样塑造了我们理解世界的方式。例如,家政机器人以仆人的身份与人类交互,便可能传递这样的含义,即社交世界是主仆的世界,其中用户是主人。所以,机器人并不是单纯的机器,它们并不独立于人类文化之外。相反,机器人被纳入了我们人类的解释活动——以一种被动的方式(我们谈论它们并理解它们),也以一种主动的方式(它们帮助塑造了我们的故事的意义)。

就其背后的哲学基础而言,这一关系性与解释学的理解进路暗示着,实在论者与工具主义者使用的关于“是”(is)的存在论语言(关于某物或某人是什么的哲学主张或哲学问题,比如机器人是什么)已变成问题。并不存在一种自在的(in itself)“东西”或“工具”,也不存在一个与人类无关的“他者”。所有的技术都由人创造并被人使用,也因此都与人类有关。这一点不仅适用于所谓的“社交”机器人。事实上,在此意义上所有的机器人都已然具备社会性与关系性,即使它们没有或尚不具备“社交性”特征。例如,工业机器人被嵌入属于特定社会工业语境的语言与技术游戏中,在这一语境中,工人在资本主义条件下已被机器取代。但对于社交机器人来说,这一社会性—关系性面向更为明显。可以说,社交机器人展现了我们的社会环境与社会文化。例如,拥有语音界面的人工智能社交机器人将会以“我们”(在家庭、社区、国家或语言共同体等中)交谈的方式来与我们交谈(而我们也会以同样的方式回应)。如果社交机器人想要变得幽默,那么它必须深挖人类已达成共识的知识、经验与技能基础。如果经由机器学习技术,机器人成功开了一个玩笑,那它一定是学会了我们人类的“游戏”。机器人的成功表现依赖于一个关系性的整体语境、一种生活形式。若要使机器人变得有趣,或更一般地说,能够建构起为人所理解的意义,我们必须以此二者为前提。社交机器人学家会让机器人模仿人类建构意义的方式,而具体方法是让机器人从它在互联网上找到的材料中进行学习——在互联网上,文本语料库能够将零散的意义具体化。在人工智能领域中,已有学者开始研究使用互联网语料库训练人工智能可能产生的伦理问题。例如,在这些语料库及我们的日常语言中,都存在着性别偏见,而这往往会被人工智能所复制(Caliskan, Bryson & Narayanan,2017)。当人工智能技术被越来越多地应用于机器人时,此类问题的数量便可能大幅增加。从人机交互层面入手考察,根据定义,人机交互涉及人类,也因此与游戏及更广阔的意义视野相关联,所以这类交互已是且总是

一种(实时的)意义建构活动。同样,在很多方面机器人都不仅仅是单纯的工具,它是意义建构活动或游戏的一环,并且依赖于更大的意义整体来实现它们的意义。再者,机器人也并非社会性层面完全的“他者”,在其与人类社会、文化的关联中,机器人也获得了许多与人类社会、文化的相似性。足够的相似性是机器人与人类共享意义的基础,也是其作为社交机器人顺利发挥作用的基础。一个全然外来的“社交”机器人缺少与人类意义的联系,我们也因此不会认为它具备社会性。

(四)对解释方式缺乏控制

说物质创造与意义建构活动是人类的(即人类会做的事),并不是说这类活动完全受人类控制。我们即使是在“纯粹”的人类情境(这一情境通常来说并不会是“纯粹”的,因为其中常常具有技术或其他事物的中介)中也无法完全对意义进行掌控。当我们与机器人进行交互时,往往会有意义涌现而出。有时这些意义是出人意料的,并不在我们的意图之内。例如,当机器人突然听上去或看上去很像人类时,我们或许会哈哈大笑,又或许会对此感到恐惧。人机交互领域的研究显示,机器人的行为常常会令人惊讶。在这种情况下,某种意义上说,机器人具有一种他者性,或至少这一互动行为以及处于互动关系中的机器人是具有他者性的。此类机器人的开发者、设计者以及用户或许会尝试对机器人产生的意义进行控制(如设计者或会尝试规避恐怖谷效应),但事实上,我们无法百分百获取解释层面的控制力。当人类与世界、与彼此接触时,总会为惊奇留下空间,总会产生我们意想不到的意义。在此意义上,如果“机器”意味着“机械”,而“机械”又意味着“可预测”,那么机器人绝不可能是单纯的机器。就像所有的交互活动一样,人机交互也能产生意图之外、意料之外的意义。问题的关键并不仅仅在于人机交互中机器人会产生预期外的行为,这一点已被记录在案(Lemaignan, Fink & Mondada et al., 2015),而在于人类也能够并将以预期外的方式解读人机间的交互行为。例如,机器人一个原本意在展现友好态度的手势,可能会被人类骤然理解为展现敌意。这样预期外的解释方式是无法被完全规避的,这本身便是我们存在于世的一种方式。但是,这样于预期外涌现而出的意义,并不能在概念层面构成完全的他者性,因为这一意义的涌现本身便依赖于人类对意义的建构与体验。我们总能将机器人的行为与我们从(其余的)人类世

界及个人社交环境中了解到的事情进行比较,总能在机器人的行为中看到它与人类行为的相似性,以及机器人与我们自身的联系性。所以机器人从未与人类完全不同,从不是完全的他者。

(五)权力

一旦从交互活动,以及意义建构、社交使用与物质构建等人类活动角度入手理解社交机器人领域,便会看到其总是与权力(power)相纠缠。社交机器人并不仅仅是工具,使用社交机器人或与之交互也并非单纯的"技术"活动。相反,它具有社会与政治层面的意义与影响,这便涉及权力问题。例如,布莱森所使用的"奴隶"(slaves)(Bryson,2010)一词便极有问题(Coeckelbergh,2015),最近布莱森本人也承认了这一点。而在卫生保健领域应用社交机器人,则可能提出关于现有卫生保健体系的政治问题:如今的卫生保健实践是如何开展的?病人被如何对待?这一开展方式与待人方式效果如何?采取这一方式背后的理由又是什么?若要在特定场景中引入社交机器人,那么它们服务于谁的利益?权力与技术是相连的,天真的工具主义者倾向于将技术与人/社会分离开来,所以忽视了这一点。他们自认为稳妥地将机器人归于"技术"一类,却无意中将关于社会与权力的"与人相关的"问题排除在视野之外。(有时工具主义者并没有那么天真,而是故意对社会关系与权力的问题视而不见。)但社交机器人也并不是"他者"或"非人类"——如果"他者"一词意味着它们仅仅是令人愉快的、权力中立的后现代文化与物质游戏中的伙伴(我相信,贡克尔本人规避了这一观点,但这仍可能是对贡克尔或哈拉维式后现代主义的一种解读方式)。如果细究哈拉维理论的政治面向,同时从马克思(Marx,1976)与福柯(Foucault,1980a)处汲取思想资源,那么我们应当指出这样一种危险,即社交机器人这一研究项目能够被嵌入支配(domination)、性别不平等、资本主义社会经济体系,以及人类互动与组织层面微观和中观的权力游戏,并对其起积极的推动作用。后人类主义者将机器人视作他异者(alterities)的幻想却会遮掩这一点。例如,机器人可被用来收集数据并出售给第三方,使用机器人可能会夺走人类的工作,机器人可能传播对待妇女的错误方式,生产机器人可能涉及对人的剥削,机器人可被大公司或国家用来炫耀自身力量,等等。这些与权力相关的问题表明,后人类主义并不导向美好的乌托邦,相反,这恰恰展现了一个敌托邦(dystopia)的现实性与可能性。社交

机器人很可能将机器人呈现为“他者”或“朋友”,但在帷幕之后(事实上这一点也并没有被完全遮盖),或许存在着操控、剥削与惩戒。关于这一点,我们可以考察一下使社交机器人得以成为可能的劳动关系。一味赞美机器人的他者性与他异性(alterity),一味欢迎将非人纳入社会的论调,可能使我们漏看更大的社会经济语境,并掩盖福柯(Foucault,1980b)所说的,“权力存在于所有类型的关系中”的事实——包括我们在开发与使用社交机器人时与他人构建的关系。同时,人们能够使用机器人从其他人身上提取数据,进而影响他们的“身体与灵魂、思想、行为,以及存在方式”(Foucault,1988)。将视野局限于机器人之他者性上,也会使我们忽视这一点。例如,社交机器人能被用于操控他人购买特定物品,或刺激他人做出其他符合特定方利益的行为,但操控方在人机互动中却是隐身的。在机器人的拟人化面具背后,躲藏着严肃的人类权力与权力关系,而它们假扮成一场俏皮的邀请,邀请我们与准他者进行权力中立的、愉快且有益的互动。关系性带来的不全是乐趣,而游戏也并不只关乎玩乐,从这一角度出发考察社交机器人(事实上还有人类自身),则打开了一只关于权力问题的潘多拉魔盒。社交机器人这一领域是关于(维特根斯坦意义上的)游戏的学问,其中一些游戏事实上十分严肃——它们是关乎权力的游戏。当然,任何数字工具中都存在权力问题,但这并不是回避社交机器人领域中权力问题的借口。

(六)第三种观点的启发:作为关系中的工具的机器人

第三种观点澄清了在机器人的使用与开发中所涉及的人类因素,使我们能够批判性地探讨人类将机器人进行拟人化的行为。正是因为社交机器人与人类及社会深深相连,我们才能够通过考察机器人嵌入其中的大整体——如关于性别的具体游戏——来识别什么是有问题的(或好的)拟人化行为,并解释其为何有问题(或为何是好的)。此外,这一立场将机器人视作意义构建的工具,并看到机器人是更大范围语境的一部分且同时是这一语境的共同构建者,由此规避了天真的工具主义式观点。机器人是我所说的——可类比更大范围的文本语境——更大范围的技术语境(con-technology)(Coeckelbergh,2018b)的一部分。正如一段文本是更大范围的文本语境的一部分,并依靠后者来确定自己的含义,技术同样也是更大范围的、充斥着意义的、人类社会关系环境中的一部分。它们扎根于这一环境,同时这一环境也塑造着它们的意义。机器人确实是工具,但它

们是关系中的工具(instruments-in-relation)——它们始终与人类,以及它们运作于其中的社会文化领域相关联。天真的工具主义暗示,人类与其工具间的关系是外在且肤浅的。但我所倡导的进路拒绝接受这一假设。我对拟人化之争的解构,依赖于批判人机间被预设的外在关系——人类和机器人以上述多种方式相互纠缠,彼此间的关系是内在且深刻的。那么,上述进路与看待人机关系的第三种观点,对评估拟人化现象而言意味着什么?上述批判性讨论又将导向怎样的规范性立场?

四、重审拟人化现象,以及作为研究项目与学科领域的社交机器人

(一)机器人既不是他者,也不是单纯的机器

首先,根据第三种观点,我们不仅需要拒斥天真的工具主义与不加批判的后人类主义,同时也需要拒斥与它们紧密相连的各类规范性立场。我的建议是,对于社交机器人领域中存在的拟人化现象与拟人化行为,我们不应谴责其僭越了机器之所是,也不应将其(错误地)视作向机器投射他者性的良机。在前文论述的许多层意义上,社交机器人——包括拟人化的社交机器人——或许会被用户体验为他者,但它们绝非绝对的他者,仿佛茕茕孑立于人类世界之外。社交机器人是人造的,并且参与了人类的意义建构过程。因此,它们并不是全然的他者,我们不能忘记这一点。①

但社交机器人也不是单纯的机器。拟人化也是人类进行意义建构的一种方式——即便机器人的设计本身并不鼓励人类对它们进行拟人化,拟人化行为也仍然会出现,因为机器人(包括社交机器人)的意义本身就总是与更广泛的人类意义建构行为,与更开阔的意义视野,也即维特根斯坦意义上的"生活形式"有关。前文呈现的进路并不能直接指导我们,应该接受还是拒斥机器人开发者/设计者或用户的拟人化行为;但其确实能够帮助我们认识到,任何关于特定人机交

① 我可以补充说,社交机器人也并不是朋友或伴侣意义上的"他者",因为我们都知道,友谊或伴侣关系之所以有价值,是因为朋友与伴侣并非由我们创造,他们具有他者性。虽然在与机器人交互的过程中,或许会有一定程度的他者性涌现,但归根结底,友谊或伴侣关系中的他者性是无法被人为创造的。

互案例的讨论,都离不开具体互动行为之上的更大意义语境,也即我们的社会与文化,以及其中所蕴含的所有面向——包括权力在内。这意味着,如果不将这种更广阔的关系性、意义建构行为,以及规范性要求纳入考量,那么工具主义立场内对拟人化现象的支持意见也是站不住脚的。

(二)重新定义社交机器人领域的目标

上文刻画的关系性与解释学进路也应被用于理解作为学科领域的社交机器人。社交机器人是一个跨学科、多学科的领域,不仅关涉组合硬件软件,打造人们所称的“机器人”(工具主义立场)或“人造他者”(artificial others)(后人类主义立场),同时也关涉人类、社会、文化,以及政治层面的意义建构(从政治意义建构的角度看,将机器人称为“他者”的说法是具有政治意味的,必须对这一说法进行质疑)。

尽管我已指出特定观点与相应的规范性立场中所存在的问题,但本文并没有得出是否应当建造拟人化机器人的具体规范性结论。人们至多能够从中得出一项警示性建议:如果你要设计此类机器人,那么请注意,你很有可能因此参与进了更广范围内的社会建构活动,而其中处处牵涉着权力问题。出于上述原因,或许有学者会总结道:还是不开发此类机器人为妙。也有学者会坚持开发此类机器人,但认为机器人学家有责任在开发时将机器人社会文化层面的意义与后果纳入考量范围。这便引出了我关于机器人学家责任的观点。

(三)机器人学家、教育者以及用户—公民的责任

就规范层面而言,承认社交机器人与机器人学的关系性、解释学以及政治面向,意味着机器人学家、计算机科学家、工程师、心理学家,以及该领域涉及的所有其他人,都同样要为他们产生的意义与社会—关系性后果负责。这一点不应被边缘化;相反,这应成为社交机器人领域的核心之一。例如,机器人硬件的设计者和机器人语音识别软件的开发者,都必须意识到,他们并不是在单纯地开发技术产品,这些产品拥有一些计划内的功能(比如,能够在家庭环境中与人类进行对话)。同时,他们也是经由机器人而介入社会游戏领域(social playing field)的社会行为者,而这一领域中的挑战并不仅仅是技术性的(比如,规避或减少性别偏见的挑战)。如此一来,社交机器人学家的责任既是确保人机交互能

够达成其目的,同时也在于确保人机交互能以对他人负责的方式展开。机器人的设计方法、编程方式等都可能产生人类预期之外的后果(包括潜在的解释学/符号学影响),所以后一项责任首先意味着,需要对用户以及任何可能受预期外后果影响的利益相关方进行回应。同时,后一项责任也意味着,需要对那些技术问题外的社会问题进行回应,并探索如何通过技术侧的改变来舒缓社会侧的问题。

为机器人在解释学/符号学层面潜在的社会影响划定准确的责任边界并非易事。我们无法完全预知未来,而一个人需要在多大程度上对那些预期外产生的意义负责,目前也仍不明确。如前所述,无论在机器人领域中,还是在人类生活与文化的其他领域中,我们在解释层面都无法达成绝对的控制。但我们应当尽可能加大控制的力度,无论是针对我们意欲产生的意义,还是对意料之外、潜在的(未来)意义。我们可以期待,机器人研究者与设计者(以及他们的公司或组织中的其他人)至少会尝试对意料之外的潜在影响进行预测与评估。而其他领域的研究者,包括人文学科与社会科学学者,也可以为此展开方法论层面的研究——如开发或移植伦理评估、场景创建的方法——从而贡献自己的力量。

考虑到机器人在社会中的嵌入性,社交机器人领域涉及的责任也包括为社交机器人在权力方面产生的意义与后果负责。机器人技术的研究人员或许并不认为自己与权力相关,因为在通常情况下,权力被认为仅与政治机构和公民间的关系有关。然而,福柯式的解读方式引导我们关注社会结构的各个层面、各个地点中无所不在的权力问题,机器人研究领域自然也包含其中。机器人的开发者、设计者,连同相关决策的制定者、各种场景下机器人的使用者一道,在塑造实践的意义、人们的故事,以及我们作为人类为自己设定的目标方面,是拥有一定程度的权力的。拟人化现象是机器人实现社交能力的支柱,所以相对其他种类的机器人,上述权力在拟人化机器人开发方面,或许具有更为重要的意义。伴随这一权力同来的,则是责任。

权力具有消积面向,可能是制约性与压迫性的。但除此之外,权力也可能是赋能性的,可以使人们做到更多的事情,增加他们的潜能。因此,在社交机器人语境中,问题便在于,我们该如何将权力授予专业技术人员,使他们能够为自己的创造物负起责任。不幸的是,传统意义上的工程师与科学家缺少完整的技能与知识储备来做到这一点,因为在过去,他们的关注重心仅是技术层面的知识与

技能。这一现状亟须改变。在方法论层面,我们需要一种不仅是多学科的,同时也是跨学科交叉的方法。这意味着,我们不仅需要计算机语言与工程层面的工具,同时也需要来自人文社科方面的工具。我们应当用关系性的眼光来审视社交机器人这一项目,并将其与对人类社会、文化的批判联系起来。我们必须做到这一点,因为如前所述,机器人研究与更大范围的社会文化环境、语言、权力等因素相关联,而对机器人的制造与使用,一方面塑造了上述关系性领域,另一方面它自身也受后者塑造。如果机器人学家接受了这一责任与挑战,并重新定义他们在这一更大整体中的位置,那么他们便能以一种对伦理与社会负责任的方式,在塑造未来社会方面发挥更重要的作用——这也是工程师们一直以来所致力的事业,从建造桥梁与铁路,到发现电力、制造手机,无不如此。

第三条进路蕴含的规范性与方法论意涵要求在机器人领域引入跨学科的研究方式,并在机器人工程师与计算机科学家教育方面进行改革。当他们开启职业生涯时,他们不仅应当愿意,更应当有能力且准备好将他们在机器人领域内的实践、开发以及设计决定与更广阔的社会背景及文化话语相联系。若要实现这一点,他们需要有能力使用相关词汇与事物(机器人)批判性地参与社会文化讨论。但是,也不能单纯地将责任全部划归给机器人学家。就用户—公民而言,他们也需要为社交机器人的未来发展担负起责任,或者更笼统地说,需要为科技与社会的未来负起责任。他们需要(经由教育)理解,他们在日常生活的不同场景、不同技术语境中遇到的社交机器人,既不是"他者",也不是单纯的机器或工具。相反,社交机器人是他们的社会与文化的一部分。也正是出于这一原因,用户—公民对社交机器人负有责任。如果我们不将机器人视作人类领域的对立面,而是作为人类领域的重要组成部分,那么机器人便可能以我们从未设想过的方式作用于人类生活。通过跨学科教育将来自不同领域、说着不同语言的人们联系在一起,或许能将我们导向对技术能力的不同理解方式,同时使我们以全新的方式理解技术作用于社会的方式,并最终在理解技术与社会两方面都达成深刻转变。这便是学生们——不仅是具有科学与工程背景的学生,同时也是具有人文社科背景的学生——需要做好的准备。

(四)一般结论

在最根本的意义上,我认为,我们需要更多的拟人化行为,而非更少。我指

的未必一定是由机器人的设计特征及互动方式所引发的拟人化行为。这一类型的拟人化行为增强并放大了机器人与其社会文化环境间已然存在且无法规避的联系,使得这一联系本身变得既令人激动又问题重重(比如,会使我们原本试图减少的特定权力关系激增)。对此,机器人的设计者与用户—公民需要进行详细的审查与讨论,他们对此类机器人的设计肩负着共同的责任。那究竟在何种意义上,我们需要对社交机器人进行更多拟人化呢?正是在认识到机器人与其解释学环境间确实存在联系、确实存在深刻的社会与语义关系的层面上,我们需要对机器人进行更多的拟人化——我们需要理解,无论是在实验室内还是在更广阔的社会层面,就像其他技术一样,社交机器人具备非常人类化的面向。在机器人的具体行为之外,社交机器人领域内的拟人化现象正展现出上述关系性与人类性——而它们提醒我们,技术一直以来都与人类相关。

参考文献

Bartneck, C., D. Kulić & E. Croft et al. 2009, "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots." *International Journal of Social Robotics* 1(1).

Bryson, J. 2010, "RobotsShould Be Slaves." in Y. Wilks ed., *Close Engagement with Artificial Companions*, Amsterdam: John Benjamins.

Caliskan, A., J. Bryson & A. Narayanan 2017, "Semantics Derived Automatically from Language Corpora Contain Human-Like Biases." *Science* 356(6334).

Coeckelbergh, M. & W. Reijers 2016, "Narrative Technologies: A Philosophical Investigation of the Narrative Capacities of Technologies by Using Ricoeur's Narrative Theory." *Human Studies* 39(3).

Coeckelbergh, M. 2011, "You, Robot: On the Linguistic Construction of Artificial Others." *AI and Society* 26(1).

Coeckelbergh, M. 2012, *Growing Moral Relations: Critique of Moral Status Ascription*, New York: Palgrave Macmillan.

Coeckelbergh, M. 2015, "The Tragedy of the Master: Automation, Vulnerability, and Distance." *Ethics and Information Technology* 17(3).

Coeckelbergh, M. 2017, *Using Words and Things: Language and Philosophy of Technology*, London: Routledge.

Coeckelbergh, M. 2018a, "How to Describe and Evaluate 'Deception' Phenomena: Recasting the Metaphysics, Ethics, and Politics of ICTs in Terms of Magic and Performance and Taking A Relational and Narrative Turn." *Ethics and Information Technology* 20(2).

Coeckelbergh, M. 2018b, "Technology Games: Using Wittgenstein for Understanding and Evaluating Technology." *Science and Engineering Ethics* 24(5).

Coeckelbergh, M. 2019, *Moved by Machines: Performance Metaphors and Philosophy of Technology*, New York: Routledge.

Darling, K. 2017, "Who's Johnny? Anthropomorphic Framing in Human – Robot: Interaction, Integration, and Policy." in P. Lin, G. Bekey & K. Abney et al. eds., *Robot Ethics 2. 0: From Autonomous Cars to Artificial Intelligence*, Oxford: Oxford University Press.

Dautenhahn, K. 2007, "Socially Intelligent Robots: Dimensions of Human – Robot Interaction." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1480).

Duffy, B. 2003, "Anthropomorphism and the Social Robot." *Robotics and Autonomous Systems* 42(3-4).

Eyssel, F. 2017, "An Experimental Psychological Perspective on Social Robotics." *Robotics and Autonomous Systems* 87.

Flusser, V. 1999, *The Shape of Things: A Philosophy of Design*, London: Reaction Books.

Fong, T., I. Nourbakhsh & K. Dautenhahn 2003, "A Survey of Socially Interactive Robots." *Robotics and Autonomous Systems* 42(3-4).

Foucault, M. 1980a, *Power/Knowledge: Selected Interviews and Other Writings 1972—1977*, New York: Pantheon Books.

Foucault, M. 1980b, *Power/Knowledge*, New York: Harvester Wheatsheaf.

Foucault, M. 1988, *Technologies of the Self: A Seminar with Michel Foucault*, Amherst: The University of Massachusetts Press.

Freud, S. 2003, *The Uncanny*, London: Penguin Classics.

Goetz, J., S. Kiesler & A. Powers 2003, "Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation." in Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication, Millbrae, CA..

Gunkel, D. 2007, *Thinking Otherwise: Philosophy, Communication*, West Lafayette: Purdue University Press.

Gunkel, D. 2017, "The Other Question: Can and Should Robots Have Rights?" *Ethics and Information Technology* 20(2).

Haraway, D. 2000, "A Cyborg Manifesto." in D. Bell & B. Kennedy eds., *The Cybercultures Reader*, London: Routledge.

Hegel, F., S. Gieselmann & A. Peters et al. 2011, "Towards A Typology of Meaningful Signals and Cues in Social Robotics." in Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication, Atlanta, GA.

Ihde, D. 1990, *Technology and the Lifeworld*, Bloomington: Indiana University Press.

Ishiguro, H. 2006, "Interactive Humanoids and Androids as Ideal Interfaces for Humans." in Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney.

Johnson, D. 2006, "Computer Systems: Moral Entities but not Moral Agents." *Ethics and Information Technology* 8(4).

Kiesler, S. & P. Hinds 2004, "Introduction to This Special Issue on Human-Robot Interaction." *Human-Computer Interaction* 19(1-2).

Kiesler, S., A. Powers & S. Fussell et al. 2008, "Anthropomorphic Interactions with A Robot and Robot-Like Agent." *Social Cognition* 26(2).

Krägeloh, C., J. Bharatharaj & S. Kutty et al. 2019, "Questionnaires to Measure Acceptability of Social Robots: A Critical Review." *Robotics* 8(4).

Kuchenbrandt, D., F. Eyssel & S. Bobinger et al. 2013, "When A Robot's Group Membership Matters: Anthropomorphization of Robots as A Function of Social Categorization." *International Journal of Social Robotics* 5(3).

Latour, B. 1993, *We Have Never Been Modern*, Cambridge, MA.: Harvard University Press.

Lemaignan, S., J. Fink & F. Mondada et al. 2015, "You're Doing It Wrong! Studying Unexpected Behaviors in Child-Robot Interaction." in A. Tapus, E. André & J. Martin et al. eds., *Social Robotics* (Vol. 9388), Cham: Springer.

MacDorman, K. & H. Ishiguro 2006, "The Uncanny Advantage of Using Androids in Cognitive and Social Science Research." *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems* 7(3).

Marx, K. 1976, *Capital*, London: Penguin Books.

Mori, M. 1970, "TheUncanny Valley." *Energy* 7.

Peca, A., M. Coeckelbergh & R. Simut et al. 2016, "Robot Enhanced Therapy for Children with Autism Disorders: Measuring Ethical Acceptability." *IEEE Technology and Society Magazine* 35(2).

Reeves, B. & C. Nass. 1996, *The Media Equation: How People Treat Computers, Television, and New*

Media Like Real People and Places, Cambridge: Cambridge University Press.

Richardson, K. 2018, "Technological Animism." in K. Swancutt & M. Mazard eds., *Animism Beyond the Soul*, New York: Berghahn Books.

Scheutz, M. 2011, "The Inherent Dangers of Unidirectional Emotional Bonds Between Humans and Social Robots." in P. Lin, K. Abney & G. Bekey eds., *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge: MIT Press.

Sparrow, R. & L. Sparrow 2006, "In the Hands of Machines? The Future of Aged Care." *Minds and Machines* 16(2).

Turkle, S. 2011, *Alone Together: Why We Expect More from Technology and Less from Each Other*, New York: Basic Books.

Złotowski, J., D. Proudfoot & K. Yogeeswaran et al. 2015, "Anthropomorphism: Opportunities and Challenges in Human-Robot Interaction." *International Journal of Social Robotics* 7(3).

计算守法主义和法律延迟的示能*

劳伦斯·迪佛　著**
暴文博　译***

摘要:延迟性是我们所知的法律的核心要素——解释法律规范并质疑其要求的能力取决于文本为公民提供的时间的空间性。当法律实践与应用进一步引入计算系统时,这些空间面临着崩塌的威胁,因为"计算守法主义"的即时性抛弃了文本天然的缓慢性。为维持该法律保护的性质,我们要明确延迟性在法律过程的哪些方面发挥着规范性作用,并确保它被如实反映在被引入的计算系统的示能上。这必然要求我们关注这些系统的设计和生产,并抵制弥漫在当代设计实践中的"效率"意识形态。

关键词:示能　效率　法律性　计算守法主义　人工智能

一、导论

文本的一个真实且首要的示能(affordance)①是表征、理解和行动之间的延迟,可以称该空间为"解释性间隙"(Diver,2018)。② 在法律领域,该间隙存在于组成法律规范的文本和文本在世界中通过改变行为的方式产生的实例化和具体化之间。为了发生规定行为,为了应用规则,必须要有解释发生的空间——正如邦科夫斯基和麦考密克(Ba′nkowski & MacCormick,2000)所言,适用规则也是解

---

* 本文原题为"Computational Legalism and the Affordance of Delay in Law",原刊于 *Journal of Cross-Disciplinary Research in Computational Law* 1(1),pp. 1-12,经授权翻译发表。

** 劳伦斯·迪佛,布鲁塞尔大学法律、科学、技术与社会研究组。

*** 暴文博,中国政法大学法学院。

① 这一词语最早由美国心理学家詹姆斯·吉布森(J. Gibson)创造,用以描述动物和自然界物质的关系,强调自然中某些要素能满足动物欲求的性质,后被唐纳德·诺曼(D. Norman)应用在设计学中,阐述物理对象因其性质给人提供的功能。例如,因为椅子可以支撑人体重量,人可以坐在上面,所以供人休息就是椅子的 affordance。本文中我参考中文世界此前的翻译惯例,将 affordance 翻译成"示能"。参见唐纳德·诺曼:《设计心理学 1:日常的设计》,小柯译,中信出版社 2015 年版。——译注

② 一个真实示能的存在无关该示能是否能被感知。

释规则。在这个空间里,时间必然会流逝。简而言之,法律主体会考虑文本对她意味着什么,用词如何在具体情况中应用,解释结果如何表明法律要求她采取(或不采取)一定行为(Rowell,2019)。①

所有这些都非即时或所予(given)的。规则含义和目的对公民来说经常是神秘的,规则和她的确切关联也会被掩盖。因为她会被"法律的生活形式"(form of life)排除在外,而且组成法律的"法律用语"(lawyer-speak)经常晦涩不明(Introna,2011)。尽管如此,法律语言确实和日常生活用语有很多重叠——外行的公民经常谈论权利、合法与非法——同时公民至少有机会学习到部分生活的法律形式,并理解它们。如果她们必须通过诉讼解释或抵抗某些东西,她们也可以咨询专家(专家在自己准备案件时也会遇到某种拓展性解释性间隙)。

本文中我讨论了有意识地为计算系统提供缓慢性或延迟性的想法。对当代生产过程和其本质来说,这一目标让人厌恶。在考虑研究延迟性在技术架构中作用的现有文献前,我将先阐述自己关于解释在法律中的作用的想法。本文的核心想法是,为促进我们所知的法律赖以存在的解释类型和评估类型,我们需要考虑如何、在哪里有意识地将延迟性植入参与法律过程的机器中。当然,计算已经在法律实践中扮演了多重角色,并且是在无数的交汇点上。存在一个由行动者、技术形式和目的构成的复杂矩阵,每人扮演的角色取决于行动者进入法律过程的时点。② 本文不是要剖析这些架构,而是要强调一些重要主题。这些主题应当被那些设计用于法律世界的计算系统和那些决定何时、何地、如何与为何使用这些系统的人铭记于心。

二、法律文本和解释学

文本的消极"物理特性"必然带来上述延迟性;从另一角度看,我们可以说文本给读者"提供"(afford)了"延迟性",也因此提供了诸如思考空间、理解和解释文本机会等其他事物(Norman,2013;Gibson,1977)。就本身而言,文本对行为

① 当然,这里预设了识别相关法律的动机和能力。这些并不是所予的,也会受到关于法律是什么或应当为何的非专业思考模式的影响。在任何情况下,对法律的事后争论必然要依赖发生在某刻的解释学间隙。

② 例如在规范性上,律师助理制作和管理文件就和诉讼律师用于预测判决的算法或自我施行的合同规范编码有着天壤之别。这种差异既体现在涉及的技术上,也体现在法律事业的规范性影响上。

可以没有任何规制效果,纵使文本在生活的法律形式里有些许法律效果。然而,公民的关注点不太可能是在特定时间内的特定法律效果。相反,对她来说合适的问题是简单的“法律要求我做什么?”,该问题的部分答案存在于法律文本之中,但法律文本不可直接强加答案。解释学间隙提供了一个空间,公民可以通过反复且多层的解释在该空间中思考答案。正如利科(Ricoeur,1977)所言,我们首先通过信心解释学(hermeneutics of faith)[①]赋予文本意涵,接受文本条款和其在表面上预设的现实。我们之后(应当)进入一个更深层次的解释性问询,即“怀疑解释学”(hermeneutics of suspicion)。在这里我们探寻文本预设了何种现实,以揭示“文本中蕴藏的权力关系、冲突和利益”(Introna,2011)。

信心和怀疑的“双重解释学”和民主社会中的法律争论密切相关。信心解释学提供了对文本的天真解读,其接受了文本的表面价值——这同过度守法主义(legalistic)类似,即对法律的文本主义解读,而相悖于法律性[②](我随后将讨论这些概念)的愿景,后者期望解释者自主且反思性地接触文本的深层含义。更深层次的接触反映在怀疑解释学上,该解释明确并质疑了文本的预设现实,这意味着对文本更合理的考虑和我们对文本的回应(默许、分歧、争论、需要求助于专家的困惑等)。法律主体(及其代表)揭开了文本面纱,以看清背后的东西。

纵使过度的怀疑解释学可能导致对文本的愤世嫉俗式解读(Introna,2011),原则上,考虑到国家对暴力的垄断,这并不坏。如果怀疑不被准许,被要求的只有信心,那么将没有深层问询的必要,也不可能或不需要争论:我们可以按照表面的文本进行。事实上,这是更强形式的守法主义背后的意识形态基础和下文讨论的计算守法主义(computational legalism)的本体论特征。如果可以接受这

① 在利科解释学中,hermeneutics of faith 和 hermeneutics of suspicion 是两个重要概念,我参考华中科技大学何卫平教授的翻译,分别将二者翻译成“信心解释学”与“怀疑解释学”。参见何卫平:《信心解释学与怀疑解释学》,《哲学研究》2017 年第 5 期。——译注

② 从语义学角度讲,legalism 和 legality 都是由形容词 legal 增加后缀变成的名词。但作者对二者有完全不同的评价,而对二者的区分也是理解本文的重点。具体说来,作者对 legalism 这种意识形态持贬义态度,认为这是一种机械适用法律、盲目追求效率、全然不顾人类价值的错误意识形态。在计算系统中保留人类的思考空间,就是要让强调思考和人性的 legality 取代错误的 legalism。这种批判和美国政治学家朱迪丝·施克莱(J. Shklar)对 legalism 的批判类似,因此,我将 legalism 翻译为“守法主义”,以呼应施克莱的《守法主义》一书与中文翻译。参见朱迪丝·施克莱:《守法主义:法、道德和政治审判》,彭亚楠译,中国政法大学出版社 2006 年版。而因为 legality 是将 legal 转变为了名词,后者的含义是某事或者某项行为符合法律要求,作者在文中也想用该词描述法律的本来特性,所以我将其译为“法律性”。还请读者留意作者在文中对两个词汇的不同态度。——译注

样的天真,那么法律自动化和随之而来的解释学间隙的崩塌将不仅是直截了当的,还将是让人向往的。因为这样的做法——浪费时间为一个既定结论留有思考空间——将毫无价值。

在时间中的空间——延迟原本会瞬间发生的过程——是促进怀疑解释学的重要部分。在计算出现前,文本给予其人类解释者的延迟性很少需要被点出;媒介在“本体论上的摩擦性”(Floridi,2013)让延迟性成为简单事实,和一个无法想象任何替代方案的深层预设(Hayles,2004)。

极度显著的文本“摩擦性”不应让我们忽视这样的事实,即延迟的示能不是所予的。人类社会多次经历了文本赋予可解释性方法的显著变迁:在手写文化下,文本被手抄保存,因此十分昂贵和稀有,文本在地理和时间上的覆盖面也受到控制其使用权的(宗教)精英限制。宗教机构代表不识字的公众解读文本,没有或者只留下很小的空间供公众思考,也因此拒绝了一部分受制于那些解释的人的怀疑性抵抗(和拒绝服从的后果)。随着印刷术的出现,普通公民有机会拥有文本并思考其含义——以此收获解释性间隙的成果——这种字面意义上的怀疑让宗教改革成为可能(Eisenstein,2012)。

采用新技术的印刷机带来的新颖示能不是所予的。同样,当我们引入和发展新的计算科技时,我们必须敏感于计算科技必然带来的示能变化。这在确定这些变化事实上是什么,以及这些变化对某些制度——例如法律(其特质在不同程度上取决于此前的示能的构造)——的意义时都是如此(Hildebrandt,2015)。

三、法律性

如上文所述,对文本持怀疑态度的想法和守法主义的愿景性概念相关。法律性认为法律解释不是冰冷的机械操作,而是涉及法律主体的推理和理由给予之事。这些主体作为有尊严、自主和能动性的人而受到尊重(Waldron,2011;Ba′nkowski,1993)。这指向了一个存在于异治(heteronomy)和无政府主义间的供思考和理性行动的空间。个人在该空间受到了足够尊重,可以在超越无条件接受其条款的情况下考虑规则。这便蕴含着对那些条款进行争论的可能。我们争论在特定时空下那些条款的含义,和其因此蕴含的权利与义务。

争辩与论争需要有进行对抗性交锋的机会,即对抗性(adversarial)辩论,该辩论让相反观点得以表达、有意义的妥协得以实现(Hildebrandt,2018)。这要求我们尊重在解释问题上有着合理分歧的每个人。当然,法律性表征的灵活空间最终不是无限的,法律真理的存在、性质和范围取决于在生活中的法律形式,即法院中被认可决定真理的先例。在那里,现存法律的普遍性可与个案中影响法律适用标准的独特混合因素相平衡。这些裁判性事实并不是用科学方法探查到的经验事实;它们产生于解释和评价的过程,尽管在生活的法律形式中提出的命题性论证也包括预测性经验事实,即证据。那种生活形式需要这样的论证,其将对一般性法律标准的解释与个案证据评估综合起来,后者从经验性原初事实(brute fact)转换为法律制度性事实(MacCormick,1992)。

在该过程中,反面考察因素也可以进入论争性战场,例如分配正义、解释性融贯或正当程序问题。诸如德沃金式整全法(Dworkin,1986)或拉德布鲁赫式正义、法律确定性与合目的性的生成性张力(Radbruch,1950)等理念也在该背景中起到了规制性的作用。这些理念在每个案件中都应被有目的地(重复)应用,并根据案件特殊情况和发生时的政治历史时刻决定将各方论证拉向这方或那方。在这种情况下,直接解释和给出的理性论证不是程序的必要之恶,我们在理想世界也不必抛弃它以直接获得法律确定性"圣杯";相反,它们是构成法律性质中规范性视野的要素,围绕着法律性愿景而建。

这种评估和适用基于价值的标准的过程,最终能让已有规则覆盖新的案件。诚如麦考密克(MacCormick,1992)在 30 年前所言:"纵使我们脑海里知道所有先例,以及迄今为止每次考察这些先例时涉及的要素和权重,我们也会在新案件中发现新的因素。"(这本质上正好描述了以机器学习为基石的法律预测系统声称要达成的目标。)我们不能,也不可能机械地应用全部程序,总是有可能通过推理在新案件中发现符合评估标准的语言、政治和证据因素:"适用这些理由是理解新信息而非适用旧信息的过程。"(MacCormick,1992:193)[①]这种灵活的法律智能在根本上与当代"人工智能"系统相矛盾。后者或许"擅长"处理范围狭窄、定义明确的任务,但不具备通过综合性经验、先前知识和当前目标来发展新颖、有效技能的能力(Chollet,2019)。鉴于法律推理的基本目标是调和人类事务

① 仅依据这个理由,机器学习(至少是迄今为止)就缺乏资格,因为机器学习只能处理过去的数据。

中的持续复杂性问题,在此意义上,法律推理的范围并不狭窄,定义也必然不明确。作为一种实践,它涉及一系列定性的技能和只是偶然相关的知识来源(隐性的、文本性的、“怀疑性的”、证据性的)。作为即时过程的一部分,如果法律推理要保持这些基本特征,它就不能被预先确定。即使在普通案件中,一个法律相关的决定也总是一个新颖、分离的推理。该决定要考虑到此前情况,并将其综合于当前案件的突出因素。立法解释和法庭论辩均是如此。

当然,在某种意义上,机器学习也可以依据数据中的过去行为模式进行推理,在新案件或法律适用中提供明显新颖的“发现”。但根据前文勾勒的规范性法律观念,这些数据(法律文本)不可能提供一个有效法律故事中的所有要素。在机器学习可进行的推理范围里,机器学习无法提出法律必须为其服从者提供的那种真正的新颖论证。例如,提出一个不同但符合语境的词语意义论证,或为了实现正义,将一项明显不相干的原则适用在新的背景里。如拉德布鲁赫(Radbruch,1950)所言,语言学解释只是一个起点——该解释只能指引法律论证之船驶离港口,此后它必须自己在公海里航行。而机器学习甚至没有能力离开港口,进行真正的语言学解释,更不用说在公海里冒险,进行真正的推理论证了。

如果一个法律论证成功融合了相关法律规则和原则、当前案件的具体情况和引导其解释的规制性理念,那么该论证将被法院认可为法律真理。当然,它也会受到可能的上诉或立法机关制定的相反规则的限制。主张这个过程不可能是机械的,不仅是对执行此种推理的机器学习的能力做出主张,而且也是做出关于该过程本质的规范性论证,即为了符合法律要求,法律论证应该先验地由人类执行。能成功——如果是暂时地——主张法律真理的生活形式有着无法被更改的制度性(MacCormick,2007),其建基于言语行为、过程正当性,并敏感于人类境况。因此,它是一种无法被更改的人类生活形式。计算主义思维对于将无可避免地改变其性质的媒介是不敏感的(这或许包含了最大程度的善意),而我们在面对计算主义思维之时,应该毫不犹豫地保护和加强法律论证的品质。至少我们必须明确,当前法律的哪些示能不可被侵犯。这应以支持法律的技术示能为基础,也就是文本(Diver,2018;Gaver,1991)。[①] 一旦我们明确了必须保护哪些示能,我们可以更好辨明,当裁判技术本身发生变化时,示能会如何变化、增强或

① 在人造物层面,这样的示能被称为“序列性”(sequential),对一项示能的使用揭示了另一项示能。

被移除(Hildebrandt,2015;Hildebrandt,2020)。

四、计算守法主义

与上述法律性的观点相反,守法主义——至少是其较强的变体(Wintgens,2012)——要求人们对成文的规则保持缄默。法律"就在那里",公民被期待遵守,而不要多想。在强守法主义下,利科的怀疑解释学完全被信心解释学替代,或者说根本就没有解释——文本含义停留在表面,没有探索(或者这种探索不被允许)能进入文本预设的世界,也就是在立法中蕴含着导致规则产生的那个政治世界。接受规则表面价值的要求源于温特根斯(Wintgens,2002)所言的"主权面纱",或横亘于法律和政治间的障碍,该障碍阻止对规则为何具有如此特质的问题进行法律探索。规则背后没有原则的运作空间,用以规制对规则的解释。政府也因此变得狭隘且"恶毒",法律则沦为"暴政的工具"(Ba′nkowski & MacCormick,2000)。公民应如自动机般行事,如算法般遵守规则(Ba′nkowski,1993)。

计算可能将这种过度守法性理念提升到新高度,它拓展和放大了温特根斯所述的(强)守法主义,并远远地超越了温特根斯的描述。计算守法主义的一个方面是代码执行的飞快速度(Diver,2021)。[①] 延迟性的缺失摧毁了解释学间隙,这不仅是因为文本(源代码)同时构成了规则和现实(Latour,1992),还因为文本的适用早已被预先确定,并在执行时被立刻施行。这种施行不受任何设计者在设计时未考虑的因素的影响。线性时间被压缩,而"用户"(公民、诉讼当事人、法律人等)未被给予任何"犹豫不决"的空间,就像法官进行审议时那样(Latour,2009)。

法律文本在缺少用现实行为回应文本要求的意愿(或由法院强制执行这些要求)时是惰性的,而代码的启用和约束,由于设计者的设置,甚至在系统运行前就有潜在效力。因此,解释学间隙崩塌的程度极深,这不一定源于恶意或有意的抹杀,而只是技术的性质所致。"讽刺的是,"利斯(Leith,1988)说道,"正是计算机程序基于规则的性质反对了法律基于规则的性质",作为代码典型特征的

① 计算守法主义的其他方面还包括规则狂热性、不透明性(opacity)、即时性、不可变性(immutability)、普遍性和私人生产性。

规则狂热性(ruleishness)[①]让其与更广阔的背景隔绝。这同依赖背景和隐性知识的法律性理念形成鲜明对比。

该过度守法主义特质并不局限于将计算机视作独立的监管“模式”——也就是说,在法律的规定范围外,代码也在构成和管理行为(Lessig,2006;Diver,2021)。风险还存在于法律工作者将计算系统视作“工具包”中的要素,或者干脆将法律业务本身外包给它们时。现实的问题是,那些受法律规制的主体是否以及在何种程度上可以发现并回应规则。在“法律科技”的背景下,也存在一种额外风险,即反思性守法主义的出现:由于隐含性地依赖于本体论上的计算“守法主义”,我们面临审议空间的崩塌,这是法律性的基石。代码约束不经意间成了法律约束。

五、作为规范性示能的延迟性

计算守法主义下执行的速度和即时性并不是所予的;同设计其他所有关于法律体系的计算系统一样,该设计总会体现出“意向性”,这种意向性部分由设计者希望该系统拥有的示能所构成(Ihde,1990)。

在系统运作中保留解释学间隙因此就成了一个示能问题,也就取决于设计者的决定(Ihde,2008)。[②] 有意提供延迟性的要求极其重要,考虑到软件工程明显对提升执行效率以服务“无缝”用户体验一事有着无处不在的关注(Storni,2014;McGeveran,2013;Ohm & Frankle,2019),这似乎预设了“低效”和“摩擦”先验地让人厌恶且背离使用者目的(Calo,2014)。但这假定太多目标和价值(工具和内在)的性质了,即那些可能或应当帮助反映在系统中的目标和价值。

重要的是,我主张提供延迟性,不是要求我们抛弃优化设计的努力,进而在计算过程中保留本可移除的延迟性。这样的方法十分任意,因为延迟性保留的范围将取决于:(1)它在某类系统中“自然”出现的位置,和(2)设计者的专业度

① 该词是作者在 rule 一词后增加“ish”与“ness”后缀形成的新名词。“ish”后缀有迷恋的意思,例如“bookish”通常表明某人喜欢读书。作者在这里使用该词是想强调代码对规则的过分迷恋,以至于忽视了其他重要之事。因此,我将该词翻译为“规则狂热性”。——译注

② 伊德(D. Ihde)反对他所称的“设计者谬误”,即认为设计者意图总是会反映在技术的使用上。他的重点是基础设施技术,这些基础在社会中起到基础性作用,也为“多重解释性”留下了空间。我不认为他的观点能很好移转到以代码为基础的个人性人造物的示能上,因为后者定义更为严格。

与责任心(标准和知识都会随何为“优化”之问而变)。在任何情况下这种“非优化”都可能同规范性目标没有多少联系,也就是适用于包含即将应用的系统的更广阔背景——例如法律体系下的“法律科技”——的目标。当这样的优化如果让人向往时,例如减少不必要的能源消耗,保留延迟性甚至是不负责任的。应对延迟性的天真方法会赦免我主张的设计者应有的规范性责任,设计者应担负有意施行延迟性的责任——如果延迟性在相关更广阔的背景下扮演了规范性角色的话,例如促进理解和可解释性。

由此产生的挑战是在以文本导向为基础规范性的法律体系里,辨明延迟性在何处是重要的(以及在何处它是不让人向往的,例如在促进正义的可及性上)。一些延迟性是有益的,其他则是有害的。一旦我们绘制了规范性的地形,我们就可以讨论这些延迟性如何且应当怎样体现在以计算系统为中介的法律实践上,以及在何种程度上这些系统的设计提供了它们。

下面几小节将讨论近期一些关于缺少“速度”的规范性价值的文献,一些更广阔的规范性目标可以通过有意降低计算速度来实现。尽管和法律实践没有明显关联,关注数字人造物设计的领域仍可以极大促进我们对“法律科技”的分析;在法律领域,人机互动问题必然会转变为法律和电脑互动的问题,而质量上和数量上的风险都更高。

(一)让人向往的低效率

欧姆和弗兰克(Ohm & Frankle,2019)提出了“让人向往的低效率”观念,该观念主张有意调节代码(其直接性)的“效率”,以保护其他可能被高效破坏的价值。计算和技术任务是“基本问题”,而其效率可能威胁到的价值则是“增强性问题”,这类问题需要“人类的判断、价值或对成功与失败的定义”。有时增强性问题需要施加给系统以所谓的“低效”,为人类留出空间,做只有人才能做的事情。作者讨论了这样的事例:华尔街的一家证券交易所引入了 38 英里(约 61 千米)长的光纤电缆,携带交易数据的光信号因此需要穿越更多的电缆,交易速度也就降低了。通过设置 350 微秒的微小延迟,高频交易被削减了,也就此避免了不稳定和市场崩溃的风险。另一个与计算法特别相关的例子是区块链应用中的工作证明。在这里,计算和时间密集性技术被混杂地应用到原本近乎一瞬间的事件中,也就是将一场交易记录在数据库中。该延迟随后帮助建立了作为人类

价值的信任和“钟表时间”。

让人惊讶的是,欧姆和弗兰克(Ohm & Frankle,2019)的分析明确去除了这样的理念,即设计一个计算系统,该系统的唯一任务就是降低计算机的运行速度,以匹配人类的处理系统。虽然他们承认此种低效形式的作用,但他们的回避让人遗憾,因为许多的“增强问题”都必然依赖“人类的处理系统”。他们将人类描述为“设备”,该设备显然类似于其他计算机的外设(如打印机或网络摄像头),而主机和外设的通信则需要软件接口(驱动程序)的协调,协调前者(快)和后者(慢)间的内部速度差异。就像“低效率”这个负面条款一样,这种分析似乎是以机器而非人为中心的,将后者视为与主机“效率”相反的“低效率”和作为增强计算机性能(而非相反)的“设备”。

如上所述,我认为应将让人向往的“低效率”看作一种故意且有益的缓慢——这是系统的规范性特征,而非可容忍的缺陷。以这种方式考虑,延迟不是机器性因其性质的不情愿妥协以服务有缺陷的人类,而是有缺陷机器的有意识妥协以服务人性。

(二)慢速计算和摩擦

弗雷泽和基钦(Fraser & Kitchin,2017;Kitchin & Fraser,2020)的“慢速计算”(slowing computing)理念反映了这种将人作为人类中心的想法。慢速计算旨在有意识地减少用户与计算架构互动中的“时间压缩、碎片化、密集化和压力”。尽管作者的重点是消费者与技术的互动(他们讨论了一些措施,比如每天在规定的时间内关闭无线网络,减少查看社交媒体和电子邮件,以及使用非“智能”设备),这个想法对在律师和公民间构成新兴“法律科技”的设备互动有着重要意义。

与此类似,麦基弗伦(McGeveran,2013)在讨论摩擦问题时,用了一种人文主义视角看待代码的即时性问题。和“低效率”一样,这个概念也有意识形态包袱,这部分源于软件工程(摩擦意味着未能优雅解决手头问题),部分源于古典自由主义经济思想(所有对“清洁”交易的障碍在先验上均不可取)。在社交媒体中,“无摩擦分享”减少了在线分享认知负担——这也意味着有机会停止思考:它变得快速、能够一键点击和一对多发送,也从而招致错误、考虑不周的分享和“错误披露”。增加分享当然符合社交媒体平台的利益,因此减少摩擦主要是

为了服务于此,而不是为了任何的用户利益。作为回应,麦基弗伦(McGeveran,2013)提出了一个“摩擦定律”,即由系统设计带来的摩擦量——除了物理系统的某些本体论限制,摩擦总是一个设计的而非先验的事实——应该是这样的,即分享的时间不低于做这件事(例如阅读文章、跑步等)的时间。

重要的是,计算倾向于完美的执行,这与通过解释进行的思考相抵牾,而作为媒介的文本必然要求这种思考,我们所知的法律也以这种思考为基础。上述观点在某种程度上是以个人(和消费者)为中心的,并不关注法律本身的论证性。但它们确实为我们指出了在“人机互动”中强调“人”的理念,特别是当所讨论的计算机涉及法律和法律服务的提供时。下一步工作是用生活的法律形式的特殊视角来思考这些理念。在这方面,有两个有用的观念可以帮助我们,那就是“争执”(tussle)和“对抗性设计”。

(三)计算架构中的“争执”

和科学、技术与社会(STS)中偏向计算机科学背景的学者一同,克拉克等人(Clark,Wroclawski & Sollins et al.,2005)描述了网络架构设计中的“争执”过程,那些与架构有利害关系的人在设计中产生了利益冲突。利益争执的结果将对系统的最终设计产生影响,甚至会超越设计者在构建“技术-社会结构”中扮演的特殊角色。

要避免架构的脆弱性。“僵硬的设计会被折断;允许变化的设计则会在压力下弯曲。”(Clark,Wroclawski & Sollins et al.,2005:466)在法律性上,这正是文本驱动的规范性所提供的在压力下弯曲的能力;一个在设计上更“僵硬”或更狂热于规则的法律体系也同样容易崩溃。那些设计技术架构的人必须平衡社会争执和技术目标,如“可扩展性、可靠性和进化能力”等设计者通常想要的目标。诚如欧姆和弗兰克的分析,我们可以理解平衡技术(基本问题)和社会(增强问题)的想法,在争执理论下,代码应让适合人类的归人类。

考虑到软件生产的性质,必须预见会出现争执的地方,并事先将其纳入设计中。这意味着(1)以结果变化为导向的设计(架构应该是灵活的,不应屈服于计算守法主义的脆弱性,也不应优先考虑某一组可能出现在结果中的利益),以及(2)将复杂的系统模块化,分隔那些可能涉及争执的元素。这是为了区隔架构选择,让其影响不在技术和人类间蔓延。作者讨论了域名系统(DNS)是如何

“相互纠缠”的,因为它既执行技术功能(在网络上命名计算机),又执行经济/法律功能(域名是商标)。就目前而言,我们可以思考一份智能合约的过度守法主义架构,或法律预测系统提供的指导性论证,和保存理性论证空间的规范性目标间的可能冲突。因为后者可以考虑在设计时未被考量的偶然性因素。

在这里,我们可以看到同法律的规范性理想间的密切联系——法律主体之间的法律争执,在较低的层次上是架构的空间争执问题;为前者提供便利需要我们以适当的预见和敏感性设计后者。特定类型的法律需要对一系列示能做出承诺,但该法律的推动者可能会发现自己与那些有不同想法的人的冲突(那些为法律实践提供技术的人,或者那些使用技术来提高“效率”,但不考虑更广泛的系统性后果的人)。冲突在技术/设计层面表现为争执,尽管因为冲突涉及法律在民主中的性质和作用,冲突的性质当然是法律或政治的。因此,为了维持传统法律意义上的“冲突”,即通过争论来平衡竞争的利益,我们必须在技术的意义上预测它。

(四)设计中的对抗主义

争执和对利益冲突的预期与宪政民主中的对抗主义(agonism)相连。对抗主义认为对抗性辩论是富有成效的,因为它能让对立的观点得以宣泄,从而实现妥协。“征求不同意见”最终还在“民主和法治的核心要义”(Hildebrandt,2018;DiSalvo,2012)中,而设计可以促成这一点。对抗主义可以不同的方式与架构设计相关联。一个系统的最终特征可能会反映在对抗性设计的结果上。[①] 但它也可以在“运行时”的地理环境中促进对抗性互动,这些互动是在利害关系人之间进行的,她们的利益和系统运行或者系统也是其中一环的过程(如法律)相关。在后一种情况下,设计者必须预见到冲突的预期;运行中的架构必须允许偶然性对抗。设计者要有意控制自己强加预定结果(或倾向于预定结果的设计)的冲动,从而在系统的运行地理中为对抗主义、为争执保留“空地”(Gaver,Beaver & Benford,2003)。[②] 这是让计算守法主义随着计算“法律性”的增长而收缩的方式之一。

根据定义(如果不是必然通过设计的话)几乎可以断言,文本驱动的规范性

① 例如建构性技术评估(constructive technology assessment)。

② 在这里,模糊性可能在设计实践中发挥重要作用。

和围绕它建立的法律程序是尊重争执和对抗性互动的:程序上的灵活、规范上的漏洞和解释上的模糊都是(法律的、民主的)利益冲突之处,正因为能够争论并获取于该问题上得到制度认可的(可能是临时的)法律真理声明,文本驱动的规范性让压力得以释放。如上所述,这不仅仅指生硬适用实质性规则,还指产生理解规则意涵的过程。正如利斯(Leith,2016)所指出的,“法律的重要之处在于双方准备争论,而非达成一致”。

在计算架构中建立对抗性冲突的空间,可以抵制反映在“算法治理术”(Rouvroy & Stiegler,2016)中的还原性和工具理性,保护民主和正义所依赖的交往理性(Pols & Spahn,2015)。这意味着需要延迟的示能,即用所谓的“低效”,来限制“技术理性”的速度和表面上的确定性,以支持偶然产生的评价与解释性空间。因此,涉及法律的计算系统应尊重构成这种空间的人类性。这意味着设计者有责任去有意抑制其系统的技术规范性,就像拳击场四周的绳索一样,以促进法律争斗中固有的人类要素。

六、结论

我们应该把延迟看作一种规范性示能,是我们所知法律的基本性质,而非某种本质上不可取、总是要被消除的东西。我们面临的挑战是如何确定它在哪些方面有助于当前的法律实践,并在必要时将其转化为计算工具设计中的必然要素,这些工具被越来越多地纳入法律体系中。这一挑战的部分任务将是说服那些有计算主义倾向的人,效率本身并不是目的。在进行这样的评估时,我们必须清楚认识正在讨论的计算系统的性质,它将在法律程序中被部署到何处,以及它将取代或增强的“人类”任务在提供法律性方面的作用。规范性的影响区别极大——律师在诉讼现场使用的预测性法律分析系统,与从事法律人类学研究的学者使用的系统有很大不同。

计算系统当然可以帮助我们克服人类的局限性,但当把它们用在本应合理维护人类能力的领域时,必须小心谨慎。如何在人类和计算机之间实现适当的互补——阿什利(Ashley,2017)所说的“认知计算”——绝非显而易见,尤其考虑到意识形态和经济上的激励措施可能更有利于更多地采用所谓的“干净”和可预测的法律计算系统。如果我们重视解释学间隙的特质以及它给予我们的解

释和“怀疑”(法律)文本的机会,那么将计算系统引入法律实践时,我们就要思考这些系统对那些特质的影响。如果要为法律性中独特的人类角度留下空间,那么我们就应对它们所依赖的示能有一个清晰的认知。思考关于时间的基本问题,以及法律技术的设计如何能够最好地反映其规范性作用,将是一个很好的开始。

参考文献

Ashley, K. 2017, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, Cambridge: Cambridge University Press.

Ba′nkowski, Z. & N. MacCormick 2000, “Legality Without Legalism.” in W. Krawietz et al. eds., *The Reasonable as Rational? On Legal Argumentation and Justification; Festschrift for Aulis Aarnio*, Berlin: Duncker & Humblot.

Ba′nkowski, Z. 1993, “Don’ t Think about It: Legalism and Legality.” in M. Karlsson, Ó. Jónsson & E. Brynjarsdóttir eds., *Rechtstheorie: Zeitschrift Für Logik, Methodenlehre, Kybernetik Und Soziologie Des Rechts*, Berlin: Duncker & Humblot.

Calo, R. 2014, “Code, Nudge, or Notice.” *Iowa Law Review* 99(2).

Chollet, F. 2019, “On the Measure of Intelligence.” ArXiv: 1911. 01547 [cs. AI].

Clark, D., J. Wroclawski & K. Sollins et al. 2005, “Tussle in Cyberspace: Defining Tomorrow’s Internet.” *IEEE/ACM Transactions on Networking* 13(3).

DiSalvo, C. 2012, *Adversarial Design: Design Thinking, Design Theory*, Cambridge: MIT Press.

Diver, L. 2018, “Law as A User: Design, Affordance, and the Technological Mediation of Norms.” *SCRIPTed* 15(1).

Diver, L. 2021, “Digisprudence: The Design of Legitimate Code.” *Law, Innovation & Technology* 13(2).

Dworkin, R. 1986, *Law’s Empire*, Cambridge: Belknap Press.

Eisenstein, E. 2012, *The Printing Revolution in Early Modern Europe*, Cambridge: Cambridge University Press.

Floridi, L. 2013, *The Ethics of Information*, Oxford: Oxford University Press.

Fraser, A. & R. Kitchin 2017, *Slow Computing*, Maynooth: Maynooth University.

Gaver, W. 1991, “Technology Affordances.” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Gaver, W., J. Beaver & S. Benford 2003, “Ambiguity as A Resource for Design.” *ACM*.

Gibson, J. 1997, "The Theory of Affordances." in R. Shaw & J. Bransford eds., *Perceiving, Acting, and Knowing: Toward An Ecological Psychology*, Mahwah: Lawrence Erlbaum Associates.

Hayles, K. 2004, "Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis." *Poetics Today* 25(1).

Hildebrandt, M. 2015, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Cheltenham: Edward Elgar Publishing.

Hildebrandt, M. 2018, "Algorithmic Regulation and the Rule of Law." *Philosophical Transactions of the Royal Society* 376(2128).

Hildebrandt, M. 2020, *Law for Computer Scientists and Other Folk*, Oxford: Oxford University Press.

Ihde, D. 1990, *Technology and the Lifeworld: From Garden to Earth*, Bloomington: Indiana University Press.

Ihde, D. 2008, *Ironic Technics*, Copenhagen: Automatic Press/VIP.

Introna, L. 2011, "Hermeneutics and Meaning-Making in Information Systems." in R. Galliers & W. Currie eds., *The Oxford Handbook of Management Information Systems: Critical Perspectives and New Directions*, Oxford: Oxford University Press.

Kitchin, R. & A. Fraser 2020, *Slow Computing: Why We Need Balanced Digital Lives*, Bristol: Bristol University Press.

Latour, B. 1992, "Where Are the Missing Masses? The Sociology of A Few Mundane Artifacts." in W. Bijker & J. Law eds., *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge: MIT Press.

Latour, B. 2009, *The Making of Law: An Ethnography of the Conseild' Etat*, Cambridge: Polity.

Leith, P. 1988, "Common Usage, Certainty and Computing." in P. Leith & P. Ingram eds., *The Jurisprudence of Orthodoxy: Queen's University Essays on H. L. A. Hart*, Abingdon: Routledge.

Leith, P. 2016, "The Rise and Fall of the Legal Expert System." *International Review of Law, Computers & Technology* 30(3).

Lessig, L. 2006, *Code Version 2. 0.* New York: Basic Books.

MacCormick, N. 1992, "Legal Deduction, Legal Predicates and Expert Systems." *International Journal for the Semiotics of Law* 5(2).

MacCormick, N. 2007, *Institutions of Law: An Essay in Legal Theory*, Oxford: Oxford University Press.

McGeveran, W. 2013, "The Law of Friction." *The University of Chicago Legal Forum*.

Norman, D. 2013, *The Design of Everyday Things*, Cambridge: MIT Press.

Ohm, P. & J. Frankle 2019, "Desirable Inefficiency." *Florida Law Review* 70.

Pols, A. & A. Spahn 2015, "Designing for the Values of Democracy and Justice." in J. van den Hoven & P. Vermass eds., *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, London: Springer.

Radbruch, G. 1950, "Legal Philosophy." in K. Wilk ed., *The Legal Philosophies of Lask, Radbruch, and Dabin*, Cambridge: Harvard University Press.

Ricoeur, P. 1977, *Freud and Philosophy: An Essay on Interpretation*, New Haven: Yale University Press.

Rouvroy, A. & B. Stiegler 2016, "The Digital Regime of Truth: From the Algorithmic Governmentality to A New Rule of Law." *La Deleuziana* 3.

Rowell, A. 2019, "Legal Knowledge, Belief, and Aspiration." *Arizona State Law Journal* 51(1).

Storni, C. 2014, "The Problem of De-Sign as Conjuring: Empowerment-in-Use and the Politics of Seams." *Proceedings of the 13th Participatory Design Conference* 14.

Waldron, J. 2011, "The Rule of Law and the Importance of Procedure." *Nomos* 50.

Wintgens, L. 2002, *Legisprudence: A New Theoretical Approach to Legislation*, Oxford: Hart Publishing.

Wintgens, L. 2012, *Legisprudence: Practical Reason in Legislation*, Abingdon: Routledge.

作为有机体的技术

——评《技术的本质:技术是什么,它是如何进化的》

马昱堃*

摘要:美国著名经济学家、复杂经济学提出者布莱恩·阿瑟的《技术的本质:技术是什么,它是如何进化的》一书,采用复杂性科学的视角,对技术及其进化过程进行了详细的剖析。阿瑟力图通过本书构建起一套关于技术本身的理论,认为技术是一个自我创生的系统。他归纳出技术的组合与递归特征,分析了技术的"组合进化"过程,将技术视为一个有机体。本文对该书主要思想进行了介绍,探讨了它带来的启发及局限性。

关键词:技术　复杂性　有机体　进化　布莱恩·阿瑟

一、引言

由布莱恩·阿瑟(Brian Arthur)写作的《技术的本质:技术是什么,它是如何进化的》(*The Nature of Technology*:*What It Is and How It Evolves*,以下简称《技术的本质》),英文版出版于2009年。这本书采用复杂性科学的视角将技术视为一个自我创生的有机体,对技术进行了"解剖学"意义上的分析,打开了技术"黑箱"。

阿瑟写作本书的目的是要回答技术及其本质是什么,其原理和特性是什么,它是如何形成、发展和进化的。他认为,过往研究虽然能帮助我们理解上述问题,但是这些研究大多是从外部视角,将技术当作独立对象来看待的(阿瑟,2018a)。为了解决他关注的这些问题,需要进入技术的内部,用"解剖学"的视角来拆解技术的要素,并最终在各项具体的技术中抽象出技术的一般性理论。

为此,他尝试打开技术的"黑箱",通过分析若干具体技术产品的发展历程

* 马昱堃,中山大学社会学与人类学学院。

来回答技术如何对现象进行编程,并一步一步组装起我们生活的这个技术世界的。本书给我们理解当代技术提供了具有启发性的视角,而书中得出的结论也起到了某种警示意义:如果技术是一种具有自我创生能力的有机体,那么在技术进化的方向上,我们施加干预的能力有限。

为了帮助读者理解作者的思想,本文首先对作者的背景做一个简要的介绍,然后回顾书中的主要思想,最后对其中思想的启发和局限进行简要讨论。

本书作者布莱恩·阿瑟是复杂经济学的提出者,美国著名经济学家和复杂系统理论学者,美国新墨西哥州圣塔菲研究所成员(唐任伍、刘洋、李楚翘,2020)。他于1999年在*Science*上首次提出了"复杂经济学"这一概念(Arthur,1999)。复杂经济学建立在复杂性科学的基础上。复杂性科学研究的是复杂系统中要素之间的相互作用如何生成整体模式,以及这种整体模式又如何导致这些要素发生变化或调整以适应整体模式的科学(阿瑟,2018b)。阿瑟将这种思想引入了经济学。

复杂经济学的核心思想是:经济的常态是非均衡的,经济活动的主体也并非经典经济学假设下的完全理性人,而是会在有限信息的条件下,不断根据自身所处情境调整自己对情境的回应(Arthur,2014;乔天宇、邱泽奇,2020)。复杂经济学主要关注经济中的涌现行为,认为传统经济学中的理性人假设存在局限性,常常不能满足实际情况(Arthur,2014)。

阿瑟在思考经济的过程中,逐渐意识到经济与技术二者是不可分割的(Arthur,2014)。经济并不是某种给定的存在,而是会持续地受技术创新、制度和安排的影响。这促使他对经济和技术之间的关系进行探讨。《技术的本质》序言中提到,他于1981年开始关注具体技术的产生和发展过程,之后开始关注经济学的复杂性问题,但于2000年之后重新开始思考技术的本质。

阿瑟为了写作本书查阅了大量技术研究的文献资料,从书中可以看出,他主要将达尔文的生物进化理论与技术进化进行比较,并且参考了库恩的(T. Kuhn)《科学革命的结构》,除此之外,他还参阅了熊彼特(J. Schumpeter)的理论(阿瑟,2018a)。本文把重心放在对本书思想的介绍和评述上,不会将他对话的这些理论一并展开。

二、主要思想

(一)技术的定义

技术的定义非常繁多,且会随着时代的发展而不断扩大内涵与外延(姜振寰,2009)。而布莱恩·阿瑟的《技术的本质》则试图从技术形态的角度,给技术下一个明确的定义。

阿瑟根据定义层次的由低到高,分别给出了关于技术的三种定义。第一种定义认为技术是实现人的目的的一种手段。这层定义最具体,可以指特定一项技术产品。在这个定义之下,作者进一步区分出了物质性的技术(例如器械)和非物质性的技术(例如算法)。它可以是一种方法、过程或者装置,例如金属冶炼方法、汽车发动机等。第二种定义认为技术是实践和元器件的集成,"是许多技术和实践构成的集合或者工具箱"。这种技术也被作者称为"技术体"(bodies of technology),它们往往是某一组相似技术的集合,例如电子技术、生物技术、数字技术等。第三种定义认为技术是可供某种文化利用的装置和工程实践的集合。这指的是总体意义上的技术(阿瑟,2018a)。在抽象层次上,第三种定义内涵最宽泛,包含了前两种。

上述三个定义对技术的界定不仅局限在工程和生产领域,也纳入了社会制度。在作者看来,技术是"指向某种目的的,被编程了的现象"。这些现象不仅包括物理性的,也包含行为性和社会性的(阿瑟,2018a)。进而,他将包括工程技术和社会制度在内的系统都视为"目的性系统"(purposed systems),它指的是"所有'实现目的的手段'的总体,既包括基于物理现象的技术,也包括基于非物理现象的技术"(阿瑟,2018a)。正是在这个意义上,作者将技术的内涵拓展到日常生活中的绝大多数活动中。于是,金融体系、法律条文均被纳入"技术"之下。通过将这些现象中的共性抽象出来,阿瑟便可以应用他的思想对所有技术现象进行分析。①

① 但阿瑟主要还是侧重于分析建立在自然现象上的技术。

(二)技术的组合与递归

阿瑟认为所有的技术都来自对此前已有技术的组合。现存的技术也构成了未来技术中的组合成分。把任何技术产品打开,能看到每一项技术的零部件本身也是一项技术产品。“一台水力发电机集中了几种主要的构件:一个用于储水的蓄水池,一个带有控制阀门和被称为水渠的进水管的进水系统,由高位能水流驱动的涡轮机,由涡轮机驱动的发电机,将输出能量转换为高压电的变压器以及排水用的泄流系统或尾水渠。”(阿瑟,2018a:32)这样一台发电机中的每一个部分都可以单独拿出来,再按照同样的拆解步骤进一步分解出它的零部件。

上述步骤会不断重复发生在技术的各个层级,这就形成了一个递归过程。将这些步骤不断重复下去,就会到达技术的源头——现象。作者认为,技术就是对自然现象进行有目的的编程。上文所说的分解过程可以在某个技术物及其零部件上不断重复。“技术具有层级结构:整体的技术是树干,主集成是枝干,次级集成是枝条,最基本的零件是更小的分枝。技术具有递归性:结构中包含某种程度的自相似组件,也就是说,技术是由不同等级的技术建构而成的。”(阿瑟,2018a:38)在这个意义上,任何一项技术都由若干个子模块构成,每一个子模块本身也是一项技术——按照这样的步骤,一项技术总是可以还原为某种自然现象。

当一系列基于某种共性现象的技术不断得到开发和利用后,就会“自然而然地”引起技术聚集,形成技术集群(或技术体)。阿瑟把这种技术集群称为“域”。一个域总是以工具箱的形式存在,为技术应用提供背景支持,并对经济产生影响(阿瑟,2018a)。

技术可以通过递归过程还原为自然现象是整本书的核心思想。技术就是沟通了人类需求和自然现象之间的一种编程表达,它体现着人类目的与客观现象之间的桥梁关系。正是在各种具体的技术中抽象出来上述组合与递归过程的基础上,作者给出了技术进化的演变机制。

(三)技术的进化机制

为了解释技术的进化过程,作者需要考察技术发展的动态过程。为此,他提出了两对概念,分别是“标准工程”与“发明过程”和“内部替换”与“结构深化”。

1. 标准工程与发明过程

第一对概念表述的是技术过程。“标准工程是执行一个新项目时,在已知可接受的原则下聚集方法和设备的过程,是对已有技术的新的计划、试制和集成过程。”(阿瑟,2018a:99)从这个定义可以看出,它与托马斯·库恩《科学革命的结构》中的“常规科学”(库恩,2012)进行了对话。①

从定义上说,标准工程是技术研发的日常活动,旨在解决一项技术产品中某一个零部件的性能问题,目的是提升技术组件的效率。对这些零部件的设计,需要设计者依据技术的各项特性做出评价和判断,并最终做出选择。整个标准工程都需要有人的参与,从而转化为一种社会组织形式:“一个新项目成功与否,它是否能形成可见的设计物,很大程度上依赖于围绕其周围的利益关系网……设计与发展是与人高度相关的组织和行动过程。”(阿瑟,2018a:103)

根本性创新(radical innovation)也被作者称为“发明过程”,它与标准工程不同。如果工程师和设计师只是对既有的技术原件进行逐个替换,那么我们一定无法从蒸汽机过渡到飞机引擎。因为标准工程解决的是局部零件的性能问题,发明过程则需要引入新的原理:“判断一个技术是否是真正的发明,我们仅需要判定对当前的目的起作用的是否是一个新的、不同以往的根本性原理即可……依据根本原理新颖性的程度可以产生一个连续统一体,而它存在于标准工程和根本的新颖性之间。”(阿瑟,2018a:121—122)新原理来自科学发展。阿瑟在书中把科学发展看作与技术发展互构的现象。科学的目的是搞清楚现象的运作机制,而技术为科学实现这一目的提供了支持:如果没有技术将猜想与现象相连,科学就不会有所发展,也就不会有新的现象进入技术了(阿瑟,2018a)。

总之,标准工程与发明过程处在一个连续统的两端。在技术产品迭代进化的过程中,这二者可能重叠或同时发生。

2. 内部替换与结构深化

第二对概念表述了技术发展的两种机制:内部替换(internal replacement)与结构深化(structural deepening)。前者指的是“用更好的部件(子技术)更换某一形成阻碍的部件”;后者则是“寻找更好的部件、材料,或者加入新组件”(阿瑟,

① 阿瑟介绍“标准工程”时在注释中提到,他不喜欢用“正常的工程”(《技术的本质》中文本译者将“常规科学”译为“正常的科学”),因为这暗示着工程中也有与科学相平行的活动,而实际上工程中并没有。所以他更倾向于使用“标准工程”。

2018a)。内部替换针对的是既有技术内部旧部件被新部件替换,而结构深化则是给技术加入更多次级系统来提升整体性能。这两种机制也可以同时发生。

这两种机制发生的根本原因是对效率的追求。如果一项技术还有效率提升空间,它就会通过上述两种机制进行优化。同样,这种机制本身也是递归性的,“技术的改进过程应该伴随着构成次级组件以及次次级组件的零部件的置换过程”(阿瑟,2018a)。换句话说,一整套技术是在整体及其各个层次上同时发生着上述两种机制。任何一个子模块或零部件的改进,都会连锁反应地使整个技术体以及它所构成的更大技术体产生更新。

当通过两种机制让技术的性能达到极限时,技术就成熟了。此时如果还需要让技术进一步取得进展,则需要注入全新的原理。作者指出,这个过程与托马斯·库恩的范式革命(库恩,2012)非常类似:每当一个新原理出现,并以内部替换和结构深化的方式注入技术后,技术的性能就开始进入迭代升级的轨道,每一项技术都在这个过程中相互“促逼”(pushed)[①]而接近各自的极限,直到技术整体性能无法提高,这时候就再次需要新原理的注入(阿瑟,2018a)。

当然,这个过程不可能发展得非常顺利。通常我们观察到的实际情况是,在新原理刚刚出现还未被引入这个周期时,旧原理依然统治着大部分市场。由于旧原理性能暂时更好、应用新原理中的成本过高以及从业者不认可等原因,新原理不会立刻被注入技术,从而产生了暂时的迟滞现象(hysteresis)。阿瑟将这种锁定(lock in)在旧原理中的状况称为自适应延伸(adaptive stretch):在新旧原理之间,人们更倾向于采用旧技术或旧原理,并通过对旧技术和旧原理的拉伸来尽可能适应新环境(阿瑟,2018a)。但旧原理最终会触碰效率的极限,从而被新原理取代。

技术就是通过上述过程和机制不断发生组合进化。可以把技术看作一张网络,每一项技术都是网络的节点。新技术进入后,可能会取代旧的技术,并且它的进入可能会创造新的组合可能性和新的需求,随之到来的是社会和经济对技术的适应(阿瑟,2018a)。

① 技术被逼迫赋予更多功能以提升性能。

(四)技术与经济的协同进化[①]

根据阿瑟对技术的定义,凡是实现某种目的的手段都可以被称作技术,并且技术不限于实体。进而就可以得出,为经济活动而设立的制度、机构也可以被称为技术:“构成经济的整套安排将包括所有制度和方法,以及所有的我们称之为技术的目的性系统,包括诊所和外科手术、市场和定价系统、贸易安排、分配系统、制度和商业,还包括金融系统、银行、监管体系以及法律系统。”(阿瑟,2018a:215)

因此,作者认为经济实际上是被技术建构起来的:“众多的技术集合在一起,创造了一种我们称之为‘经济’的东西。经济从它的技术中浮现,不断从它的技术中创造自己,并且决定哪种新技术将会进入其中……经济是技术的一种表达,并随这些技术的进化而进化。”(阿瑟,2018a:214)在这个意义上说,整个经济都由技术塑造:“经济浮现或萌发于它的技术当中;它意味着经济不仅会随着技术的变化而重新适应,它还随着技术的变化而继续构成和重构;它还意味着经济的特征(形式和结构)也将随着它的技术的变化而变化。”(阿瑟,2018a:216)

当经济活动中的各个主体遭遇新技术时,经济会表现出非均衡性:这些主体会因“小规模经济活动中的变化而向外传播,使商业组织和制度发生变化,最终使社会自身发生变化”(阿瑟,2018a)。作者将技术决定经济的论断推向了极致:新技术创造新产业,为了与新产业相适配,就会产生新制度。随着生产效率得到进一步优化,这一过程会循环下去。并且,整个过程本身具有技术性质(阿瑟,2018a)。这个循环过程实际上并不会停滞,新问题、新需求、新技术、新原理会不断出现,并被加入这个过程本身,从而让“经济总是存在于永恒变化之中,存在于一种永恒的新颖性之中。它永远存在于自我创造的过程之中”(阿瑟,2018a)。

① 阿瑟是经济学家,不过他对技术的兴趣与他对经济的兴趣并重。就本书而言,他把关注重点放在了技术上,对技术与经济关系的表述主要见于书的第八章和第十章。有关阿瑟本人的经济与技术思想评述需要再有一篇论文的篇幅。感兴趣的读者可以查阅参考文献中唐任伍等人(2020)的论文。

（五）技术性与生物性相互转化

在本书的最后，作者进入了一个更宏观的视角：技术本身具有生物属性，反之亦然。一方面，数字化时代拓展了技术组合的可能性：只要将数字模块加入技术中，将实体编码为数据，就可以将其与整个数字技术体系关联起来，从而融入“数字技术域”中。另一方面，对生物技术的研发，例如基因编辑、神经科学等领域的发展，也在模糊生物与技术的界限。“从概念上看，生物学正在变成技术。从实际上看，则技术正在成为生物学。”（阿瑟，2018a：233）

当越来越多的技术通过数字化关联起来的时候，就形成了一张覆盖生活方方面面的技术网络。作为一个整体的技术网络本身，已经具备了“感知环境并适当反应”的能力，从而表现出了某种程度的认知能力。当技术的触角进一步伸向生物现象并将其编码，整个技术就可能形成“自构成、自优化、具有认知能力，还能自集成、自修复以及自保护”的系统（阿瑟，2018a）。

阿瑟的论述暗示了我们无法再用单纯的机械视角来看待当前的技术发展。虽然技术的“机械性依然是中心议题，但是我们现在认识到，随着机械变得具有互动性并复杂起来，它们所揭示的世界也成了一个复合体。它们是开放的、进化的，并且表现出了无法从某个部分预测未来的涌现性质（emergent properties）。我们所倾向的观点不再是一种纯粹的秩序，它是一个整体，一个有机的整体，而且是不完美的”（阿瑟，2018a：236）。

三、讨论与反思

评价这样一本尝试从人类实现自身目的的所有手段中抽象出共性的著作无疑非常困难，因为它涉及的思想触及科学技术社会学、技术哲学和经济学，横跨了自然科学、社会科学和工程技术，并且作者本人还开创了复杂经济学这一分支。甚至有评论家将本书与库恩的《科学革命的结构》并列（赵阵，2015）。本文对《技术的本质》的述评尽量把握全书的内核，但限于笔者能力和篇幅，无法涵盖整本书的所有精髓。

毫无疑问，作者将复杂性科学思想应用到对技术的分析中，提供了非常具有启发性的视角：技术在每一个微观尺度上更新自身的同时，整个技术都在协同进

化,并持续地塑造着经济和整个社会,进一步影响塑造着人类的生活环境。基于这一点,作者指出技术创新总是作为解决问题的方案出现,这些方案离不开长期的技术积累,以及着迷于解决问题的具体的人。这些启发对于创新政策的制定都有参考价值。然而,本书也存在着一定局限。

(一)本书弱化了人的劳动是技术得以进化的动力因

将技术视为一种自我创造的有机体的视角也在一定程度上排除了人的能动性。整个技术体在局部和整体的层次上都同时发生着进化,而每一步进化都意味着整体的性能升级。但这一过程不可能是自然发生的——如果把人的体力和脑力劳动从作者所定义的技术中去除,那么这样的技术显然只是静态的。换句话说,技术之所以"表现得"是一个有机整体,是因为人的劳动和选择不可缺少(Denton,2012)。虽然阿瑟也提到了设计与选择在技术发展中扮演的角色,但这种选择也被限定在了追求效率的约束条件下。

作者对技术的定义和分析,不可避免地把技术视为一种"类有机体"。他在书中从两方面出发给出了这一结论。第一个方面还没有将这一结论推到极致,而是强调了视角的转换:"所有的技术在某种意义上都同时具有机械性和生物性。如果你从上至下地检视一项技术,你可以将其看作一种互相连接的元器件的安排,这些元器件为达到某个目的而互相制约、互相啮合。因此技术变成了像钟表一样的发条装置,它变得具有机械性了。但是,如果你在思想中将它从下往上进行检视的话,想象这些零部件是如何聚集到一起的,你则会将它们看成一个整体,一个完整的器官,一个更高等的,具有功能性、目的性的整体。它变成了一个机体。因此,一项技术具有机械性还是生物性取决于你的立场。"(阿瑟,2018a:232)在这个意义上,视角或者立场是把技术视为机械结构还是有机体的原因。

第二个方面更加彻底:"技术完全是生物性的。技术具备能使我们联想到生物的某些属性:当它们感知环境并产生反应,当它们变得可以自组装、自构成、自修复并且能够'认知'的时候,它们就越来越像生物了。技术越复杂、越'高技术',就越具有生物性。"(阿瑟,2018a:232)作者在提到技术可以自我建构的时候,按照他对技术的定义,其实是在说实现目标的手段,或者说工程实践的集合等具有了内生动力。但是,也正是因为工程实践中的人对技术进行有目的的选

择才使技术成为人们所见到的模样。而这种选择并不仅仅是技术性的,也不完全是随机性的。

总之,作者认为技术是一个能够自我进化或者“自创生”(autopoiesis)的实体。无论是对技术的改进需求,还是技术的生长动力,均来自技术内部。在这个意义上,技术被摆在了分析的第一位。这样的视角转换在具有启发性的同时,也在一定程度上忽视了劳动的因素、意义的因素和制度的因素,毕竟作为物理存在的技术体和作为法律条文的制度体系自身不会有任何意识,之所以它们能够体现出一种发展的样态,是因为人的活动为整个体系注入了动力。而人的活动本身,并不能全部归结到自然属性。相反,人的活动具有选择性,而这种选择性,以及人对目的的规定,都构成了技术的动力因(罗斯琦,2021)。

(二)本书较难解释不同时期不同国家的技术发展状况差异

把技术视为一种具有自我建构能力的有机体,其实并不能回答为什么在历史的不同时期,不同国家之间的技术存在发展差异。作者在书中表示,技术发展可以用“因果性金字塔”来类比:每一个发明都由更靠近金字塔底座的知识和需求支撑,并且这个过程可以不断追溯到人类捕获的现象。所以“这也意味着当必要性和需求的碎片都一一铺垫到位时,一项发明就将显露”(阿瑟,2018a)。但是他没有进一步说明这种必要性和需求是如何被规定的,并以怎样的方式进入技术发展过程当中的。

阿瑟在对创新与国家竞争力进行论述的时候,承认了文化起到的作用:“真正前沿的技术,那些处于边缘的复杂技术并不是源于知识,而是源于别的东西,我将它们称为‘深奥的手艺’(deep craft)。深奥的手艺不只是知识,它是一套认知体系……它整体地来自某种信念的共有文化,蕴含着共同经历的某种无法言明的文化。”(阿瑟,2018a:179)对这种文化进行详细的分析毕竟不是阿瑟本书的目的,但从本书结论中也能看到,他并没有把这种文化因素安排在技术的自我创生过程的重要位置。

他也强调了历史的重要性:“如果技术由于某种偶然以不同的序列出现,那建立在这些技术之上的技术也会不同。技术是历史的产物。它还告诉我们,技术的价值不仅在于可以用它做什么,而且在于它进一步可能导致什么。”(阿瑟,2018a:190)虽然偶然性因素的确存在,但我们并不能得出技术发展的时空差异

性是纯粹偶然的。事实上,作者所描述的技术发展状态更多地发生于近现代——只有当我们开始以追逐效率为目的,并且技术开始与现代科学交织在一起时,才会持续产生新的需求来创造技术发展的空间。而这涉及针对现代性的讨论,本文不会在这个主题上进一步延伸。

(三)人类在技术发展的进程中真的束手无策吗?

作者采用了将技术视为一个有机体的视角——把技术视为社会发展的第一推动力,并且技术进化的动力是内生的。正如前文所言,这种视角弱化了人的主观能动性,也可能成为替技术产生的负面后果进行辩护的理由。然而,如果我们在对技术的思考中完全忽略掉社会因素,那么整个技术图景就变得不全面了。有学者指出,思考科学技术艰难的原因有二:科学主义和技术进步论。前者主张科学技术的中立性,后者则认为技术的发展是人类自由意志无法干预的。但技术本身就具有社会性和政治性(克莱曼,2009)。拉图尔也指出,人与技术一经结合,就形成了新的结合体,二者原本的状态就发生了变化——这种结合体应当被视为新的行动者(郭明哲,2008)。

芒福德(L. Mumford)则更加直接地指出:"技术与文明作为一个整体是人类的,有意识或无意识的选择、智能活动和奋斗的结果……它们虽然无法控制,但也不是外部的。选择本身体现在社会的小的增量变化、瞬息之间的决定以及暴风雨般的巨变;一个人如不能在机器的发展中看到这种选择,那么当选择的结果累加起来产生了一个合成的变化时,他就无法真正理解这种变化,只能归之于外部的或非人力控制的因素了。"(芒福德,2009:9)

不过,也正如阿瑟所指出的那样,数字化时代的到来,以及人工智能、量子计算、生物科学和神经科学的新进展,不断向着将人自身技术化的方向过渡。这些新的进展毕竟创造了利用它们的可能性。整个社会在面临新技术开辟出的可能性空间[①]时,到底有多大的选择余地?这似乎取决于我们的立场:是将技术视为一个有着内生动力的有机体,还是把它看作人类在经济和社会制度中实现各种目的的手段总和?针对前者,我们能做的似乎不多;但如果采用后一种视角,我们仍然能对技术的发展方向产生影响。

---

① 阿瑟在书中用了另外一个名词:"机会利基"(opportunity niches)。这个概念主要是在阿瑟阐述技术与经济关系时出现的。文中使用"可能性空间"是取出这一概念在技术进化方面的内涵。

参考文献

阿瑟,布莱恩,2018a,《技术的本质》,曹东溟、王健译,杭州:浙江人民出版社。

阿瑟,布莱恩,2018b,《复杂经济学:经济思想的新框架》,贾拥民译,杭州:浙江人民出版社。

郭明哲,2008,《行动者网络理论(ANT)——布鲁诺·拉图尔科学哲学研究》,复旦大学哲学学院博士学位论文。

姜振寰,2009,《技术哲学概论》,北京:人民出版社。

克莱曼,丹尼尔,2009,《科学技术在社会中:从生物技术到互联网》,张敦敏译,北京:商务印书馆。

库恩,托马斯,2012,《科学革命的结构》第 4 版,金吾伦、胡新和译,北京:北京大学出版社。

罗斯琦,2021,《超越"劳动终结论":"替代"结果还是"互补"过程?——关于技术与劳动的理论辨析》,《社会发展研究》第 8 期。

芒福德,刘易斯,2009,《技术与文明》,陈允明、王克仁、李华山译,北京:中国建筑工业出版社。

乔天宇、邱泽奇,2020,《复杂性研究与拓展社会学边界的机会》,《社会学研究》第 2 期。

唐任伍、刘洋、李楚翘,2020,《布莱恩·阿瑟对复杂经济学的贡献——科睿唯安"引文桂冠"经济学奖得主学术贡献评介》,《经济学动态》第 3 期。

赵阵,2015,《探寻技术的本质与进化逻辑——布莱恩·阿瑟技术思想研究》,《自然辩证法研究》第 10 期。

Arthur, B. 1999, "Complexity and the Economy." *Science* 284.

Arthur, B. 2014, "Complexity Economics: A Different Framework for Economic Thought." in *Complexity and the Economy*, Oxford: Oxford University Press.

Denton P. 2012, "Review of *The Nature of Technology: What It Is and How It Evolves*, by W. Brian Arthur." *Essays in Philosophy* 13.

平台资本主义的过去、现在与未来*

——评尼克·斯尔尼塞克的《平台资本主义》

郭浩田**

摘要:在《平台资本主义》一书中,作者分析了平台资本主义的历史,即第二次世界大战之后的三次经济危机如何与平台资本主义的兴起相关联。作者接着定义了“平台”,阐明了平台的特征并区分了平台的种类。作者分析了平台发展的趋势以及存在的问题,并对未来进行了展望。该书最大的贡献在于其历史分析将资本主义的不同阶段勾连起来,揭示了资本主义的普遍规律。该书在预测平台资本主义的未来时略显论述不足,但对未来的研究具有较强的启发性。

关键词:互联网　平台　资本主义　数据

数字经济、人工智能、大数据等话题已经成为诸多学科的研究热点,以“平台”作为核心的相关研究在社会学、传播学、法学、公共管理等学科中也占据了重要的地位。斯尔尼塞克(N. Srnicek)的《平台资本主义》(*Platform Capitalism*)一书为平台研究提供了一种政治经济学的分析视角,平台是这本书的起点,但作者分析的落点着重于资本主义。

资本主义危机是不可避免的,并且环环相扣。过往的危机酝酿了当下的繁荣,而当下的繁荣又隐含着未来的危机。平台资本主义的兴起,是第二次世界大战以后三次经济危机的结果,也是下一次危机可能爆发的起点。本文将遵循作者的章节安排,首先回顾平台资本主义的过去,即战后资本主义的经济危机;接着探讨平台资本主义的当下,厘清平台的概念与分类;然后思考平台资本主义的问题与未来的可能性;最后探讨《平台资本主义》一书的贡献,以及对未来研究的启示。

* 本文系国家建设高水平大学公派研究生项目(项目批准号:201808060448)的阶段性研究成果。

** 郭浩田,剑桥大学社会学系。

一、战后的资本主义经济危机

本书的第一章是作者关于资本主义经济危机的历史分析。在介绍平台资本主义之前,斯尔尼塞克首先对第二次世界大战之后的世界经济史进行了政治经济学的分析。他认为当下平台资本主义的兴起与战后的三次经济危机——分别是 20 世纪 70 年代、90 年代和 2008 年的经济危机——紧密相关,正是这三次经济危机为数字经济的兴起奠定了基础,并左右了其发展方向(Srnicek,2017)。

(一)20 世纪 70 年代的经济衰退

第二次世界大战之后,欧美和东亚部分地区在战后重建之中迎来了经济的高速增长。这一时期的高收入经济体有两个值得观察的特点,分别是商业模式和就业性质(Srnicek,2017)。战后的美国制造业主导了全球经济,美国工厂大规模采用福特式的生产线,工人们遵循着泰勒主义的生产原则。不断壮大的工人群体因其相似的处境而组成了工会,促进了劳资双方的集体谈判并提高了工人的福利。

20 世纪五六十年代,日本和德国的制造业开始兴起,其发展受益于劳动力价格较低、熟练程度较高、汇率优势以及政府和银行的支持(尤其是日本)。20 世纪 70 年代,由于石油危机和越南战争等因素,美国的经济呈现了衰退的态势。日本和德国的竞争加剧了美国制造业利润率的下滑,美国企业开始寻求一些解决之道,例如采用日本丰田式的生产模式,削减劳动力并打击工会,以及将大量的工作外包到其他国家。这样的全球外包产业链的形成受益于通信技术的发展,也与 20 世纪 90 年代的互联网繁荣和泡沫息息相关。

(二)20 世纪 90 年代的繁荣与泡沫

在 1985 年的广场协议之后,日元对美元升值,美国的制造业迎来了短暂的健康成长。但好景不长,国际市场对于日本经济的担忧又使得日元贬值,美国以及其他西方主要经济体的制造业重回疲态。

与之相反的是互联网经济,在 20 世纪 90 年代和 21 世纪初呈现出蓬勃发展的景象。作者认为,这一阶段的本质是互联网的商业化,其内在驱动力是大量风

险资本的金融投机行为(Srnicek,2017)。丹尼尔·席勒(Schiller,1999)也在《数字资本主义》一书中对互联网的商业化进行过详细的分析。在制造业前景晦暗不明的状况下,风险资本瞄准了互联网产业。在这一时期,美国有5万多家公司进入互联网领域,吸纳了大概2 560亿美元的资本(Srnicek,2017)。这些互联网公司缺乏明确的商业模式,也没有稳定的盈利,它们在这一阶段的主要目标是抢占用户群体。

同一时期,与互联网相关的基础设施也获得了大量投资,尤其是在一些低收入的经济体当中,电子通信产业是吸收外国投资最高的产业。在全球范围内,大量的通信基础设施纷纷建立,包括光缆、数据存储设备和服务器。这大大降低了全球产业链的通信成本,方便了协调分工,也促进了美国制造业的向外转移。

1997年,金融风暴席卷亚洲,也影响了美国的经济。为了延缓经济的崩盘,美国联邦储备委员会(简称"美联储")开始急速调低利率,实行超级宽松的货币政策,向市场里注入了大量的流动性资金(Srnicek,2017)。这样的货币政策改变了美国经济增长的路径,由传统依赖制造业竞争力的提高,转变为从其他行业中寻求利润。这样的货币政策促进了互联网行业投资的进一步成长,即使在互联网产业泡沫化以及"9·11"事件之后,宽松的货币政策也并未停止。

此外,美联储的干预措施也包括降低抵押贷款利率。在宽松的货币政策之下,大量资本开始寻求新的投资点,而次级抵押贷款市场正是这样的一个高风险、高收益行业。这些因素为2008年的次贷危机做了铺垫。

(三)2008年的金融危机

2006年前后,美国的房地产价格迎来了一个转折点,房价的下降影响到了总体经济,尤其是美国的家庭财富,这影响了他们偿还债务的能力。全面性的次级抵押贷款危机于2008年爆发,以雷曼兄弟公司的倒台作为标志。此时,美国的金融市场已经与抵押贷款市场紧密地联结在一起,崩跌的房价使得许多次级抵押贷款成为烂账,金融市场一片愁云惨淡,美联储不得不出手救援。

美联储救市的方法依然是注入大量的资金,同时也取得了部分银行的权益,高额的私人债务因而转化成了公共债务。为确保美元的流动性,美联储将联邦基金利率调降至接近于0。同时,美联储也加紧同各个发达国家的中央银行的合作,共同开启了一个低利率的环境(Srnicek,2017)。

次贷危机之后,发达国家政府面临着高额债务,财政政策日趋紧缩,对于经济的刺激更加依赖于货币政策。宽松的货币政策使得金融投资的回报率降低,因此资本会更加倾向于寻求那些高风险、高回报的行业,数字产业即是其中之一。

近些年来,伴随着量化宽松政策,诸多科技企业将资金转移到“离岸避税天堂”。这使得科技企业有了充足的资金投资人工智能之类的风险产品,但这种避税行为挤压了政府的收入,加剧了财政政策的紧缩,所以政府也更加依赖于货币政策。

要理解平台资本主义的兴起,除了理解资本市场的危机之外,还要分析劳动力市场的就业状况。作者认为,伴随着世界共产主义在 20 世纪 80 年代末 90 年代初的衰落,待就业大军的规模在不断地扩大,劳动力的无产化进一步加深,许多人不得不依赖于不稳定、非正式的工作(Srnicek,2017)。2008 年的次贷危机又导致了失业人口增加,尤其是长期失业人口的上升。长期失业人口的债务压力更大,更容易被迫接受不稳定、非正式的工作。这些失业的劳动力,与逐利的资本一起,为平台资本主义的兴起创造了条件。

二、平台的特征与类型

2010 年前后,伴随着自动化、人工智能、物联网、大数据等技术的进步,一批互联网公司随之兴起。它们带来了新的商业模式,被冠以各种各样的名称,例如零工经济(gig economy)、共享经济(sharing economy)、注意力经济(attention economy),等等。一些观点认为,在这样的经济模式中,劳动力越来越依赖于知识(knowledge),传统的产业工人正在被知识工人取代,即便是物质化的商品当中也蕴含着越来越多的非物质化要素。尽管作者承认这样的经济模式有一定新的成分,但他认为这些概念没有反映出问题的本质。这些公司的经济模式没有脱离资本主义的范畴,它们的特殊点在于将数据(data)作为一种原材料。在本书的第二章中,作者首先区分了数据与知识的定义,接着阐述了平台的概念与特征,最后对平台进行了分类。

(一)作为原材料的数据

作者认为,数据是关于某事发生的信息(information that something happened),而知识是关于某事为何发生的信息(information that why something happened);数据可能会涉及知识,但这并不一定(Srnicek,2017)。数据并不能脱离物质基础而存在,数据的收集和储存都要依赖于一定的设备。在进行数据分析之前,还要对数据进行清理和标准化处理。因此,作者将数据视为一种原材料(raw material),如同石油一样,需要提取、精炼、加工,最终用于不同目的(Srnicek,2017)。

在资本主义的不同历史阶段,数据一直在被收集和使用着,只不过程度不同。在福特主义生产模式中,大量的数据丢失了,只有少量的数据被用于辅助生产。在精益生产模式中,全球供应链的建立加强了对数据的依赖,但仍缺乏对大量数据的提取和使用。这些既有的生产模式不能很好地利用数据这种原材料,然而信息技术的发展大幅降低了数据的采集和分析的成本。亚马逊、脸书、谷歌等一批企业通过对数据的高效利用,产生了新的商业模式;通用电气、西门子等传统的加工制造企业,也借助信息技术拓展了对数据的利用,向新的商业模式靠拢(Srnicek,2017)。

(二)平台的特征

在对数据的概念进行界定之后,作者紧接着定义了平台。平台是使得两个或多个群组可以互动的数字基础设施,作为中介的平台沟通了不同的群体,包括顾客、广告商、服务提供商、生产者、供应者,甚至非人的客体(Srnicek,2017)。平台通常还会有一些数字工具,用户可以借此创造属于自己的产品、服务和市场。因此,作为用户活动发生的界面和用户之间互动的中介,平台具备了获取和记录数据的优势(Srnicek,2017)。基于此定义,作者分析了平台的四个重要特征(Srnicek,2017)。

第一,平台的范围远超互联网公司或科技公司。只要是数字互动发生的地方,都可以成为平台运行的基础。这一特征与下一节中平台的分类有关。

第二,平台产生并依赖网络效应(network effect)。当越来越多的人使用一个平台时,对于其他人来说,这个平台的价值就越来越大。平台用户数目的增长

可以吸引更多的使用者,并扩大平台收集的数据量,这些数据又可以经过分析用于改善平台的服务。基于网络效应,垄断的种子也扎根在了平台的基本特性之中。

第三,平台经常采用交叉补贴的策略(cross-subsidisation)。平台往往通过一些免费、低价的商品或者服务来吸引大量的用户,再通过一些较高价格的商品或服务获取利润,并对前者进行补贴。例如谷歌通过广告服务盈利,并补贴搜索服务。

第四,虽然从表面上来看平台只是一个供不同用户开展数字互动的空间,但实际上平台包含了一种政治规则,这个规则的制定者是平台的拥有者。尽管如此,这些既定规则还是允许不同用户生产一些意料之外的产品,可能超乎平台拥有者的预期,甚至撼动他们的地位。

(三)平台的分类

正如作者所说的平台的第一个特征,传统的制造业和新兴的互联网企业都可以成为平台企业。作者区分了五类平台:第一类是广告平台,这类企业借助数据分析来投放广告,例如谷歌;第二类是云平台,这类企业提供云存储和云计算服务,例如 IBM 和亚马逊;第三类是工业平台,通过工业互联网来辅助生产制造过程,例如通用电气和西门子;第四类是产品平台,通过出租自己的产品来收取费用,例如罗尔斯·罗伊斯公司和 Spotify;第五类是精益平台,这些企业基本没有固定资产,大多是通过联结服务提供者和消费者来收取费用,例如 Uber 和 AirBnB(Srnicek,2017)。这种新的商业模式与资本主义长期利润率下降的趋势以及周期性的经济危机互相交织。精益生产和及时生产的模式在 20 世纪 70 年代之后一直在发展,而现在平台企业的商业模式正是同一个趋势的延伸。

三、平台资本主义的问题

第三章的题目是“大型平台战争”,这个“大型”既可以指平台的规模,又可以指战争的范围。伴随着平台的不断扩大,平台与平台之间、平台与用户之间、平台与监管者之间都会产生矛盾,这些矛盾对平台未来的发展产生了深远影响。

(一)平台与平台

基于网络效应的平台天然地带有垄断的倾向,平台企业会想方设法地提高数据的数量和丰富程度。平台企业常用的一种手段是建立封闭的生态系统,例如苹果和安卓两套手机操作系统,只能在不同的设备上使用;通用电气和西门子之间的工业互联网也不互通。这样一来,先发企业具有很大的优势,但这并非牢不可破,大量的资本仍然可能造就新的优势平台。

在平台资本主义当中,企业的竞争力不仅是降低产品成本的能力,更是对数据利用的能力。因此,对于平台与平台之间的冲突,作者指出了四个趋势:第一,对数据提取和利用的扩大,越来越多的企业投资了收集、存储、分析数据的基础设施;第二,占据关键位置,作者运用了“根茎式”的比喻,指出平台将自身置于关键位置,例如互联网的入口,因此许多大型平台企业会投入自然语言界面市场;第三,市场融合,平台企业会涉足越来越多的产品领域,变得越来越相似;第四,生态系统封闭,平台企业会在寻求封闭软件系统的基础上,寻求硬件系统的封闭,但这有可能带来互联网的碎片化(Srnicek,2017)。

为了争夺用户,许多平台企业消耗大量资金打击竞争对手,这些企业的繁荣是建立在资本的狂热之上的。国际和国内的经验均印证了这一点,例如各个网约车公司之间的竞争、各个短视频网站之间的竞争。这股平台战争的热潮伴随着资本的降温已经渐渐消退,许多平台企业将面临转型问题。作者预测,一部分企业将会转为提供奢侈的或急需的服务(Srnicek,2017)。

(二)平台与工人

尽管平台提供了竞争与控制的新方式,但最后决定平台未来的仍然是利润率。在作者看来,平台没有解决利润率下降的问题,只是吸走了制造业的利润。尤其是在精益平台的商业模式之中,工人和工会的反抗会大大增加平台的运营成本,例如 Uber 司机对企业的抗议以及美国和欧洲各国的立法确认了工人与 Uber 的雇佣关系。最基本的工人权利一旦被确认,这种生产的模型便无法持续了。

(三)平台与国家

作者在本章最后简要探讨了作为平台监管者的国家的角色。作者认为国家拥有控制平台的权力,可以通过一系列的法律保障用户的隐私、保护工人的权益,打击平台企业的避税行为(Srnicek,2017)。作者同时提出,国家可以建设公共平台,但这样的公共平台需要独立于国家监控(Srnicek,2017)。国家也可以推动平台的后资本主义转向,利用数据来分配资源,促进公众参与,以期产生更深远的技术进步(Srnicek,2017)。

四、理论贡献与研究启示

(一)理论贡献

本书最大的理论贡献在于其历史分析。全书仅有一百多页,单是关于三次资本主义经济危机的分析,便占据了将近三分之一的篇幅。也正是这一章的论述,让全书的基调确定在了“资本主义”,而不是平台上。作者在这一章的写作也让看似不相干的几个历史事件连接了起来,通过对三次经济危机之间关系的分析,以及三次经济危机与平台资本主义兴起之间关系的分析,揭示了一个普遍的道理:技术进步了,但资本主义的生产关系仍然不变。亦即,平台资本主义的兴起,尽管有技术进步的因素,但更重要的是资本主义发展的进程。

本书同时批判了平台资本主义的兴起对劳动力的影响,具有一定的理论贡献。作者认为,历次经济危机带来了剩余劳动力,这些人更依赖于不稳定的工作,而平台企业,尤其是精益平台高度依赖于这些劳动力。尽管平台企业将自己提供的工作冠以“自由”等一系列动人的标签,但标签之下的剥削关系并未改变,劳动力的实际生存状况甚至面临着恶化的趋势。

本书对数据和知识的定义以及区分、对平台的定义以及分类有一定的理论贡献。斯尔尼塞克并不是最早使用“平台”一词指代互联网和科技公司的学者,格里斯佩(T. Gillespie)早在 2010 年的研究中便探讨过平台的政治学,从牛津英语词典中梳理过平台的不同含义(Gillespie,2010)。许多学者对于平台的探讨,多集中于平台为不同用户的沟通提供了数字化基础设施上(Helmond,2015;

Schwarz,2017;Hicks,2017)。斯尔尼塞克接受了这一技术化的定义,但也抓住了数据这一关键特征,给出了政治经济学层面的一种定义,即平台以数据作为生产的原材料,而不是生产工具。本书对平台的五种分类囊括了传统的制造业与新兴的互联网产业,尽管研究者更多使用的是作为平台的后者,但前者的存在符合作者对于资本主义历史进程的延续性分析,让作者的理论体系更加完备。

当然,本书也有一定的不足之处,例如对于平台的五种分类并不完善,各个分类之间可能存在交叉情况。作者用较大的篇幅探讨了平台资本主义的历史与当下,在对平台资本主义的未来展望上略显不足。但这本一百多页的书不是一部翔实的理论著述或者实证研究,更像是一则启发性的研究,下面将简单探讨本书对既有和未来研究的启发。

(二)研究启示

《平台资本主义》一书对于研究的启发有三个层面:一是基于平台资本主义的研究;二是探讨平台资本主义与平台其他主义之间的关系;三是探讨平台资本主义与其他资本主义之间的关系。

第一个层面的研究,也是本土化比较充分的研究。斯尔尼塞克在书中简短地提到了中国的互联网平台可能是西方平台发展的一种未来,因为像微信这样的平台几乎囊括了所有的生活场景。国内外的一些学者关注到了这类超级平台,探讨了这些平台的形成以及相关的治理措施(Plantin & de Seta,2019;de Kloet,Poell & Zeng et al.,2019;张志安、冉桢,2022;吕鹏、周旅军、范晓光,2022);也有一些学者关注平台如何剥削劳动者(孙萍,2019)。在这一层面,比较缺少从经济史角度梳理中国的货币政策与互联网产业发展之间关系的研究。

第二个层面的研究是关于平台资本主义与平台其他主义之间关系的研究。这一类研究的关注点在于平台的所有权,在平台资本主义当中,作为技术架构的平台被企业所掌控,而如果所有权转移到用户或者集体手中,便会带来关于平台合作主义(platform cooperativism)或平台社会主义(platform socialism)的探讨。前者的研究例如肖尔茨(Scholz,2014;Scholz,2016)、帕帕迪米特罗普洛斯(Papadimitropoulos,2021),后者的研究例如马尔杜恩(Muldoon,2022)、马云志和王寅(2022)。

第三个层面是关于平台资本主义与其他资本主义之间关系的研究。这一类

研究的关注点在于资本主义经济中的不同主体,例如中小企业或者政府。在中国的语境下,有学者关注到了小资本主义与平台资本主义之间的关系(Zhang,2020),也有学者关注到了国家与平台资本主义之间的关系(Rolf & Schindler,2023)。探讨平台资本主义与国家、公众间的关系,涉及政治经济学、管理学、传播学、社会学、经济学等多学科领域,是未来研究的重要议题。

总的来说,《平台资本主义》一书的理论视角值得我们学习和借鉴。后续更多的研究,或立足于平台,或立足于资本主义,或立足于二者的结合点,均可对平台资本主义的历史进行回顾,对其当下进行批判讨论,并对可能的未来进行展望。

参考文献

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。

马云志、王寅,2022,《平台资本主义批判和重构平台社会主义——数字时代对平台经济的新运思》,《河北学刊》第 1 期。

孙萍,2019,《"算法逻辑"下的数字劳动—— 一项对平台经济下外卖送餐员的研究》,《思想战线》第 6 期。

张志安、冉桢,2022,《中国互联网平台治理——路径、效果与特征》,《新闻与写作》第 5 期。

de Kloet, J. , T. Poell & G. Zeng et al. 2019, "The Platformization of Chinese Society: Infrastructure, Governance, and Practice." *Chinese Journal of Communication* 12(3).

Gillespie, T. 2010, "The Politics of 'Platforms'." *New Media & Society* 12(3).

Helmond, A. 2015, "The Platformization of the Web: Making Web Data Platform Ready." *Social Media+Society* 1(2).

Hicks, M. 2017, *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing*, Cambridge: MIT Press.

Muldoon, J. 2022, *Platform Socialism: How to Reclaim Our Digital Future from Big Tech*, London: Pluto Press.

Papadimitropoulos, E. 2021, "Platform Capitalism, Platform Cooperativism, and the Commons." *Rethinking Marxism* 33(2).

Plantin, J. & G. de Seta 2019, "WeChat as Infrastructure: The Techno-Nationalist Shaping of Chinese Digital Platforms." *Chinese Journal of Communication* 12(3).

Rolf, S. & S. Schindler 2023, "The US-China Rivalry and the Emergence of State Platform

Capitalism." *Environment and Planning A: Economy and Space* (online first).

Schiller, D. 1999, *Digital Capitalism: Networking the Global Market System*, Cambridge: MIT Press.

Scholz, T. 2014, "Platform Cooperativism vs. the Sharing Economy." *Big Data & Civic Engagement* 47.

Scholz, T. 2016, *Platform Cooperativism: Challenging the Corporate Sharing Economy*, New York: Rosa Luxemburg Foundation.

Schwarz, A. 2017, "Platform Logic: An Interdisciplinary Approach to the Platform-Based Economy." *Policy & Internet* 9(4).

Srnicek, N. 2017, *Platform Capitalism*, Cambridge: Polity.

Zhang, L. 2020, "When Platform Capitalism Meets Petty Capitalism in China: Alibaba and An Integrated Approach to Platformization." *Iternational Journal of Communication* 14(21).

推进“智慧城市”研究进程的一部力作

——评《智慧城市:大数据、互联网时代的城市治理》第5版*

包涵川　向秋婵**

摘要:金江军教授所著《智慧城市:大数据、互联网时代的城市治理》一书,对于推进智慧城市建设研究具有至关重要的意义。该书分为五部分,首先介绍了信息化和城市化的时代背景,进而阐述了国内外智慧城市建设进展,然后以互联网、区块链等为例,剖析了智慧城市建设的相关技术,再陈述了智慧城市建设的五大构件,包括“新基建”、智慧政府、智慧经济、智慧社会和智慧城市发展环境,最后强调领导干部应树立信息化决策思维,在学术研究当中凸显了前沿性和系统性。基于《智慧城市》一书,未来应厘清智慧城市建设和城市智慧治理的学科关系,兼顾城市智慧治理的“工具”和“目标”导向,在关注技术问题的同时深化技术运用的研究。

关键词:《智慧城市》　智慧城市建设　智慧治理　治理技术

一、引言:《智慧城市》与智慧治理研究的新进展

治理技术的革新不断引领着社会治理现代化。新时代背景下,以物联网、云计算和大数据等新一代信息技术为主要抓手,智慧城市建设成为当前社会治理的重点。建设智慧城市是推进城市治理体系和治理能力现代化的必然路径,是顺应数字化发展形势的关键举措。习近平总书记在党的二十大上强调:“加强城市基础设施建设,打造宜居、韧性、智慧城市。”并提出要加快“网络强国”和“数字中国”建设,加快“推进市域社会治理现代化,提高市域社会治理能力”。

* 本文系国家社会科学基金青年项目“‘党员下沉基层’背景下的社区治理共同体建设机制研究”(项目批准号:21CZZ008)的阶段性研究成果。

** 包涵川,长安大学人文学院;向秋婵,长安大学人文学院。

为此,党和国家以及各地“十四五”规划均把智慧城市建设作为未来城市发展重心。一系列政策的颁布实施旨在加快打造智慧城市,以提升城市管理和社会治理的数字化、智能化和精准化水平。在此基础上,智慧城市建设还构成了各地治理改革和治理创新的重要内容,成为中国特色社会治理的亮点和优势。

从公共治理学术研究出发,智慧城市建设成了学术界关注的核心热点。金江军教授作为中央党校政法教研部副研究员和北京大学智慧城市研究中心副主任,长期从事数字政府、智慧城市等方面研究,旨在以信息化和数字化推进新型智慧城市建设和城市现代化治理发展脉络,为领导干部和相关管理人员提供现代化治理和信息化治理的新思路。金江军教授编写的《智慧城市:大数据、互联网时代的城市治理》第5版(以下简称《智慧城市》)于2021年由电子工业出版社正式出版,该书以丰富的经验素材和深厚的理论基础,系统论述了智慧城市建设的主要内容和实践机制,较为完整地构建了智慧城市研究的学术体系。具体来看,该书系统阐述了中国新型智慧城市建设体系架构,详细解读了中国智慧城市建设的创新性政策,分析了中国各地智慧城市建设的关键问题,并以此为基础提供了切合中国具体实际的建设方案,特别是明确提出了领导干部在城市治理中树立和强化信息化思维的核心意义。正如作者开篇指出:“城镇化是中国经济发展的最大引擎。加快建设新型智慧城市,是新型城镇化的重要内容。”(金江军,2021:5)总的来看,该书概念明确、条理清晰、内容丰富,为读者全面把握智慧城市理论和实践的内在机理提供了深层次的学术思考。

本文拟对《智慧城市》做出评述,首先简要介绍其主要内容,并结合学术界的相关探讨对其进行评述,再阐述该书在多个方面对于智慧城市理论的贡献,最后结合书中内容与当前实践进行发散性思考。

二、《智慧城市》的主要内容和学术贡献

(一)《智慧城市》的主要内容

《智慧城市》一书开篇梳理了中国城市化和信息化的发展历程,进而提出了“智慧城市”的概念,并将其界定为通过广泛采用物联网、云计算、移动互联网、大数据等新一代信息技术,提高城市规划、建设、管理、服务、生产、生活的自动

化、智能化水平,使城市运转更高效、更敏捷、更低碳。该书指出,物联网、云计算、移动互联网、大数据、人工智能和区块链等技术是打造智慧城市的六大关键信息技术。而“新一代信息基础设施”、智慧政府、智慧经济、智慧社会和智慧城市发展环境是构成智慧城市的五大部分(图 1)。

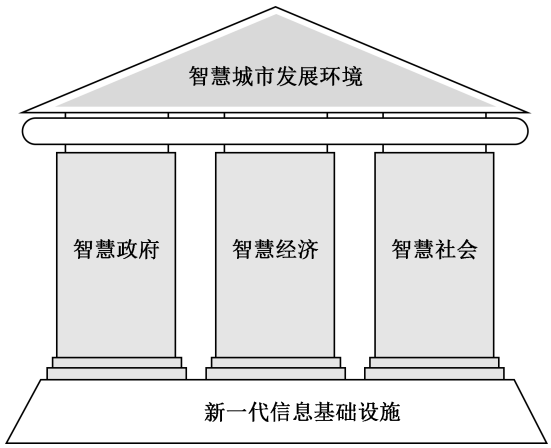


图 1 智慧城市的五大组成部分

资料来源:金江军(2021:10)。

该书共 14 章,主要围绕智慧城市的六大关键技术和五大组成部分依次展开,行文脉络清晰。总的来看,按照各章内容和结构的差异,该书可以被分为五大部分。

第一部分首先介绍了智慧城市的发展背景——中国城镇化和城市信息化接力融合发展。金江军教授认为,城市信息化相继经历了数字城市、信息城市、智慧城市和互联城市的发展阶段。该部分紧接着分析了智慧城市五大组成部分和六大关键技术的体系构架。作者指出,城市建设的实践活动必须有科学理论作为支撑,于是与智慧城市密切相关的基础学科城市信息学应运而生。在金江军教授看来,对于智慧城市建设的研究,既是学术不断结合社会发展的实际需要,又是新兴学科发展的必然要求。

第二部分介绍了国内外智慧城市的发展情况。比如,作者论述了各个国家对于相关技术的具体运用:美国采取逐个突破重点领域的策略,着力打造智能化方针系统 Urbansim;韩国推出 u-Korea 发展战略并提出具体实施方案 u-City,不断加强信息基础设施建设和移动互联网的应用,使信息技术渗透到韩国城市的

各个方面;中国相继制定智慧城市建设政策方案,在融合多种技术的同时不断推进城市治理现代化进程。最后,该部分指出了中国面临的现实问题,包括在智慧城市建设上存在理论研究滞后于实践、CPU 等关键技术和产品储备不足以及缺乏各部门间的横向协调机制的问题。作者在分析这些问题的基础上,有针对性地提出要加强基础设施建设,开展智慧城市创新应用,发展智慧城市相关产业和理顺体制机制。实际上,相关领域已经成为各个地方政府施政的重要着眼点。

第三部分详细阐释了为建设智慧城市提供技术保障的科技创新,包括物联网、云计算、移动互联网、大数据技术、人工智能和区块链这六大关键技术。在信息技术融合运用的基础上,作者分析了六大关键技术的相关政策和应用现状,并提出了相应的发展对策。

第一,物联网技术。《智慧城市》一书首先探讨了广泛运用的物联网技术,介绍了由该技术衍生而来的传感器、二维码、射频识别技术和 M2M(Machine to Machine)技术。作者认为物联网技术仍有着极大的应用价值等待挖掘,尤其在信息化和工业化(“两化”)和电子政务领域的应用仍存在较大的提升空间。

第二,云计算技术作为信息时代蓬勃发展的一种信息技术也受到了关注。书中说明,作为一种实现共享软硬件资源的网络计算模式,云计算技术具有快速、弹性、按需服务等优点。“云计算的资源是虚拟化的、可以动态扩展的”,用户根据自己的实际需求获取相应的资源和服务,“使人们可以像用水、用电一样使用计算资源”(金江军,2021)。由此出发,该书着重强调云计算技术应发挥促进政务信息化从电子政府到云政府转变的功能,而政务信息化的进步亦能够推动整体意义上国家治理科学化的不断发展。

第三,移动互联网技术在建设智慧城市中也发挥着至关重要的作用。《智慧城市》一书在分析移动互联网技术应用现状的基础上,指出移动治理技术的三个发展方向:一是加强与各个电信运营商的合作;二是制定适用的行政法律法规;三是选择正确的移动电子政务作为切入点。以政府信息为中心的电子政务正在向以人民或用户为中心的数字政府转变。实际上,学术界已经发现,数字政府是电子政务的继承和深化,而移动互联网技术在此方面发挥了至关重要的作用(翟云,2022)。

第四,大数据技术。《智慧城市》指出,大数据技术通过建立集中、共享数据资源的平台,能够高效便捷地处理日益复杂、零散的信息。在介绍大数据技术的

基础上,作者分析了大数据在社会安全综合治理、税务管理等领域的应用情况。在此基础上,该书也指出了当前存在的一系列问题,并认为政府仍需在掌握关键技术、开放公共数据资源、制定相适应的法律法规和建立“以人民为中心”的政府数据网站这四个方面做出努力。随着政府和社会组织的运行越来越依赖于大数据技术,当前数据安全和个人隐私等问题层出不穷。作者强调,在注重公开信息资源的同时,数字安全的现实困境也不容忽视。

第五,人工智能技术。作为近年来最重要的科技成果之一,人工智能在城市治理中崭露头角。金江军教授认为,中国在人工智能技术发展当中既存在优势,也有着相对不足之处,而未来应从产业支持、政策扶持、供需对接、平台建设和促进人工智能与实体经济融合这五个方面入手,持续推进智慧城市建设进程。而在具体的技术工具方面,《智慧城市》一书指明,机器人制造、“新硬件”是人工智能应用的重点发展方向。实际上,在现实治理实践中,“政治机器人”已经开始初步运用,并大量参与到公共治理活动当中,这构成了城市管理领域的前沿(张梦、虞崇胜,2022)。

第六,区块链这一新兴信息技术以其去中心化和数据难以窜改的特征,在金融、公共服务、物联网和物流等领域得到充分应用。通过实地调查,作者发现区块链技术已经逐步融入政务服务和社会生活中,并以佛山市禅城区的实践为例做出了分析。禅城区推出了“区块链+”政务服务、“区块链+”共享社区和“区块链+”社会矫正等新模式,构建“自信+他信+你信”的信用机制。金江军教授认为,禅城区在创新区块链运用模式方面给其他地区树立了典范,具有较强的借鉴意义。

在《智慧城市》的第四部分,即书中第9~13章中,金江军教授向读者系统介绍了智慧城市的五大构成要件,即“新基建”、智慧政府、智慧经济、智慧社会和智慧城市发展环境。

首先,“新基建”在智慧城市中的作用犹如房屋的地基,必须扎稳打牢。“新基建”为智慧城市、数字经济、智慧生活等发展提供动能。在此基础上,《智慧城市》一书指明了新基建的发展路径,政府要协同企业从搭建城市骨干网和宽带城域网处着手,实现城市无线网络全覆盖目标的同时推进电信网、广播电视网和互联网“三网融合”。

其次,新型智慧城市要在筑牢“新基建”的基础上稳步推进智慧政府的建

设。作者认为:“新型信息基础设施是‘新基建’的核心内容。‘高速、泛在、集约、安全’的城市新型信息基础设施是城市基础设施的重要组成部分,是智慧城市的重要支撑。”(金江军,2021:141)作者据此评介了诸多案例,如美国 2012 年发布的电子政务战略、德国主导的“数字德国 2015”等。在此基础上,作者有针对性地提出以下建议:中国应建设集中、统一的政府云,大力发展移动电子政务,推进“数据大集中”和进一步促进政府信息公开。实际上,当前的研究与实践早已指出,促进信息集成和共享、推动民众广泛参与是智慧政府建设的关键(卫鑫、陈星宇,2020)。

最后,该书对智慧经济、智慧社会和智慧城市发展环境做出了分析,其中包含着独特而深刻的见解。作者着重论述了发展智慧经济方面的可行举措,提出需在智能制造技术、智慧产业、智慧企业、共享经济和智能产品这五个方面做出努力,继而融合运用到智慧教育、智慧医疗、智慧社区、智慧家庭和智慧旅游领域。该书最后借助 PEST 模型(宏观环境分析)解析了中国智慧城市发展环境,包括相关的智慧城市政策、稳步推进的数字经济、信息化时代的到来和新一代技术的快速发展等方面。

《智慧城市》的第五部分创造性地提出了城市领导干部应树立信息化决策思维,强调从信息化角度去思考、解决城市发展所面临的问题(金江军,2021)。新时代以来,信息化思维与新时代的发展和公共政策创新相契合,打开了城市治理的新思维,筑牢了将政府纳入数字化轨道的思想基础。在当前时代,具备数字素养是公务员胜任数字政府场景与任务的关键(张红春、杨欢,2023)。进而言之,政府要持续提升干部队伍数字意识和数字素养。在此方面,该部分探讨了信息化对统筹“四个全面”战略布局、贯彻五大发展理念和推进“五位一体”总体布局的促进作用,从政策理论根源说明信息化思维的树立方式,表明城市领导干部和管理人员要具备领导决策、城市安全、经济发展、政府履职的信息化思维。

实际上,在推进社会治理现代化的进程中,中央和各地已经意识到树立信息化思维的重要性。科学化和信息化决策思维亦在很大程度上改变了一些领导干部“拍脑袋决策,拍胸脯蛮干”的行为方式。各级干部借助信息化手段及时地收集信息、了解情况并进行科学决策,用信息化思维发现问题、解决问题。实践当中,相关思维在各个领域得到了充分运用。比如,税收大数据通过整合跨地区、跨行业的收入信息,精准推算出相关人员偷逃税款的情况,重点打击了非法收入

行为。金江军教授基于国内外的丰富治理实践,最后再次强调树立信息化思维的意义,将其作为全方位提升城市治理体系与治理能力的现代化水平的关键进路。

(二)《智慧城市》的学术贡献

结合上述内容,《智慧城市》一书在智慧治理和城市治理等领域做出了较为突出的学术贡献,具体包括以下两个方面。

一方面,《智慧城市》做出了前沿性探索。相比于其他研究,城市信息化是兼具技术性和前沿性的话题,这要求相关著述必须紧跟科学技术发展的最新成果,并对其在城市治理领域中的崭新实践进行回应。《智慧城市》一书凭借广阔的视野,对于诸多新兴技术做出了介绍,并结合国内外多个实践案例对技术运用进行了评介。这既为智慧城市建设的学术发展注入了丰富的经验素材,又凭借内容的丰富多样性在理论研究当中平添了阅读趣味。

另一方面,《智慧城市》做出了系统性研究。目前,关于数据治理、信息治理和技术治理等领域的探讨已经屡见不鲜。但是,较少有研究将智慧城市建设的多种维度综合起来加以分析。《智慧城市》一书在此方面做出了突破,其主要内容几乎涵盖了城市治理当中的各种技术,并展现了智慧城市建设的时代背景、治理环境、技术运用、实践机制和干部思维,涉及智慧城市建设的公共政策、经济运行和社会基础等诸多重要领域,表现出了较强的系统性和全面性,从而将不同学科和不同领域有机综合起来,将智慧城市的研究不断深化。

三、智慧城市建设研究的发展进路:对《智慧城市》的延伸思考

可以看到,金江军教授所著的《智慧城市》不仅契合了当前蓬勃发展的智慧城市建设趋势,还从学术视角出发对相关实践进行了深入解读。下文将根据《智慧城市》一书的主要内容,对当前智慧城市的相关研究做出延伸性思考。

从学科发展视角观之,厘清智慧城市建设和城市智慧治理的学科关系是未来智慧城市建设研究的奠基性工作。国内外相关话题的研究大多围绕两个维度展开,一是如何通过新兴技术建设智慧城市,本质上是城市科学的重要领域。

《智慧城市》一书便属于此种类型,它紧紧围绕城市化和信息化两大主线,对城市建设问题进行了探讨。二是如何在城市当中通过新技术优化治理体系,本质上是公共管理学的重要议题,体现为诸多关于城市智慧治理的研究,其着眼于城市治理有机体的实现。在未来的学术研究当中,应不断融合两种学科之间的优势,将城市建设和治理研究结合起来。

从对待城市治理的“技术”的态度和研究路径来看,城市智慧治理研究既需要坚持工具主义的技术化路径,也需要关注目标和价值导向的智慧“善治”。《智慧城市》一书不仅提出信息化是合理指导城镇工作的必然要求,亦强调了作为治理途径的“城市信息化”之意义。实际上,如何在城市治理中看待新兴技术,学界有着不同的观点,分别表现在究竟是“以技术为目标”的治理,还是“以技术为工具”的治理,前者要求形成信息化的治理思维,而后者则仅将其视为一项达成善治目标的工具(郁建兴、高翔、王诗宗等,2023)。未来应更进一步辨析两种视角的各自优势,探讨智慧城市建设的基础性问题,而《智慧城市》一书在此方面做出了重要的先导性探索,在上述两个方面均提出了创见性思考。

从拓展既有智慧城市的研究域和研究对象来看,未来城市智慧治理研究要在关注技术问题的同时,深化对于技术运用的研究。智慧城市建设研究不仅要重视各项现代技术,更要关注技术在实践中的运用过程。城市治理是重要的实践场域,没有相应的制度和人员行动,任何技术都不会转化为实际效应。实际上,在许多地方,城市运行检测网络已经较为成熟,但经常出现许多人员盯着数据“大屏”却无法拿出治理方案的困境。《智慧城市》重视此方面的现实困境,并指出“政策环境”在智慧城市建设中的意义,更将“信息化思维”作为该书的落脚点。依循既有探索,未来应更加重视优化对于技术运用的研究,关注各地在治理中采取的机制,并提出切实可行的优化方案,从而提升城市智慧治理能力。

四、结语

智慧城市议题具有重大的实践意义和理论价值,既能够推动治理体系的不断完善,又能够在危机情境下显著提升城市的韧性水平(武永超,2021),不断缓解愈发显著的“城市病”。当前地方政府的横向竞争,在相当程度上也是智慧治理能力的竞争,这构成了城市综合实力的重要方面。有鉴于此,金江军教授的

《智慧城市》指出了深刻理解智慧城市系统结构和内在机理的核心意义,为智慧城市运作规律和原理的探索,以及城市智慧治理政策网络的优化做出了重要贡献。

从《智慧城市》当中可以看出,相关话题既是学术发展的前沿领域,又是学科交叉领域,在技术研究的基础上涉及公共管理学、城市规划、经济学和社会学等多种学科。正如本文所指出的,该书既是近年来关于智慧城市建设的系统论述,又是对于各类技术及其在治理领域中运用过程的有力解读,更是在信息技术不断发展背景下对国家治理的深入思考。未来应进一步以《智慧城市》的出版为契机,顺应学科发展的时代趋势,整合不同学科的优势资源,不断将相关研究引向深入。

参考文献

金江军,2021,《智慧城市:大数据、互联网时代的城市治理》第 5 版,北京:电子工业出版社。

卫鑫、陈星宇,2020,《智慧政府的功能定位及建设路径探究》,《中国行政管理》第 7 期。

武永超,2021,《智慧城市建设能够提升城市韧性吗?——一项准自然实验》,《公共行政评论》第 4 期。

郁建兴、高翔、王诗宗等,2023,《数字时代的公共管理研究范式革命》,《管理世界》第 1 期。

翟云,2022,《数字政府替代电子政务了吗?——基于政务信息化与治理现代化的分野》,《中国行政管理》第 2 期。

张红春、杨欢,2023,《数字政府背景下的公务员数字素养框架——一个概念模型》,《电子政务》第 1 期。

张梦、虞崇胜,2022,《智能时代政治机器人的初步应用、挑战及风险治理》,《华中科技大学学报》(社会科学版)第 5 期。

数智化时代的智慧养老:概念、发展及前瞻

——对智慧养老探索者左美云教授的访谈

左美云　张承蒙*

中国人民大学二级教授、信息学院副院长、智慧养老研究所所长左美云教授是国内较早从信息管理学科视角关注智慧养老、智慧康养医养技术迭代的学者,他指出智慧养老的理想模式应朝线上线下协同、助老孝老用老一起发力、人机结合、虚实结合方向发展,致力于倡导养老数智化转型的概念和实践,并多次接受中央电视台和中国教育电视台的专访。

一、当我们在说“智慧养老”时,我们在说什么?

(一)智慧养老:内涵与模式的建构路径

张承蒙:《智慧养老:内涵与模式》这本著作凝聚了您大量心血。这本2018年出版的书,从五年后的今天来看,仍然可以给我们很多启发。您决定出版这本书时,是考虑到促进学术、服务实践,还是为了进一步引起争鸣,推动这一领域的发展?

左美云:从“内涵与模式”的名字来看,这是一本比较学术的书。实际上我不知道这本书后来会卖得这么好,我知道清华大学出版社出版该书一年内就连续印了四次。至于我们团队提出来“智慧养老”这个词,也不是为了引起争鸣。我们是想对它的内涵做一些界定,归纳一些模式,希望给学术界提供一块铺路石。这是我们在这个领域的第一本学术专著,我们首先想做的是让大家对“智慧养老”有一个共同理解。

之前有过“智能养老”的概念,跟“智慧养老”其实是两个不同的阶段,跟我

* 左美云,中国人民大学智慧养老研究所;张承蒙,北京大学人口研究所,《智能社会研究》青年编委。

们的国情发展有关。“智能养老”更多强调物联网、传感器等用于监控的技术和设备,是刚开始的阶段,我们要去适应技术、适应设备。随着社会发展和技术进步,技术设备的性能提升,我们开始“关注人”。“智慧养老”就是要“以老年人为中心”(左美云,2018)。比如充电这样一个小事,要让老年人隔一两天就充一次电对他们来讲是很难的。要想让老年人接受产品和设备,那么一定要以老年人为中心。智慧养老的目的,是让老年人融入数智化社会并能享受到其发展的红利,其中包括老年人的数字生活,打造老龄友好的数智化社会。我们界定“智慧养老”(smart senior care,SSC)的概念,是指利用信息技术等现代科技技术(如互联网、社交网、物联网、5G、移动计算、大数据、云计算、人工智能、区块链、数字孪生等),围绕老人的生活起居、安全保障、医疗卫生、保健康复、娱乐休闲、学习分享等方面支持老年人的生活服务和管理,对涉老信息自动监测、预警,甚至主动处置,实现这些技术与老年人的友好、自主式、个性化智能交互,一方面提升老年人的生活质量,另一方面利用好老年人的经验智慧,使智慧科技和智慧老人相得益彰,帮助老年人更便捷地实现社会参与,尽可能地增强独立性,最终使老年人过得更幸福、更有尊严、更有价值。智慧养老包括三个方面的含义,分别是智慧助老、智慧孝老和智慧用老。其中,“智慧助老”和“智慧孝老”属于“为老”(for seniors),“智慧助老”主要是物质的支持,“智慧孝老”主要是精神的支持;“智慧用老”则是靠老(by seniors),即充分利用好老人的经验、知识和技能。

我们希望能够把我们的一些学术成果,以及观察到的一些实践归纳,放到《智慧养老:内涵与模式》里,因而这本书学术和实践两方面的内容都有。我们走访调研了很多地方,在书里有一些学术的探索,也有我们对模式的归纳,包括一些国内外模式的归纳。基本上国内早期做智慧养老的企业,跟我们都有联系。我们把他们一些好的做法做了归纳。2018 年这本书出版以后,提供了一个讨论的基础。

“争鸣”的话,我们 2011 年提出了“智慧养老”的概念,到了 2014 年得到了山东财经大学管理科学与工程学院院长(现副校长)张新教授的肯定和热情支持,我们一起发起了全国性的“智慧养老与智慧医疗论坛”,这个论坛起到了争鸣和传播的作用,以学术活动推动研究、凝聚学者。另外在我们整个老年学研究领域,我觉得氛围比较好,相互之间协奏、共鸣,齐心协力在做一些事情。老年学领域的杜鹏教授、陈功教授都比较欢迎我们的研究,他们也提出过“重视科技创

新在养老服务领域的应用,建立健全智慧养老标准”(杜鹏,2022)、“以技术支撑智慧养老、促进养老服务体系优化”(陈功、赵新阳、索浩宇,2021)等观点。中国老年学和老年医学学会教学研究委员会的“中国老年学学科建设研讨会”每年也都会设置“智慧养老”相关论坛,对“智慧养老”传播起到了推动作用。

张承蒙:您在《智慧养老:内涵与模式》一书中,将自己的研究转向归因为“做有同情心的学术”和“发现学术蓝海”。您觉得从您的学科基础和研究专长来讲,“同情心”怎样体现在研究中?

左美云:从大的方面来讲,当代大学在做好人才培养、科学研究之外,还有一个很重要的职能,即服务社会。习近平总书记多次叮嘱科研工作者要“把论文写在祖国的大地上”(习近平,2020)。中国人民大学一直有着始终“与党和国家同呼吸共命运”的使命。作为科研人员,我们也经常考虑我们国家的痛点、难点问题是什么,而人口老龄化,无论对于世界还是中国,都是一个巨大的挑战,当然也孕育着巨大的机会。我们认为老年人群值得学术界去关注。

从微观层面看,家家都有老人,人人都会变老,随着年龄的增长,老年人也面临着更多的不利因素。尤其当面临一些公共卫生危机时,老年人会显得更加脆弱,需要更多的照护和陪伴。“有同情心的学术”是受到了著名管理学家徐淑英教授的启发,她倡导“我们要对那些受邀参与调研的人抱有同情心(compassion)”(徐淑英,2015)。我们十多年前就发现,随着数字化的快速发展,各类智慧技术层出不穷,但许多没有考虑到老年人的生理和心理特征,加重了老年人的认知负担与抵触情绪,“数字鸿沟”问题日渐凸显,老年人对于社交媒体和移动支付工具的使用不太熟练,往往在社会生活中遇到碰壁的情况。我们之前在组织学生参与面向老年人的志愿服务活动时,也发现老年人心理比较敏感和独特,他们会很爱惜这些电子产品,以至于不太敢去使用。我们发现了这样的一些问题,所以认为应当去关注和关心如何用我们掌握的信息技术协助老年人的日常生活。

张承蒙:接着刚才的话题,您在《智慧养老:内涵与模式》中指出“信息技术发展对老年人的影响”是您发现的“学术蓝海”,这对于学术的价值是什么?

左美云:从学术研究上来讲,2007 年我作为中国人民大学信息学院经济信

息管理系主任,受邀赴新加坡国立大学信息学院(School of Computing)访问,和几位同行老师畅谈信息系统学科未来的发展,主要谈到了三个领域:老年人、农村、医疗。我们当时认为这三个领域的研究还非常少,甚至属于空白,尤其是农村和老年人的"数字鸿沟"或者说"信息劣势"是很明显的。2007年,企业信息化研究,即"企业怎么建设信息系统、怎么使用信息系统",已经到了一个研究的"红海"。当时关于企业电子商务和政府电子政务的研究也有很多了,所以大家就觉得要找研究还比较少的领域,就是上面的三个内容。我们观察到了老年人这样一个特定群体,是一片学术的蓝海。现在关于智慧医疗的研究非常多,但在2007年,我们考虑到医院里的职工素质都很高,所以觉得这个可能很快会有人去做。而农村的信息化由于要经常出京调研,可能相对城市老年人来讲会比较难介入。所以最后我们选择了城市的老年人作为主要研究对象,进入了这片学术蓝海。

张承蒙:我们想在此基础上再了解一下,关于《智慧养老:内涵与模式》这本著作,您的写作初衷跟刚才您讲到的"发现学术蓝海"和"做有同情心的学术"有关吗?

左美云:首先从学术成果的价值来说,要能经受得起时间考验。若作品发表多年以后还能具有影响力,说明作品中的思想能够"立得住"。其次我经常讲,这本书"十年磨一剑",花了我的黄金十年,也是我带领的团队十年集体心血的结晶。我们在写作的过程中,也做了一些沉淀和梳理,将那些我们认为能"站得住"的内容写到书里。我的写作初衷,跟刚才说的一样,"板凳甘坐十年冷"。我之前做企业信息化研究,也在学术界和产业界提出了一些新的观点,会有声音说:"你做企业信息化研究做得挺好的,你为什么去研究养老的问题?"那个时候的老年人话题,大家都感觉有点"夕阳"的意思。那段时间完全是坐冷板凳。我要去找街道工作人员调研,去公园里边找老人调研,到养老院找照护人员和老人做调研、做访谈。老年人其实对计算机和互联网不是那么有自信心,也不是很愿意接待我们。当时从信息技术的角度去谈,老年人经常说:"我们这儿没有什么信息化""你讲的东西对我没用"。所以去找老年人做访谈是挺难的,那个时候我们师生访谈回来经常会有挫败感。但是我们愈挫愈奋,因为目标是清晰的。如今全社会都关注老年人话题,老年人也感觉受尊重了。现在我们去调研,老年

人都很热情,也很愿意接受调研。想起刚开始做这一领域的时候,我们经常碰壁,但是我们还是坚信“信息社会的光芒会照耀到每个人”,要考虑怎么去帮助弱势群体、怎么覆盖到不同层次的人群。所以我们经常说坚守初心,做有同情心的学术。

(二)从管理信息系统走向智慧养老:学科话语体系廓清

张承蒙:从您的学科背景来讲,您觉得管理科学和工程学科带给您的是什么?或者说,这个学科为您研究老年学、智慧养老提供了哪些便利?

左美云:我博士期间学习的是管理科学工程学科的二级学科——管理信息系统。管理信息系统有三个主要研究方向:一是研究信息系统的采纳和使用行为,跟心理学和行为科学有一定关系,从行为的角度去研究“怎样使用一个系统”;二是研究和设计信息系统,即怎样设计一个信息系统,跟计算机和软件开发有很强的相关性;三是随着大数据的发展,基于系统产生的数据,对数据进行分析和预测,找到管理和决策的优化路径。从学科视角来讲,管理信息系统(或者信息管理与信息系统)涉及两方面的学科:一是管理科学,涉及系统的优化;二是工程管理和应用,涉及具体的产品设计和实现。上述三个研究方向和两个学科基础,这五个方面的学科属性都会对做养老研究有一些启示。

比如,从信息系统采纳和使用方面来讲,在接触信息技术设备和技术的时候,老年人的行为特征是怎样的?老年人群有什么样的行为,为什么会接受和采纳某一个具体品类的设备和技术?为什么不接受某一个具体品类的技术?采纳和不采纳的老年人各有什么样的行为和特点?我们做研究特别希望能解决$y=F(x)$的问题,找到F的规律(模型或机制),其中y是因变量,比如采纳意愿、持续使用率、满意度等,x是自变量,既可能包括人口学、心理行为学等因素,也可能涉及产品功能有用、易用等因素。信息系统采纳和使用的类似研究对于智慧养老的研究会有借鉴和帮助。在做智慧养老之前,我们比较多的是研究企业信息化、电子商务,从企业信息化的研究中关于“个人信息设备的使用”这方面的内容,转移到了老年人这一特定人群上。

从设计科学来说,面向用户进行人机交互的设计时考虑到适老化,需要关注如何面向老年人设计可穿戴设备、手机应用等信息技术产品。管理科学比较强调优化、数据分析,我们就会对老年人行为做建模,基于数据分析老年人会有什

么样的使用行为。比如,我们会关注:在老年人比较频繁地使用一些购物软件购买产品的过程中,点击行为流是怎么样的?第一步做什么?第二步做什么?我们也可以根据老年人的用户画像做一些产品推荐和老年用户流失预测。对于我来说,主要还是从自己学科的研究范式、研究方法出发来关注智慧养老问题。

二、智慧养老,作为一项社会议程

(一)概念变革背后的社会现实

张承蒙:在您看来,老年人群是否认同“智慧养老”相关的一系列技术和理念?

左美云:如果没有跟老年人讲清楚“智慧养老”的含义,老年人就会感觉听不懂。智慧养老中的“智慧”一词跟我们平常用到的“智慧”含义不完全一样。我们约定俗成的“智慧”就是“老年人的经验、知识”,但这里的“智慧”更多讲的是“智慧技术”以及“合理使用这种技术”。所以如果没给老年人讲清楚,可能就会被老年人认为是一种噱头或者一种学术标签。

每次当我把“智慧养老”带来的美好的老年生活呈现给老年人时,我一般会从场景上讲,在从早上起床开始到晚上睡觉整个过程中,每个环节如果有了技术加持,生活会有多好。这样老年人都会觉得这是“我需要的”,“智慧养老就是我们未来应该选择的一个模式”。我的微信里边也有一些老年朋友,他们跟我一直保持着联系,这些老人是懂智慧养老的,他们觉得这些技术很便利,体验是很好的。所以,要把智慧养老的作用和影响给老年人描绘出来。

另外还有一个很重要的因素,就是让老年人意识到未来的养老模式是这样的。比如今年春节我跟我父母在一起的时候,我就跟他们讲:“你们这一代人是最幸福的一代人。为什么呢?因为你们有四个子女。如果你们病了,我们四个人都能来照看。而我只有一个儿子,我就不能够完全指望他。我可能更多是指望机器人,指望智能设备。”所以我们未来面对的就是这种人力资源、服务人员的短缺,面临着无人养老的困境。我在家里也有些一直在用的人工智能助手,我可能会比一般的老年人更清楚,知道这些数字的“虚拟人”或物理的“机器人”很快就要来到我们身边,并且已经在影响我们的生活。当把这些东西讲给老年人

听的时候,老年人就会感觉智慧养老不是噱头,是我们要去追求的。当然,是不是能够做到,需要我们学术界和产业界一起去努力。

张承蒙:您觉得现在的老年人能够明白“智慧养老”的含义,并且能够如科研者所期望的那样使用技术吗?老年人作为“智慧养老”的受益者,我们怎么做才能让他们真正受益?

左美云:现在很多做智慧养老的企业,当下想赢利还挺难的。他们现在大部分都是靠政府购买智慧养老的服务,维持在生存线上。当然也有一小部分开始从盈亏平衡转向开始赢利。这里有个比较大的问题,如果政府或者企业家对智慧养老没有正确的认识,认为“普惠”的智慧养老就是给所有老年人都提供同样的设备、产品,就会造成资源的错配和闲置。我们应该给老年人提供普及性的信息服务或者养老服务,但提供的方式应该因人而异,因为不同的老年人具有不同的信息素养和信息能力,应该提供不同的设备给这些老年人。有的老年人喜欢看电视,就提供电视接入相关养老服务的方式;有的老年人喜欢用智能手机,就通过智能手机提供相关的养老产品和服务。普惠跟合理要结合起来,否则就是“叫好不叫座”,大家说起来都支持,但是要想赢利就很难。另外,不能只考虑老年人怎么“智慧”、合理地使用技术,而应当考虑科研者与厂商怎么样去合理、“智慧”地提供技术。老年人作为服务接受方,要让他们很合理地使用技术有点难,特别是这一代老年人。“智慧养老”中“智慧”的英文叫 smart,就含有“灵巧”“聪明”的意思。作为服务的提供方,应更多考虑怎样给老年人提供技术。

张承蒙:是否有让您印象深刻的具体智慧养老案例,这一案例对您的研究有什么影响?

左美云:我这里讲一个关于智能床垫的有趣例子。有一个女婿给老人装了一个智能床垫,他就可以在手机上看到老丈人晚上的睡眠情况。女婿发现他老丈人总是在晚上 12 点左右就要起来坐一会儿,坐半小时之后又睡了。女婿就去问老丈人为什么每夜要起来坐一会儿。老丈人说最近开了一种药,晚上 9 点来钟服药,到了 12 点以后就比较难受,要坐一会儿,抚摸心脏就好了。女婿知道后第二天就带着老丈人去找大夫调药,后续没再出现这个情况。再过了一阵子,女婿发现手机上啥也看不着了,去家里看他老丈人,就发现智能床垫的电源被老丈

人拔掉了,已经采集不到信息了。这个案例给了我们正面的启示,也给了一个反面的启示:正面的启示是智能床垫起到了健康监测的作用,反面的启示是老人意识到被监测了,不知情的情况下隐私没有得到重视,索性把电源拔掉了。我们经常在团队里讲:一是智能床垫确实可以起到监测、预警的作用;二是在安装这些产品的时候,老年人是否有知情权,他的隐私问题是否得到关注。这对"智慧养老"而言是一个很有趣的案例!

(二)技术驱动带来的观念重构

张承蒙:我们经常会提到技术的社会影响力(social implications of technology),讲到智慧养老的时候,主要还是将其用于满足我们老年人日益增长的美好生活需要。我们想了解的是:我们现在智慧养老的现实需求是不是清晰的?智慧养老供需之间的矛盾,是由于技术的不足,还是市场的不足?

左美云:我认为智慧养老肯定是一个经济现象,也是一个社会现象。从经济上来讲,智慧养老本身就是一个产业,即智慧健康养老产业。工信部、民政部、国家卫健委等部委发布了《智慧健康养老产品及服务推广目录》,并且也在评选智慧健康养老产业的基地,所以从经济上来讲,智慧养老是我们国家未来新的经济增长点。我们之前的研究发现,政府对于智慧健康养老发展的推动,主要通过建立示范试点进行(左美云、刘浏、尚进,2021)。比如,山东省就把打造智慧健康养老产业作为一个很重要的经济增长点。从全社会来讲,智慧养老会推动整个银发经济以及老龄产业、老龄行业的发展。

回到问题本身,现在的智慧养老,需求的类型和层次是清楚的。谈需求,我们一般会按照马斯洛需求层次理论的五个层次去做分类,即生理需求、安全需求、情感需求、尊重需求和自我实现需求。马斯洛的需求层次理论已经得到广泛的认可,是各项行为科学研究的基础理论之一。我们可以借助马斯洛的这一理论,推导出老年人的需求。但是落实到具体场景的需求,是不够明确的。我们一直呼吁的是,智慧养老一定要落实到老年人的具体生活和应用的场景上。2020年,上海市就做了一个很好的事情。上海市民政局联合市经信委发布了12个智慧养老应用场景需求,包含照护服务类、安全防护类、情感关爱类等不同类别的场景。国家层面也相继发布了《国务院办公厅印发关于切实解决老年人运用智能技术困难实施方案的通知》《全国老龄办关于开展"智慧助老"行动的通知》等

文件,文件中也划分了消费、出行、就医、办事等场景。但在每一个具体的场景下,老年人会有什么样的行为特点,有什么需求?这个目前还不够明确。原因有两方面:一是不会用智能设备的老年人不懂技术,很难表达需求;二是企业端产品市场调研没有下沉到老年人中去。智慧养老产品琳琅满目,但是做出的产品是不是适合老年人使用,我觉得需求识别问题还是没有得到很好的解决,没有得到充分的理解。从市场上来讲,一是现在企业竞争多、合作弱,数据、接口标准不统一;二是各个产品生产厂商没有实现体系化建设,形成产业联盟才能更好地满足地老年人对智慧养老的需求。

张承蒙:智慧养老与科技适老化的关系是什么?您曾提出要在"界面适老"的基础上,进一步研究"交互适老"和"共融适老"(人机共融)(左美云,2023),您是否可以大体描述一下交互适老和共融适老的场景?

左美云:在我们管理信息系统里边有一个很重要的研究,就是人和机之间的关系(human-computer interaction,HCI)。这个"I"最早是 interface,强调界面。之前工信部主导的《互联网应用适老化及无障碍改造专项行动方案》,大部分都还停留在"关注界面"的维度,像"大字体、大音量、大图标、大按键"所谓的"四大"等。我们倡导的是要重视老年人与智能产品的"交互",即 interaction,做到"交互适老"。要研究老年人使用设备或应用程序时,他的第一步做了什么,第二步要做什么,对这种步骤和过程进行动态的适老化,包括一些类似操作出错的提醒内容。从老年人的心理特点来说,老年人喜欢那种一个一个功能、类似"开抽屉"的交互过程,就是老年人打开这个"抽屉"发现是这一个功能,打开另一个"抽屉"发现是另一个功能。所以每一个功能,比如出行的功能是不是能做成一个"抽屉",消费的功能做成另一个"抽屉",功能使用完以后就把对应的"抽屉"放进去。人机共融是比较高层次的人机互动,即 integration,做到"共融适老"。老年学上经常讲要"保护身体残存机能",如果这些产品一下就把老年人的所有功能都替代了,就不利于老年人的"成功"老化。因而比较好的设计是合理分工老人和智慧技术的任务,既做到对老年人生活起到支持作用,又能维持或调动老人残余的功能。

张承蒙:您觉得最近兴起的 ChatGPT 技术,对智慧养老是否有启发和借鉴,

可以用到智慧养老的哪些方面?

左美云:我们前期探讨过元宇宙、人工智能等新兴科技在智慧健康养老领域适用性的问题,我们认为可以通过创新元宇宙等技术的接入方式,在需求识别基础上促进健康养老发展(左美云、沈原燕杭、刘妃,2023)。ChatGPT 这个工具我们在关注,并且我们也在用它做一些研究,就是面向老年人的智能问答系统。ChatGPT 是一个问答系统,国内也已经有很多类似的聊天机器人产品。现在 ChatGPT 的用户有许多程序员和学生,他们写程序、写论文在用,但我觉得在智慧养老领域最好还是回到 ChatGPT 的本来面目上,就是聊天。老年人很容易感受到孤独,如果 ChatGPT 或者类似工具,能够帮助老年人做一些精神陪伴、情感上的交流,我觉得是非常好的。今后的 ChatGPT 不应该仅仅是智能问答,未来也可以帮助老年人做一些智能创造,老年人可以通过表达比如"我希望画一个什么样的画",老年人说一段要求和大概的意境描述,ChatGPT 就能产出一幅画。能说这幅画就是 ChatGPT 做的吗?我觉得不完全是,老年人的描述内容也很重要。管理学领域里边有个词,我觉得在这里就可以用,叫"价值共创"(value co-creation)。老年人和智能产品共同创造价值的过程,使得老年人的创意得以实现,让老年人感觉到自我价值的创造。还有很多老年人喜欢唱歌、画画、写诗、作词,虽然自己本来的能力不一定行,但如果老年人说一段话、描述一个场景,ChatGPT 就可以生成出来,这个生成出来的内容,就是老年人和 ChatGPT 共同创造的产物。我觉得这种 ChatGPT 和类似产品,在智慧用老、智慧助老方面,可以帮助老年人在情感上得到满足,还可以帮助其实现价值。ChatGPT 和物理上的机器人在智慧养老领域未来会走得更远,而不仅仅是问答上。

三、智慧养老的挑战和破局

(一)智慧养老的挑战:需求识别和隐私边界

张承蒙:您在文章里曾指出了智慧养老在设备层面、平台层面和技术层面的三重挑战,您觉得最主要的挑战是哪一方面?

左美云:实际上更大的挑战还是我前面说的需求不清晰、体系化不完整。这三个层面的挑战都是比较具体的,我们可以去给设备厂商、平台和技术研发人员

提建议,他们应该注意什么问题。

设备层面主要就是需求,一定要去注意、了解老年人在数智化时代的需求到底是什么,针对他们的使用习惯、使用行为和心理去研制设备产品,我觉得这是最关键的。平台层面,现在我们经常说平台空心化很严重,现在做了很多智慧养老线上平台,但是没有充足的服务资源,现有服务资源之间的融合也不太好,就导致这种平台出现各种问题。技术层面的问题都是可以去解决的,只要发现了技术问题就可以去创新、解决。而设备层面与需求侧紧密联系,平台跟产业体系化发展、供应侧有关系,也涉及供需双方的问题,当然平台归根结底还是要以老年人为中心,要去了解老年人的行为心理和需求。

张承蒙:智慧养老的发展过程中,您觉得是否存在隐私边界和服务冗余的冲突?如何去协调平衡?相关企业对老年人个人信息的收集,是否会增加服务的不确定性?

左美云:我觉得关于隐私的问题已经讨论得很多了,因为一般人的隐私问题老年人都存在。而隐私研究现在又非常热门,也非常多,包括隐私计算现在发展也很快。随着国家政策法律加强,隐私本身可能不会是一个特别大的问题,可以通过一些制度安排和技术手段去解决。有两个另外的问题,我觉得要关注:老年人的隐私问题和技术伦理问题。老年人的隐私问题跟一般人群的隐私还是有所不同的,不同之处在于随着年龄的增长,老年人可能对隐私采取一种让渡的方式来换取服务,年龄增长使得老年人可能倾向于享受服务比保护隐私更重要。比如,原来不接受安装家庭监控的高龄老人,随着年龄的增长,以及其他朋辈的影响,可能会同意安装家庭监控。就智慧养老发展中的技术伦理而言,公正和平等是不同的概念。平等指所有人都可以享受一样的信息服务,但有一些应用程序老年人不一定能使用,因此我们更要追求老年人“公正”地享有信息服务。我们也有文章提到,要意识到智慧助老和适老化不能是“一阵风”,每代人都有每代人的“数字鸿沟”问题,要给跟不上时代的人留“一扇窗”(左美云,2023)。当然,我们最近也提出来,要关注老年人在使用智能产品过程中“网络成瘾”(沈原燕杭、左美云、李逸博,2022)的问题,未来老年网民网络成瘾带来的社会问题需要关注,我认为可通过子女的数字反哺和政府的有序引导来解决,政府可以通过鼓励老年人走出网络、多参与线下活动,在处理老年网民网络成瘾方面有所

作为。

（二）智慧养老的破局：他山之石与本土探索

张承蒙：智慧养老发展方面，您觉得国际上是否有很好的经验？我看您之前提到了芬兰在国家层面推进智慧养老的进展，与我国比较而言，他们的亮点在哪里？我国的智慧养老政策实践上还有哪些改进空间？

左美云：芬兰、挪威等北欧国家和英国，做的智慧养老都不错。他们有一个特别大的好处，就是他们国家人口规模都比较小，容易推广一些项目。像芬兰，他们的线上线下一体化服务做得比较好，已经建立了全国性的医疗数据协作中心，老年人数据可以打通使用。就我国智慧养老发展而言，上海的智慧健康养老发展比较快，可以将个人专属健康档案授权给医疗机构使用，这个很好，尤其是在老年人健康服务上。政策上我们经常讲“有为政府”，就需要政府统筹各个部门，打通智慧养老领域相关的数据孤岛，这是政府可以去做的。我们一直倡导能够有一个国家层面统一的平台，可以将老龄事业和产业相关的数据、资源统筹起来，这将会极大地推动智慧养老的快速发展。

四、前瞻愿景：面向数智化时代的智慧养老

张承蒙：您觉得智慧养老发展的方向可能是什么？就智慧养老的方向而言，是线下服务更重要，还是数据技术发展更重要？

左美云：我认为线上和线下服务是一体两面，缺一不可。数智化技术能够将智慧养老推向规模经济、范围经济和长尾经济，使智慧养老产业有一个很好的发展。国民养老属于微利的行业，长期来看，如果不实现规模经济、范围经济和长尾经济这些模式，企业很难发展。智慧养老体现在“以老年人为中心”，所以线下服务很重要。我认为线下服务里面两类人才很重要：一是对智慧养老模式、内涵熟悉的领导人才；二是年轻的后备人才，包括服务人才、管理人才。很多学校在开设养老服务管理、智慧养老或智慧健康养老专业，吸引年轻人加入。我认为这里智慧养老相关的课程就很重要，要让大学生感觉到，做这个行业跟在那些“IT 大厂”工作的人是一样的，也能用到最新的数智化知识和技能，他们就会感兴趣。这两方面人才的准备很重要。此外，智慧养老发展也需要建立一个全国

性的平台，不同的地域可以根据地域特征选择相应的模块，使得老年人流动、子女随迁带来的问题得到解决。

张承蒙：基于您的研究经验和智慧养老领域研究的想象力，对将来要加入本领域的研究者，您有什么建议？

左美云：老年人群作为一个特定人群，所有的学科都能找到发展机会。传统的老年学，可以跟其他学科交叉，找到新的研究方向。关于“智慧助老”的研究现在比较丰富了，希望大家多多关注“智慧用老”的研究，我们这些年除了“智慧助老”之外，“智慧用老”的研究也在增加。前面我们也讲到，“智慧用老”即充分利用好老人的经验、知识和技能（左美云，2018），因为国家现在非常重视开发老年人力资源，数智化的技术在开发老年人力资源方面怎样发挥作用，值得进一步思考。然后就是“智慧孝老”这一块，我觉得也可以做些研究和实践。其实老年人的精神空虚、抑郁问题比较严重，我们曾经借鉴了陈功教授的孝老模型（陈功，2009），做了一个“智慧孝老”的理论模型，探讨如何将信息技术手段与中国传统观念和文化相结合，对老年人的精神生活进行支持，我们的“智慧孝老”模型包含九个方面内容：供养老人（供老）、照料老人（料老）、陪伴老人（伴老）、顺从老人（顺老）、尊敬老人（敬老）、忍耐老人（耐老）、祭奠老人（祭老）、不啃老、不扰老。因为“孝”是中国独特的传统文化，“智慧孝老”可以促进代际交流，满足老年人的物质和精神双重需求，中国的学者应该加强“智慧孝老”的研究，在国际上形成我们的特色，发出我们的声音。如果我们做出来一些智慧孝老的产品和应用，很容易在国际上有影响力。智慧养老和医养、康养结合也是一个很重要的方向。中国老年人健康预期寿命与预期寿命之间相差 8~9 年，经常讲“长寿不健康”。怎样让老年人又长寿又健康，智慧技术可以发挥很大的作用。另外，用仿真平台去模拟一些与养老相关的养老政策、保险政策效果，可以产生很高的实践价值。最后要强调的是，做智慧养老领域的研究，一定要重视调研、体验，在了解老年人群需求的基础上进行研究，借助老龄实验室等让普通人群也能有老龄体验，这样才能明白老年人的需求所在，进而开发相应的智慧养老产品和服务。

参考文献

陈功,2009,《社会变迁中的养老和孝观念研究》,北京:中国社会出版社。

陈功、赵新阳、索浩宇,2021,《“十四五”时期养老服务高质量发展的机遇和挑战》,《行政管理改革》第 3 期。

杜鹏,2022,《积极应对人口老龄化,推进中国式现代化》,《人口与发展》第 6 期。

沈原燕杭、左美云、李逸博,2022,《短视频平台上的“银发网红”——研究机会与相关政策》,《北京理工大学学报》(社会科学版)第 6 期。

习近平,2020,《在经济社会领域专家座谈会上的讲话》,《人民日报》8 月 25 日,第 2 版。

徐淑英,2015,《科学精神和对社会负责的学术》,《管理世界》第 1 期。

左美云,2018,《智慧养老:内涵与模式》,北京:清华大学出版社。

左美云,2023,《老年友好型社会导向的智慧技术适老化》,《人民论坛·学术前沿》第 1 期。

左美云、刘浏、尚进,2021,《从国家政策看智慧健康养老发展脉络》,《中国信息界》第 1 期。

左美云、沈原燕杭、刘妃,2023,《智慧健康养老领域接入元宇宙的方式与问题研究》,《科技智囊》第 1 期。