

# 重访计算社会科学：从范式创新到交叉学科

严宇 方鹿敏 孟天广

**摘要：**第四次科技革命驱动人类社会迈入数字时代，整个人类文明正在经历快速、大尺度和深层次的数字化转型。计算社会科学致力于形成交叉学科，从本体论、方法论和认识论层次为理解数字文明语境下新生社会经济现象构建概念、理论和方法体系。计算社会科学构想自提出以来经历了三大转变：一是研究范式从聚焦数据驱动发展为理论驱动与数据驱动相融合，重新重视因果关系；二是技术方法从聚焦各类数据挖掘方法到推进数据挖掘、仿真模拟、复杂系统方法并存的多样化研究方法；三是从聚焦论转变为从本体论和认识论构建新概念、新理论以解释数字时代的社会演变。我国计算社会科学迅速发展，已经形成了数字经济、计算政治学、计算社会学、计算传播学和计算法学等新兴交叉学科，未来应加快计算社会科学交叉学科建设以形成独特学术、话语和学科体系，为人类社会进入数字文明时代贡献原创理论和范式。

**关键词：**计算社会科学；研究范式；交叉学科；学科体系

## 一、引言：构建数字时代的社会科学

第四次科技革命驱动着人类社会快速进入数字时代，整个人类文明正在经历快速、大尺度和深层次的数字化转型，这为社会科学的知识生产和学科发展带来了本体论、方法论和认识论层次的全新问题，成为数字时代社会科学理论范式与方法范式迭代的时代机遇<sup>①</sup>。譬如，如何理解人类社会数字化转型的模式和规律，如何认识数字时代的新生经济社会现象，如何驱动经济社会数字化发展和国家治理数字化转型，这都成为数字时代社会科学发展的基本命题，为学术共同体推进计算社会科学交叉学科的发展提供了学理依据和实践需求。

作为前沿交叉学科，“计算社会科学”(Computational Social Science)的发展历程概括了社会

**作者简介：**严宇，清华大学社会科学学院博士后(北京 100084)；方鹿敏，清华大学社会科学学院博士后(北京 100084)；孟天广(通讯作者)，清华大学计算社会科学与国家治理实验室副主任、清华大学社会科学学院院长聘副教授、博士生导师(北京 100084)。

**基金项目：**本文系国家社会科学基金重大项目“基于大数据的智能化社会治理监测、评估与应对策略研究”(18ZDA110)的阶段性成果。

<sup>①</sup> 江小涓、黄颖轩：《数字时代的市场秩序、市场监管与平台治理》，《经济研究》2021年第12期；王天夫：《数字时代的社会变迁与社会研究》，《中国社会科学》2021年第12期。

科学在数字时代所呈现出的新发展、新路径和新范式<sup>①</sup>。计算社会科学从萌芽到兴起的历程体现了人类社会迈入数字时代社会科学知识生产的必然趋势,其在过去十多年间的快速发展为社会科学研究范式创新和理论迭代提供了重大贡献。计算社会科学的兴起既是数字时代社会科学创新发展的必然趋势,更是社会科学与数据科学、计算科学交叉融合的产物。第四次科技革命的爆发起源于学科交叉融合,反过来进一步驱动着经由学科交叉融合,以构建支撑数字文明的知识生产和人才积累的基础性作用。国内外学界深刻认识到将数据科学和计算科学引入到社会科学研究的重要意义,计算社会科学也因此成为国际范围内社会科学形成新范式、新理论和新知识的前沿领域,利用大数据、机器学习、仿真模拟等新兴方法对复杂的人类行为及社会运行进行跨学科研究,显著提高了人们理解、分析和预测社会规律的能力。

自 2000 年以来,国内外顶尖高校和科研机构都陆续建立了计算社会科学的研究机构和教学单位,并投入大量人力、物力和财力资源,推进计算社会科学的科学研究和人才培养,形成了丰硕的成果。为了推进学科交叉,《科学》杂志专门设置了“计算社会科学”专栏发表高质量研究成果。根据《自然》杂志的统计,自 2014 年以来,共有 4144 篇计算社会科学相关论文正式发表。根据图 1,2014 年以来计算社会科学相关论文呈逐年上升趋势,而且充分呈现了这一领域的文理交叉属性,譬如社会科学、计算机、物理、数学和工程学科都贡献该领域的知识积累,而社会科学领域贡献的文献比例将近三分之一。

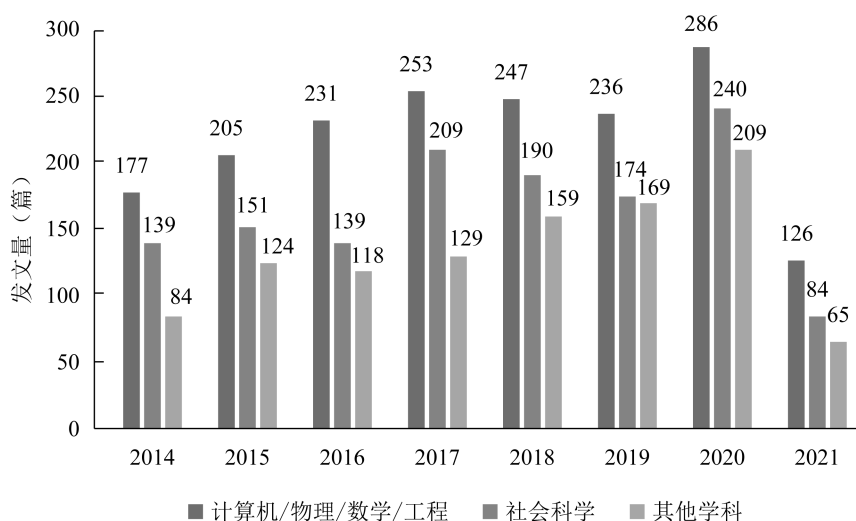


图 1 2014 年以来计算社会科学相关论文分学科数量<sup>②</sup>

本文聚焦 2009 至 2022 年全球范围内计算社会科学的发展历程,结合数字时代社会科学研究范式与理论体系的发展趋势,系统阐述计算社会科学经由研究范式创新推进数字时代社会科学理论体系迭代,进而依托文理交叉推进“新文科”建设的学科建设价值。数字时代呼唤全新的社会科学,迫切要求社会科学对于理解数字文明语境下新生社会经济现象形成本体论、方法论和认识论层面的概念和理论体系,因而计算社会科学是全球范围内学术、话语和学科体系竞争

<sup>①</sup> 张小劲、孟天广:《论计算社会科学的缘起、发展与创新范式》,《理论探索》2017 年第 6 期;David Lazer, et al., “Computational Social Science,” *Science*, vol.323, no.5915(2009), pp. 721-723.

<sup>②</sup> 数据来源为《自然》杂志网站, <https://dash.springernature.com/nature-insights-css/>, 访问时间为 2022 年 4 月 30 日。图 1 为作者自制。

的重要领域,这为数字化转型先发国家的中国提供了前所未有的学科引领机遇,因此应该加快计算社会科学交叉学科建设,为国际社会进入数字文明贡献原创理论和范式。

## 二、计算社会科学的新进展:延续与突破

作为数据科学、计算科学和社会科学融合形成的交叉领域,计算社会科学近年来吸引了大量学术关注。然而,目前对于如何定义计算社会科学,学界尚未形成共识;甚至有学者提出对于正在蓬勃发展、尚未定型的计算社会科学这一新兴学科,急于对其下定义未必是件好事<sup>①</sup>。

在充分借鉴国内外相关文献的基础上<sup>②</sup>,本文尝试对计算社会科学的核心研究内容进行界定。本文认为现阶段,计算社会科学主要借助当前日益强大的计算能力和不断迭代优化的算法,对与社会系统相关的大数据进行分析和挖掘,从而获取有关人类行为和社会运行规律的知识。这一界定指明了计算社会科学计算什么、如何计算和算出什么,换言之,就是该领域研究的对象、过程和目标。

当前,我们通常讨论的计算社会科学,其计算的是与人类行为和社会运行相关的海量数据,不仅包括近年来常常出现在学术研究中的社交媒体数据,如微博、脸书、推特上的发文,还包括公众在日常生活中留下的各类文字性和图像性的数字痕迹,包括搜索记录、购买记录、短视频、照片等等。上述数据构成了计算社会科学的研究基础,为研究社会系统提供了信息基础。

为实现对海量数据的计算,计算社会科学必须依赖于强大的算力和优化的算法。算力,即计算能力、运算力,指的是计算机的硬件和软件如何高效地执行算法,实现数据输入到结果输出的转化。随着CPU和GPU芯片的升级优化,海量数据的处理速度与效果越来越好,这为计算社会科学的蓬勃发展提供了硬件支持。算法广义上讲是一种规则,用于将输入转化为输出<sup>③</sup>。算力和算法的结合才能将海量数据中的信息提取出来,进而生成有关社会系统的有效知识。

计算社会科学研究的目的在于获得有关社会系统运行规律的知识。当然,这一目标并非计算社会科学所独有,传统社会科学<sup>④</sup>同样致力于发现复杂的社会规律。但不同之处有两点:一是受制于数据和研究方法的限制,传统社会科学无法通过有限工具全面理解复杂的社会系统,但得益于海量数据的积累和算力算法的持续发展,计算社会科学为全面理解复杂社会系统创造了可能<sup>⑤</sup>;二是传统社会科学尤其强调对社会运行过程中因果关联及其机制的追求<sup>⑥</sup>,工具变量、倾向值匹配、断点回归、实验法等方法无不体现传统社会科学对因果规律的追求,但对于计算社会科学而言,因果性并非唯一和必须的,相关性同样能够为理解社会运行提供有益知识,甚至能启发新理论。

① 譬如普林斯顿大学的马修·萨尔加尼克(Matthew Salganik)。

② Claudio Cioffi-Revilla, *Introduction to Computational Social Science: Principles and Applications* (Second Edition), New York: Springer, 2017, p. 2; David Lazer, et al., "Computational Social Science," *Science*, vol. 323, no. 5915 (2009), pp. 721-723; David Lazer, et al., "Computational Social Science: Obstacles and Opportunities," *Science*, vol. 369, no. 6507 (2020), pp. 1060-62.

③ Nicholas Diakopoulos, "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures," *Digital Journalism*, vol. 3, no. 3 (2015), pp. 398-415.

④ "传统社会科学"这一名称的提出主要是针对计算社会科学而言,从时间上将二者加以划分。目前主流的观点是将以定量分析和定性分析为主的社会科学研究视为传统社会科学。

⑤ David Lazer, "Social Science, Today," *Science*, vol. 359, no. 6371 (2018), p. 42.

⑥ 孟天广:《政治科学视角下的大数据方法与因果推论》,《政治学研究》2018年第3期。

综上所述,计算社会科学作为多学科交叉融合的产物,以“计算”为核心,为大数据时代下社会科学的发展能提供了新范式和新路径,为加深我们对社会系统运行规律的全面认知提供了新可能。在此意义上,计算社会科学常常被视为新兴学科。然而,计算社会科学并非与传统社会科学毫无关联或截然不同,前者更应被视为是后者的延续和突破。

### (一)计算社会科学的延续

2009年大卫·拉泽尔(David Lazer)及其合作者发表的《计算社会科学》<sup>①</sup>常常被视为是计算社会科学的“诞生”;但本文认为更加准确的说法应该是拉泽尔等人重新定义或开启了计算社会科学蓬勃发展的新阶段。2009年之前就有学者提出“计算社会科学”一词,同样是基于数据进行推论和开展社会科学研究<sup>②</sup>。而拉泽尔等人提出的计算社会科学则是在大数据背景下,海量数据为理解和还原社会运行规律带来了新可能。这逐渐成为我们通常所讨论的计算社会科学,或称之为“大数据+”的计算社会科学。

本文认为计算社会科学是传统社会科学的延续,主要体现在以下两点:

一是从研究内容上看,计算社会科学仍然是以现实世界(包括微观层面的个体、中观层面的组织和宏观层面的社会)为研究对象,获取描述性、解释性和预测性知识为主的经验研究<sup>③</sup>。而这可以追溯到社会科学的起源,或者是社会科学的“前计算时期”,譬如古希腊学者亚里士多德对政体的比较分析。

二是从研究方法上看,计算社会科学与传统社会科学(尤其是其中的定量分析)存在紧密联系。近代社会学家涂尔干已经广泛使用统计学方法对大量数据展开分析,探究自杀行为背后的社会规律。17世纪中叶,现代统计学的诞生,及其随后在社会科学各学科的广泛应用推动了定量社会科学的诞生和快速发展<sup>④</sup>。随着定量社会科学逐渐被学界和实务界所接纳和运用,一方面数据分析方法和技术在不断优化和创新,另一方面用“数据说话和说理”的理念逐步建立,这些都为计算社会科学进入快速发展的新阶段奠定了坚实的理念、方法和技术基础。

### (二)计算社会科学的突破

近年来,大数据的快速发开展开了计算社会科学蓬勃发展的新阶段,实现了对传统社会科学的突破,这主要体现在以下三个方面。

一是数据不同。体现在数据的规模、结构、来源等方面。传统社会科学主要使用的是小数据(抽样数据),其具有结构化程度高、数据更新速度慢的特点;数据主要依靠研究者根据研究问题和理论设计研究方案加以采集,因此也被称之为是“设计的数据”。然而,目前计算社会科学主要依赖的是大数据(或者说是“近似的总体数据”),其具有非结构化、形式复杂、数据更新速度快等特点;数据的生成虽然仍依赖于数字化设备设施的记录和存储,但这一过程并不依赖于或受制于研究者的主观意愿,而是对社会运行的全过程、全时段记录,因此也被称之为“发现的数据”<sup>⑤</sup>。对于小数据和大数据的差异,本文认为关键点不在于数据的体量、结构或来源,而在于是

① David Lazer, et al., “Computational Social Science,” *Science*, vol.323, no.5915(2009), pp. 721-723.

② Thomas Schelling, *Micromotives and Macrobehavior*. New York: WW Norton & Company, 2006, pp. 136-166;黄璜:《社会科学研究中“基于主体建模”方法评述》,《国外社会科学》2010年第5期。

③ Uwe Engel, et al., *Handbook of Computational Social Science, Volume 1: Theory, Case Studies and Ethics*, Oxfordshire: Taylor & Francis, 2021, p. 19.

④ 张小劲、孟天广:《论计算社会科学的缘起、发展与创新范式》,《理论探索》2017年第6期。

⑤ Matthew Salganik, *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press, 2018, p. 83.

否能够服务于研究目标——理解社会系统的运行规律。小数据体现的是传统社会科学研究中理论的重要性和研究者的主导地位,而大数据强调的是社会自发生成数据的价值。由此可见,小数据和大数据都可以实现对复杂社会系统的理解,但实现路径不同,而这也引出了第二个差异。

二是研究视角不同。传统社会科学是理论驱动研究,而计算社会科学是数据驱动研究。长期以来,传统社会科学采用的是“假设—验证—结论”的基本程序<sup>①</sup>,强调理论驱动范式,重在统计推论和逻辑推理,譬如政治学的行为主义范式<sup>②</sup>。然而,计算社会科学强调的是利用数据挖掘技术获取海量数据背后存在的模式和规律<sup>③</sup>。因此,如果将抽象理论和具体实践视为二元对立的客体,那么传统社会科学采用的是从理论到实践的自上而下路径,而计算社会科学采用的是从实践到理论的自下而上路径。当然,即便是强调数据驱动的计算社会科学,也并非无视理论,而应该将其视为检验理论、证伪理论、发展新理论的大数据路径。

三是研究重点不同。传统社会科学重视因果性,而计算社会科学更重视相关性。在传统社会科学研究中,因果性(Causality)被视为社会科学研究的圣杯<sup>④</sup>,譬如加里·金(Gary King)等学者在其经典著作《社会科学中的研究设计》就提出“科学研究”以描述性和因果性推论为研究目的。而对于计算社会科学而言,其主要目标在于从大数据中挖掘出反映社会运行的模式、关系或趋势,将非结构化、复杂的数据转化为结构化、可理解的社会知识,因此早期计算社会科学研究强调相关性而非因果性,简言之经由大数据分析发现相关性即构成知识。然而,近年来计算社会科学的研究范式对上述观点形成反思,再次强调因果性与大数据、机器学习方法的融合,或者理解机器学习的因果性对于知识成产至关重要<sup>⑤</sup>。可以说,一方面相关性作为因果性的基础使得计算社会科学研究能够利用数据驱动范式去检验因果性知识,另一方面因果性可以帮助研究者理解复杂社会系统中特定关系的具体因果机制,进而帮助研究者优化算法模型提升知识发现能力。

### 三、计算社会科学的研究范式创新

作为新兴交叉学科,当前计算社会科学主要包括四大研究方法:自动信息提取(Automated Information Extraction)、社会网络分析(Social Network Analysis, SNA)、社会复杂系统分析(Social Complexity)和社会仿真建模(Social Simulation Modeling)<sup>⑥</sup>。每个方法都含有一系列概念、理论和技术,但这并不意味着它们各自为战、毫无交叉和对话。事实上,当涉及到具体研究问题时,上述领域的交叉并不少见。譬如,针对恐怖组织的组织网络分析,既会涉及到社会网络分析,又会涉及到自动信息提取。

首先,自动信息提取指的是运用算法从数据中提取有关社会运行的信息。虽然这里的数据

① 苏毓淞、刘江锐:《计算社会科学与研究范式之争:理论的终结?》,《复旦学报(社会科学版)》2021年第2期。

② 张小劲、孟天广:《论计算社会科学的缘起、发展与创新范式》,《理论探索》2017年第6期。

③ David Lazer, et al., “Computational Social Science,” *Science*, vol.323, no.5915(2009), pp. 721-723.

④ 左才:《政治学研究中的因果关系:四种不同的理解视角》,《国外理论动态》2017年第1期。

⑤ 孟天广:《政治科学视角下的大数据方法与因果推论》,《政治学研究》2018年第3期。

⑥ Claudio Cioffi-Revilla, *Introduction to Computational Social Science: Principles and Applications* (Second Edition), New York: Springer, 2017, pp. 12-18.

可以是任何类型、规模的数据,但是在计算社会科学研究中常常分析的是非结构化、形式复杂的大数据,包括文字、图片、视频、音频等。就目前已发表的学术研究而言,针对文本数据的自动信息提取相对较多,从中发展出主题分析、语义分析、情感分析等路径。譬如,针对推特、脸书等社交媒体数据的大数据分析,为我们进一步理解国际社会(尤其是美国)虚假信息和政治极化现象日益严重的现象提供了新的启发<sup>①</sup>。而在我国,涉及到自动信息提取的研究也已经在政府回应性、社会经济地位等主题陆续出现<sup>②</sup>。

其次,社会网络分析旨在分析社会网络内各主体互动关系以及由此产生的关系结构和属性。如同计算社会科学的历史,社会网络分析也并非是全新领域,其在社会科学主要学科(包括政治学、经济学、社会学、心理学和人类学)已有多年的发展历史,可以追溯到1950年代<sup>③</sup>。然而,得益于大数据的持续积累和计算能力的显著提升,社会网络分析近十年来发展迅猛,成为计算社会科学交叉学科中最引人注目的领域。一方面,大数据的出现将社会网络内的主体、距离、互动关系等要素数字化,从而为计算社会网络提供了坚实的数据支撑;另一方面,算力的提升为计算海量数据提供了技术支撑,从而充分挖掘出数据所承载的信息,形成对社会网络的新知识。目前,该领域研究一方面对数字时代社会网络自身的属性和特点<sup>④</sup>展开了分析,另一方面也对社会网络的影响后果进行了有益探索,包括对个体行为<sup>⑤</sup>、政治动员<sup>⑥</sup>和经济发展<sup>⑦</sup>的影响等方面。

再次,社会复杂系统分析旨在对社会系统内个体、组织等主体之间的相互作用及其形成的复杂系统,涉及政治、经济、社会、生态等领域或子系统。类似物理学对宇宙起源的研究,社会复杂系统分析寻求的是对人类社会及其文明的理解,换言之,即在不同权威类型下人类社会是如何建立和维持的<sup>⑧</sup>。在此意义上,社会复杂系统分析与政治学中关于政体类型及其变迁的研究类似,不同之处在于此前对社会系统的研究主要依赖对历史文献和遗迹的定性分析与对小数据的定量分析,而计算社会科学的出现推动其他复杂系统模型的出现,譬如神经网络建模、基于主体建模等。

---

① Nir Grinberg, et al., "Fake news on Twitter during the 2016 US Presidential Election," *Science*, vol. 363, no. 6425 (2019), pp. 374-378; Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik, "Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media," *Political Communication*, vol. 38, no. 1-2 (2021), pp. 98-139.

② Junyan Jiang, Tianguang Meng, and Qing Zhang, "From Internet to Social Safety Net: The Policy Consequences of Online Participation in China," *Governance*, vol. 32, no. 3, 2019, pp. 531-46; Tianguang Meng and Zesen Yang, "Variety of Responsive Institutions and Quality of Responsiveness in Cyber China," *China Review*, vol. 20, no. 3, 2020, pp. 13-42; 么晓明等:《大数据驱动的社会经济地位分析研究综述》,《计算机科学》2022年第4期。

③ Uwe Engel, et al., *Handbook of Computational Social Science, Volume 1: Theory, Case Studies and Ethics*, Oxfordshire: Taylor & Francis, 2021, p. 23.

④ Patrick Park, Joshua Blumenstock, and Michael Macy, "The Strength of Long-range Ties in Population-scale Social Networks", *Science*, vol. 362, no. 6421 (2018), pp. 1410-1413. Aili Asikainen, et al., "Cumulative Effects of Triadic Closure and Homophily in Social Networks," *Science Advances*, vol. 6, no. 19 (2020): eaax7310.

⑤ Sinan Aral and Dylan Walker, "Identifying Influential and Susceptible Members of Social Networks," *Science*, vol. 337, no. 6092 (2012), pp. 337-341.

⑥ Robert Bond, et al., "A 61-million-person Experiment in Social Influence and Political Mobilization," *Nature*, vol. 489, no. 7415 (2012), pp. 295-298.

⑦ Nathan Eagle, Michael Macy, and Rob Claxton, "Network Diversity and Economic Development," *Science*, vol. 328, no. 5981 (2010), pp. 1029-1031.

⑧ Claudio Cioffi-Revilla, *Introduction to Computational Social Science: Principles and Applications* (Second Edition), p. 205; Robert Carneiro, "A Theory of the Origin of the State: Traditional Theories of State Origins are Considered and Rejected in favor of a New Ecological Hypothesis," *Science*, vol. 169, no. 3947 (1970), pp. 733 - 738.

最后,社会仿真建模的核心内容在于通过构建虚拟世界来探究社会复杂系统。由此可见,社会仿真建模和复杂社会系统分析存在紧密联系,后者为前者提供了新的研究路径。不同于以往研究使用历史文献、民族志、小数据的方式来研究社会复杂系统,社会仿真建模将现实世界的各类要素公式化,建立与现实相对的虚拟世界,并通过改变参数来模拟各类场景,获取更多社会系统运行相关的信息和知识。譬如,社会仿真建模常被用于社会冲突研究<sup>①</sup>(包括集体行动、叛乱、内战、国家间冲突、恐怖主义行动等),通过改变参与主体的行为、所处环境等参数来推演相应的社会后果,加深对社会系统在不同情境下运行状态的理解。

#### 四、计算社会科学驱动交叉学科图谱

以“大数据+”形式出现的计算社会科学,有效推动了社会科学多个研究领域与大数据相关概念和技术的结合,催生了计算政治学、计算社会学、计算传播学、计算法学等学科<sup>②</sup>。但不同学科领域在“可计算性”和“计算化”程度上存在一定差异,前者强调了一个领域所涉及的实际研究问题在多大程度上可以通过信息处理范式(information-processing paradigm)下的社会计算来解决<sup>③</sup>,后者则表明该领域在多大程度上获取并利用了所需数据以适应自身研究发展的需要。对于社会科学的细分领域而言,“可计算”属性极大影响着每一个领域向大数据方法转型的内生驱动力,进而影响到该领域“计算化”的发展水平。

在政治学领域,社会计算引发了一场深刻的研究革命。当前,大数据时代政治学研究的“可计算性”主要体现在两个维度上:一方面,政策文本、网络表达、新闻图片、视频等为观察和分析各类政治现象提供了海量有效的自然数据;另一方面,大数据也塑造了新的政治现象,将政治参与、政府治理等拓展至虚拟空间,进一步重构了政治主体的行为模式和国家—社会关系<sup>④</sup>。相较于定量研究所使用的小数据,大数据更多体现了观察对象自觉自然的实践行为,因此许多政治学者认为,大数据可以为选举、社会抗争、政治决策、人口迁徙等研究提供更为真实和直接的观察资料,使研究结果更加贴近政治现实<sup>⑤</sup>。对于虚拟空间新的政治现象所带来的一系列国家—社会关系的变化,大数据提供了更多的素材来追踪其中的影响过程并获得有意义的见解<sup>⑥</sup>。另外,不可否认的是,数字政府、网络政治参与等新的政治现象本身即是一种海量数据的载体,除催生新的研究问题和方向之外,还在一定程度上成为连接大数据与既有政治学理论的桥梁,将虚拟空间的政治现象进一步系统化为完整的研究领域。

传统的“计算社会学”(computational sociology)特指使用模拟仿真来预测社会行为和发展

<sup>①</sup> Frédéric Amblard, et al., "Analyzing Social Conflict Via Computational Social Simulation: A Review of Approaches," *Complex Societal Dynamics: Security Challenges and Opportunities*, vol. 75(2010), pp. 126-142.

<sup>②</sup> 张小劲、孟天广:《论计算社会科学的缘起、发展与创新范式》,《理论探索》2017年第6期。

<sup>③</sup> Claudio Cioffi-Revilla, *Introduction to Computational Social Science: Principles and Applications*, (Second Edition), New York: Springer, 2017, p. 2.

<sup>④</sup> 孟天广、郭凤林:《大数据政治学:新信息时代的政治现象及其探析路径》,《国外理论动态》2015年第1期。

<sup>⑤</sup> Stephen Ansolabehere and Eitan Hersh, "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate," *Political Analysis*, vol.20, no.4 (2012), pp. 437-459; Pierre F. Landry and Mingming Shen, "Reaching Migrants in Survey Research: The Use of the Global Positioning System to Reduce Coverage Bias in China," *Political Analysis*, vol.13, no.1(2005), pp. 1-22.

<sup>⑥</sup> Johann Höchtl, Peter Parycek, and Ralph Schöllhammer, "Big Data in the Policy Cycle: Policy Decision Making in the Digital Era," *Journal of Organizational Computing and Electronic Commerce*, vol.26, no.1-2 (2016), pp. 147-169.

轨迹,其基础是高度经验化、现实化的模型<sup>①</sup>。社会计算为社会学研究者提供了可以更为深入观察分析复杂人类行为和社会进程的海量数据和技术方法,形成了被称为“新计算社会学”的大数据社会学。由于社会学研究涵盖面广、研究层次多样,文本内容、人际网络、地理定位、经济行为等均可成为社会学研究的自然数据<sup>②</sup>。当代计算机科学和互联网技术的发展,极大增加了此类数据的可获得性,这些过去难以挖掘的数据不仅为描述性和解释性社会研究提供了更全面的实证基础,也为探索未知社会现象提供了无限可能<sup>③</sup>。因此,社会计算已被越来越多地用于社会学各个分支方向的研究。研究者注意到,推特等社交媒体已经成为当代集体行动的重要组织工具,并且提供了海量的实时数据集,可以帮助研究者了解相关信息如何在社交网络中传播并最终推动集体行动的生成<sup>④</sup>。社会计算的进步也使得对文本数据的搜集更为高效,知识社会学和文化社会学研究者通过文本分析的计算技术考察学科和社会文化变迁的过程,并试图从新的视角找到两者之间的关联性<sup>⑤</sup>。相较于计算传播学,新计算社会学在将研究结果发展为新的系统性理论或重构经典社会概念和解释机制方面更加积极<sup>⑥</sup>,大数据、经验和理论的结合也为研究者进一步观察社会结构中的各类系统性变化提供了更有力的逻辑支撑。

由于大量的原始社会数据集以文字、音频和视频等形式存储于媒体中,在过去十年里,计算社会科学研究范式为传播学领域带来了新的研究视角,在国内外形成了具有共识性的“计算传播学”(computational communication science)研究<sup>⑦</sup>。广义的传播学是一门研究人类进行社会信息传递的学科,随着互联网时代的兴起和社交媒体平台的迅速发展,海量的数据集包含在诸如微博、推特、短视频和电子邮件等通讯工具中,这类数据作为广泛的加工处理信息承载了人类信息交换更自然的实践模式以及所形成的新的传播效果,为计算社会科学在传播学领域的应用提供了前提。传统的定性和定量研究方法已无法满足传播学领域新的研究需求,社会计算对理解信息传播复杂性的工具性意义不仅实现了对海量数据的抓取和有效利用,也带来了对数据进行分析和应用的技术革命<sup>⑧</sup>。例如,主题模型(topic modeling)和文本聚类(document clustering)已见于对社交媒体用户海量复杂的交流内容进行分类、理解和阐释的研究中<sup>⑨</sup>。尽管计算

① Michael Macy and Robert Willer, "From Factors to Actors: Computational Sociology and Agent-Based Modeling," *Annual Review of Sociology*, vol.28, no.1(2002), pp. 143-166.

② 罗玮、罗教讲:《新计算社会学:大数据时代的社会学研究》,《社会学研究》2015年第3期。

③ 张文宏:《大数据时代社会学研究的机遇和挑战》,《社会科学辑刊》2018年第4期。

④ Sandra González-Bailón and Ning Wang, "Networked Discontent: The Anatomy of Protest Campaigns in Social Media," *Social Networks*, vol.44 (2016), pp. 95-104; Zeynep Tufekci and Christopher Wilson, "Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square," *Journal of Communication*, vol.62, no.2(2012), pp. 363-379.

⑤ Christopher Andrew Bail, "Combining Natural Language Processing and Network Analysis to Examine How Advocacy Organizations Stimulate Conversation on Social Media," *Proceedings of the National Academy of Sciences*, vol.113, no.42(2016), pp. 11823-11828; Eliza Evans, Charles Gomez and Daniel McFarland, "Measuring Paradigmaticness of Disciplines Using Text," *Sociological Science*, vol.3(2016), pp. 757-778.

⑥ Achim Edelman, et al., "Computational Social Science and Sociology," *Annual Review of Sociology*, vol.46 (2020), pp. 61-81.

⑦ 张伦:《计算传播学范式对传播效果研究的机遇与挑战》,《新闻与写作》2020年第5期; Martin Hilbert, et al., "Computational Communication Science: A Methodological Catalyzer for a Maturing Discipline," *International Journal of Communication*, vol.13(2019), pp. 3912-3934.

⑧ Wouter van Atteveldt and Tai-Quan Peng, "When Communication Meets Computation: Opportunities, Challenges, and Pitfalls in Computational Communication Science," *Communication Methods and Measures*, vol.12, no. 2-3(2018), pp. 81-92.

⑨ Stephan A. Curiskis, et al., "An Evaluation of Document Clustering and Topic Modelling in Two Online Social Networks: Twitter and Reddit," *Information Processing & Management*, vol.57, no. 2(2020), pp. 1-21.



社会科学为传播学带来了新的发展趋势和研究空间,有学者指出,计算传播学面临着一个更隐蔽的挑战,即如何将大数据和研究结果与建构新的理论机制衔接起来,形成理论驱动型研究。由于人类社会信息传播路径日趋复杂化,若仅将社会计算视为一种研究工具,许多通过对大数据计算所得到的人类特定的传播行为则无法被系统化,难以推进传播学议题的进一步创新<sup>①</sup>。

与上述三个社会科学领域不同的是,法学学界长期争辩学科的本体性质,即法学究竟属于规范分析的教义学还是经验分析的社会科学<sup>②</sup>。“计算法学”(computational jurisprudence)将大数据引入法学研究,旨在建立一种新的学科框架,即作为注释中心的法律文本以大数据形式存在(legal text as data),受到教义学的推崇,进而在工具层面构成法学实证研究的基础<sup>③</sup>。尽管计算法学的形成得益于近20年来人工智能和大数据处理技术的快速发展,但法律文本在被作为大数据使用之前,已经被视为是一种“代码”(law as code)。这些代码基于法律文本中的形式逻辑生成,直接被计算机系统处理,并用于法律推理实践。相比之下,将法律作为大数据(law as data)打破了固定机械的处理模式,研究者可以超越法律条文和法律文本中的固定逻辑,找到形式逻辑之外更深层次的规律,为法学研究提供新的洞见<sup>④</sup>。此外,视频、语音记录、电子邮件等非结构化数据,经算法处理后形成结构化数据,而司法实践也越来越愿意采纳这些结构化数据<sup>⑤</sup>。当然,社会计算在司法实践中的应用亦可能与传统的原则和标准产生冲突,甚至重构法理学研究中对价值原则的阐释。

## 五、结语:作为交叉学科的计算社会科学

自2009年计算社会科学作为交叉学科构想被引入社会科学研究以来,一场“范式革命”已悄然而至。回顾2009年提出计算社会科学构想以来的发展历程,计算社会科学正在经历下列三大转变以驱动学科建设<sup>⑥</sup>:一是计算社会科学的研究范式从聚焦数据驱动发展为理论驱动与数据驱动相融合,从相关关系即知识发展为大数据因果推论、AI因果分析;二是计算社会科学的技术方法从聚焦各类数据挖掘方法到推进数据挖掘、仿真模拟、复杂系统方法并存的多样化研究方法;三是计算社会科学直面数字时代的社会科学本体论和认识论转变,围绕数字时代的新生研究对象如数字经济、数字社会和数字政府等,形成新概念、新理论以解释数字时代的社会构造机理、社会经济运行机制、社会经济演化规律和国家治理模式的系统性知识。当然,计算社会科学虽然是集合数据科学、计算科学和社会科学等学科范式形成的新兴交叉学科,但其并不是与以往社会科学理论与研究范式的断裂,而是社会科学学科体系在数字时代的延续和突破。

数字时代呼唤全新的社会科学,迫切要求社会科学对于理解数字文明语境下新生社会经济

① 徐明华、冯亚凡:《社会计算视域下传播学研究的嬗变与反思》,《现代传播(中国传媒大学学报)》2017年第12期;Yannis Theocharis and Andreas Jungherr, “Computational Social Science and the Study of Political Communication,” *Political Communication*, vol.38, no.1-2 (2021), pp. 1-22.

② 陈兴良:《法学知识的演进与分化——以社科法学和法教义法学为视角》,《中国法律评论》2021年第4期。

③ 王伟:《从计算社会科学到计算法学:重新检视因果关系科学之范畴》,《网络信息法学研究》2020年第1期。

④ 张妮、徐静村:《计算法学:法律与人工智能的交叉研究》,《现代法学》2019年第6期;Jens Frankenreiter and Michael A. Livermore, “Computational Methods in Legal Analysis,” *Annual Review of Law and Social Science*, vol.16(2020), pp. 39-57.

⑤ 张妮、徐静村:《计算法学:法律与人工智能的交叉研究》,《现代法学》2019年第6期。

⑥ David Lazer, et al., “Computational Social Science: Obstacles and Opportunities,” *Science*, vol. 369, no. 6507 (2020), pp. 1060-1062.

现象形成本体论、方法论和认识论层面的概念、理论和方法体系。一方面,计算社会科学是社会科学未来发展的必然趋势。计算社会科学成为国际范围内社会科学形成新范式、新理论和新知识的前沿领域,利用大数据、机器学习等新兴方法对复杂的人类行为及社会运行进行跨学科研究,提高了人们理解、分析和预测社会规律的能力,因此构成国家“新文科”建设的基础领域。另一方面,计算社会科学致力于培养服务国家重大战略的急需人才。十九大提出“数字中国”“智慧社会”和“网络强国”等科技驱动战略,“十四五规划”明确我国要“加快数字化发展”,统筹推进数字经济、数字政府和数字社会发展,尤其是提出具备数字素养的人才是数字化发展的关键,这为计算社会科学交叉学科人才培养提出国家需求。

我国计算社会科学在过去十余年里快速发展,在持续推进一系列概念、理论和研究方法的同时,推动了计算思维在社会科学各领域的应用和发展,形成了数字经济、计算政治学、计算社会学、计算传播学和计算法学等新兴交叉学科<sup>①</sup>,有力地支撑了社会经济的数字化发展和国家治理的数字化转型。有鉴于此,我国应加快计算社会科学交叉学科建设,通过构建全国共享的数字科研基础设施、优化学科目录体系、建设文理交叉重点实验室、加快培养复合型数字人才和协同育人机制,基于我国数字化转型的先发实践和文理交叉的学科基础,形成计算社会科学的学术、话语和学科体系,为国际社会进入数字文明时代贡献原创理论和范式。

[责任编辑 贾乐耀]

---

<sup>①</sup> 刘涛雄、徐晓飞:《互联网搜索行为能帮助我们预测宏观经济吗?》,《经济研究》2015年第12期;罗玮、罗教讲:《新计算社会学:大数据时代的社会学研究》,《社会学研究》2015年第3期;孟天广:《政治科学视角下的大数据方法与因果推论》,《政治学研究》2018年第3期;申卫星、刘云:《法学研究新范式:计算法学的内涵、范畴与方法》,《法学研究》2020年第5期;王成军:《计算传播学的起源、概念和应用》,《编辑学刊》2016年第3期。