

大数据技术与家庭治理研究创新方法*

冯帅帅^{1,2}

(1.湖南师范大学公共管理学院;

2.长沙新一代人工智能伦理治理与公共政策实验室, 湖南 长沙 410081)

[摘要]自党的十九届四中全会明确了家庭在基层治理和国家治理现代化中的基础性地位以来,学术界开展家庭治理研究的积极性得到进一步激发。在社会转型和现代信息技术的渗透下,家庭治理的内涵和外延发生了重大变化,倒逼家庭治理研究方法的适时更新。本文提出将计算社会科学方法引入家庭治理研究领域,重新发掘家庭治理研究的重要价值。家庭治理研究的计算范式转换具有一定的合理性和可行性:合理性表现在开展计算家庭治理研究的背后有一些潜在因素在起推动作用,包括新文科建设呼吁和社会发展需求,而技术手段多样、数据资源丰富和研究内容匹配则为开展计算家庭治理研究提供了可行性。家庭治理研究引入计算方法源于新文科建设的要求,需要重新塑造家庭研究的基础理路,同时,我国人口结构改变衍生的少子化、老龄化家庭结构改变,使得家庭治理研究必须改变以适应社会发展需求。当前,计算社会科学的技术手段的多样性、数据资料的丰富性、研究内容的匹配性,也使得计算社会科学引入家庭治理研究具备可行性。计算方法在家庭治理研究中的实践应用,包括社会大数据计算方法、社会仿真模拟方法等,这些新方法为传统家庭治理研究提供了技术便利,同时也拓宽了传统家庭研究领域。

[关键词]家庭治理 大数据技术 计算社会科学 实践应用 未来探索

[中图分类号] C913.1 [文献标识码] A [文章编号] 2096-983X(2025)03-0121-11

一、引言

党的十九届四中全会指出,要“坚持和完善共建共治共享的社会治理制度”,并明确了家庭治理在国家治理现代化中的重要价值,提出要“注重发挥家庭家教家风在基层社会治理中的重要作用”。这是以习近平同志为核心的党中央紧扣“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”这一主题,从推进党和国家事业发展全局出发做出的重大决策^[1]。中国素有重视家庭作用的优良传统和

丰富的家庭治理经验,但如何发掘历史传统中的优秀思想和经验,并将其同中国的现实发展需求相结合,以服务于国家治理现代化目标,这是一个亟需解决的时代课题。党的十九大报告指出,“中国特色社会主义进入新时代,我国社会主要矛盾已经转化为人民日益增长的美好生活需要和不平衡不充分的发展之间的矛盾”。我们把今天称为新时代,其重要表征是中国已经成为即将完成伟大复兴的社会主义大国,其经济、政治、文化、生态、科技、军事发展和社会治理能力,均取得了显著的成就。与此同时,

收稿日期:2024-11-12;修回日期:2025-03-10

*基金项目:湖南省社会科学基金项目“教育‘双减’背景下的家长教育素养测量及提升路径研究”(22YBQ031)

作者简介:冯帅帅,博士,讲师,主要从事计算社会学研究。

大数据、云计算、人工智能等高新科技的迅速发展,推动社会进入到了一个前所未有的全新状态。处在这个社会阶段的家庭治理,也毫无疑问表现出不同的特征:第一,家庭治理的主体(家庭成员、国家与社区、市场和社会主体)都发生了重大蜕变;第二,家庭目标和功能实现的内外环境也同样发生了剧烈变革,家庭单元展现出了新的面貌;第三,新技术的发展使得家庭治理的研究视域和可资利用的研究方法更是得到了极大拓展,尤其是在大数据技术对整个社会科学研究的冲击之下。

库恩提出,常规科学在发展到一定阶段后,既有范式无法解决某些“反常”问题,那么科学家共同体就会开发出新的研究范式来取代旧范式,这就是科学研究的“范式革命”^[2]。随着互联网、云计算、分布式存储、人工智能等硬件技术的日新月异和社交网站(如论坛、微博、微信)、电商购物、智能手机的飞速普及和发展,海量数据正在被源源不断地生产出来。大数据时代的到来势必会引发一场巨大的变革,不单是商业服务的拓展与社会治理的革新,以海量、连续、密集数据为基础的新科学范式也已初露端倪^[3]。鉴于庞大数据的人工制品属性^[4],社会科学理所应当被推到了科学变革的前沿,“计算社会科学”新方法应运而生^[5]。它致力于通过计算的方式,从跨平台、多样化、海量性的数据中挖掘出有益价值,为拓展、深化甚至革新对个体行为、群体交往、组织结构乃至整个社会运行规律的认识开辟出一条新路径^[6],这就是社会科学研究的计算范式转换。本文尝试将大数据技术和计算社会科学方法引入家庭治理研究领域,借此深入发掘家庭治理研究的重要价值。

总体而言,本研究的主要工作有三点:一是概述既往家庭治理研究的基本内涵和局限;二是分析家庭治理研究引入计算社会科学方法的合理性和可行性;三是对计算方法在当前家庭治理研究领域的实践应用、优势特征和将来探索展开论述。

二、家庭治理研究的内涵、局限与创新方法

在社会学中,家庭被称为社会的细胞和人类社会里最基本的生活单元,因为“人们的养生送死,几千年来主要是在家庭这个社会细胞里进行的”^[7]。家庭这个社会基本单元的组成、运行和发展,有其内在的机制和规律。运用科学有效的方法,解决家庭组成、运行和发展中出现的各种问题,从而使家庭的组成合情合法,家庭运行温馨和谐,家庭发展顺畅有序。这就是“家庭治理”。本文提出“家庭治理”概念的原因有三个方面。第一,中国素有的家庭治理传统。在中国古代,家国观念紧密相连,家庭不仅是维持社会与人口再生产的基础,也是基层政治的组成部分。孟子有云:“天下之本在国,国之本在家,家之本在身。”^{[8](P167)}要有效进行社会治理,首要任务是强化家庭治理。“一室不扫,何以扫天下”,这一流传已久的道理同样强调了家国相通的理念。第二,传统治理理论的局限性。传统政治学将“治理”限定于国家、政府或公共组织对公共事务的管理,强调集体利益和权力制衡。这种定义忽视了社会微观单元(如家庭、企业)的治理需求。20世纪后期,“治理”概念已突破公共领域边界,延伸至企业治理、社区治理、网络治理等领域。家庭作为社会基本单元,其内部规则、权力结构与冲突协调机制,本质上符合治理的核心逻辑——通过规则与互动建立秩序。第三,现代家庭公私领域的模糊性。一方面,家庭问题的公共化趋势愈加明显,当代家庭已并非完全封闭的私人领域,如家庭教育影响社会人力资本、家庭暴力触发公权力干预、赡养义务由法律强制规定。这些问题都需要家庭研究视角的适时更新。另一方面,在市场化 and 理性主义观念的冲击下,家庭成员的“私欲”(如资源分配、代际权力)也在向公共性发生转化,家庭成员的“私欲”需通过协商、责任分配等治理手段调和,否则可能演变为社会冲突(如遗产纠纷、抚养权争夺)。

当代家庭治理的主体具有明显地复合特性。其一,家庭成员是家庭治理的基本主体。家庭成员不仅是构成家庭和家庭治理的必备要素和基本力量,而且家庭治理其他主体只有通过家庭成员这个基本主体,才能够发挥其影响家庭运行的作用。其二,社会关系主体。人与家庭都是生活在社会关系网络之中,以家庭成员为中心形成该家庭具有“差序格局”特性的社会关系网络,这些亲近或信任程度由强到弱的各种社会关系,都会对家庭治理目标的实现产生强弱不一的影响力,从而成为家庭治理的社会关系主体。其三,国家与社区主体。家庭既是社会的细胞,也是国家的细胞。国家通过其治理水平、经济实力、意识形态特别是社会运行制度的设计,从宏观上对家庭治理的实现发挥影响作用。这种影响有的比较间接,人们感受不明显,有的则非常直接,如国家的户籍制度、独生子女制度和放宽生育政策设计,对家庭结构的影响则非常直接,人们感受非常明显和深刻。社区作为国家治理的基层单位,则是通过其与家庭的直接往来完成家庭治理主体功能的。其四,市场主体。随着社会的发展,家庭服务市场业已形成,为家庭提供各种有偿服务,由此形成了影响家庭治理功能实现的市场主体。

家庭治理经验和家庭治理研究是一对相互联系又有区分的概念,后者是学术共同体基于一定的理论和方法对琐碎而庞杂的家庭治理经验进行总结、提炼,以及展开推论的一类学术活动。长久以来,婚姻家庭问题和家庭治理研究一直是社会学学者的重要关照领域,而以家庭治理研究为核心内容的家庭社会学一度成为社会学最重要的分支学科之一。就内容而言,社会学关注家庭治理研究的主题非常丰富,包括但不限于:家庭生活^[9]、家庭行为^[10]、家庭关系^[11]、家庭教育^[12]、家庭生计^[13]和家政社会服务^[14],以及上述内容在社会变迁和新技术冲击下的时代演变特征^[15-16]等。家庭治理研究的方法也比较多样,如个案深描、定量建模、历史文献、国际与区域比较等。伴随着国家在制度层面对家庭

治理问题的重视,学术界对家庭治理的讨论也开始逐渐超越家庭空间单元而同基层社会治理^[17]、现代社会建设^[18]和性别平等^[19]等中观议题勾连起来,极大丰富和拓展了家庭治理的内涵和外延,也进一步激发了学术界开展家庭治理研究的积极性。

总体来看,虽然家庭治理研究已经取得较大成效,但仍存在部分潜在问题和缺陷。如杭苏红^[20]基于中国家庭社会学发展的历史回顾指出,尽管70年来中国家庭社会学在研究议题、研究方法、研究视角等各个方面均取得了卓越成就,但仍有以下瓶颈问题需要克服:一方面,无论是对中国历史传统中优秀的家庭治理思想还是西方的家庭相关理论,目前都缺乏足够的梳理和反思;另一方面,在跨学科视角和创新方法的采纳上,现有研究还存在明显不足,“仍需要进一步加强与其他学科和专业的交流与相互砥砺,从而激活学科内在的创新性与活力”。正是基于这样的思考,本文提出将计算社会科学方法引入家庭治理研究领域。

需要说明的是,本文只是倡导将计算方法作为家庭治理研究路径之一,而非唯一。方法只是研究者出于探究科学规律的目的,收集和分析资料的路径选择而已,其本身并无高下之分,具体方法的选择要根据选题、客观条件和研究者喜好而定。我们呼吁家庭治理研究引入计算方法是因为:第一,家庭治理研究本身具有明显的跨学科属性,这一点与计算社会科学方法的跨学科特征相同^[21]。家庭在变化环境中遇到的复杂问题需要跨学科的视角,单一视角不足以解释家庭问题的复杂性^[22]。跨学科优势可以让家庭治理研究者“超越学科边界,整合理论、政策和实践,领导学术界、高等教育、公民社会和其他部门联合起来共同解决人类家庭生活问题”^[23],也更容易接受因计算范式转换而可能造成的观念冲击。第二,以个案深描和统计建模为代表的传统方法在横截面讨论家庭生活、家庭互动和家庭结构等议题上均有突出优势,如费孝通^[7]借助参与式观察方法对花蓝瑶

和江村家庭组织的生动刻画,王跃生^[24]基于全国人口普查数据对中国城乡家庭结构的分析。但是,受制于研究思路和方法的局限,现有家庭治理研究在对家庭问题的长时段和多类别观察上一直未有进展。这正是计算社会科学方法的优势所在。

具体而言,所谓家庭治理研究的计算范式转换,就是运用计算社会科学方法对家庭治理的研究对象进行分析,从而科学地总结和探索家庭活动的内在规律。大数据时代到来,海量数据、分布式运算和编程技术的发展为学者开展家庭治理研究提供了新的视角和方法。一方面,互联网(包括移动互联网)和普遍的传感器等数据采集设施正在汇集大量数据,使得对家庭治理研究对象的分析可以突破原有的数据和内容限制;另一方面,分布式运算、计算机编程等技术的进步则为创新家庭治理研究的内容和方法提供了算力和算法支持。

三、家庭治理研究引入计算方法的合理性与可行性

恩格斯在《自然辩证法》中提出了历史唯物主义的著名命题:历史的发展进程是必然性与偶然性的辩证统一。这一命题同样适用于解释家庭治理研究的计算范式转换上来。任何一个新方法的提出都是为了解决传统方法所力有不逮的问题,这是新方法被提出的必然性。但解决该问题的新方法可以有很多,为什么是方法A而不是B被提出,原因在于当前所具有的各项条件与方法A的亲合性更高,此为新方法提出的偶然性。从这个角度来说,家庭治理研究的计算范式转换也是必然性和偶然性的辩证统一。为避免过度阐释的麻烦,这里我们用合理性和可行性分别代替必要性和偶然性。

(一) 家庭治理研究引入计算方法的合理性

1. 新文科建设呼吁

2018年8月,在全国教育大会召开前,中共中央明确提出“高等教育要努力发展新工科、

新医科、新农科、新文科”。2019年4月,教育部、中央政法委、科技部等13个部门联合启动“六卓越一拔尖”计划2.0版,全面推进新工科、新医科、新农科、新文科建设。新文科的概念逐渐引起各界广泛关注。2020年11月,由教育部新文科建设工作组主办的新文科建设工作会议在山东大学(威海)召开,发布《新文科建设宣言》,对新文科建设作出全面部署。会议指出,文科教育是培养自信心、自豪感、自主性,产生影响力、感召力、塑造力,形成国家民族文化自觉的主战场主阵地主渠道。新文科是相对于传统文科而言的,其中,打破专业壁垒和学科障碍,是新文科建设的关键所在^[25]。具体来说,就是要做到文科内部融通、文理交叉融合,突破固有专业藩篱,充分把握学科所涉对象的规律性,通过研究认识和解决学科所面对的复杂社会问题,为国家和社会发展提供支撑和指导。作为人文学科的重要组成部分,家庭社会学和家庭治理研究当然要积极响应国家号召,打破局限,用更加广阔的视角,对家庭研究的基础理路(理论与方法)进行重新塑造,彰显中国特色、中国风格和中国气派。

2. 社会发展需求

迫切的社会需求是家庭治理研究理论和方法创新的重要时代背景。我国正处在一个高速变迁的社会转型当中,很多具有深远影响的社会现象对创新家庭治理研究提出了极为迫切的现实需求,具体包括:第一,人口老龄化。全国老龄办常务副主任王建军指出,预计到2025年,我国60岁及以上老年人口数将达到3亿人,占总人口的1/5;到2033年将突破4亿人,占总人口的1/4左右;而到2050年前后将达到4.87亿人,约占总人口的1/3。^[26]老龄化对家庭结构的冲击、养老问题、鳏寡孤独老人照料、家庭代际关系等都是家庭治理研究的重要主题。第二,家庭少子化。最新发布的第七次人口普查数据显示,2020年中国新生儿人口为1200万人,比2019年下降了18%,几乎成为新中国有记录以来新生儿人口最低的一年。少子化背景

下的家庭生育意愿、家庭结构、育儿照料、亲子关系、生育政策的效果预测等问题也亟需家庭治理研究的快速跟进。第三,家务劳动社会化。随着市场化和城市化进程的加快,双职工家庭占比越来越高,虽然现代家庭的经济条件因此得到较大改善,但也对家庭关系和家庭生活造成一定程度的影响。在这一背景下,越来越多的“有钱没闲”的现代家庭开始诉诸市场寻求帮助。家务劳动社会化对现代家庭的冲击,以及随之延伸而来的家政工、家政市场等主题都是新时期家庭治理研究应重点关照的内容。

(二) 家庭治理研究引入计算方法的可行性

1. 技术手段多样

计算社会科学方法是大数据时代信息技术、社会计算和社会科学相互融合产生的新的跨学科研究方法。发展到今天,计算社会科学方法已经积累了相当丰富的技术手段。第一,社会信息自动提取和分析技术。随着深度学习、迁移学习、神经网络等新算法的提出,社会信息有效提取和分析技术取得重大突破。过去研究者一般将问卷数据、历史信息、古籍翻阅、新闻报纸、音频广播等作为研究资料的主要来源,且一般是手动完成的,现在这类资料,乃至其他新型信息资料的获取基本可依靠计算工具来实现。如龚为纲等人^[27]对GDELT全球新闻媒体信息的调用与分析,黄荣贵^[28]对微博数据的分析等均系社会信息自动提取与分析技术的创新应用。第二,社会网络分析技术。网络分析技术是由社会理论和应用与形式数学、统计学和计算机技术有机结合而来的结果^[29]。近年来,社会网络分析技术吸引了国内外来自计算机科学和社会科学等多个学科专家的关注,包括人际网络^[30]、组织网络^[31]、学术合作网络^[32]等。第三,社会仿真模拟技术。复杂适应系统(complex adaptive system)理论认为,人类社会是一个典型的复杂自适应系统,“主体异质性、主体适应性、因素交互性、作用关系非线性,导致了主体行为、系统状态和系统演变的多样性与不确定

性”^[33],这些特征使得传统的线性解释模型陷入瓶颈,社会仿真模拟技术由此提出。常见的社会仿真模型包括:系统动力学模型(SDM)、元胞自动机(CA)、基于方程建模(EBM)、基于主体建模(ABM)和ABM高阶形态的多主体模型(MAM)。就现实应用而言,社会仿真模拟技术在社会互动流动^[34]、国家安全^[35]、传播效果^[36]等研究领域均拥有不错的应用效果。上述这些新工具为开展计算范式下的家庭治理研究提供了必要的技术基础。

2. 数据资源丰富

拉泽尔等人^[5]指出,我们在日常生活中的电子邮件、移动通讯、公共交通、信用卡购物、看病就医、网页浏览、点子购物、视频监控、社交软件等这些不为传统社会科学研究者所重视的“数字踪迹”信息正在大量的生产和聚集,这些长时段、大规模、连续性的“新数据”类型既“可以描绘出个人和群体行为的综合图景”,也有可能“会改变我们对于生活、组织和社会的理解”。因此,无论是商业运营、政府管理,还是开展社会科学研究(尤其是量化研究),这些数据集无疑拥有巨大的价值。传统定量社会科学横向研究案例、纵向研究变量,且通常假设观察值之间具有独立性,而计算社会科学的关注则涵盖了文本、位置、运动、网络、图像以及视频多种内容,并应用统计模型来获取数据中的多重交互依赖性^[37]。丰富的数据资源为研究者开展社会科学研究带来了重大的影响,家庭治理研究自然也包含在内。可支撑开展计算家庭治理研究的数据集相当丰富,如家庭就医记录、家庭电子消费、家庭财政、家庭健康、社交媒体互动、家庭出行等,甚至微博、微信、论坛等网络空间涉及家庭主题讨论的文本等皆可开展分析。近年来,大小数据交互论证和融合使用的研究路径逐渐得到重视。因此,不少传统成熟的家庭专项调查数据库(如西南财经政法大学的“中国家庭金融调查项目”、浙江大学的“中国家庭大数据库”等)和其他大型数据库中涵盖的家庭领域数据也可以为开展计算家庭

治理研究提供数据支撑。

3. 研究内容匹配

家庭治理研究的很多内容非常适宜采用计算的方法来开展,可以较好的弥补传统方法的不足。以家庭生育预测研究为例。2015年10月,中共十八届中央委员会第五次全体会议公报指出:全面实施一对夫妇可生育两个孩子政策,积极开展应对人口老龄化行动。在二胎政策出台前夕,关于“全面放开二胎是否能够有效刺激家庭生育决策和行为”的政策效果问题引发学界热议。一部分学者对此持较为积极的态度。如翟振武等人^[38]基于全国人口普查数据的分析预测认为,符合二胎政策条件的目标人群规模较大,且二胎生育意愿仍处于较高水平,因此全面放开二胎政策将会明显改善我国总人口未来进入负增长的趋势,延缓人口老龄化进程。而另一些学者则对此呈怀疑态度。庄亚儿等人^[39]基于2013年开展的全国生育意愿调查数据发现,在一孩家庭生育二胎的时间安排上,“未确定”的比例高达59.0%,因此他们预测,“全国范围内不太可能出现出生堆积现象”。

从上述研究可以发现,同样是基于传统量化方法的预测性研究,对同一对象的研究结论却存在较大差距,这与传统量化研究的“过度拟合”(overfitting)^①问题相关。计算社会科学方法最新研究进展表明,机器学习(ML)建模方法在开展预测研究^[40]、处理异质性^[41]和应对模型过拟合问题^[42]上具有突出优势。遗憾的是,目前借助机器学习建模技术开展家庭治理量化研究的成果还比较少,但随着计算方法的普及,我们对机器学习范式在家庭治理研究领域(特别是预测研究)的运用前景充满期待。

四、计算方法在家庭治理研究中的 的实践应用、优势特征与未来探索

从20世纪90年代到现在,计算社会科学已形成颇为成熟的方法体系,其中,社会大数据计算(又称大数据挖掘与分析)、社会仿真模拟和互联网社会实验是公认的三大计算社会科学方法^[43]。就国内外相关研究成果而言,已有的计算家庭治理研究主要借助前两大方法予以展开。

(一) 计算家庭治理研究的实践应用

1. 社会大数据计算方法在家庭治理研究领域的运用

互联网大数据类型一般存在多源、异构、多样、高维等特点,无法直接为研究者提供数据支持,需要借助计算手段,将非结构化和半结构数据转换为直接可为研究者使用的结构化数据。大数据挖掘与分析技术为拓展家庭治理研究视野提供了便利,如林寒和罗教讲^[44]以微舆情大数据库为支撑,就全面二孩政策在网络空间中所引发舆情走势、媒体活跃度和所涉舆情主体的关注特征展开分析。冯利伟^[45]基于Web of Science数据库全面梳理了1986—2017三十年间发表于外文重点期刊上的关于“工作—家庭冲突”(work-family conflict)的学术讨论,并以可视化和计算方法从发文数量、发文期刊、发文学者来源地、学术属性、高引与共引文献、核心作者群、重要研究机构、关键词共现与战略坐标等分析维度予以呈现。国外学者马尔米、吉奥尼斯和索林^[46]基于大数据挖掘和分析技术对家谱网络(genealogical networks)展开了研究,并开发了自动推断大规模家谱网络的“原则化概率机器学习方法”,作者将之命名为BinClass。为证明BinClass的适用性,作者以

^①2016年9月, Lever、Krzyszynski和Altman三人在Nature Methods杂志上联合署名发表了一篇题为《模型选择与过度拟合》(Model selection and overfitting)的文章,专门讨论科学研究中的过度拟合问题。在传统OLS回归建模中,回归模型的参数估计是通过最小化损失函数得到的,即以最小化观察值与预测值之间的垂直平方距离为核心目标,基于此提升模型的拟合度。置言之,OLS建模技术的本质是通过控制偏差来试图降低模型总体的预测误差。但是,随着模型中预测变量的增多,模型的方差值会因此而增大,导致解析模型只能对本数据集拥有很好的拟合效果,但却并不适用于其它现实样本,预测能力低。

某位家谱学家收集的由116,640个个体组成的长时段家谱网络数据和瑞典HisKi项目^①的人口记录数据为支撑进行了检验。验证结果显示, BinClass具有61.6%的基础预测准确率。该研究还发现, 人们的婚配选择与社会经济地位密切相关, 也即我们常说的“门当户对”, 并且在长达150年的时间跨度内, 这种现象保持在相当稳定的状态。

2. 社会仿真模拟方法在家庭治理研究领域的运用

目前, 仿真模拟方法在家庭治理研究中的应用较为广泛。如罗莎莉和詹姆斯^[47]提出“基于主体的多代家庭模拟”构想, 期望借此探究多代家庭在长时段变迁中的演变特征, 如家庭规模、家庭关系、家庭与社会环境的互动等。类似的, 埃利安等人^[48]采用人群模拟(crowd simulation)方法对家庭单元的主体互动展开模拟, 将父母、青少年、儿童和老人等主体引入模型, 并设置相关的参数来管理他们的行为, 最终在参数调整和模型修正的基础上提出了一个“家庭关系良好的家庭行为基础模型”。杨等人^[49]关注到女性角色在家庭中的重要作用, 并基于ABM方法分析了明清时期女性在精英家族中的功能体现。研究发现, 母亲、父亲、叔父和叔母的综合影响对于维持成功的家族规范具有重要意义。此外, 仿真模拟方法在家庭行为决策^[50]、工作—家庭冲突的动态演化^[51]等议题上均有不错的尝试。

(二) 计算家庭治理研究的优势与未来探索

通过对国内外文献的梳理可以发现, 计算方法之于家庭治理研究具有突出优势。其一, 可以拓宽传统家庭治理研究视域。以微观家庭主体互动研究为例, 以往学者对家庭主体互动的关注大多采用观察法和个案深描法开展研究, 其分析也往往只能建立在互动结果发生之后, 缺乏对“另一种可能”的关注。仿真模拟

方法的应用为研究者分析家庭互动结果的“其他可能”提供了技术基础。其开展步骤被王飞跃^[52]总结为ACP方法: 第一步是通过算法架构人工社会(artificial societies), 架构基础既可以源自现实经验, 也可以来自研究者的启发性知识; 第二步是开展计算实验(computational experiments), 即将计算机变为社会实验室, 研究者可以充分发挥想象力, 通过参数调整观察不同参数配置下的主体行动结果; 第三步是平行执行(parallel execution), 通过现实与人工系统的互动, 平行执行, 实现对复杂社会问题的控制和管理。其二, 为研究传统议题提供便利。以上文列举的家谱网络研究为例, 家谱网络研究在过去是非常琐碎和耗时, 需要研究者手动翻阅大量的文字资料和记录。但借助大数据挖掘与分析技术, 能够大幅压缩工作量, 提高研究效率。其三, 提升量化研究结论的泛化能力和预测精准度。社会科学传统量化研究的泛化能力较差, 常常出现因数据不同、变量处理方式不同和变量选取数量不同而导致结论明显相悖。哪怕是在复杂模型层出不穷的今天, 传统量化研究的这一问题仍没有得到很好解决。而机器学习算法建模技术的交叉验证可以在一定程度上缓解该问题。机器学习算法可以看作是在经验训练的指导下检索大量候选程序, 最终选取具备最佳拟合效果和优化性能模型的程序, 可以帮助研究者识别和筛选关键变量, 提升模型简洁度, 拔高量化模型的泛化能力。

总体上, 尽管学术界现已积累了不少计算家庭治理研究成果, 但具有以下特征。第一, 国外学者成果发表数量远多于国内学者。第二, 从学科归属来看, 自然学科成果居多, 人文学科较少。第三, 大部分研究的层次还停留在相对粗浅的层面, 基本上都是技术专家对家庭治理研究内容的初步探索, 缺乏进一步的理论梳理与提升。我们认为, 立足于“新文科”发展定位

^①“HisKi”项目诞生于1980年代, 旨在对芬兰教区手写的登记簿进行数字索引。数据库包含大约500万条出生记录和500万条死亡、婚姻和移民记录。除了父母的姓名和职业之外, 每个出生记录通常还包含孩子的姓名、出生地点和出生日期。HisKi数据集可在<http://hiski.genealogia.fi/hiski?en>上公开获得。

和中国社会需求,未来国内学者开展计算家庭治理研究获可从以下内容展开重点尝试。

第一,基于计算方法对中国历史古籍中的家庭治理思想开展知识谱系梳理。中国的家庭治理历经种种变迁。从《周易·家人》伊始,关于家庭治理的著述就不绝如缕。到春秋战国的《大学》《孝经》都包含家庭治理思想,再到南北朝和隋唐时期的《颜氏家训》《女论语》等,形成我国传统家庭治理领域的经典文献。宋元明清时期的《家范》《童蒙须知》《三字经》《治家格言》等等,都是古代家庭治理的主要依据。这类传统的家庭领域的著述尽管其思想部分已不适用于现代社会,但是能够为研究当代中国家庭的由来及其思想提供依据。随着云存储、移动互联网、大数据和人工智能新技术发展,古籍的数字化工作有了突破性进展。比较成熟的古籍数据包括:爱如生典海数字平台、中华经典古籍库、书同文古籍数据库、国学迷古籍数据库、鼎秀古籍全文检索平台等。在不断更新升级的自然语言处理技术支持下,这些优秀的古籍数据为计算家庭治理研究提供基础资料。

第二,家庭生育政策研究。在中国低生育率和老龄化程度日趋加深双重背景下,关注家庭生育意愿具有重要的学术和政策价值。计算范式转换下的家庭生育研究可在以下议题中展开:其一,探究家庭生育意愿的关键变量。家庭生育意愿是复杂因素共同作用下的家庭行为决策,相关研究已指出,家庭收入、生育政策、社保水平、家庭支持等诸多宏观、微观因素均对家庭生育意愿具有显著的影响^[53-54]。假使在某种特殊情境下需要通过制定某项政策激发生育积极性,那么找寻关键变量的方法就变得极为重要。事实上,机器学习建模技术在选元、分类、聚类等诸多领域拥有相当优势,能够快速识别和筛选出关键变量,提升模型的预测精度^[55]。其二,数据驱动下的家庭生育决策模拟研究。借助传统量化研究得出回归系数后,以上述系数数值为基础构建家庭生育决策模拟系统,通过

参数调整观察平行人工世界的动态演化,进而为现实世界提供指导。

第三,计算方法在家务劳动社会化研究中也具有突出优势。作为家庭治理研究的重要组成部分,近来,由家务劳动社会化延展出来的“家政工”问题成为学术界的热点话题,学术讨论散布在职业社会学、产业社会学、组织社会学,以及以个案方法为基础的劳动过程研究等诸多领域。实际上,计算社会科学方法也可比较好地用于家务劳动社会化问题研究中来:其一,大数据挖掘与分析技术的应用。如为应对我国家政市场的服务类型不匹配、服务层次不匹配、地域供需不匹配和季节性供需失衡问题,孔德财等人^[56]基于家政服务人员和雇主的服务技能、服务时间和多指标评价等信息要素建构了一套以家政服务人员 and 雇主的最大双边满意度为目标“家政服务人员与雇主的双边匹配”算法模型,专门用于解决需求匹配问题。未来在海量家政业务数据聚集下,基于供需数据汇总、用户画像、特征工程、相似度计算、排序建模、过滤建模等算法架构下的“家政服务算法推荐系统”也可以实现。除上述内容外,大数据挖掘与分析方法在家政行业诚信研究、家政服务的规范与标准化建设研究等议题均拥有不错的应用前景。其二,仿真模拟方法的应用。家政计算机仿真模拟方法在家政市场研究领域的应用前景也非常广阔,宏观模拟如家政服务供需平衡研究、家政行业健康业态研究、家政企业发展过程研究等;微观模拟如家政服务人员的行为激励与行为控制研究、家政雇主与家政服务从业人员的互动行为模式研究、家政雇主与家政服务从业人员信任关系建立机制研究等,都拥有良好应用技术和条件。

当然,数据伦理和合法性也是使用计算社会科学方法介入家庭治理研究必须重点考虑的问题。在开展计算家庭治理研究时,数据伦理与合法性不仅是使用计算社会科学方法不可或缺的前提,更是确保研究严谨性、公正性和可持续性的基石。随着大数据、人工智能等技术

的飞速发展,计算社会科学方法为家庭治理研究提供了前所未有的洞察力和精确度,但同时也带来了前所未有的伦理与法律挑战。其一,数据伦理要求研究者在收集、处理家庭数据时,必须坚守知情同意原则,确保每位家庭成员都明确知晓其数据将被如何使用,并在自愿的基础上给予授权。这意味着,研究设计之初就应构建透明、清晰的沟通机制,让家庭成员充分理解研究目的、数据保护措施以及可能的风险,从而建立起基于信任的合作关系。此外,研究者还需持续关注数据使用的道德边界,避免过度采集敏感信息,尤其是涉及个人隐私、家庭关系等高度敏感领域的的数据,确保研究活动不会侵犯家庭成员的基本权利。

表1 家庭大数据分类管理建议

数据类型	使用边界	监管强度
家庭行为数据	限于科学研究(匿名化处理)	高(需伦理审查)
家庭成员生物数据	符合相关要求后可用于科学研究	最高(法律禁止泄露)
家庭经济数据	使用需符合《个人信息保护法》	中(成员知情同意+授权)

其二,合法性是保障家庭治理研究顺利进行的关键。这要求研究者必须严格遵守国家和地区关于数据保护、隐私权的法律法规,确保所有数据处理活动都在法律框架内进行。包括但不限于,研究前需获得必要的审批或许可,实施过程中采取适当的技术和管理措施保护数据安全,以及研究成果发布时遵守版权和数据共享的相关规定。在跨国或跨地区合作中,还需特别注意不同法律体系间的差异,确保研究活动的合规性。进一步地,面对数据偏见和歧视性使用的潜在风险,研究者应采取积极措施加以预防和纠正。这包括在数据收集阶段就注重多样性和代表性,确保样本能够全面反映不同家庭背景、社会经济地位、文化背景等维度的特征;在数据分析阶段,运用统计方法识别并减轻潜在的偏见影响;在结果解读和传播时,保持客观、中立,避免对特定群体造成误解或伤害。

参考文献:

- [1]马焱. 家庭家教家风:创新基层社会治理体系的新视角——兼论新时代妇联组织的家庭工作[J]. 妇女研究论丛, 2020(6): 58-63.
- [2]库恩. 科学革命的结构(第四版)[M]. 胡新,译. 北京:北京大学出版社, 2012: 88.
- [3]HETT, TANSLEY S, TOLLE K. The fourth paradigm: Data-intensive scientific discovery[J]. Proceedings of the IEEE, 2011, 99(8): 1334-1337.
- [4]EPSTEIN J M, AXTELL R. Growing artificial societies: Social science from the bottom up[M]. Cambridge: The MIT Press, 1996: 134.
- [5]LAZER D, PENTLAND A, Adamic L, et al. Computational social science[J]. Science, 2009, 323(5915): 721-723.
- [6]乔菲-雷维利亚. 计算社会科学:原则与应用[M]. 梁君英,译. 杭州:浙江大学出版社, 2019: 8.
- [7]费孝通. 社会学讲义[M]. 上海:华东师范大学出版社, 2019: 64.
- [8]杨伯峻. 孟子译注[M]. 北京:中华书局, 1960.
- [9]风笑天,易松国. 城市居民家庭生活质量:指标及其结构[J]. 社会学研究, 2000(4): 107-118.
- [10]杨菊华,李路路. 代际互动与家庭凝聚力——东亚国家和地区比较研究[J]. 社会学研究, 2009(3): 26-53.
- [11]石金群. 转型期家庭代际关系流变:机制、逻辑与张力[J]. 社会学研究, 2016(6): 191-213.
- [12]卢春天,李一飞,陈玲. 情感投入还是经济支持:对家庭教育投资的实证分析[J]. 社会发展研究, 2019(1): 49-67.
- [13]高功敬,陈岱云,梁丽霞. 中国城市贫困家庭生计资本指标测量及现状分析[J]. 济南大学学报(社会科学版), 2016(3): 101-119.
- [14]吴心越. 市场化的照顾工作:性别、阶层与亲密关系劳动[J]. 社会学评论, 2019(1): 75-86.
- [15]於嘉,谢宇. 社会变迁与初婚影响因素的变化[J]. 社会学研究, 2013(4): 1-25.
- [16]孙信茹,王东林. “抖音之家”:新技术与家庭互动的文化阐释[J]. 新闻大学, 2021(10): 58-75.
- [17]党日红,李明舜. 构建基层社会治理新格局须纳入家庭视角[J]. 妇女研究论丛, 2019(6): 12-15.
- [18]南钢. 家庭治理:现代社会治理的重要基石[J]. 中国浦东干部学院学报, 2015(3): 128-132.
- [19]吴兰丽. 国家治理·家庭建设·性别平等——第二届湖北高校性别平等论坛综述[J]. 妇女研究论丛, 2016(2): 122-124.
- [20]杭苏红. 经验、历史与方法:家庭社会学七十年

- 回顾[J]. 中华女子学院学报, 2021(4): 55-62.
- [21]罗教讲, 张东驰. 大数据时代的计算社会科学与学术话语体系重构[J]. 吉首大学学报(社会科学版), 2018(2): 9-16.
- [22]HORN MJ, NICKOLS S Y. Interdisciplinary research: Have we lost our focus?[J]. Family & Consumer Sciences Research Journal, 2010, 11(1): 9-14.
- [23]MCGREGOR S L T. Name changes and future-proofing the profession: Human sciences as a name?[J]. International Journal of Home Economics, 2010, 3(1): 20-37.
- [24]王跃生. 中国城乡家庭结构变动分析——基于2010年人口普查数据[J]. 中国社会科学, 2013(12): 60-77.
- [25]樊丽明, 杨灿明, 马骁, 等. 新文科建设的内涵与发展路径(笔谈)[J]. 中国高教研究, 2019(10): 10-13.
- [26]到2050年老年人将占我国总人口约三分之一[EB/OL]. (2018-07-19)[2024-10-14]. https://www.gov.cn/xinwen/2018-07/19/content_5307839.htm.
- [27]龚为纲, 朱萌, 张赛, 等. 媒介霸权、文化圈群与东方主义话语的全球传播——以舆情大数据GDELT中的涉华舆情为例[J]. 社会学研究, 2019(5): 138-164.
- [28]黄荣贵. 网络场域、文化认同与劳工关注社群——基于话题模型与社群侦测的大数据分析[J]. 社会, 2017(2): 26-50.
- [29]沃瑟曼, 福斯特. 社会网络分析: 方法与应用[M]. 陈禹, 孙彩虹, 译. 北京: 中国人民大学出版社, 2012: 47.
- [30]宋恩梅, 左慧慧. 新浪微博中的“权威”与“人气”: 以社会网络分析为方法[J]. 图书情报知识, 2012(3): 43-54.
- [31]黄荣贵, 桂勇, 孙小逸. 微博空间组织间网络结构及其形成机制: 以环保NGO为例[J]. 社会, 2014(3): 37-60.
- [32]KEUCHENIUS A, TORNBERG P, UITERMARK J. Adoption and adaptation: A computational case study of the spread of Granovetter's weak ties hypothesis[J]. Social Networks, 2021(66): 10-25.
- [33]罗俊. 计算·模拟·实验: 计算社会科学的三大研究方法[J]. 学术论坛, 2020(1): 35-49.
- [34]冯泽鲲. 身份、情感与约束: 本土社会地位等级形成的ABM实验[J]. 甘肃社会科学, 2021(1): 216-224.
- [35]WALDROP M M. Free agents[J]. Science, 2018, 360(6385): 144-147.
- [36]李凤翔, 罗教讲. 计算社会科学视角: 媒体传播效果的计算机模拟研究[J]. 学术论坛, 2018(4): 15-27.
- [37]LAZER D, PENTLAND A, WATTS D J, et al. Computational social science: Obstacles and opportunities[J]. Science, 2020, 369(6507): 1060-1062.
- [38]翟振武, 张现苓, 靳永爱. 立即全面放开二胎政策的人口学后果分析[J]. 人口研究, 2015(2): 3-17.
- [39]庄亚儿, 姜玉, 王志理, 等. 当前我国城乡居民的生育意愿——基于2013年全国生育意愿调查[J]. 人口研究, 2014(3): 3-13.
- [40]陈云松, 吴晓刚, 胡安宁, 等. 社会预测: 基于机器学习的研究新范式[J]. 社会学研究, 2020(3): 94-117.
- [41]胡安宁, 吴晓刚, 陈云松. 处理效应异质性分析——机器学习方法带来的机遇与挑战[J]. 社会学研究, 2021(1): 91-114.
- [42]ZHANG J X, FENG S S. Machine learning Modeling: A new way to do quantitative research in social sciences in the era of AI[J]. Journal of Web Engineering, 2021, 20(2): 281-302.
- [43]罗玮, 罗教讲. 新计算社会学: 大数据时代的社会学研究[J]. 社会学研究, 2015(3): 222-246.
- [44]林寒, 罗教讲. 全面二孩政策的舆情解读——基于微舆情大数据的分析[J]. 情报杂志, 2016(7): 66-72.
- [45]冯利伟. 国外工作家庭冲突研究态势——基于文献计量的分析[J]. 经济管理, 2018(4): 187-208.
- [46]MALMIE, GIONIS A, SOLIN A. Computationally inferred genealogical networks uncover long-term trends in assortative mating[C/OL]. 2018 World Wide Web Conference, 2018: 883-892. <https://doi.org/10.1145/3178876.318613>.
- [47]OTTERS R V, HOLLANDER J F. The times they are a' changing: Multigenerational family simulations[J]. Marriage and Family Review, 2018, 54(2): 183-208.
- [48]ELIYAN L, HALABI O, SALEH M. Modeling family behavior in crowd simulation[C/OL]. 2018 International Conference on Computer and Applications, 2018: 66-70. <https://ieeexplore.ieee.org/document/8460299>.
- [49]YANG C, KURAHASHI S, ONO I, et al. Pattern-oriented inverse simulation for analyzing social problems: Family strategies in civil service examination in imperial China[J]. Advances in Complex Systems, 2012, 15(7): 1801-1808.
- [50]OTTERS R V, HOLLANDER J F. Leaving home and boomerang decisions: A family simulation protocol[J]. Marriage and Family Review, 2015, 51(1): 39-58.
- [51]WU G, DUAN K, JIAN Z, et al. System dynamics model and simulation of employee

work-family conflict in the construction industry[J]. *International Journal of Environmental Research and Public Health*, 2016, 13(11): 1059-1063.

[52]王飞跃. 平行控制: 数据驱动的计算控制方法[J]. *自动化学报*, 2013(4): 293-302.

[53]王天宇, 彭晓博. 社会保障对生育意愿的影响: 来自新型农村合作医疗的证据[J]. *经济研究*, 2015(2): 103-117.

[54]许琪. 性别公平理论在中国成立吗?——家务劳动分工、隔代养育与女性的生育行为[J]. *江苏社会科学*,

2021(4): 47-58.

[55]ATHEY S, IMBENS G. Recursive partitioning for heterogeneous causal effects[J]. *Proceedings of the National Academy of Sciences*, 2016, 113(27): 7353-7360.

[56]孔德财, 姜艳萍, 纪楠. 家政服务人员与雇主的双边匹配模型[J]. *东北大学学报(自然科学版)*, 2015(11): 1168-1172.

【责任编辑 史敏】

Innovative Methods in Big Data Technology and Family Governance Research

FENG Shuashuai

Abstract: Since the 19th Central Committee of the Communist Party of China's Fourth Plenary Session emphasized the fundamental role of the family in grassroots governance and the modernization of national governance, the enthusiasm of academic circles for conducting family governance research has been further stimulated. Amidst social transformation and the infiltration of modern information technology, significant changes have occurred in the connotation and extension of family governance, necessitating adaptive updates in research methods for family governance. This paper proposes introducing computational social science methods into the field of family governance research, with the hope of rediscovering the important value of family governance research. The paradigm shift towards computational approaches in family governance research exhibits certain rationality and feasibility. Rationality lies in the underlying factors driving the development of computational family governance research, including the call for new liberal arts construction and societal development needs, while the diversity of technical means, abundance of data resources, and alignment of research content provide the feasibility for conducting computational family governance research. The study reviews and summarizes practical applications of computational family governance research in academic circles both domestically and internationally, and discusses the advantageous characteristics of introducing computational social science methods into family governance research, as well as future research content suitable for the Chinese context.

Keywords: family governance; big data technology; computational social science; practical applications; future explorations