

■ 教育数字化

DOI:10.15998/j.cnki.issn1673-8012.2023.05.006

数据密集型社会科学发展的经验及启示： 一个国际案例的分析

张海生¹, 公钦正²

(1. 重庆文理学院 期刊社, 重庆 402160; 2. 中国人民大学 教育学院, 北京 100872)

摘要:面对未来发展的不确定性,以及技术发展在学科建设及其研究范式的高渗透性,高校尤其是以人文社会科学见长的高校更应该主动迎接并积极运用数据密集型科学发现的优势,充分考虑数据驱动下的知识发现对于传统社会科学学科建制及其研究范式的重大影响、冲击乃至颠覆作用。一方面,数据密集型科学的发现及其在社会科学领域的渗透与应用,会带来数据密集型科学与社会科学的交叉融合发展,推动社会科学的学科建设方向由传统走向现代、由经验转向科学;另一方面,借助数据密集型科学发现基于数据驱动的研究优势能够有效提升社会科学研究的速度、效率和效果,进一步推动社会科学研究范式的转型,且人文主义与科学主义相结合的混合研究将成为社会科学研究的主流范式。以耶鲁大学为研究个案,着重分析其发展数据密集型社会科学的愿景目标和实践举措。研究发现,耶鲁大学通过发展并应用数据密集型社会科学的方法,拟实现借助数据密集型科学的强大力量不断赋能并推动社会科学研究范式转型、通过建立社会科学研究社区促进跨学科研究与协作、有选择地改善基础服务并强化制度保障与治理支持3个维度的愿景目标。为了实现这些愿景目标,耶鲁大学通过创建数据密集型社会科学中心、了解其他大学和行业卓越机构的发展趋势、在大学层面建立委员会等实践举措,实现重塑组织结构、调整组织行为和促进跨学科研究的整合效应;通过加强基础数据密集型课程群建设、增强数据密集型社会科学发展的多样性和包容性,以及教师年度会议和委员会等制度保障和治理支持,不断优化数据密集型社会科学的人才培养目标,推动课程教学的创新发展和科学评估。由此,耶鲁大学不仅推动了数据密集型科学与社会科学的知识交叉融合发展,还加快了数据密集型科学在社会科学研究中的应用。我国高校在未来的数据密集型社会科学实践发展中,需要主动转变社会科学发展的传统理念,主动运用数据密集型科学

修回日期:20230615

基金项目:教育部人文社会科学研究青年基金项目“人工智能、制度嵌入与大学治理效能提升研究”(22YJC880107);重庆市高等教育学会高等教育科学研究课题“人工智能赋能大学治理变革研究”(CQGJ21B080)

作者简介:张海生,男,安徽临泉人,重庆文理学院期刊社编辑,管理学博士,主要从事高等教育理论与管理和期刊编辑出版研究。

通信作者:公钦正,男,山东临沂人,中国人民大学教育学院博士生,主要从事高等教育理论与管理和比较高等教育研究。

引用格式:张海生,公钦正.数据密集型社会科学发展的经验及启示:一个国际案例的分析[J].重庆高教研究,2023,11(5):62-73.

Citation format:ZHANG Haisheng, GONG Qinzheng. Experiences and enlightenment from the development of data-intensive social sciences: an analysis of an international case [J]. Chongqing higher education research, 2023, 11(5): 6273.

发现赋能社会科学转型发展,加强社会科学和数据密集型科学之间的深度融合,建立与数据密集型社会科学相匹配的人才培养目标与跨学科课程体系,注重数据密集型社会科学发展的整体规划与适时动态调整。

关键词:数据密集型社会科学;学科建设;科学研究范式;跨学科研究

[中图分类号]G644 [文献标志码]A [文章编号]16738012(2023)05006212

一、问题提出

科学技术大繁荣的时代一定是社会大发展的时代。以大数据、人工智能为代表的新一轮科技革命的兴起,为社会各领域的发展注入新动能。为了应对新一轮科技革命尤其是数据驱动的科学发现(数据密集型科学发现)的挑战,全球知名高校纷纷通过转变发展理念、重塑组织结构、优化人才培养目标等方式,希冀在新一轮的教育、科技和人才竞争中掌握先机、保持一流地位,如哈佛大学成立定量社会科学研究所(Institute for Quantitative Social Science at Haman)、斯坦福大学成立人工智能研究所(Stanford Institute for Human-Centered Artificial Intelligence)、麻省理工学院成立史瓦茨曼计算机学院(Schwarzman College of Computing)、加州大学伯克利分校提出包容性智能倡议(UC Berkeley Inclusive Intelligence Initiative)等。中国也不例外,在国家和政府层面,相继发布了《新一代人工智能发展规划》《中国互联网发展报告(2021)》等政策或报告,旨在加快创新型国家和世界科技强国的建设步伐;在高校层面,各高校纷纷成立与数据密集型科学密切相关的学院或研究中心,开设与数据密集型科学密切相关的专业,加快开展与数据密集型科学密切相关的人才培养和科学研究工作。

那么,什么是数据密集型科学呢?作为科学方法革命历程中的“第四范式”(the fourth paradigm)^[1],数据密集型科学是21世纪科学研究最重要的发现,其主要聚焦于数据密集系统和科学交流^[2]。随着信息技术的发展,尤其是大数据技术在科学研究中的应用和拓展,研究者不仅可以通过对广泛数据的实时动态监测与分析来应对难以解决或不可触及的科学问题,更是把数据作为科学研究的对象和工具,基于数据来思考、设计和实施科学研究。因此,数据密集型的知识发现就此诞生^[3]。可见,数据密集型科学的出现,不仅带来科研工作者对科研基础设施、科学数据管理和新型科学交流体系等需求的新变化^[4],还带来了科研范式的转变——由“科学始于观察”转变为“科学始于数据”^[5]。同时,数据密集型科学的出现使得“应用信息手段解决信息时代的各种问题”(包括人文社会科学)不仅成为“可能”,而且已经是“必须”和“不可避免”^[6]——也迫使社会科学在学科建设和研究过程中采用数据密集型科学计算的方式开展学科基本问题的研究,以及由其交叉融合所带来的学科的新研究问题、新研究内容甚至新研究领域的产生^[7]。

数据密集型科学与社会科学的相遇,已经带来情报学、图书馆学、教育学等领域学科建设方向和科学研究范式的变迁,产生了数据科学/计算科学与传统社会科学深度交叉融合后的计算社会科学^[8-12]。实践中,技术的发展对人类知识生产版图和范围带来的冲击与要求,使惯常于“精工细作”的社会科学同时在研究对象和研究方法上迎来新的挑战。面对未来发展的不确定性,以及技术发展在学科建设及其研究范式的高渗透性,高校尤其是以人文社会科学见长的高校更应该主动迎接并积极运用数据密集型科学发现的优势,充分考虑数据驱动下的知识发现对于传统社会科学学科建制(尤其是跨学科/交叉学科/超学科发展)及其研究范式的重大影响、冲击乃至颠覆作用。

作为一所以人文教育著称的世界顶尖大学,耶鲁大学的社会科学在世界排行榜上长期居于领先地位,其社会科学的发展动向和改革举措也因此备受关注^[13]。近年来,为了巩固在社会科学领域的

优势地位,耶鲁大学第二十三任校长彼得·沙洛维(Peter Salovey)将数据密集型/数据驱动(data-intensive/data-driven)的社会科学发展作为学校范围内的学术优先事项(university-wide academic priorities)之一^[14],这是耶鲁大学应对新一轮科技革命的主要举措,也在一定程度上反映了社会科学在以文科见长的世界一流大学发展中的最新变革趋势。为此,本研究以耶鲁大学作为典型案例进行深入探究,分析其数据密集型社会科学发展的愿景目标、具体举措、实践特征及其对我国社会科学转型发展的启示意义。

二、数据密集型科学发现对于社会科学发展的双重价值

科技与社会的关系向来是科技哲学热衷于讨论的关键议题之一,特别是在科技日新月异的今天,任何一项科技的突破与应用尤其是数据密集型科学发现在高等教育领域的渗透,对传统社会科学的多方面发展(包括学科组织结构、学科建制和研究范式等)均带来直接影响和全新挑战。总体而言,数据密集型科学发现及其在社会科学领域的渗透与应用,一方面会带来数据密集型科学与社会科学的交叉融合发展,推动社会科学的学科建设方向发生转变——由传统走向现代、由经验转向科学,并由此带来组织结构、学科目标和人才培养方案的调整与优化;另一方面,以数据驱动的研究范式会进一步推动社会科学研究范式的转型——由侧重人文主义转向科学主义,人文主义研究范式与科学主义研究范式的混合研究将成为社会科学领域研究的主流方式,旨在借助数据密集型科学研究范式的优势来提升社会科学研究的速度、效率和效果^[15]。按照马克斯·韦伯关于理性的二分法来解释^①,数据密集型科学发现对社会科学的学科建设及其研究范式兼具价值理性和工具理性双重作用。其中,数据密集型科学发现及其对社会科学的学科建设方向所起到的引领作用,更接近于马克斯·韦伯关于价值合理性行为(价值理性)的界定,而数据密集型科学发现及其对社会科学研究范式所赋予的新动能,则更接近于其关于工具合理性行为(工具理性)的界定。由此,借由马克斯·韦伯关于理性二分法的分析框架,可以对耶鲁大学发展数据密集型社会科学的科学性、必要性和可行性做出合理性解释。需要指出的是,尽管数据密集型科学发现在社会科学领域的渗透与融入,其本身就会对传统的社会科学的学科建设及其科学研究兼具工具理性和价值理性,但在实践中我们还需要警惕价值理性的异化和工具理性的僭越,使二者达至动态平衡,从而为社会科学的高质量发展服务。

(一)数据密集型科学发现对于社会科学发展的工具理性

从立足于存在和实践的知识论立场出发,数据密集型科学发现将为社会科学的知识生产与再生产活动提供研究理念及研究工具的支持。一方面,数据驱动的科学发现带来社会科学知识生产与再生产方式的操作技术突破。正如印刷术的突破改变了早期欧洲的社会轮廓和学术空间一样,技术革新使材料、技术、工艺等知识生产工具快速进步,而新材料、新技术、新工艺又为社会科学知识的扩充开辟了更为广阔的发展前景。数据密集型科学发现的兴起及其广泛应用正是得益于存储材料、计算机技术和硬件工艺的进步,使其被运用在社会科学海量知识的云储存、虚拟化等操作中,这不仅为社会科学领域的研究者提供基于数据驱动的更强洞察力和更精确的决策力,也有效锻炼了社会科学研

^① 马克斯·韦伯把理性分为价值理性和工具理性。价值合理性行为是对存在应该是什么的认识,立足于信念、理想的合理性,是指绝对不计后果地遵从某些价值信念而行为,即为了实现理想可以不惜一切代价,将后果与责任归咎于难以改变的客观条件、命运的安排或者邪恶势力的阻碍;工具合理性行为是对存在是什么的认识,是指以能够计算和预测后果为条件来实现目的的行为,即基于目的合理性,只对后果负责,以完全可以预计和计算的方式追求目的、手段、条件和结果相一致。同时,价值理性和工具理性并不是相互排斥的,二者是相互整合和统一的。关于马克斯·韦伯的价值理性和工具理性解读详情可参见:马克斯·韦伯. 经济与社会(上卷)[M]. 北京:商务印书馆,1997;苏国勋. 理性化及其限制:韦伯思想引论[M]. 上海:上海人民出版社,1988;王彩云,郑超. 价值理性和工具理性及其方法论意义:基于马克斯·韦伯的理性二分法[J]. 济南大学学报(社会科学版),2014,24(2):4853.

究者主动运用大数据技术、云计算技术解决社会科学研究问题的能力,即通过密集型科学发现的应用,提高研究者在单位时间内挖掘、处理海量信息资料的速度、效率和效果。实践中,大数据技术已经在古籍数字化、决策智能化等方面发挥了不可替代的作用,极大地改变了社会科学的知识存储形态和问题解决方式,推动社会科学研究范式的转型。可见,数据密集型科学在社会科学知识生产与再生产过程中以操作技术的面貌出现,为人文精神和社会秩序的合理性与科学性提供了“实验事实”“实证判据”的证明^[16],从而进一步拓展了社会科学领域知识生产与再生产的疆域。另一方面,数据密集型科学发现还带来社会科学研究者开展科学研究工作理念的转变。尽管社会科学的知识体系是以人和社会为中心的,但其创造过程却不能只依靠灵感、想象和直觉,还需要理性、秩序、逻辑、方法等“科学因素”,这是数据密集型科学发现为社会科学研究者开展基于数据驱动的科学研究工作所带来的理念创新和重要影响。社会科学研究也需要借助科学的研究工具尤其是数据驱动的研究方法,通过科学的理性去描述、解释并验证相关的社会现象和社会问题,探究其发生演变机制,促进人们对社会的科学认识,从科学的角度引发人们对某些问题的“再思考”,并在这种思考中主动运用数据密集型科学工具的优势,找寻解决现实复杂社会问题的新方向、新途径。

(二)数据密集型科学发现对于社会科学发展的价值理性

从根本上看,数据密集型科学发现不仅在社会科学研究中发挥工具价值,还能凭借其以自然科学为主体的知识体系及其研究范式的高渗透性,推动其与传统社会科学在知识层面上的交叉融合发展,推动社会科学学科建制的跨学科乃至超学科发展。所以,从这个视角来看,数据密集型科学发现对于社会科学的发展除了发挥作为知识生产与再生产载体的工具理性外,还具有在知识层面上推动与社会科学相互促进、交融与共生的价值理性^①——在很大程度上,也就发挥着引领社会科学学科建设发展方向的积极作用——集中体现为数据密集型科学能有效推动社会科学的学科发展方向由单一学科情境转向多学科交叉融合发展的应用情境和适应情境^[17],并在深度交叉融合发展的过程中找寻社会科学学科建设新的增长点和生长点,进而带来不同类型、不同层次高校社会科学学科建设的特色化和差异化发展。

另外,数据密集型科学发现对于社会科学发展的价值理性还在于推动社会科学的应用性属性与科学化水平。一方面,将数据密集型科学与社会科学的学科建设及其科学研究相结合,会进一步增强社会科学学科建设的应用趋向,推动社会科学的学科建设、人才培养及科学研究活动主动与世界科技前沿、经济主战场、国家重大需求、人民生命健康相结合^[18],实现社会科学学科建设和研究对于社会治理体系和治理能力现代化所产生的更为广阔的社会效益和社会效用。另一方面,借由数据驱动的算法优势,拓展社会科学的研究范围和科学化水平。数据密集型科学在人类社会的广泛应用不仅改变了客观物质世界的运行逻辑,同样也为社会科学研究的范式转换和人们精神世界的完善带来新的挑战。大数据、人工智能等技术的渗透与应用,迫切需要传统社会科学在伦理价值、制度规范和人际关系层面对已经出现和将要出现的相关问题予以探究和解答,这就对社会科学领域的研究者开辟了新的知识拓展空间和研究范围。与此同时,还要求社会科学领域的研究者们以积极开放的姿态主动掌握数据密集型科学发现的相关理论性知识和工具性知识,在把握其知识特性和工具优势的基础上实现对社会科学研究的科学化证明。

然而,无论是推动学科建设走向跨学科性还是提升学科建设及科学研究的科学化水平,都需要打破传统社会科学学科建设的发展理念和组织形式,通过强化制度保障与治理支持(包括重塑发展愿

^① 在强调跨学科、交叉学科乃至超学科融合发展的今天,就算是传统的人文学科和社会科学,也避免不了科学技术尤其是数据驱动对其学科建设和研究的影响。而且,在同一所大学当中,不同学科领域之间的融合尤其是人文社会科学与科学技术的融合,实际上更具有地缘优势和合作优势。

景、调整组织结构、改变组织行为、优化培养目标和课程体系等)来实现社会科学学科建制的成熟度及其科学研究的科学化水平。

三、耶鲁大学数据密集型社会科学发展的愿景目标与实践举措

(一)耶鲁大学数据密集型社会科学发展的愿景目标

2016年末,沙洛维校长将数据密集型社会科学发展确定为耶鲁大学的最高学术优先事项,并为此成立了学校数据密集型社会科学委员会(the University-wide Committee on Data-Intensive Social Sciences,简称DISSC),负责为整个学校数据密集型社会科学的发展提供政策建议和服务咨询,旨在通过发展并应用数据密集型社会科学的方法推动耶鲁大学在促进前沿知识生产、解决重要政策问题方面做出更大贡献。在上述发展理念的指导下,耶鲁大学通过发展数据密集型社会科学旨在实现以下3个方面的愿景目标。

1. 借助数据密集型科学的强大力量不断赋能并推动社会科学研究范式转型

大数据、人工智能等技术的快速发展,为社会科学研究范式的变革提供了数据基础和现实条件。在此背景下,大学需要对技术进步带来的变化进行有效回应,并重新评估数据密集型科学发展对于社会科学研究所带来的技术渗透和资源影响。所幸的是,耶鲁大学十分敏锐地意识到这种技术变革对于社会科学研究范式变革的重要意义和现实价值。例如,DISSC认为,社会科学也越来越需要用到自然科学中常用的基础设施和研究工具,而这些基础设施和研究工具往往更依赖于专业/高技能的研究人员,而不仅仅是有形的基础设施,因此耶鲁大学发展数据密集型社会科学需要建立在致力于支持数据密集型社会科学研究的大量教育资源和基础设施的基础之上,包括为此新成立的托宾经济政策中心(Tobin Center for Economic Policy)、社会政策研究中心(Institution for Social Policy Studies)、麦克米伦中心(MacMillan Center)、经济增长中心(Economic Growth Center)和考尔斯基基金会(the Cowles Foundation)等,以及用来支持数据密集型研究所必须的重要基础设施,如人口普查数据中心(FSRC)、耶鲁大学计算研究中心(YCRC)等^[19]。可见,耶鲁大学发展数据密集型社会科学的愿景之一就是希冀借由数据驱动(数据技术进步及其在社会科学领域的渗透与应用)来促进研究人员始终保持并拥有最先进的研究基础设施和工具,在培养社会科学研究人员的创造力和生产力的同时,让其以无法预测的方式(尤其是以数据驱动和多学科交叉融合的方式)推进研究、改变世界,并由此帮助耶鲁大学吸引和留住世界上最优秀的研究者,进而达至并不断拓展社会科学研究前沿、维持其社会科学世界领先地位的远景目标^[19]。

2. 通过建立社会科学研究社区促进跨学科研究与协作

随着现实问题的复杂性程度不断加剧,单一学科的知识很难有效解决现实问题,跨学科、跨领域的共同协作就变得十分重要。为了推进数据密集型社会科学的发展,耶鲁大学不仅将数据密集型社会科学置于优先发展战略的高位,还通过密切联系大学社区,促进跨部门、跨院系和跨学科的协调。为此,耶鲁大学通过建立社会科学研究社区,来促进各部门和学院之间的协调与整合,实现不同学科之间的交叉融合发展。具体而言,耶鲁大学通过积极利用学校数据密集型社会科学领域专业知识的质量和广度,支持不同院系和学科之间实现研究的密切关联,提升各个院系、学科和部门之间的开放度,助力并支持跨部门和跨学院的跨学科知识的生产。事实上,专业知识领域之间是相互联通的,不应被人为地分割,而应在大学校园内随处可见和可访问,即社会科学不仅应出现在文理学院,而且还要广泛分布于耶鲁大学的诸多学院,包括法学院、公共卫生学院和医学院等^[19]。由此可见,在跨学科发展实践中,社会科学实际上为建立不同院校和学科间的知识联系提供了绝佳的途径和契机,也为各学科之间的交叉融合、分工协作和协同创新提供了坚实的基础。

3. 有选择地改善基础服务并强化制度保障与治理支持

耶鲁大学发展数据密集型社会科学的战略规划既不是盲目的也不是冒进的,更不是采取平均主义,而是在充分调研的基础上审慎决策,遵循“有所为有所不为”的发展理念。一方面通过有选择性地优先确定一些可以加强基础服务的领域,如对社会科学研究人员的需求给予更多关注,包括IT支持、对研究人员及学生的统计咨询与培训等。耶鲁大学认为,这些基础服务一旦被大量研究人员(包括教师)和学生使用,即使是轻微地改进也会产生显著的效益。另一方面,鉴于未来需求的不确定性和不可预料性,数据密集型社会科学发展需要因时而异、适时调整,以确保耶鲁大学的教育资源与时代同步发展。为此,耶鲁大学不仅注重有效的教师投入和管理,还强调有必要在大学层面进行治理整合,如通过年度教师集体会议、各种委员会等制度设计和保障,不断完善数据密集型社会科学发展的治理支持^[19]。

(二) 耶鲁大学数据密集型社会科学发展的实践举措

为了在新一轮科技革命时代继续维持其社会科学在全球范围内的卓越领先地位,耶鲁大学将数据密集(data-intensive)确立为学校社会科学发展的新方向与新使命。在此发展愿景与目标的指导下,耶鲁大学围绕社会科学的组织结构与行为、课程与教学等领域计划或已经实施了一系列实践举措。耶鲁大学系列实践举措的颁布和实施并不是随心所欲的,而是在研究了同行机构做法(包括了解它们如何支持数据密集型社会科学的研究和教育)、征求相关领域和学校教职员工的反馈意见、以及分析学校当前的优质资源及其可以加强的领域的基础上,才审慎地颁布实施了统领学校数据密集型社会科学发展的专门报告——《学校数据密集型社会科学委员会报告》(*Report of the University-Wide Data-Intensive Social Science Committee*)(以下简称《报告》)。《报告》指出,耶鲁大学在发展数据密集型社会科学的过程中需要在组织结构与机构整合、课程与教学项目等方面进行一系列改革与探索^[20],集中体现在以下两个方面:一是通过成立校级信息技术中心、建立校级委员会、定期举办面向全校教师的会议和论坛、对大学组织机构及其行为进行调整和优化等方式,促进社会科学和数据科学的学科交叉与知识交融;二是通过重塑人才培养目标和课程结构体系,将与科学技术原理及应用相关的知识纳入通识教育模块传授给学生,使他们无论在知识上、思维上,还是实际应用技能上,都能为更进一步深入研究做好充分准备,由此也可以看出耶鲁大学数据密集型科学与社会科学学科建设及研究之间的深刻关联。

1. 重塑组织结构,调整组织行为,促进跨学科研究的整合效应

知识变革带动组织结构转型和组织行为调整,以数据源激增、计算能力增强和新分析技术为特征的数据驱动下的技术发展及应用正在改变社会科学的研究方式,社会科学也因此越来越需要类似于自然科学研究中常用的核心设施、最先进的物理设施以及人力基础设施,以实现借由基础设施来扩展其学科研究前沿的目的。为此,耶鲁大学在社会科学现有优势地位的基础上确立了向数据密集型转型的发展方向。相应的,学校内部的组织结构及其行为也需要重组和调整,以满足数据密集型社会科学对核心设施与技术的现实需求。

首先,通过创建数据密集型社会科学中心,为全校师生提供公共社区。耶鲁大学正在筹建一个能够在全校范围内为研究者提供基础设施并使他们能够安全获取和使用新数据资源、了解数据密集型社会科学可用的事件、信息和资源的数据密集型社会科学中心,进而为所有在数据密集型社会科学的不同方法及应用方面有所研究和关注的师生建立起公共社区。为此,《报告》确定了数据密集型社会科学中心的6项职能:第一,通过与专家团队共同建立安全的数据库资源,以获取、计算和存储那些关键却未能有效使用的数据,使社会科学领域的研究人员可以基于解决重要政策问题而取得领先于同行的成果;第二,提供大量专项基金,以促进跨部门、跨院系、跨学科的联系和研究协作,为人员交流、

数据获取、硬件使用或其他项目所需费用提供资金支持;第三,提供研究咨询和数据科学服务,为学生和教师在项目规划、编程指导、研究设计和分析工具咨询等方面提供不同程度的支持;第四,将外联和门户网站作为一个信息库,向社会科学研究人员介绍耶鲁大学的数据科学资源,引导他们通过校园网访问并获取他们所需的各类资源;第五,通过研讨会、会议等形式,访问学者和研究人员可以围绕数据、方法以及数据密集型社会科学的创新研究设计及应用建立社区并促进智力交流与合作;第六,信息技术团队对数据和计算密集型社会科学的信息技术提供支持与服务^[21]。总之,该中心是耶鲁大学为数据密集型社会科学所做的影响最大且最直观的改革探索与实践。

其次,通过了解其他大学和行业卓越机构的发展趋势,适时调整组织目标和组织行为。为了使学校社会科学领域的教师与学生有能力超越他们所在学科的研究前沿、在软件开发和硬件建设方面及时赶上科技创新发展浪潮,耶鲁大学通过与其他大学和行业卓越机构的互动与交流,为不同社会科学领域的学科建设和知识生产创造机会。《报告》认为,耶鲁大学有必要关注并了解其他世界顶尖大学数据密集型社会科学的建设动向与实践举措,从他人的经验与教训中习得有助于耶鲁大学正在实施的政策和项目,以更好地促进由耶鲁大学院校研究和战略分析办公室(Yale's Office of Strategic Analysis and Institutional Research)来负责推进的数据密集型社会科学的研究和教学^[21]。该办公室通过每年发布包括其他顶尖大学的数据密集型社会科学研究和教学进展及其重要基础设施投资与数据政策变化,以及同行机构正在启动或停止的关键项目内容的“创新和经验教训”报告,实现对全球数据密集型社会科学研究领域前沿进展的了解。与此同时,“创新和经验教训”报告还会发给耶鲁大学各二级学院院长、研究中心主任和任课教师,以便他们基于每年的报告进一步调整发展目标和研究方向,并根据新的发展目标和愿景,采取更有针对性和更有效益的行动方案。

再次,通过在大学层面建立委员会,整合并提升学校各部门、院系和学科间的协同效率。数据密集型社会科学研究所需的基础设施需求表明,跨学科和跨院系的合作将持续增长。基于这一判断,耶鲁大学通过建立稳定的运作机制,以促进整个大学范围内涉及数据密集型社会科学的各个研究中心、部门和院系之间的有效沟通和协调。为此,耶鲁大学建立了一个每年召开两次会议的委员会,委员会成员由社会科学中心主任、系主任和其他与数据密集型社会科学研究相关的学术领袖组成,成员规模控制在15~20人。这种制度化的会议使在这一领域负有重大责任的院系领导者以及重要学者聚集在一起,分享关于数据密集型社会科学的关键信息和发展趋势,通过整合并协调各部门、院系及学科之间的发展计划,为学校及其相关部门的科学决策提供决策与服务咨询。

2. 优化培养目标,开发课程项目,强化制度保障和治理支持

耶鲁大学认为,随着新技能和新工具的开发,以及数据密集型社会科学研究设计的改进,耶鲁大学数据密集型社会科学的人才培养目标也应该随着新技能和新工具以及数据密集型社会科学研究设计的变化而加以动态调整和优化,以帮助和引导耶鲁大学所有学生养成识别经验推理的优势和缺点的能力,提出关于经验主张的探索性问题,并在形成意见和决策时更加明智地使用定量证据。为此,耶鲁大学必须为数据密集型社会科学所需要的关键思想和方法提供容易理解且规范、系统的基础课程,为学生提供能够获得学科前沿知识的方法和途径,包括那些在耶鲁大学接受教育但还没有大量接触过统计数据、编程或高级数学的学生,定期更新教学和学习目标,并通过沟通、协调、创新和评估等方式,科学评估学生的理解水平、掌握程度以及学业进展情况。

第一,加强基础数据密集型课程群建设。课程是育人的载体。耶鲁大学为了实现其数据密集型社会科学人才培养的目标,在课程建设方面采取了一系列措施,具体包括:(1)在耶鲁大学课程体系中建立数据密集型课程模块,包括关于研究设计的基本见解、关于定量证据的推理,以及在信念形成与决策中使用定量证据等,真正为学生提供多种数据分析的应用机会。(2)在数据密集型社会科学

领域开设“标志性”跨学科讲座课程,以吸引各个学科背景的本科生选课,使之初步接触严格的实证科学调查、研究设计原则、证据质量评估、使用数据决策等“数据导向”的课程内容^[21]。(3)开发耶鲁数据(Yale data)系列课程。这一系列课程以加州大学伯克利分校流行的数据课程为模型(于2019年春季在耶鲁大学面向非统计学、数学或计算机科学背景的学生正式推出),旨在形成一门广泛涉及人文学科、社会科学和自然科学的入门级数据科学课程^[22]。(4)设立“通往量化社会科学高级成就”的专项计划,为本科一年级和二年级学生建立一条“通往定量社会科学高级成就的途径”(a pathway to advanced achievement in quantitative social science),使他们学习一系列旨在为未来从事定量社会科学专业研究做好准备的课程。该专项计划的具体内容包括让学生广泛而准确地介绍定量社会科学中的工具和应用,以及让本科生设计和执行优秀的高级项目等,从而为他们进入研究生院或与教授一起从事前沿项目研究奠定基础^[23]。

第二,增强数据密集型社会科学发展的多样性和包容性。这主要体现在托宾经济政策中心(Tobin Center for Economic Policy)的博士预科项目(pre-doctoral program)实施范围的不断拓展。经济学是开展数据密集型社会科学的主要社会科学领域之一。目前,托宾经济政策中心的经济学博士预科项目专门用来支持与政策相关的经济学研究,旨在为考虑攻读经济学或相关学科博士学位的学士或硕士提供高质量教育培训的同时,为导师做一至两年的全职研究助理^[24]。《报告》进一步建议,将该博士预科项目扩展到其他学科,使参与该项目的研究人员有机会与经济学以外的其他数据密集型社会科学领域(如心理学、政治学、语言学等)的教师合作,从而促进耶鲁大学数据密集型社会科学学科发展的多样性和包容性^[21]。

第三,通过教师年度会议和各种委员会等制度保障,推动课程的创新发展和科学评估。一方面,耶鲁大学开设了许多统计学、方法论方面的入门课程,这些课程往往涵盖重叠的统计概念与技术——仅根据学科重点(如经济、政治)而有所侧重,并没有确保教师之间协调整合的制度保障。为此,耶鲁大学每年组织所有开设统计学和研究设计入门课程的教师召开一次集体会议,旨在为该课程的内容调整、授课技术、合作教学以及经验交流提供制度保障,为更好地满足学生的发展需求、提升跨学科人员的合作效率,以及创新教学理念提供更好的支持。另一方面,耶鲁大学通过设立课程委员会来评估学生的定量思维和能力。从2009届开始,耶鲁学院(耶鲁大学的本科生院)的学生被要求在人文、自然科学和社会科学3个学科领域中选择两门课程,并满足外语、写作(两门课程)和定量推理(两门课程)的技能要求^[25],但其中许多课程本质上并不是数据密集型科学领域的。鉴于计算科学和数据密集型科学的分析越来越重要,耶鲁大学本科生学院院长通过在学校层面组建一个委员会来评估学生如何才能获得足够的定量推理(quantitative reasoning)知识基础和水平。然而,该委员会的课程评估标准和基线水平并不是自我设定的,而是建立在研究和学习其他同行机构(如哈佛大学^①)评估本科生在接触数据分析和统计推理要求的基础上,实现课程的创新发展和科学评估。可见,耶鲁大学对数据密集型社会科学的课程教学效果实施的是定期的动态评估,即通过不断地沟通、协调、创新等活动,加强对学生对数据密集型科学知识的理解水平、掌握程度以及学业进展情况的过程评估和持续评估。

由以上的分析可知,耶鲁大学在数据密集型社会科学学科建设及研究方面至少有以下两个方面的贡献:一是推动数据科学与社会科学的知识交叉融合发展;二是加速数据科学在社会科学研究中的

^① 例如,哈佛大学近期就实施了一种基于数据(quantitative reasoning with data, QRD)要求的定量推理,取代过去经验推理和数学推理(empirical and mathematical reasoning)的要求。新的QRD要求的目的是确保本科生达到一个涉及数学、统计和计算方法的定量基础水平,使他们能够批判性地思考数据,因为它要应用于文理学院的所有调查领域。详情请参见 <https://www.harvardmagazine.com/2019/04/harvard-course-requirements-quantitative-reasoning>。

实践应用。

四、对我国数据密集型社会科学发展的启示

习近平总书记强调:“坚持和发展中国特色社会主义,必须高度重视哲学社会科学,结合中国特色社会主义伟大实践,加快构建中国特色哲学社会科学。”^[26]繁荣发展具有中国特色的哲学社会科学,必须坚持扎根中国、关怀人类、立足当代、面向未来的发展思路,紧跟时代潮流,同时发挥好科学技术特别是大数据、人工智能等数据科学在社会科学研究中的工具理性和价值理性。因此,结合数据密集型社会科学发展的耶鲁大学实践,我们认为,在中国高校开展数据密集型社会科学的学科建设及科学研究活动的实践中,以下4点经验值得关注和借鉴:

(一)转变社会科学发展的传统理念,主动运用数据密集型科学赋能社会科学转型发展

不可否认,数据密集型科学作为应用于社会科学研究中的操作技术,不仅能够帮助研究者掌握前沿、系统的数据与信息,保持在同行中领先的研究能力,还能使其产出的研究成果在规模、质量等方面始终处于该领域的先进水平。与此同时,基于数据驱动的发展与应用,在实践中还能有效促进科学主义研究范式在社会科学研究中的应用,拓展社会科学进行知识生产与再生产的范围和维度,推动科学主义研究范式与人文主义研究范式的深度融合。从耶鲁大学发展数据密集型社会科学的实践探索中可以看出,对数据科学工具理性的充分利用不仅能够加快社会科学研究转型发展,帮助社会科学领域的研究者抢占学科前沿成果产出的先机,也将在思维逻辑和研究设计上使研究者更加理性、更具创新性。只有掌握先机、抢占先机,才能为学科的持续发展和特色发展提供源源不断的动力支持,才能不断提升中国特色社会科学的现实影响力和国家话语权。因此,必须转变传统社会科学的学科建设理念,突破单一研究范式的限制,主动将数据科学思维纳入社会科学学科建设及其科学研究过程中,借助数据技术的强大力量来实现社会科学研究范式的科学化转型,使科学化的研究工具也能为社会科学研究所用,提高社会科学研究效率。例如,在人才培养环节,可以加强对数据密集型科学基础知识的系统化设计与安排,面向全校本科生开设数据密集型科学必修课程;在教学环节,着重提升学生对数据密集型科学知识的理解水平和应用能力;在知识生产环节,设置数据密集型科学方法研究中心,为从事社会科学研究教师们提供培训支持和咨询服务。

(二)加强社会科学与数据密集型科学之间的深度融合发展

学科是人为划分的知识体系,但知识本身对于客观世界而言是完整统一的,人类探索世界、认识世界的重要机会和主要场所往往集中在不同学科的交叉融合之处^[27]。换言之,不同学科的研究者从自身专长出发认识和关照其他学科,都能有效地拓展其所在学科领域的研究范畴,催生新的研究领域。然而,当下的学科建设往往将学科交叉融理解为学科合作,很多院校的做法是将不同学科的知识进行生产线式简单对接和组合,而没有真正从多学科的融合视角对现实复杂问题进行通盘考虑,从而导致知识更新的速度难以匹配社会发展的需求,新的知识体系的形成依然任重道远^[28]。借鉴耶鲁大学的做法,建议在大学层面为学校数据密集型社会科学的发展设立专门的数据研究与应用中心,成立校级学术委员会,并在此基础上以固定周期为时间单位通过举办制度化的会议和论坛,供不同学科领域的教师和学生就与科学技术、社会科学等相关的问题展开理论层面的深入交流和探讨,推动数据密集型科学与社会科学的知识交叉融合。此外,还应坚持问题导向,不断消除学科间的隔阂,加强学科间、院系间不同研究领域的合作,并通过治理机制的整合创新促进跨学科研究协同效用的最大化发挥。

(三)建立与数据密集型社会科学相匹配的人才培养目标与跨学科课程体系

在数据密集型科学赋能背景下,社会科学人才培养的目标必须与其学科建设的未来方向及其时

代使命相契合。结合社会发展需求和技术变革要求,适时调整社会科学人才培养目标和课程结构体系,比如在人才培养方案中主动融入并加强对数据密集型科学思维、原理、技能和方法的训练,打破传统社会科学的人才培养模式只局限于以单一学科专业为支点的封闭“小圈子”,以面向未来的前瞻性思维在更为宽广的学科维度和知识体系中培养人才^[29]。具体而言,就是要树立学生中心、问题导向、交叉复合、创新驱动的人才培养新理念和跨学科课程教学新实践。比如,可以将数学、统计学、信息技术等支撑大数据技术理解与应用的基础学科知识作为固定内容纳入本科生的通识教育体系,加强对学生的逻辑思维锻炼,提高学生的数据信息素养,进而提升学生对于未来工作世界的市场竞争力和环境适应性。通过加大对跨学科课程模块的设置、提高实践教学的比重、加强文科实验室建设等实践,加强课程资源平台的体系建设,提高服务师生需求的现实能力,为师生的个性发展和特色发展服务。此外,大学还有责任在软件和硬件上为教师和学生提供必要乃至卓越的支持,使数据密集型科学发现的工具价值在研究者、学习者的使用中得到正视和发挥。人才培养不能仅局限于传授学生以知识,更为重要的是要培养学生理解复杂环境和创造新知识的能力。事实上,思维和方法的训练远比知识传授本身重要。

(四)注重数据密集型社会科学发展的整体规划与动态调整

数据密集型科学的发展虽然为传统社会科学的发展注入了新动能,但技术的变革依然存在很大的不确定性,未来发展需求也存在很大的不可预料性。为此,在我国社会科学转型发展过程中,不能眉毛胡子一把抓,盲目扩张,而应坚持有所为有所不为的建设原则,根据中国特色哲学社会科学体系建设的战略目标来决定哪些领域应当作为优先支持事项,哪些领域需要根据未来发展变化做出适时调整。同时,在此基础上还应不断提炼中国社会科学发展的经验和做法,使之形成理论、形成话语、形成体系,建构属于我国数据密集型社会科学发展的学术体系、话语体系和学科体系,形成中国特色哲学社会科学的自主知识体系,提升中国文化的影响力、感召力和塑造力。事实上,社会科学和数据密集型科学的融合发展本就是一个动态调整的过程,需要综合考量技术发展、国家战略意图、社会发展需求、学科传统优势等多方面的因素。对于大学而言,就是要主动融合学校办学历史传统、区域资源优势、学科办学水平等,不断提炼数据密集型社会科学发展的特色领域,走差异化、特色化的学科建设与研究之路。

参考文献:

- [1] HEY T, TANSLEY S, TOLLE K. The fourth paradigm: data-intensive scientific discovery [M]. Redmond, Washington: Microsoft Research, 2009:413.
- [2] 郭娇,吴寒天.“大数据时代”的因果推断:教育政策评估的新路径[J].重庆高教研究,2022,10(4):3948.
- [3] HEY T, TANSLEY S, TOLLE K. 第四范式:数据密集型科学发现[M].潘教峰,张晓林,等译.北京:科学出版社,2012:译者的话.
- [4] 黄鑫,邓仲华.数据密集型科学研究的需求分析与保障[J].情报理论与实践,2017,40(2):6670,79.
- [5] 黄欣荣.数据密集型科学发现及其哲学问题[J].自然辩证法研究,2015,31(11):4854.
- [6] 汤建民.情报计量学视野下的当代人文社科发展背景描绘:一个人文社科领域数据密集型科学发现的探索性案例[J].情报资料工作,2016(1):1721.
- [7] 周晓英.情报学进展系列论文之七 数据密集型科学研究范式的兴起与情报学的应对[J].情报资料工作,2012(2):511.
- [8] 邓胜利,付少雄.计算科学与信息科学融合下的图书情报学科建设:基于德雷塞尔大学计算与信息学院实践[J].情报资料工作,2019,40(3):8087.
- [9] 刘艳红,罗健.数据密集型科学环境下的情报服务与发展[J].图书与情报,2013(6):105-108.
- [10] 郑永和,严晓梅,王晶莹,等.计算教育论纲:立场、范式与体系[J].华东师范大学学报(教育科学版),2020,38(6):419.

- [11] 王战军, 乔刚. 大数据驱动的教育研究新范式[J]. 北京大学教育评论, 2018, 16(1): 179-185.
- [12] 张婧婧, 封晨. 多学科视域下教育密集型研究的机遇与挑战: 基于美国《数据密集型教育研究》的解读[J]. 复旦教育论坛, 2017, 15(5): 99-105.
- [13] 公钦正, 周光礼, 张海生. 知识转型与组织应对: 文科见长大学改进发展研究——以耶鲁大学为例[J]. 现代大学教育, 2022, 38(6): 6574, 113.
- [14] SALOVEY P. Fall Update on Academic Strategy [EB/OL]. (2017-09-15) [2023-05-03]. <https://provost.yale.edu/sites/default/files/files/Fall%20Update%20on%20Academic%20Strategy.pdf>.
- [15] 张海生. 人工智能赋能学科建设: 解释模型与逻辑解构[J]. 高校教育管理, 2023, 17(3): 4250, 75.
- [16] 陈勇. 科学精神与人文精神关系探析[J]. 自然辩证法研究, 1997(1): 2328.
- [17] 张海生. 知识生产模式转型与新工科建设的基本策略[J]. 中国高校科技, 2018(9): 4751.
- [18] 李立国, 张海生. 学科治理的三重逻辑及其互动机制[J]. 教育发展研究, 2022, 42(3): 4450.
- [19] Yale University. University-Wide Academic [EB/OL]. [2023-05-02]. <https://provost.yale.edu/university-wide-academic-priorities>.
- [20] DISSC. Data-Intensive Social Science Task Force Recommendations [EB/OL]. (2021-02-18) [2023-05-06]. <https://provost.yale.edu/sites/default/files/files/data%20intensive%20social%20science%20center%20task%20force%20report.pdf>.
- [21] DISSC. Yale University Report of the University-Wide Data-Intensive Social Science Committee [EB/OL]. [2023-05-04]. https://provost.yale.edu/sites/default/files/files/DISSC%20Final%20Report_2_7_20.pdf.
- [22] Yale News. Why take YData? Because data science shouldn't be a 'black box' [EB/OL]. (2018-12-10) [2023-02-04]. <https://news.yale.edu/2018/12/10/why-take-ydata-because-data-science-shouldnt-be-black-box>.
- [23] Yale College. Yale College Programs of Study Fall and Spring Terms 2020-2021 [EB/OL]. (2020-06-15) [2023-02-04]. <http://catalog.yale.edu/pdf/ycps-2020-2021.pdf>.
- [24] Tobin Center for Economic Policy. Pre-Doctoral Fellows [EB/OL]. [2023-02-04]. <https://tobin.yale.edu/people/pre-doctoral-fellows>.
- [25] Yale College. Qr Courses [EB/OL]. [2022-05-05]. <https://science.yalecollege.yale.edu/academics/faculty-resources/qr-courses-without-prerequisite/qr-courses>.
- [26] 习近平. 在哲学社会科学工作座谈会上的讲话 [N]. 人民日报, 2016-05-19(02).
- [27] 公钦正, 张海生. 新文科人才培养何以实现: 基于中英两所大学 PPE 专业人才培养方案的案例研究 [J]. 重庆高教研究, 2022, 10(2): 87-98.
- [28] 王冬冬. 在守正创新中推进新文科建设 [N]. 中国社会科学报, 2022-06-07(08).
- [29] 李志生. 新技术革命背景下高质量文科人才培养的思考 [J]. 新文科教育研究, 2022(1): 2232, 144-142.

(编辑: 杨慷慨 校对: 张海生)

Experiences and Enlightenment from the Development of Data-Intensive Social Sciences: An Analysis of an International Case

ZHANG Haisheng¹, GONG Qinzheng²

(1. *Journal of Chongqing University of Arts and Sciences, Chongqing 402160, China;*

2. *School of Education, Renmin University of China, Beijing 100872, China*)

Abstract: Faced with the uncertainty of future development and the high permeability of technological development in disciplinary construction and research paradigms, universities, especially those specializing in humanities and social sciences, should actively embrace and utilize the advantages of data-intensive scientific discovery, fully considering the significant influence, impact, and even subversive effect of data-driven knowledge discovery on the traditional social science disciplinary structure and research paradigms. The discovery of data-intensive science and its penetration and application in the field of social sciences will, on the

one hand, bring about the cross integration and development of data-intensive science and social sciences, promoting the disciplinary construction direction of social sciences from traditional to modern and from experience to science; on the other hand, the advantage of data-driven research based on data intensive scientific discovery can effectively improve the speed, efficiency and effect of social science research, further promote the transformation of social science research paradigm, and the hybrid research combining humanism and scientism will become the mainstream paradigm of social science research. Taking Yale University as a research case, the study focuses on analyzing its vision, goals, and practical measures for developing data-intensive social sciences. Research has found that Yale University through the development and application of data-intensive social sciences, aims to continuously empower and promote the transformation of social science research paradigms through the powerful power of data-intensive sciences, and promote interdisciplinary research and collaboration through the establishment of social science research communities. The vision goal is to selectively improve basic services and strengthen institutional guarantees and governance support in three dimensions. In order to achieve the vision goals, Yale University has achieved the integration effect of reshaping the organizational structure, adjusting organizational behavior and promoting interdisciplinary research by creating a data intensive social science center, understanding the development trend of other universities and industry excellence institutions, establishing committees at the university level and other practical measures; by strengthening the construction of basic data-intensive curriculum groups, enhancing the diversity and inclusiveness of data-intensive social science development, and ensuring institutional guarantees and governance support such as annual teacher meetings and committees, the talent cultivation goals of data-intensive social science should be continuously optimized, promoting innovative development and effective evaluation of curriculum teaching. As a result, Yale University not only promotes the cross integration and development of knowledge between data-intensive sciences and social sciences, but also accelerates the application of data-intensive sciences in social science research. In the future development practice of data-intensive social sciences in Chinese universities, it is necessary to actively transform the traditional concept of social science development and use data-intensive scientific discoveries to empower the transformation and development of social sciences. It is necessary not only to strengthen the deep integration and development between social sciences and data-intensive sciences, but also to establish talent cultivation goals and interdisciplinary curriculum systems that match the development of data-intensive social sciences, also to pay attention to the overall planning and timely dynamic adjustment of the development of data-intensive social sciences.

Key words: data-intensive social sciences; discipline construction; scientific research paradigm; interdisciplinary study