

大数据与舆论研究的“知识图景”： 基于“人与技术”分析框架

张 宁 唐嘉仪

摘 要:“大数据与舆论”的研究本质上就是围绕“人”(舆论)与“技术”(大数据)的关系研究。为提供一个兼具现实意义与学术价值,同时又能促进国内外学术对话的分析框架,文章提出“人与技术”的分析框架,围绕这一框架探讨国内外学者在2013—2018年间关于大数据与舆论研究的“知识图景”,对文献进行内容分析和计量分析。研究发现:从“人-技术”的视角出发,在国内的学术研究中,大数据与舆论研究重点体现为“技术对人的影响”,其中“舆论引导”和“舆论挖掘”是学术话语建构的两个重点议题。而在国外的学术研究中,学者们则更关注大数据对“舆论挖掘”和“舆论预测”的作用。研究提出,大数据与舆论研究应该坚持“以人为本”的根本导向,切实提高大数据在舆论研究中的价值和作用。

关键词:大数据;舆论;人与技术;舆论引导;舆论监测

中图分类号:G206 **文献标识码:**A **文章编号:**2096-5443(2020)03-0087-13

项目基金:教育部哲学社会科学研究重大课题攻关项目(16JZD006);广东省广州市大数据与公共传播人文社会科学重点研究基地、中山大学互联网与治理中心研究成果

一、问题意识与研究设计

(一)问题意识与研究缘起

舆论的概念最早可追溯到洛克所著的 *An Essay Concerning Human Understanding* 一书。李普曼在《舆论学》中称舆论是“他人脑海中的图像——关于自身、关于别人、关于他们的需求、意图和人际关系的图像”^[1]。陈力丹结合中国实际,对舆论进行如下定义:“舆论是公众关于现实社会以及社会中的各种现象、问题所表达的信念、态度、意见和情绪表现的总和。具有相对的一致性、强烈程度和持续性,对社会发展及有关事态的进程产生影响。其中混杂着理智和非理智的成分”^[2]。

技术的更新换代在不断改变舆论的环境和形态,同时也加大了舆论研究的难度。首先,互联网作为民间公共领域的重要组成部分,在形成公共舆论方面的作用和地位日益突出。社会化媒体所构建的舆论空间表现出“复合式场域的特征”^[3]。有研究者指出,“利用 Twitter 等社交媒体来搜集信息并作为新闻来制造舆论存在着信息真实性风险”^[4],网络舆论监督借助于网络平台的交互性使分散的舆论产生共振合力,强大舆论压力倒逼政府改革现象时有发生^[5]。其次,诸如个性化新闻系统等算法推荐技术的普及可能使公民无法接触到必要的信息从而做出明智的行动和决策。更有甚者认为,新媒体环境下的社会互动过程,将导致以心理群体或舆论群体形式存在的社会公众出现意见极化^[6]。网络的去个人化特质有利于群体认同的形成,扩大群间的差异,使得网络舆论容易朝原有偏向的方向继续移动,最终呈现出极端的观点^[7]。第三,过去,对舆论和图像的信息和表达方式加以控制,一直是国家权力的维系工具,“在大众传媒时代尤其完善无比”^[8],然而随着社交媒体等现代传播技术的兴起与发展,舆论监测和管理的难度在不断加大。舆论的形式当前也超越了传统的主要以文

字为载体的存在方式,近年来网络“表情包”的出现和风靡使得越来越多网民通过这种非文字的表达方式进行意见表达,而诸如此类的“通过对娱乐符号加工的方式,使得当下的公众重新掌握了一种新型的话语权,并以此形成舆论”^[9],在权力场域的争夺中获取一席之地,也使得舆论监测和挖掘的难度大大增加。

正如卡斯特所宣称的,“技术就是社会,若无技术工具,社会也无法被了解或再现”^[10]。大数据作为一项技术,其开发和应用为舆论研究带来了全新的思路和可能性。卡斯特也曾预言,“计算机的功能之一,是为了某种特殊的目的而搜集、加工和使用大量的个人信息”^[8]。据麦肯锡全球研究所称,“大数据是指数据集的大小超出典型数据库软件工具捕获、存储、管理和分析的能力”^[11]。大数据在信息挖掘、情感分析、语义分析、实时追踪等方面均有独特的属性与应用潜力,大数据的视野与思维使得舆论研究有了新的范式和思路。如何引入大数据思维来构建现代舆论研究体系,提升舆论研究的质量与水平,拓展舆论研究的深度和广度,是业界和学界共同关注的焦点议题。

有学者指出,“公众是舆论的主体”^[12],要想研究舆论,离不开对舆论主体的研究。在大数据时代,“公众既是大数据资料库信息的主要来源,同时也是数据库的使用主体”^[13]。曼海姆相信,技术的方法一旦被运用到人类社会,则会“引起人本身的深刻变化”^[14],人与技术的关系随着技术的发展不断地被形塑和改变。正如艾卢尔所说,“技术是被人们从后面推动的”^[15],笔者认为,作为技术的一种,大数据的出现也改变了舆论研究的范式,而“大数据与舆论”研究在本质上就是围绕“人”(舆论)与“技术”(大数据)的结构关系研究。

从研究现状来看,尽管已有学者对大数据应用于舆论的研究现状进行梳理,但这些梳理大多数是综述类的回顾^[16],或是对研究现状进行宏观的陈述^[17],尚未有研究通过文献计量的方法对大数据与舆论研究的具体情况进行分析,而且分析思路较为零散,没有提供较为清晰的分析框架。有学者认为,在技术与人的互动关系中,最基本的关系问题是“控制与被控制的问题”^[18]。我们认为,“大数据与舆论”研究从本质上看是关于“人和技术”关系的反思。

为提供一个兼具现实意义与学术价值,同时又能促进国内外学术对话的研究框架,本文创造性地提出“人-技术”的分析框架,围绕此研究框架,重点关注在现有的大数据与舆论研究的学术话语中,人和技术的结构关系如何被认知和呈现。梳理学者在现有的文献中如何讨论和认知人与技术的关系,一方面可以帮助我们了解目前这一重要的结构关系在学术话语中如何被呈现和理解——了解“大数据时代,技术如何影响人和舆论”“人是否会在大数据时代对技术产生反作用”等问题的思考在学术研究中如何被解答;另一方面可以发现现有关于大数据与舆论研究的不足与限制,寻找人与技术的互动关系中尚未被学术研究深入讨论的重要问题,为后续的研究提供有价值的反思。

通过文献内容分析和文献计量分析的方法,本文以国外 SSCI 核心数据库和国内 CSSCI 核心数据库中的相关文献为分析对象,紧扣“人-技术”的分析框架,从结构特征和焦点叙事两方面对国内外大数据与舆论研究的进展进行梳理,探寻在现有的大数据与舆论学术话语中,学者们围绕人与技术的关系提出哪些学术主张。本次研究同时涵盖国外 SSCI 数据库和国内 CSSCI 数据库文献样本,既可以同时了解不同地区学者的研究进展和特点,同时可以进行横向对比,直观地比较国内外学者在这一领域的研究差异,并在此基础上对现有的研究进行述评,为国内学者未来开展相关领域的研究提出展望。

(二) 研究方法 with 编码说明

本次研究采用文献计量和内容分析的综合方法,不拘泥于传统的内容编码手段,从“人-技术”的框架出发,对 2013—2018 年间国内外大数据与舆论研究的焦点议题进行梳理,并在相关计算机软件(统计软件 R 和社会网络分析软件 Ucinet)的辅助下对样本文献的标题和摘要高频词进行统计,以及对关键词之间的共现关系进行分析。

在文献样本方面,笔者以“big data”和“public opinion”为关键词在 Web of Science 数据库中进行

主题检索,同时以“大数据”和“舆论”为关键词在中国知网数据库进行主题检索。选择以“主题”进行检索的原因主要是用“关键词”和“标题”进行检索的覆盖率过低,可能会遗漏大量文献,而以“全文”进行检索则发现检索出无关文献过多,故综合考虑下选择以“主题”作为检索条件。喻国明等人在2013年提出,“借助大数据,舆情研究的视角将更加多元化”^[19],为验证这一研究设想,以及了解经过五年发展后大数据与舆论研究的整体态势,本次研究将样本文献检索范围限定在2013—2018年间,检索时间为2019年3月,共得到相关英文文献(SSCI核心数据库)82篇,中文核心文献(CSSCI核心数据库)110篇,剔除无关文献后,最终样本文献量为182篇,其中英文文献75篇,中文文献107篇。需要说明的是,本次研究从“人-技术”的分析框架出发,探寻以“大数据和舆论”作为研究焦点的文献如何对人和技术的结构关系进行定义和描述,因此分析样本中不仅涵盖了使用大数据技术作为舆论研究方法的文献,也包括了以其他研究方法讨论大数据和舆论研究关系的文献材料。

本次研究的“知识图景”分析框架主要分为结构特征和焦点叙事两方面,分析编码设计如下:

表1 文献研究编码表

知识图景	细分内容	编码设计
结构特征	人与技术关系认识 ^a	(1)技术可以对公众舆论进行挖掘 (2)大数据技术可以引导公众舆论 (3)大数据改变和影响公众舆论 (4)技术与公众舆论相互作用 (5)大数据可以预测人的舆论和行为 (6)其他
	关键词共现	借助社会网络分析软件 Ucinet 完成
焦点叙事	研究焦点 ^b	(1)大数据与社会治理 (2)公众舆论与舆论引导 (3)大数据在舆论研究中的应用 (4)大数据与舆论监测 (5)大数据与传播改革 (6)大数据与公共传播 (7)大数据引起的问题 (8)其他议题
	标题、摘要高频词	借助统计软件 R 编程完成

^a 由于一篇文献所提到的“人-技术”关系可能不止一个,故在实际编码中“人与技术的关系”的总体数量可能高于文献总数;^b 由于一篇文献所关注的核心问题可能不止一个,故在实际编码中焦点话题的总体数量可能高于文献总数。

为了检验编码条目的信度,本研究随机抽取了中英文文献各20篇,由两位编码员独立完成文献编码,并通过内容分析相互同意及信度检验公式进行测量(Holsti, 1969)。经编码检验,英文文献信度检测结果为:焦点话题82%，“人-技术的关系”81%;中文文献中信度检测结果为:焦点话题84%，“人-技术的关系”80%。对未一致的编码部分,两位编码员进行二次讨论和再编码,以保证编码的一致性和统一性。根据社会研究方法的要求,一般来说内容分析的信度达到80%就符合要求,由检验结果可得,本研究在类目分类及指标设计,以及具体的编码上都比较符合信度标准,具有较高的可信度。

二、知识图景：大数据与舆论研究的进展描述

(一) 大数据与舆论研究的结构特征

1. “人与技术”：预测还是引导？

在大数据的介入下,舆论研究中人与技术的关系被重新建构和理解。通过文献内容分析,我们发现在国内外大数据与舆论研究中,人与技术的关系可以被大致归纳为六大面向:“技术可以对公众舆论进行挖掘”“大数据技术可以引导公众舆论”“大数据改变和影响公众舆论”“技术与公众舆论相互作用”“大数据可以预测人的舆论和行为”和“其他”。人与技术不同的结构关系在 SSCI 和 CSSCI 文献中的分布情况如图 1 所示。具体来看,从“人-技术”的结构关系视角出发,国内学术研究更强调大数据对公众舆论的引导和挖掘作用,而国外学者则更多地从大数据对公众舆论的挖掘和预测角度切入:

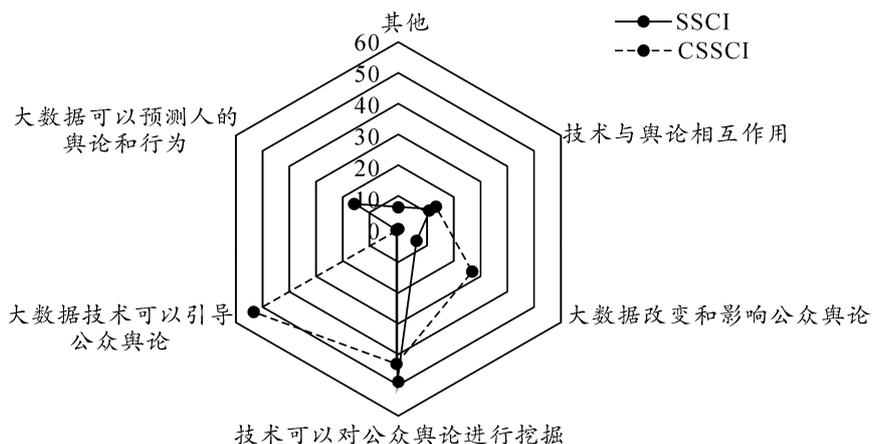


图1 国内外学术研究中“人与技术”结构关系的分布对比

在国内大数据与舆论研究话语中,学者们对人与技术的关系建构主要体现为“技术对人的影响”,具体细分成两个重点——“舆论引导”和“舆论挖掘”。在“舆论引导”中,学术研究重点讨论了两大关键点:其一,在社交媒体传播时代,信息量呈现指数级增长,因此也容易造成舆论风险,尤其是意识形态方面的建设工作难度加大,大数据这项技术被认为应该充分应用在意识形态教育^[20-22]、文化思想塑造^[23-24]、社会动员^[17]等领域;其二,在风险社会中,谣言管理问题的重要性凸显,大数据技术可以帮助政府和社会管理者实现高效的谣言治理和舆情回应^[25-27],提高风险事件中的舆论引导能力,并为政府传播带来更多有效引导舆论的传播手段和渠道。而在“舆论挖掘”中,“人”被理解成可以通过人工智能、算法等大数据工具识别、分析和预测的对象。有学者提到,大数据时代数据使用的关键是“挖掘数据的潜在价值、实现数据重组的创新价值”^[28]。在大数据技术的帮助下,社会管理者可以识别“重点风险人群”^[29],并对其行为、社会网络、人口特征进行综合分析^[30],为舆论预测提供科学的方法,降低社会舆论风险,而这种应用潜力不仅对国内公众有效,还可能被应用在国际舆论的研判与传播工作中^[31-32]。笔者认为,国内学者在开展大数据与舆论研究时,对人与技术的关系描述存在某种“简化论”的理解现状,过分强调技术对人的作用,而对人如何反作用于技术方面的思考不够。

而在国外研究中,学者们更关注大数据技术对公众舆论的“挖掘”和“预测”作用,同时强调了技术与公众舆论的相互影响关系,认为大数据技术在影响公众舆论的同时,舆论本身也在反作用于大数据技术的发展和具体应用进程。从大数据对公众及其舆论进行挖掘的角度来看,国外学者的研究重点提及了三个方面:一是通过大数据挖掘公众舆论背后的心态和情感,从而对社会运动,尤其是政

治运动中的公众舆论进行研究^[33-34];二是在商业领域中,借助大数据的舆论挖掘功能寻找营销和商业机会^[35-36];三是从方法论的角度着眼,介绍通过何种方式具体实现大数据在舆论信息收集和挖掘方面的功能^[37-38]。而关于大数据在公众舆论方面的预测功能,国外学者的研究则以政治运动为主要的讨论方向。具体来看,在政治运动的领域,有学者指出,通过大数据信息收集、模型建构和验证,可以在政治选举等政治行动中预测公众的舆论走向甚至行为举动^[39-40]。也有部分研究针对大数据在舆论和行为预测方面的具体操作方法^[41]和存在的问题与困境进行分析^[42]。

从发文数量的年度发展趋势来看,自 2013 年开始,国内外有关“大数据与舆论”的相关研究都处在稳定、快速、持续增长的状态,且国内外研究从数量的增长情况来看呈现出较为一致的趋势,尤其是在 2015 年之后更是呈现出近乎“井喷式”的发文趋势。可以说“大数据与舆论”的研究正成为目前国内外学术界的一个关注焦点,笔者相信,这一焦点议题将在未来持续吸引学术界的关注和讨论,引发更多具体的讨论话题和发展方向。

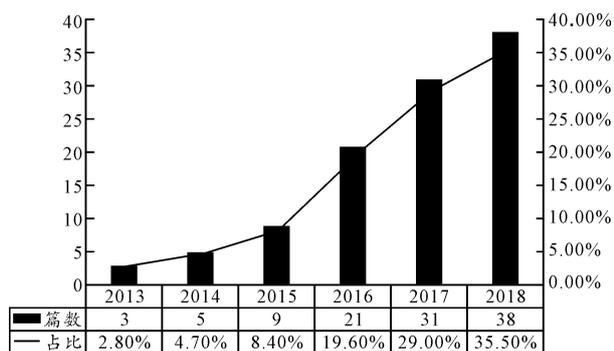


图 2 2013—2018 年 CSCI 论文发文数量与占比趋势

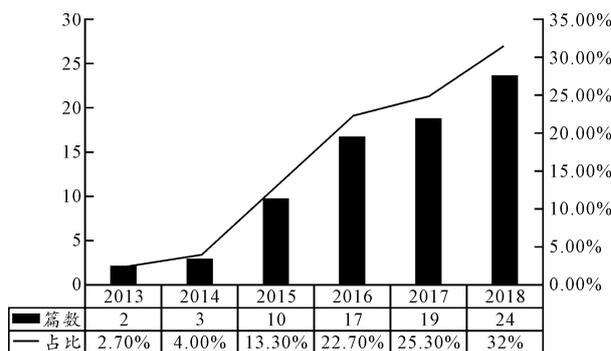


图 3 2013—2018 年 SSCI 论文发文数量与占比趋势

此外,笔者从纵向的时间维度开展分析,对 2013—2018 年间国内外大数据与舆论研究中人与技术的关系定义情况分布进行统计(图 4)。如图所示,进入 2015 年以来,大数据与舆论学术研究的数量激增,和前文分析一致,“舆论挖掘”和“舆论引导”是这几年学术研究中围绕人与技术关系所提出的核心观点。其中“舆论挖掘”在 2016 和 2018 年数量最多,而“舆论引导”则在 2017 年出现较为集中。此外,关于大数据与舆论引导、大数据对公众舆论的影响以及大数据对公众舆论的挖掘这三方面的研究数量在整体上均呈现出上升的发展趋势。

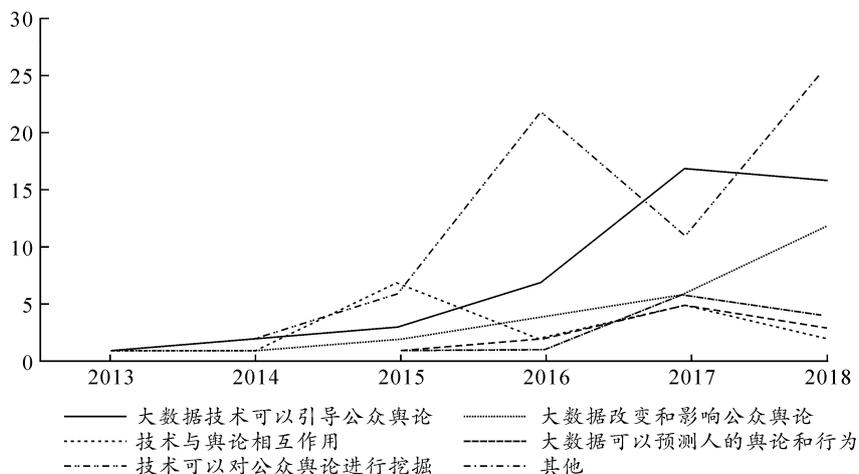


图 4 国内外学术研究中“人与技术”结构关系的年份分布

无论是国内还是国外的研究,我们从文献关键词共现分析中都可以发现,围绕社交媒体开展大数据与舆论研究,是国内外当前的研究热点。从“人-技术”的视角出发,在大数据时代,舆论研究的“人”容易被简化成“网民”尤其是“社交媒体用户”这一类型主体。一方面这显示了随着社交媒体不断的发展和普及,社交媒体对社会舆论的影响不断深入;另一方面也凸显了现有的学术研究在开展大数据与舆论研究时对其他类型社会主体的舆论关注有所欠缺,大数据与舆论研究似乎已经逐渐等同于大数据与社交媒体舆论研究,这是国内外学术研究普遍存在的问题和特征。

(二)大数据与舆论研究的焦点叙事

1. 研究议题:实践价值驱动的问题意识

如图7示,在八个主要的研究焦点中,国内学者最关注的“公众舆论与舆论引导”“大数据与传播改革”和“大数据与社会治理”三大议题,其中舆论引导和传播改革的重要性尤为突出。具体来看,在“公众舆论与舆论引导”问题下,从“人-技术”的视角出发,国内学术研究的焦点又大致分为三大重点:一是从风险社会视角出发,关注突发事件发生后大数据对公众舆论引导的作用^[43];二是结合重点人群,讨论大数据如何对特定人群(如青年、大学生)进行意识形态教育和舆论引导工作^[44-45];三是突出方法论的研究导向,从实践指导的层面讨论大数据时代公众舆论的引导工作有哪些工作思路,包括司法层面^[46]、教育层面^[47]、媒体层面^[48]等。

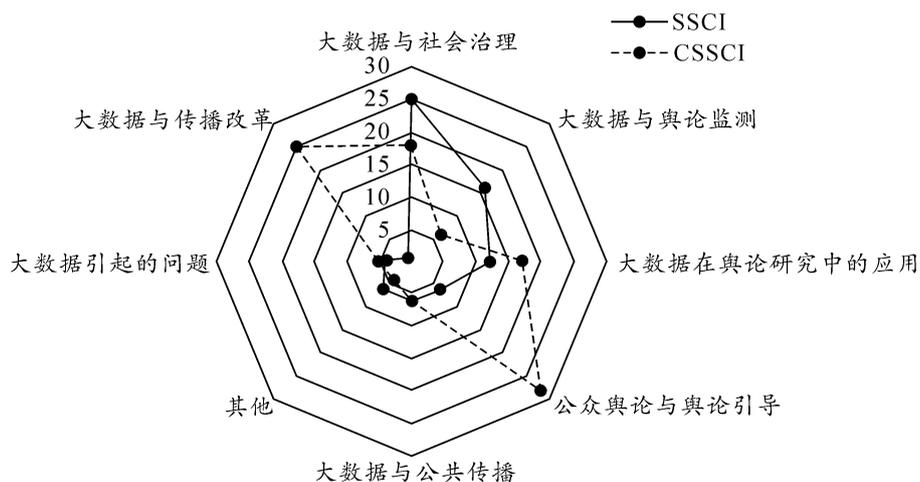


图7 国内外学术研究中研究议题的分布对比

而相比之下,国外学者对大数据与舆论的研究则以社会治理作为焦点问题。同时,国外学者在探讨“大数据与社会治理”方面的问题时涉及的具体议题较为广泛。在这些大数据与舆论的相关研究中,除了强调大数据与政府治理相关的舆论问题^[39],还重点提及了大数据在医疗和公共卫生问题治理^[49-50]和环境风险治理方面的应用问题^[51]。此外,“大数据与舆论监测”和“大数据在舆论研究中的应用”也是国外学者普遍重视的关键议题。前者偏向于讨论通过大数据这项技术,对公众舆论背后的社会态度和社会情绪进行分析,尤其是社交媒体数据承载的公众舆论意义^[34]。后者则呈现出较强的反思性和批判性,从两个方面反思大数据在舆论研究中的应用问题:一是反思大数据作为一种方法本身存在的有限性,强调在现行条件下大数据对公众舆论的研究仍有一定的困境和限制,认为大数据并不能解决舆论研究中的全部问题^[52];二是将大数据对公众舆论的研究与传统舆论研究方法进行对比,在提出大数据在舆论研究方面具有优势的同时,强调将大数据舆论研究与其他方法相结合的重要性^[53-54]。

总的来看,国内外学者在大数据与舆论的相关研究中,主要提出的研究问题呈现出强烈的实践

解”^[62]。大数据与舆论研究经历过去几年的发展,已经初步形成一个整体体系,进入了一个高速发展的研究阶段,在相关领域开拓出丰富而多元的研究议题,且中西方学者的研究分别呈现出各自明显的倾向性和选择性。总的来说,从“人-技术”的视角出发,当前国内核心论文中关于“大数据与舆论”的研究呈现出两大本土化特征:第一,尽管国外学者的研究也会提及大数据对公众舆论的影响,但几乎没有直接提及“舆论引导”的概念,而在国内的“大数据与舆论”研究中,“舆论引导”则被放置在极其重要的位置。受到社会转型时期特定的社会背景和社会发展现状,以及长期以来中国社会文化和舆论生态环境等综合因素的影响,笔者认为,“大数据对公众舆论的引导”将在未来相当长一段时间内继续成为国内大数据与舆论研究的重点问题。第二,从整体上看,大数据与舆论的相关研究主要服务于政府的舆论管理需求,互联网思维的“受众中心”“以人为本”在目前国内“大数据与舆论”研究中仍没有充分体现,结合国内相关研究的两个重点议题——舆论引导和社会治理可以发现,在学术话语体系中,我们可以发现大数据更主要被理解和定义为可以帮助政府、企业等相关机构识别、发现、跟踪、监控、分析公众的工具,而缺少对大数据时代公众舆论如何反作用于大数据这项技术的探讨和论述。

从“人-技术”的视角出发,国外学者的研究整体上也呈现出两大偏向:首先,国外的大数据与舆论研究有更多批判性和反思性的讨论,包括大数据在预测和挖掘方面的局限性、公众数据隐私和数据收集的伦理道德、公众对大数据这项技术的态度和认知等。国外学者在进行相关研究时不仅从技术对人的影响角度切入,也会关切人本身如何对技术产生反身性的影响。其次,过去几年国外“大数据与舆论”的应用在商业领域中的相关研究中备受关注,相比起国内学者更倾向于从政府治理的视角出发讨论如何使用大数据进行舆论方面的应用探讨,国外学者们的研究提出更多与商业和营销领域相关的问题,这一点在国内的相关研究中少有出现。

总的来说,国内学者相比之下更强调大数据对公众及其舆论的改变、影响和引导,国外学者则更多地聚焦于大数据在预测、预判、模拟公众舆论、情感甚至行为方面的潜在功能,尤其是在政治选举、金融市场、商业银行方面的公众研究,是国外学者在“大数据与舆论”研究所重点关注的视角。尽管国内外“大数据与舆论”研究存在一定的差异和偏向,但是在对“人与技术”的关系理解上,国内外学术话语建构也表现出一定的相似性,主要可以概括为三方面:第一,在研究对象上,对网络舆论尤其是社交媒体高度关注。文献分析显示,当前学术研究中利用大数据进行的舆论研究大多与社交媒体相关联,国内的微博、微信,国外的 Twitter、Facebook 和 Google+ 都是学者们重点关注的案例。在这些学术研究中,“社交媒体舆论”在很大程度上已经被等同于“网络舆论”,进而等同于“公众舆论”。一方面这体现了当前社交媒体在移动互联网社交时代的重要地位;另一方面也让我们警惕这样的研究取向是否会忽略了其他语义下的公众舆论,造成研究与现实的差异甚至“鸿沟”。第二,国内外研究均重点提到了大数据对舆论挖掘的重要性。通过数据挖掘,获取社会管理、金融经济、文化传播、国际关系、医疗健康等多领域的关键数据信息,是目前国内外学者们共同提及的关于大数据时代人与技术关系的重要视角。同时,意见和舆论挖掘也常常与社会情绪分析相联系,重点探讨大数据作为一种信息分析工具如何对舆论背后的情感意义进行挖掘。略有不同的是,国内学者在相关论述上更多以应用前景的描述性分析为主,而国外学者的研究则重点提及方法论,即通过何种技术工具可以实现相关挖掘功能的介绍。第三,在学术关怀上,从“人-技术”的视角出发,大数据与社会治理是国内外学者共同关注的焦点议题。国内外相关研究都不谋而合地认为大数据在舆论研究中可以实现不同方面的社会治理功能。具体来看,国外学者更多地提及大数据在医疗健康、公共卫生、和平稳定、环境保护等方面的治理应用;而国内学者的研究则更多联系到金融经济、教育管理、政府改革等问题,现实导向性强是现有大数据与舆论研究的普遍特征。

四、研究结论:大数据与舆论研究中应该坚持“人本”导向

正如有学者所强调的,“互联网思维的核心要义是用户思维”^[63],笔者提出:大数据时代的舆论研究应该体现“以人为本”的理念和导向。基于对样本文献的梳理和分析,结合中外学术对话的比较结果,笔者认为,从“人-技术”的视角出发,未来国内关于“大数据与舆论”的研究存在几个重要的发展方向和有待解答的关键议题:

其一,对大数据将如何被应用在舆论研究这一领域的问题,国内学者缺乏对存在的问题进行批判式的思考,即使在一些带有反思性的研究中,也大多数是针对现存的问题提出“隔靴搔痒式”或“千篇一律式”的对策和建议,缺乏对大数据与舆论领域存在的问题进行探讨,也较少结合中国国情提出本土化、在地化的对策和建议。其中“数据隐私”和“人才培养”这两个问题已经被不少国外学者所关注,但是在中文样本文献中,对这些问题的讨论仅有少数。如刁生富和赵亚萍反思了大数据时代个人隐私和自由权利可能会受到侵犯的客观现实,提出在大数据传播时代法律应该明确对“隐私”的边界进行划分的观点。张志安和邹禹同也认为大数据时代“被遗忘权”引入中国具有极强的社会现实意义。谢耘耕和万旋傲认为大数据要充分应用在舆论研究的领域中,一个无法避免的话题是数据分析型人才培养的现实问题。这些研究为我们提供了一部分反思大数据与舆论研究现状的思路,启示我们未来国内大数据与舆论的学术研究应更多考虑技术给社会带来的多重影响,

其二,当前国内大数据与舆论研究的焦点议题之一是大数据在舆论监测和引导方面的应用,但是对如何利用大数据这项技术对公众舆论背后的社会心态进行解剖则仍待进一步研究。公众通过发表言论形成公众舆论,社会管理者利用大数据了解这些舆论背后的态度和诉求后做出回应并满足其中的合理诉求,实际上是公众在借助大数据行使自身合法权利的一种新手段。从研究现状来看,关于大数据工具如何在收集公众舆论的基础上分析社会心态和社会情绪等问题,从而发现和解决社会问题的讨论,国内学者的研究仍相对停留在“纸上谈兵”的阶段。笔者认为,未来国内学者在讨论大数据与舆论的相关问题时,理应在关注舆论监测和引导的基础上,进一步探寻如何借助数据对人类情感情绪和社会行为进行预测、分析和反思,从而实现大数据对社会心态的研究分析。

其三,国内学者对大数据与舆论相关议题的研究需坚持以人为本的中心要旨,进一步突出“以技术服务人”理念,坚持“技术向善”“数据向善”而不是“以技术控制人”的中心要旨,避免陷入技术决定论的片面思维模式,使学术研究更具有社会温度和人文情怀。一如有学者所指出的,“技术是人功能的延伸,被寄予了为人服务、为人谋利的情感态度”^[64]。然而当前国内学者更多把大数据与公众舆论之间的关系定义为技术(大数据)对人(舆论)产生影响,忽略了人在这种互动关系中的能动作用,关于人如何反作用于技术则缺乏深入讨论。例如,笔者认为,在“监测”“挖掘”和“引导”之间至少还有“回应”的环节,从“人-技术”的视角出发,对大数据与舆论问题展开的研究还有更丰富和多元的思维方式和学术关怀。

本次研究以2013—2018年间英文SSCI核心数据库和中文CSSCI核心数据库中关于大数据与舆论的文献作为分析样本,结合内容分析和文献计量的研究方法,从“人-技术”的分析框架出发,关注现有的大数据与舆论的相关研究中,人和技术的结构关系如何被呈现,并通过关键词共现分析、标题和摘要高频词等统计,描绘出国内外大数据与舆论研究的知识图景概况。在本次研究中,存在两点局限需要做出说明:其一,考虑到篇幅和研究精力的限制,部分国内外非核心数据库中的论文尚未包含在本次研究的文献库中,抽样时间的限定也会对文献样本的全面程度有一定影响。在未来的研究中,可以采集来源更广泛而完整的文献样本,从而扩大文献样本的覆盖面,提高研究结论的普适性与指导意义。其二,在研究方法上,本文引入了统计软件R和社会网络分析软件Ucinet,充实了文本分析的方法论意义。但是对计量结果的分析主要还是依赖于研究者本身对数据结果的认知和理解,未来随着数据分析和信息挖掘手段的不断进步和完善,我们相信可以进一步利用更多不同的研究方法

获得更加客观、科学的文献分析和研究效果。

参考文献:

- [1] 沃尔特·李普曼. 舆论学. 林珊译. 北京: 华夏出版社, 1989: 3.
- [2] 陈力丹. 舆论学——舆论导向研究. 上海: 上海交通大学出版社, 2012: 16.
- [3] 陈福平, 许丹红. 观点与链接: 在线社交网络中的群体政治极化 一个微观行为的解释框架. 社会, 2017, 4: 217-240.
- [4] 任孟山, 朱振明. 试论伊朗“Twitter 革命”中社会媒体的政治传播功能. 国际新闻界, 2009, 9: 24-28.
- [5] 陈涛. 关于网络行动中情感动员研究的文献综述及理论反思. 东南传播, 2016, 2: 83-86.
- [6] 夏倩芳, 原永涛. 从群体极化到公众极化: 极化研究的进路与转向. 新闻与传播研究, 2017, 6: 6-33+127.
- [7] 杨洗. 社会化媒体舆论的极化和共识——以“广州区伯嫖娼”之新浪微博数据为例. 新闻与传播研究, 2016, 2: 66-79.
- [8] 曼纽尔·卡斯特. 认同的力量(第二版). 曹荣湘译. 北京: 社会科学文献出版社, 2006: 309, 334.
- [9] 师曾志, 金锦萍. 新媒介赋权: 国家与社会的协同演进. 北京: 社会科学文献出版社, 2013: 103.
- [10] 曼纽尔·卡斯特. 网络社会的崛起. 夏铸九等译. 北京: 社会科学文献出版社, 2001: 6.
- [11] J. Manyika, M. Chui, B. Brown, et al. Bigdata: The Next Frontier For Innovation, Competition, And Productivity. Mckinsey Global Institute, Seoul, South Korea, Tech. Rep., 2011.
- [12] 韩运荣, 喻国明. 舆论学原理、方法与应用. 北京: 中国传媒大学出版社, 2013: 56.
- [13] 李璐. 理解大数据的温度——以麦克卢汉“冷热”媒介理论为基础//李西建, 金惠敏. 美学麦克卢汉: 媒介研究新维度论集. 北京: 商务印书馆, 2017: 302.
- [14] 卡尔·曼海姆. 重建时代的人与社会: 现代社会结构研究. 张旅平译. 南京: 译林出版社, 2011: 187.
- [15] 雅克·艾卢尔. 技术社会//安东尼·吉登斯. 现代性的后果. 田禾译. 南京: 译林出版社, 2016: 143.
- [16] 曾凡斌. 大数据应用于舆论研究的现状与反思. 现代传播-中国传媒大学学报, 2017, 2: 132-138.
- [17] 李彪, 郑满宁. 社交媒体时代的网络舆情——生态变化及舆情研究现状、趋势. 新闻记者, 2014, 1: 36-41.
- [18] 余志为. 论新媒介时代的“媒介控制”. 编辑之友, 2015, 9: 52-55.
- [19] 喻国明, 王斌, 李彪等. 传播学研究: 大数据时代的新范式. 新闻记者, 2013, 6: 22-27.
- [20] 杨建义. 历史虚无主义的网络传播与应对. 党政视野, 2016, 12: 48-48.
- [21] 张志安, 聂鑫. 青少年网络权益保护的主体责任与多元共治. 传媒观察, 2018, 11: 2+7-14.
- [22] 郑洁. 大数据时代我国意识形态安全面临的机遇、挑战与对策. 教学与研究, 2017, 11: 63-70.
- [23] 葛慧君. 深入实施创新驱动战略, 加快传统媒体融合发展. 传媒, 2017, 10: 3.
- [24] 莫远明, 黄江华. AI+IP+TT 视野下的数字出版融合发展研究. 出版广角, 2018, 1: 23-25.
- [25] 高志华, 蔡钰. 政府回应网络舆论问题的探讨. 行政管理改革, 2018, 3: 69-72.
- [26] 任中义. 中国政治文化安全与网络谣言治理. 郑州大学学报(哲学社会科学版), 2018, 1: 81-83.
- [27] 郑保卫, 李鹏. 大数据时代的政府传播模式变革与理念创新. 现代传播(中国传媒大学学报), 2016, 12: 26-30.
- [28] 孙丽杰, 李春华. 大数据环境下网络舆情管理方法研究. 思想政治教育研究, 2017, 1: 124-129.
- [29] 黄炜, 余辉, 李岳峰. 国内网络反恐研究的现状、问题和展望. 现代图书情报技术, 2016, 11: 4-13.
- [30] 余秀才. 网络舆情研究中的大数据技术使用与问题. 新闻大学, 2017, 2: 117-123+156-157.
- [31] 蒋洛丹. 国家形象传播中的大数据技术应用. 中州学刊, 2018, 11: 168-172.
- [32] 万旋傲, 刘璐. 新时期的中国国际舆论生态: 前沿议题与动态走向——首届中国国际舆论学年会暨海丝国际舆情研讨会综述. 新媒体与社会, 2017, 1: 123-132.
- [33] D. Beck. Medical Students and Controversial Ethical Issues: Results from The Multicenter Study SBAME. BMC Medical Ethics 2014, 15: 85-92.
- [34] J. Nelson, S. Quinn, B. Swedberg et al. Geovisual Analytics Approach to Exploring Public Political Discourse on Twitter. ISPRS International Journal of Geo-Information, 2015, 1: 337-366.
- [35] J. R. Glass, G. H. Kruse, S. A. Miller. Socioeconomic Considerations of The Commercial Weathervane Scallop Fishery off Alaska Using SWOT Analysis. Ocean & Coastal Management, 2015, 1: 154-165.

- [36] I. Zheludev, R. Smith, T. Aste. When Can Social Media Lead Financial Markets? *Scientific Report*, 2014, 4:4213-3424.
- [37] D. Blazquez, J. Domenech. Big Data Sources and Methods for Social and Economic Analyses. *Technological Forecasting and Social Change*, 2018, 130:99-113.
- [38] T. S. Tomeny, C. J. Vargo, S El-Toukhy. Geographic and Demographic Correlates of Autism-Related Anti-Vaccine Beliefs on Twitter, 2009-15. *Social Science & Medicine*, 2017, 191: 168-175.
- [40] T. Yasseri, J. Bright. Wikipedia Traffic Data and Electoral Prediction: Towards Theoretically Informed Models. *EPJ Data Science*, 2016, 1:22-36.
- [42] Z. Xie, G. N. Liu, J. J. Wu, et al. Social Media Would Not Lie: Prediction of the 2016 Taiwan Election via Online Heterogeneous Data. *EPJ Data Science*, 2018:7-32.
- [43] G. Zhu, C. Huang, B. Hu, et al. Autonomy in Individual Behavior Under Multimedia Information. *Multimedia Tools and Applications*, 2016, 22:14433-14449.
- [44] A. Ecker. Estimating Policy Positions Using Social Network Data: Cross-Validating Position Estimates of Political Parties and Individual Legislators in the Polish Parliament. *Social Science Computer Review* 2017, 1: 53-67.
- [45] 丁晓蔚,高淑萍. 大数据与重大公共危机事件舆情研判——基于对天津港爆炸事件相关舆情信息的分析. *中国出版*, 2016, 22:26-30.
- [46] 操菊华,康存辉. 媒体融合环境下发挥思政课舆论引导作用的路径分析. *思想政治教育研究*, 2018, 5:101-105.
- [47] 王冠宇. 封闭与开放: 社交媒体环境下青年群体的网络舆论引导策略探究. *探索*, 2018, 4:179-186.
- [48] 林文学,张伟. 以司法方式加强社会主义核心价值观建设的方法论. *法律适用*, 2018, 19:34-41.
- [49] 谭可. 论网络新媒体马克思主义大众化传播的三重维度. *湖南社会科学*, 2013, 6: 289-291.
- [50] 邢丽梅,顾理平. 媒体新常态中传统主流媒体的舆论引导. *江苏社会科学*, 2015, 6: 268-272.
- [51] R. Hays, G. D. White. The Care. Data Consensus? A Qualitative Analysis of Opinions Expressed on Twitter. *Bmc Public Health*, 2015, 1:838-850.
- [52] C. Y. Mao, W. S. Xiao. A Comprehensive Algorithm for Evaluating Node Influences in Social Networks Based on Preference Analysis and Random Walk. *Complexity*, 2018, 3:1-16.
- [53] M. S. Berry, N. P. Nickerson, E.C. Metcalf. Using Spatial, Economic, And Ecological Opinion Data to Inform Gray Wolf Conservation. *Wildlife Society Bulletin*, 2016, 3:554-563.
- [54] S. Lee. Space at Your Fingertips: Assessing the Public's Interest in Space Activities. *Space Policy*, 2015, 11:39-46.
- [55] A. Ceron, F. Negri. The "Social Side" of Public Policy: Monitoring Online Public Opinion and Its Mobilization During the Policy Cycle. *Policy & Internet*, 2016, 2:131-147.
- [56] P. Miller. Is There A Future for Surveys?. *Public Opinion Quarterly*, 2017, Special Issue: 205-212.
- [57] 李良荣,袁鸣徽. 2017年中国网络媒体的基本格局和态势. *新闻记者*, 2018, 1:44-49.
- [58] M. Palomino, T. Taylor, A Gãker, et al. The Online Dissemination of Nature-Health Concepts: Lessons from Sentiment Analysis of Social Media Relating to "Nature-Deficit Disorder". *Int J Environ Res Public Health*, 2016, 13, 1-23.
- [59] Y. J. Fang, X. Chen, S. Zheng, et al. Modelling Propagation of Public Opinions on Microblogging Big Data Using Sentiment Analysis and Compartmental Models. *International Journal on Semantic Web and Information Systems*, 2017, 1: 11-27
- [60] Y. Rao, H. Xie, J. Li, et al. Social Emotion Classification of Short Text Via Topic-Level Maximum Entropy Model. *Information & Management*, 2016, 8:978-986.
- [61] J. Mikal, S. Hurst, M. Conway. Ethical Issues in Using Twitter for Population-Level Depression Monitoring: A Qualitative Study. *BMC Medical Ethics*, 2016, 2:22-32.
- [62] S. M. Jang, P. S. Hart. Polarized Frames On "Climate Change" And "Global Warming" Across Countries and States: Evidence from Twitter Big Data. *Global Environmental Change*, 2015, 32:11-17.
- [63] Q. Luo, X. Zhai. "I Will Never Go to Hong Kong Again!" How The Secondary Crisis Communication Of "Occupy Central" on Weibo Shifted To A Tourism Boycott. *Tourism Management*, 2017, 62:159-172.
- [64] 高亮华. 人文主义视野中的技术. 北京: 中国社会科学出版社, 1996: 14.
- [65] 陈长, 黄华龙, 曹渝. 大数据治理视角下网络舆论传播生态治理研究. *出版发行研究*, 2018, 9:15+21-24.

[66] 刁生富,赵亚萍.论透明化时代的隐私权保护.河南师范大学学报(哲学社会科学版),2017,6:71-77.

Knowledge Scene of Big Data and Public Opinion Research: based on an Analytical Framework of "Human and Technology"

Zhang Ning, Tang Jiayi (Sun Yat-sen University)

Abstract: The study of "big data and public opinion" essentially focuses on the relationship between "technology" (big data) and "human" (public opinion). In order to provide an analytical framework with both practical significance and academic value, and at the same time to promote academic dialogue among different research, this paper takes research on big data and public opinion in the period of 2013-2018 as examples. It is found that from the perspective of "human-technology", in domestic academic research, big data and public opinion research focuses on the impact of technology on people, among which public opinion guidance and "public opinion mining" are two key issues in the construction of academic discourse. In foreign academic research, scholars pay more attention to the role of big data in public opinion mining and public opinion prediction. The research proposes that the research of big data and public opinion should adhere to the fundamental guidance of human-oriented, and effectively improve the value and role of big data in public opinion research.

Key Words: big data; opinion; human-technology; public opinion guidance; public opinion monitoring

■收稿日期:2019-10-09

■作者单位:张宁,中山大学传播与设计学院;广东广州 510006

唐嘉仪,中山大学传播与设计学院

■责任编辑:汪晓清