

DOI: 10.12120/bjutsxb202106001

大数据与人工智能背景下的网络舆情治理： 作用、风险和路径

李明德^{1,2}, 邝岩³

(1. 西安交通大学 传播内容智能理解研究中心, 陕西 西安 710049;

2. 西安交通大学 新闻与新媒体学院, 陕西 西安 710049;

3. 西安交通大学 马克思主义学院, 陕西 西安 710049)

摘要: 大数据技术和人工智能技术作为强有力的信息化手段, 为网络舆情治理工作提供了全新的信息资源、技术手段和治理范式, 在一定程度上实现了网络舆情治理的全面化、自动化、科学化、精准化和个性化, 增强了网络舆情治理的效能; 同时, 大数据与人工智能在网络舆情治理中仍存在着算法呈现舆情与真实社会民意存在偏差、舆情预测干涉和情绪分析复杂性导致舆情误判、虚假新闻和信息茧房问题误导舆情引导决策等问题。在此基础上, 针对网络舆情治理面临的隐私伦理、数据垄断、数据滥用信息管理风险等新挑战, 提出了具体的解决措施: 优化技术升级, 推进网络舆情治理综合化技术发展; 完善舆情治理法制和理念, 实现舆情数据管理应用机制升级; 壮大人才队伍, 培养复合型舆情治理专业人才; 加强平台自律, 提升公众的网络媒介素养和算法素养。

关键词: 国家治理; 跨学科融合; 网络舆情治理; 大数据; 人工智能

中图分类号: G206; TP39

文献标志码: A

文章编号: 1671-0398(2021)06-0001-10

大数据与人工智能作为新一轮科技革命和产业变革的重要驱动力量, 二者密不可分, 正在走向一体化发展, 并将对国家治理体系和治理能力现代化产生深刻影响。网络舆情治理既是互联网时代国家治理的重要组成部分, 能够反映一段时期内社会心态和社会形势的特征, 为社会治理提供依据和基础; 同时也是根据特定的需要, 针对各种网络舆情存在的问题, 对互联网中的数据进行监测和收集, 并对相关的数据信息进行多层次的思维加工和综合分析研究, 为舆情引导提供指导的过程。当今世界, 信息技术创新日新月异, 互联网是社会舆情传播的主要媒介之一, 网络舆情呈现出新的特点和发展趋势。在此背景下, 大数据、人工智能等技术为网络舆情治理工作提供了全新的信息资源、方法与范式, 也为网络舆情治理工作带来了新的挑战 and 风险。

收稿日期: 2021-05-29

基金项目: 教育部哲学社会科学研究重大课题攻关项目(18JZD022); 陕西省软科学研究计划资助项目(2021KRM182); 传播内容认知国家重点实验室开放课题(20K02)

作者简介: 李明德(1959—), 男, 西安交通大学新闻与新媒体学院院长, 传播内容智能理解研究中心教授, 博士生导师;

邝岩(1992—), 女, 西安交通大学马克思主义学院博士研究生。

一、大数据与人工智能在网络舆情治理中的作用

近年来,大数据和人工智能技术在网络舆情治理中的应用日趋成熟和广泛,推动了网络舆情治理的全面化、自动化、科学化、精准化,增强了网络舆情治理各个环节的效能。

(一) 实现网络舆情监测的全面化和自动化

随着大数据时代网络信息的爆炸性增长,传统的网络舆情监测方法很难在海量信息中进行最优化的抽样以反映舆情的整体特征,进而影响后续舆情分析与引导的科学性。大数据借助挖掘技术和人工智能语义、音视频识别技术,丰富了网络舆情的来源和数据类型,拓宽了网络舆情监测的渠道,使互联网上海量的数据集成为有价值的舆情信息资源,提升了网络舆情监测的效能。第一,网络舆情监测人员借助人工智能、传感器、物联网等技术和设备,可以对多元舆情主体进行智能识别、分类,将网络文本、音视频、表情、传感数据等都纳入舆情监测的范围。第二,人工智能实现了网络舆情监测目标从内容向“内容+关系+情感”的转向,而传统的网络舆情监测只重视舆情主体在舆情事件中进行的话语表达,抓取的舆情数据主要集中于文本内容层面。随着数据挖掘、传感技术、语义识别、情感分析等技术的发展,人工智能以其自动识别、感知获取、自动筛选的优势,能够自动采集网络上的结构化数据和非结构化数据,利用向量机法、贝叶斯法、随机场法等机器学习算法对网络舆情数据进行情感倾向性分析,不仅可以监测话语背后的社会心理、动机、情感诉求等因素,还可以探究舆情传播中各传播节点之间的社会网络关系,实现对互联网用户的情感心理和社会关系的多维度监测。第三,大数据和人工智能舆情监测技术还能够实现全过程的网络舆情动态监测,依托感知智能技术和数据流挖掘技术实时处理高速传播的大量舆情数据,把握网络舆情的产生来源和传播趋势,实现全网热点舆情的实时监测,使网络舆情监测工作更加全面和及时。

(二) 助力网络舆情研判实现预测和决策的科学化

我们对网络舆情数据进行收集和监测的目标在于,发现舆情热点、分析舆情信息、进行舆情预测。在突发事件中,专业人员的舆情分析研判能够帮助舆情主体进行舆情危机处理和舆情回应,指导相关主体采取提前公开信息、征集群众意见、持续舆情监测等措施,增强舆情主体防范和化解舆情危机的能力。在大数据时代,互联网上每天产生数以亿计的数据,大数据和人工智能技术通过对海量数据进行抓取、研究和判断,深入挖掘数据之间的相关性,能够实现网络舆情的提前预警,并对舆情研判和治理决策进行验证分析。第一,网络舆情研判人员使用大数据挖掘技术挖掘、分析海量网络信息背后蕴含的舆情价值和数据之间的相关关系,建立网络舆情演变模型,预判舆情趋势的走向,能够将网络舆情治理的时间点提前到舆论爆发阶段的初始时间点;第二,运用机器学习、虚拟现实智能建模技术和人工神经网络预测模型等人工智能技术,可以对网络舆情的风险、演变规律、民意承受能力等情况进行加工和分析推理,生成舆情评估和预警报告,对舆情的发展态势进行仿真和预测;第三,借助认知智能技术能够实现网络舆情与适用决策的逻辑推演,在基于人工智能的网络舆情研判系统中,通过大量历史训练集训练而设定基准函数后,将需要研判的网络舆情数据和决策方案输入处理系统,可以对不同决策方案的可行性和执行效果进行试验模拟,从而为相关舆情主体进一步调整、优化决策方案提供思路和依据,助力舆情预测和决策的科学化^[1]。

(三) 推动网络舆情引导的精准化和个性化

在大数据和人工智能时代,新闻生产行业和公众信息需求的多样性变革推动网络舆情治理方式的转变,促使网络舆情引导走向精准化和个性化,具体体现在网络舆情引导内容生产、内容呈现、内容分发等环节的全新升级。在网络舆情引导内容生产环节,依托于人工智能和大数据挖掘技术

的写稿机器人,能够从海量资讯中瞬时筛选出热点话题,采用机器学习方法进行自动化新闻写作,快速生成新闻报道,不仅提高了新闻生产的工作效率,同时也提高了舆情引导主体主动发声的效率。例如,2017年,四川九寨沟发生7.0级地震,中国地震台网的机器人仅用时25秒就自动编写了包括震中地形、热力人口、周边村镇/县区等内容的地震新闻报道^①,从而快速完成了舆情引导的信息生产工作,有效解决了传统媒体在新闻传播中的延误、遗漏或人为失误等问题。在网络舆情引导内容呈现环节,大数据和人工智能技术与虚拟现实(VR)、增强现实(AR)、及可穿戴设备等结合,通过VR、AR等技术模拟虚拟环境,为受众提供视觉、听觉、触觉三维式的感官体验,丰富了网络舆情引导内容的呈现方式,使舆情引导内容的呈现方式从“平面化”转向“立体化”,以更强的“沉浸感”助力主流意识形态的传播,增强舆情引导工作的感染力和影响力。在网络舆情引导智能分发环节,人工智能和大数据技术能够快捷地收集网民信息,形成包括网民兴趣偏好、阅读习惯、地理位置等信息的数据档案,对其进行“精准画像”。在网络舆情事件的讨论中,不同的网民往往持有不同价值倾向的观点和诉求,政府部门和媒体可以依据网民画像,利用智能推荐算法为网民设置议程,向特定人群进行更加精准化的内容推送,为他们提供权威、及时、个性化的信息和数据服务,从而发挥政府部门和主流媒体的传播优势,提高网络舆情引导的精度。

二、大数据与人工智能在网络舆情治理中引发的风险

当前,大数据和人工智能技术的发展与应用还处于不断探索中,在网络舆情治理中仍存在着一些局限性,使网络舆情治理面临新的风险与挑战。

(一) 舆情监测盲区导致算法呈现舆情与真实社会民意存在偏差

目前,我们利用大数据和人工智能算法实现的网络舆情监测的全面性是相对意义上的。本文所指的全面性是相对于传统的网络舆情监测方法,大数据和人工智能大大拓宽了舆情监测的范围。在实际应用中,大数据与人工智能背景下的网络舆情监测仍然存在盲区,主要存在三个方面的原因。第一,大数据技术的不完善产生了舆情监测的盲区。现有的舆情监测技术与全面性要求、客户需求还存在很大差距,例如网络评论区中的缩写文字、拼音、表情包等信息提取技术还不够成熟,造成此类信息难以监测,技术与需求之间的差距导致舆情监测提取到的数据精准率不够高、数据不全面。第二,网络平台的监测准入性问题也导致了舆情监测盲区的产生。一些网络平台,如微信、QQ等为用户提供了隐私设置功能,用户设置为不可见的微信朋友圈、QQ空间等成为网络舆情监测的盲区;同时,微信公众号文章评论区的内容是经账号主体筛选后而公开的,这意味着网络舆情监测难以获得全部的评论内容。由于这些舆情监测盲区中含有大量丰富的舆情信息,以及平台设置问题,因而无法进行监测。第三,我国相关部门、舆情监测公司等主体与网络平台没能达成数据共享,当其进行舆情监测时未能获得平台的全部接口权限,故而不能对网络平台数据进行全面挖掘。

此外,由于研究人员利用大数据挖掘技术进行网络舆情监测时,尤为关注舆情数据来源的整体性,使得舆情监测获取的数据呈现体量巨大的特点,因而在进行舆情研究时无法分析全部数据,仅是着眼于数据之间的相关性,通过当下时间段和历史时间段内的相关性数据,挖掘舆情数据背后的民意反映、社会网络关系和预测舆情演变走势。但是,相关性不等于绝对性,谷歌流感趋势预测的失误、大数据预测选举结果的多样性等,表明人们通过大数据相关性分析得出的舆情样态不完全等

^① 2017年8月8日,《四川阿坝州九寨沟县发生7.0级地震》, <https://mp.weixin.qq.com/s/qHf2ln1sFwfikZyOyZ8yRA>。

同于真实的社会民意。由此可见,由于网络舆情监测盲区和相关性分析的局限性,通过大数据和人工智能技术获取的网络舆情只能体现微观真实、现象真实,而真实民意是宏观真实、本质真实。因而,如果我们将两者完全等同,则会出现以偏概全、以点带面的舆情认知风险。

(二) 舆情预测干涉和情绪分析复杂性引致舆情误判风险

1927年,量子物理学的创始人之一维尔纳·海森堡(Werner Heisenberg)在《关于量子论的运动学和力学的直觉内容》一文中提出了“测不准原理”(在量子世界中,测量粒子位置,必然会影响到粒子的速度),揭示了在量子尺度的微距世界中的“测量即干涉”现象。这一“测量即干涉”现象也出现在人工智能和大数据分析中。人们在利用大数据和人工智能技术对网络舆情进行监测和分析的过程中,监测和分析技术的介入可能会干扰网络舆情的表现和演化,进而导致舆情误判。例如,一些社交媒体会利用技术自动抓取用户的搜索热点或用户提及的关键词并进行热点推荐,但这些自动计算的“热点”有时是与真实舆情热点无关,或者营销主体有意为之的虚假信息,而且这些“热点”经过推荐后会引发用户进一步的点击行为,扰乱了真实舆情的发展走势。因此,如果我们过分依赖大数据和人工智能进行舆情分析,则可能被噪音数据所误导,难以做出符合具体环境的舆情治理决策。

此外,大数据与人工智能背景下的网络舆情研判还存在情绪分析偏误现象。不仅网络舆情中包含了民众的情绪、情感和意愿,而且网民的情绪表达推动并影响着网络舆情的发生、扩散、变动和衰减的全过程。虽然机器学习、自然语言处理、卷积神经网络分析等信息科学情绪分析技术通过爬取网民在社交平台上所发表的言论和评论,对其进行情绪特征分析和情绪演化阶段划分,为从非结构化的网络内容中提取用户情绪提供了途径;但是由于情绪分析涉及舆论学、心理学、计算机科学等多个研究领域,以及情绪本身的高度复杂性、汉语的语义模糊性以及新型网络词汇的不断涌现,导致情绪分析处理技术的应用效果仍不够理想^[2]。这些困难为网络舆情的情绪分析带来挑战,情绪分析偏误问题会使监测主体不能准确把握舆情的发展程度和传播方向,进而导致舆情误判等风险。

(三) 虚假新闻和信息茧房问题误导舆情引导决策

写稿机器人能够提高舆情引导的效率,但是写稿机器人从海量资讯中迅速筛选出热点话题生成新闻报道的技术逻辑与流程严谨的传统媒体议程设置的方式迥然不同。这意味着,使用写稿机器人进行信息生产,会将一部分议程设置的权利让渡给机器与用户,可能会导致新闻报道的“议程偏离”和虚假信息问题。例如,如果我们没有对写稿机器人设置严格的信息抓取规则或者未对其生成内容设立人工审查机制,则可能会出现写稿机器人抓取到谣言信息进行报道的现象,或者写稿机器人被利用成为“机器人水军”,在网络上营造虚假舆论,误导舆情引导决策。此外,如果我们利用写稿机器人进行新闻报道,由于写稿机器人还不能像资深媒体人那样对新闻事件发表个人的深度看法,所以可能使真正关涉公共利益的重要信息被边缘化,阻碍媒体议程设置功能的发挥,进而削弱媒体的新闻舆论引导力。

人工智能推荐算法和大数据用户画像技术的应用在提高了舆情引导精准性的同时,还导致了信息茧房问题。“信息茧房”(Information Cocoons)概念由美国知名学者桑斯坦(Cass. R. Sunstein)在2006年出版的著作《信息乌托邦:众人如何生产知识》中提出,“信息茧房”是指在信息领域,人们关注的信息会被自己的兴趣所引导,久而久之将自身桎梏于像蚕茧一般的“茧房”中的现象^[3]。人工智能算法推荐技术在新闻分发中容易自动过滤掉网民不感兴趣、不熟悉的信息内容,加剧了“信息茧房”效应。某些内容分发平台出于对流量利益的追求,在一定程度上纵容了算法推荐技术对“标题党”“泛娱乐化”等吸引眼球内容的传播,造成一些自媒体对舆情事件的报道和解读从追求流量的目标出发,导致舆论失焦。互联网用户长期处于“信息茧房”中,将很容易受片面信息的影

响,形成一个相对封闭的舆论场域,而一旦主流意识形态被拒斥在这个场域之外,那么身处其中的成员便很难与外部的主流意识形态实现互动,正确的舆情引导将很难发挥效用^[4]。此外,在人工智能环境下,信息内容在生产中存在大量的内容“搬运”“洗稿”现象,带来原创优质内容减少、内容质量良莠不齐和版权侵权等问题,冲击网络舆情引导的生态环境。

(四) 隐私伦理、数据滥用、数据垄断等产生信息管理风险

现今,各类搜索引擎、社交网站、舆情监测平台能够随时随地记录网民在网络上的任何信息,而且通过服务器记录的人们的网络数据日志已成为继自然资源、金钱、人力等社会资源之后的新数据资源。同时,大数据挖掘技术也使人们的一切行为都暴露在“第三只眼”的监测之下,并留下一条永远存在的数据足迹,产生了隐私伦理、数据滥用、数据垄断等信息管理问题。

在互联网时代,只要人们使用了网络设备,其在网上的所有行为都可能被监控并永久存储。这些每个人产生的小数据被有关公司汇聚到一起,经过分析、处理形成大数据网络舆情信息资源,而这些由个人行为产生的经过汇聚而形成的大数据,是否存在所有权问题;每一个个体是否应当具备个人数据被谁处理、经过怎样处理的知情权;数据被相关公司采集后,有关公司是否永久地拥有用户数据的保存权,在处理数据形成舆情分析报告时,是否会侵犯个体的隐私。这些问题都构成了网络舆情治理的隐私伦理问题^[5]。与隐私伦理问题相伴而来的的问题是,人们因数据的过度搜集和使用而产生的数据滥用问题,从现实情况看,应用程序(App)强制授权、超范围收集个人信息等现象仍大量存在,脸书(Facebook)数据丑闻、大数据“杀熟”等数据权利越界事件频繁发生等。这些都表明,用户数据权属规制和数据安全问题已成为不容回避的数据管理风险问题。

目前,社会发展除了隐私伦理和数据滥用问题,在数据的收集、使用过程中,还出现了数据收集公司的数据垄断问题。中国人民大学网络与移动数据管理实验室(WAMDM)基于约3000万用户数据和约30万APP数据,对当前移动应用场景下用户隐私数据被收集的情况进行调研分析,发布了《2019年度中国隐私风险指数分析报告》。该分析报告显示,前10%的权限数据收集者可获取99%的权限数据^[6]。这表明,社会常见的“二八定律”在网络世界中的反映更为严重,互联网数据垄断态势居高不下。此外,在网络舆情引导工作中,一些企业和部门缺乏数据共享理念,导致有些公共数据存在数据分散、数据质量评价无统一标准、数据使用率低等问题,无法达到信息的互联互通、整合利用,从而削弱了数据为相关部门和主流媒体舆情引导提供决策参考的作用。

三、大数据与人工智能技术在网络舆情治理中的困境

(一) 技术困境

1. 跨学科的方法融合的困境

大数据和人工智能网络舆情治理是一个融合多学科的领域,在实践应用中需要计算机技术、新闻学、管理学、情报学等多学科的方法交叉,这就决定了大数据和人工智能舆情治理具有横断程度广、涉及的学科门类多、学科融合领域深的特点。一方面,这些特点均给网络舆情治理的科学理论的构建和完善带来了困难;另一方面,除学科理论困境外,大数据和人工智能网络舆情治理领域的技术理论体系也需要完善。例如,网络舆情监测中的测量干涉问题、写稿机器人的虚假新闻问题,以及算法推荐技术的信息茧房等,其发展完善需要多领域学科的技术发展。我们要想提升大数据和人工智能在网络舆情治理中的应用效果,还需要进一步构建和完善舆情治理的跨学科理论体系和技术。

2. 舆情分析技术的更新

本文通过分析我国近年来发生的热点舆情事件发现,在各类社交媒体平台上,公众的话语体系

经常出现更新,表现出舆情发布主体多样、舆情传播渠道多元、舆情发展趋势不确定性强、舆情意见表达新颖等特征。例如,在互联网上每年会流行一些“新词”或“旧词新解”,而且这些流行词的语义经常发生变化,甚至在不同的语境中存在不同的含义。这些特征的具体体现方式千变万化,反映的社会心态难以把握,导致网络舆情事件的话语体系和表征的不可控性和隐蔽性大大增强,为机器及时识别、发现舆情信息带来了巨大挑战。如果我们单纯使用一个关键词或者使用过往舆情案例的经验表达模式等搜索方式进行舆情监测和分析,即很容易造成误判或信息遗漏。因而,网络舆情话语体系的更迭和舆情信息的隐蔽性会导致基于大数据和人工智能的网络舆情监测系统舆情发现的速度变慢,舆情预警效能和分析研判的准确性也会随之大大削弱。

3. 传统“小数据”的“大数据傲慢”

“大数据傲慢”^[7]是指研究人员在利用大数据进行研究时,认为已掌握的数据即是总体,大数据可以完全替代而不是补充传统的数据收集和分析。“大数据傲慢”思维方式是一种研究误区,研究人员认为应用大数据方法计算出的“统计学相关性”能直接取代事物之间真实的因果和联系,从而过度应用这种技术,就会导致研究结果出现偏差^[8]。这种思维方法会给技术应用带来困难,因为使用大数据技术挖掘得到的数据信息只能为研究人员建立事物之间的相关关系提供依据,却无法说明事物之间的因果关系。也就是说,技术只能帮助人类搜集更多事件信息,却无法针对相关事件给出确切的解释,舆情事件发生及演变的原因最终仍需舆情工作者自行判断。如果研究人员用数据之间的因果关系代替了相关关系,就可能造成舆情分析失真。

(二) 社会困境

1. 技术应用的资金和制度保障力度尚待提升

我们应用大数据和人工智能进行网络舆情治理时,需要及时研发或引进最新的信息收集和分析技术系统,从而需要雄厚的资金和人力保障。大数据与人工智能舆情分析系统愈是先进,所需要的研发成本越高,对人才的要求也越高。然而在实际工作中,部分相关部门、机构和企业不具备与其实际所需的网络舆情分析系统相匹配的资金和专业人才力量,在面对负面舆情时还是采取传统的舆情收集与处置办法,导致网络舆情治理工作陷入被动局面。另外,部分地区和机构在处理关于本地区和本机构的负面舆情时,只注重对舆情信息的封、删、堵,缺乏上下级和部门间的相互联动,使得应用技术进行舆情治理不能得到有效的贯彻落实,给网络舆情治理工作带来一定难度。

2. 媒体工作者应积极倡导正能量

大数据和人工智能技术量化并重构了信息传播效果的评价体系,点击量、评论量、转发量等数据,成为信息的传播效果评估指标,反映了某条信息的传播力、影响力^[9],使得部分媒体工作者在评价目标和经济利益驱使下,倾向于生产能够吸引用户眼球的内容,在一定程度上促进了虚假新闻、“标题党”信息、“泛娱乐化”内容的泛滥,不利于主流舆论引导内容的传播。在此信息环境下,部分公众由于缺乏媒介素养,容易被智能推荐算法等信息传播技术裹挟,从而消解其主体认知和思维能力,沉溺于碎片化、浅层化、娱乐化的信息接收方式中,甚至无视、解构主流意识形态传播内容与价值,增加了正确网络舆情引导内容的传播难度。

3. 舆情治理和数据管理体系亟需完善

目前,我国人工智能与大数据支撑下的网络舆情治理的监管与应用仍处在探索阶段,相应的数据收集、数据存储、数据使用监管体系还不完善。在行业发展方面,大数据与人工智能舆情治理应该遵循的行业标准尚未完全建立起来,大部分媒体、舆情监测机构未能及时公开舆情数据的采集和分析过程,相关部门、媒体、舆情监测机构之间也未能建立起数据共享共治的理念。在公民隐私保护方面,国家相关部门对于网络平台过度收集用户信息等行为陆续出台了相应的法律法规,但是相应规章还需要更细化、更加密切结合互联网产业实际发展的新动向,并且在实际操作中需要进一步

细化监管要求、规范监管流程和举措。

四、大数据与人工智能背景下网络舆情治理的优化路径

本文通过梳理大数据与人工智能在网络舆情治理中引发的风险,分析其风险产生的原因,并从技术、制度、媒体、公众等层面提出网络舆情治理的优化路径。

(一) 优化技术升级,推进网络舆情治理综合化技术发展

1. 以小数据补充大数据,关注舆情信息的整体性、混杂性

大数据和小数据拥有各自的优势与劣势。我们在使用小数据进行舆情分析时,能够以相对较低的成本明确定义变量、控制数据生成机制,但缺点是数据量少且代表性不高;而大数据具有数据体量大、收集时间短、数据类型丰富等优点,但也存在数据价值密度低、数据噪音高的缺点。我们在网络舆情治理中应以小数据补充大数据,综合利用两者的优势开展治理,可以大大提高数据监测的准确性和预测的精确度。例如,研究人员可以使用大数据挖掘技术对热点事件网络舆情数据进行收集,再结合小数据信息,如网络意见领袖、热点网络评论等数据,以提高网络舆情分析的全面性和准确性。

2. 关注人机结合,将传统的舆情治理方法与大数据、人工智能方法相融合

相比于传统的舆情治理方法,大数据和人工智能方法是伴随互联网和移动终端高度发展的结果,对于治理当代互联网中混杂性极强的舆情数据具有显著优势。但是,舆情预测干涉和分析偏误问题表明,大数据和人工智能方法仍存在缺陷,其所揭示的规律对适用范围的时效性和地域性要求较高,人工智能机器还不能完全代替研究人员进行舆情分析和决策。因此,我们在舆情治理中不应割裂大数据和人工智能方法与传统的科学研究方法,而应当有机融合大数据、人工智能技术和物联网、云计算、传统舆情研判技术,关注专家分析与人工智能的协同,将人的价值与人工智能的优势相结合,提高网络舆情数据分析和决策预判结果的准确度。

3. 以大数据和人工智能算法模型和运算逻辑为技术引擎,建立综合性网络舆情治理模型

针对信息茧房和虚假新闻问题,我们应当发展多元化与个性化并存的智能算法推荐技术,扩展算法推荐的信息范围,在原有的内容推荐、信息协同过滤推荐、信息流行度推荐基础上,将信息内容品质、主流价值观引领力、算法问责等指标嵌入算法的设计中,以纠正算法推荐带来的信息窄化和群体极化效应,加强主流意识形态对舆情治理技术的价值引领。同时,研究人员应利用深度学习、话题聚类、自然语言处理、人工神经网络、区块链等技术升级算法模型,建立更加科学高效的网络舆情治理模型,优化网络水军检测方法,精准识别机器人水军,采用技术手段对虚假网络舆情进行监测和干预。

(二) 完善法制理念,实现舆情数据管理应用机制升级

1. 推进数据资源的交流共享,实现多元主体协同治理

在当今的网络舆情传播环境中,我们要想破解隐私伦理、数据滥用、数据垄断等难题,必须善用大数据和人工智能技术,凝聚相关部门、主流媒体、信息平台的力量,实现治理的优势。第一,我国相关部门应建立包括基础标准、技术标准、管理标准、数据共享和交换标准在内的大数据标准体系,完善大数据使用标准和规范,为构建共享、共建的智能化数据交流平台和数据行业立法提供依据。第二,除了互联网舆情治理领域,金融、经济、教育等领域都在使用大数据及人工智能技术辅助决策、制定政策。我国应建立统一、标准化的数据接口,可以提高行业领域间数据资源的使用效率,实现高效的数据传输和信息交换,促进数据共享。第三,我国推出数据信息管理相关法律法规,明确各级部门、机构、企业的数据收集、使用范围和数据提供义务,并积极突破数据壁垒,实现公共数据

资源的有效使用和规范管理。

2. 运用法律和技术手段加强个人信息保护,健全隐私伦理约束机制

大数据挖掘技术拓展了网络舆情监测数据的来源。在当今互联网时代,网络舆情监测数据的来源包括大量用户生成内容、搜索引擎日志数据、移动终端数据等。数据来源的广泛增加了用户隐私泄露的风险。2016年,中华人民共和国第十二届全国人民代表大会常务委员会第二十四次会议通过的《中华人民共和国网络安全法》明确提出:“网络运营者应当对其收集的用户信息严格保密,并建立健全用户信息保护制度。”^[10]但是随着技术的发展和实际应用状况的快速更迭,侵犯用户隐私的现象仍普遍存在。大数据与人工智能时代的网络舆情治理应做到兼顾高度智能化与隐私安全性的平衡,国家必须进一步研究、制订和出台关于信息保护和技术规制的法律法规和管理办法,以法律条文形式明确信息搜集和智能算法的技术伦理、技术规范、技术标准、数据收集范围、法律主体、权利边界等,确定监管机构和部门,用法治手段把算法技术及其运用完全纳入法治化的框架。此外,相关部门和平台还应当健全隐私保护的道德伦理约束机制,加强媒体平台和技术开发商的伦理建设,大力培育和发展行业自律组织,建立由多主体参与的监督体系,用道德评价补充法律规范,形成隐私保护的伦理自律。

(三)壮大人才队伍,培养复合型舆情治理专业人才

1. 培养掌握大数据和人工智能技术的复合型人才

大数据和人工智能网络舆情治理是技术驱动的研究领域,人才在应用研究中占据极重要的地位,国家应该进一步对大数据、人工智能人才培养给予政策和资金支持,进一步优化人才培养体系、加强人才培养过程管理。一方面,相关部门应当统筹科研、教育、媒体等机构的力量,建立舆情治理的相关学科,制定学科发展方案,加强各学科人才的交叉培养;另一方面,可以建立专门的高等职业技术教育学院,开展大数据和人工智能网络舆情治理教育和职业技能培训,打造一批精品人才培养基地,重点培养技术人员的计算机、新媒体、传播学等方面的知识,为网络舆情治理提供智力支持。

2. 培养具有网络舆情治理“数据素养”的复合型人才

研究人员利用大数据和人工智能技术进行网络舆情治理,除了需要具备良好的专业技术素养,还要具备网络新闻传播知识和数据素养。“数据素养”包括“对数据的敏感性,数据收集、处理、分析、判断和利用的能力,尊重数据伦理、保证数据准确、安全和隐私的修养”^[11]。这就要求网络舆情治理人员除了需要具备基础的数据处理能力、舆情治理技术和网络舆情分析能力,还应当具备数据伦理和算法伦理修养,在技术研发和网络舆情治理的过程中,秉持对数据进行规范性收集和应用的理念。

3. 培养具备高度理想信念和熟悉宣传业务的复合型人才

网络舆情是“社会的皮肤”,具有反映和整合社会意见的作用,网络舆情治理工作不仅是进行社会治理的重要环节,也是维护意识形态安全的迫切需求,具有一定的意识形态属性。网络舆情治理人员必须要树立高度的理想信念,坚定政治立场,做好网络舆论引导和意识形态引领工作。为此,一方面,宣传部门、高校等机构要积极引进和培养数据与宣传业务双向精通的复合型人才,吸纳不同学科专业人才,加强思政教育。另一方面,政府部门、主流媒体等机构可以通过吸收有影响力的自媒体意见领袖进入网络舆情治理管理体系,对其进行马克思主义指导下的网络舆情引导培育,利用算法推荐、音视频等大数据与人工智能技术,主动占据网络舆情治理传播阵地。

(四)加强平台自律,提升公众的网络媒介素养和算法素养

1. 加强媒体工作者的伦理道德建设,形成网络信息平台的伦理自律

人们在利用大数据和人工智能技术能够为网络舆情治理带来便利,为公众提供更加风清气朗的网络空间,但同时也带来网络内容泛“娱乐化”、虚假信息泛滥等问题。媒体面对新的技术和媒

介环境,需要构建一个与之相匹配的媒介伦理体系。第一,媒体工作者要摒弃逐利心态,强化专业素养和社会责任意识,在使用数据技术进行信息发布和推送时发挥把关人作用,确保信息发布的准确性和客观性,并以公共价值和公众利益为重心,提升主流舆论引导力。第二,媒体工作者和信息平台要筑牢新闻伦理意识,在收集、使用和存储数据时,确保所收集和发布的信息不会侵犯他人的隐私权等权利。第三,我们要发展网络平台的行业自律组织,尽快制定健全的道德伦理约束机制,防止公共领域的失守和媒介伦理的失范。相关部门可以以《中国互联网行业自律公约》为基础,鼓励、支持并推动相关媒体组织制定网络信息发布和管理自律公约,用具体化、操作化、规范化的自律规范体系,形成网络舆情治理领域自我约束、自我净化的道德约束机制。

2. 提升公众的媒介素养,理性看待大数据与人工智能算法

媒介素养既包括公众熟练使用各类媒介的能力,也包括判断、评价、理解、制造不同形式与内容的媒介信息的能力^[12]。算法素养是指用户在面对算法和算法筛选的信息时所表现出来的态度、知识、技巧以及能力的总和^[13]。在人人都可以通过媒体平台发声的背景下,不只是媒体工作者要具有媒介和算法素养,每个用户也要有媒介和算法素养。为此,针对大数据和人工智能背景下的机器人水军、信息茧房和虚假新闻等现象,相关部门和组织要培育公众形成甄别信息真伪的能力,通过各种媒体和平台对用户进行算法技术普及教育,以及算法推荐的技术逻辑和运行机理教育,使其认识到算法的“技术中立”并不意味着“价值中立”,有效防范公众被算法技术裹挟或操纵。此外,主流媒体等舆情主体还要引导网民积极理性地参与网络公共事件讨论,使其进行理性的舆论表达,促进社会共识的形成。

五、结语

总体而言,大数据和人工智能技术的兴起,使得网络舆情数据的生成和获取日益普遍,推动了网络舆情监测的全面化和自动化、助力了舆情研判和决策的科学化、实现了网络舆情引导的精准化和个性化,为社会治理提供了依据和基础。同时,大数据和人工智能技术的发展和应用还处于不断探索和发展中,由于存在跨学科的方法融合困境、忽视传统“小数据”的“大数据傲慢”、公众的媒介素养缺乏、舆情治理和数据管理体系不够完善等技术和社会层面的原因,网络舆情治理工作面临着新的风险与挑战。因此,我们需要推进网络舆情治理综合化技术发展,推动人机结合,将传统的舆情治理方法与大数据、人工智能方法相融合,建立综合性网络舆情治理模型;实现多元主体共同参与、协同治理,推进数据资源的交流共享,健全隐私伦理约束机制;培养精通大数据和人工智能技术、具备网络舆情治理数据素养、具备高度理想信念和熟悉宣传业务的复合型人才;加强媒体工作者和网络平台的伦理自律建设,提升公众的媒介素养和算法素养,从最大程度上凝聚共识,优化网络信息文明建设。

参考文献:

- [1] 安娜,林建成. 人工智能在网络舆情治理中的现实问题与应对策略[J]. 思想理论教育, 2020(12): 91-95.
- [2] 汪行东,胡志方. 大数据舆情技术缺陷与对策分析[J]. 今传媒, 2019, 27(11): 55-56.
- [3] 桑斯坦. 信息乌托邦:众人如何生产知识[M]. 毕竟悦,译. 北京:法律出版社, 2008: 8.
- [4] 范洁,张志丹. 人工智能时代意识形态工作面临的机遇与挑战[J]. 南通大学学报(社会科学版), 2020, 36(5): 1-8.
- [5] 邝岩. 大数据挖掘技术在网络舆情监测中的应用研究[D]. 北京:北京理工大学, 2016: 62-63.
- [6] 泰伯智库顾问委员孟小峰团队. 中国隐私风险指数分析报告[EB/OL]. (2020-04-27)[2021-03-29].

- <http://www.tiuchina.com/yjbg/346.html>.
- [7] LAZER D, KENNEDY R, KING G, et al. The parable of google flu: traps in big data analysis[J]. Science, 2014, 343(6176): 1203-1205.
- [8] 邹晓辉, 朱闻斐, 杨磊, 等. 谷歌流感预测——大数据在公共卫生领域的尝试[J]. 中华预防医学杂志, 2015, 49(6): 581-584.
- [9] 蒋成贵. 算法推荐模式下网络空间主流价值观建设[J]. 中国特色社会主义研究, 2019(3): 66-72.
- [10] 中华人民共和国网络安全法[EB/OL]. (2016-11-07)[2021-03-29]. http://www.gov.cn/xinwen/2016-11/07/content_5129723.htm.
- [11] 刘彩娥. 国内高校数据素养教育的问题与对策[J]. 北京工业大学学报(社会科学版), 2015(3): 76-80.
- [12] 王一芄, 卢晓云. 全媒体环境下新闻受众媒介素养研究——基于受众媒介使用行为的实证分析[J]. 视听, 2017(11): 16-17.
- [13] 张林. 智能算法推荐的意识形态风险及其治理[J]. 探索, 2021(1): 176-188.

Network Public Opinion Management in the Context of Big Data and Artificial Intelligence: Value, Risk and Path Exploration

LI Mingde^{1,2}, KUANG Yan³

- (1. Research Center for Intelligent Understanding of Communication Content, Xi'an Jiaotong University, Xi'an 710049, China;
2. School of Journalism and New Media, Xi'an Jiaotong University, Xi'an 710049, China;
3. School of Marxism, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: As a powerful means of information, big data technology and artificial intelligence technology are moving toward integration. They provide brand-new information resources, technical means and governance paradigms for the network public opinion management. And to a certain extent, they have realized the comprehensive, automatic, scientific, precise and personalized management of network public opinion. However, there still exist some challenges like deviations between the public opinion presented by the algorithm and the real public opinion, public opinion misjudging risks, false news, information cocoon rooms, public opinion misleading in decision-making, privacy ethics, data monopoly, and risks of data abuse in information management. In order to solve the challenges, some measures for improvement are concluded in the paper.

Key words: state governance; interdisciplinary fusion; network public opinion management; big data; artificial intelligence

(责任编辑:冯蓉)