

舆情研判的大数据赋能：风险与对策

李沛阳

西交利物浦大学, 江苏 苏州 215123

摘要：网络舆情是一种极具特殊意义的呈现形态，在大数据时代背景下彰显出多元特性。大数据技术的优势为研判和治理网络舆情提供较好的工具支撑，但是爆炸式增长的数据信息势必影响网络舆情研判的精准度。新时期网络舆情治理面临严峻的问题与考验，其治理路径包含：精细化舆情数据库建设、多元化研究方法融合、深化舆情工作者的培训活动等，旨在发挥其正向作用，提升舆情研判的精准度和成效。

关键词：舆情研判；大数据；对策

Big Data Empowerment for Public Opinion Research and Judgment: Risks and Countermeasures

Li Peiyang

Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123

Abstract： Online public opinion is a form of presentation with special significance, which manifests its multiple characteristics under the background of big data era. The advantages of big data technology provide better tool support for the research, judgment and management of network public opinion, but the explosive growth of data information is bound to affect the accuracy of the research and judgment of network public opinion. In the new era, the governance of online public opinion is faced with severe problems and tests. Its governance approaches include: the construction of fine public opinion database, the integration of diversified research methods, and the deepening of training activities for public opinion workers, so as to play a positive role and improve the accuracy and effectiveness of public opinion research and judgment.

Keywords： public opinion research and judgment; big data; countermeasures

一、大数据视域下网络舆情传播形态

大数据也可称之为海量数据，描述了一个系统，包含了庞大、复杂且难以通过现有技术获得和分析这些系统中所有信息的能力^[1-2]。大数据视域下的网络舆情传播态势呈现多元化特性：一是注重信息多元化，信息技术的发展势必会诱发网络舆情的变革，在此过程中会呈现多元态势。伴随人类思想观念的更迭，使得人们社会价值观念也显示不同样态，进一步促进网络舆情多样化。二是舆论信息开放化，基于大数据时代背景下的网络舆情更加开放，受网络互动和网络用户匿名等因素的影响，网民可随时在网络中发表自身的观点，直接造成网络舆情数据的爆发式增长。三是舆论信息传播范围广且传播速度快，围绕大数据背景下，网络舆情传播速度十分迅速，尤其是遇到部分突发事件，网民极易带入情绪发表言论，很容易刺激网民评论和转载传播，也会使得社会环境产生不稳定因素。四是在时代背景加持下，网络舆情更注重小微力量，在网络中，由一两段文字或是一张图片就会产生巨大的社会反响，更甚至会出现公共事件或是危机事件。

二、网络舆情大数据治理存在的问题

(一) 易受外界干扰，造成舆情研判精准度失真

相较于小数据而言，大数据技术以普查的形式收集了大量的

舆情信息，使得舆情状况变得更加全面、完整。然而，由于水军、社交机器人等虚假信息发布者的存在，这些原始数据的可靠性受到了一定程度的影响^[3-4]。噪声的存在会对舆情分析的结果产生负面影响，而且由于大量的重复数据，不仅增加了数据误差的风险，还极大地降低了分析的效率，使得技术赋能变得更加困难。随着人工智能技术的飞速进步，智能机器人已经变得越来越重要，它们可以更高效地制造、分享、转载各种信息。然而，由于智能机器人的灵活性以及强大的传播能力，也可用来制造出虚假或者进行恶劣的宣传。利用深度学习技术，机器人水军不仅能够超越传统的工具思考和属性，而且还能够从“类人”特征中提取出更多独到之处，比如，能够自动地分析和比较真实用户之间的互动，并且能够自动地将其中的某些观点转化成更多的内容，从而更好地控制和影响网络环境。隐性地改变着某些人的想法和看法^[5-6]。以机器人水军为例，虚假的信息泛滥成灾，占领着各种媒介，严重损害着网络舆论的秩序，削弱了公民的发声权。这些无关紧要的信息，使得大规模的谣传无从谈起，从而给政府的政治和经济管理工作造成了沉重的负担，并且给社会的公正和正义造成了潜移默化的影响。

(二) 信息获取界限模糊，大数据安全体系亟待完善

针对数据信息资源的整合，可以从网络平台中获取相关“原料”支撑，以便人们能够有效地收集、挖掘和分析各类数字资源。然而，在利用这些数据的同时，也必须建立完善的技术规范和法律保

障，以防止信息的安全性受到威胁。分布式爬虫技术的普遍使用，已经渗透到互联网的各个方面^[7-8]。然而，伴随将个人隐私信息转换成公开的，信息的安全范围也变得越来越不明确。由于技术“黑箱”的大数据算法的可解释性相当低，因此，管控机构无力实施对于收集和处置权限和安全范围的重要控制。随着科学技术的发展，网络舆情管理者需要更好地识别和管理各种数据，包括但不限于明确数据的公共性和可靠性、严格遵守相关法律法规、维护和管理个人信息的完整性。同时，由于云平台是一种重要的数据存储介质，因此也需要采取有效的措施来防止和解决可能出现的数据泄露和攻击。

（三）治理理念存在偏差，信息孤岛等不良倾向产生

科学技术的发展，作为舆情治理主体的政府部门，受益于“重技术、轻规律”的趋势，也更倾向于“信息孤岛”的变动，从而获得了较高的舆情治理和指导水平。然而，要想真正实施有效的舆情治理，就必须加深对信息科学发展趋势的理解，并采取有力措施，防止出现舆情失控的情况。随着科学发展的进步，大数据技术的发展已经超越了以往的范畴，它的关键是数据的可持续性和可用性，而目前，由于多种主客观原因，许多舆论管控机构仍然停留在数据的收集、传输、处置和交流的层级上^[9-10]。“信息孤岛”的现象源自于人们对于外部技术的依赖性，此种依赖性造成人们对于技术创新变革和多元化影响的认知缺失。过于崇拜技术会造成不同主体间失去自我意识，削减共识性原则，继而加剧不同群体间的问题点，不仅会对社会发展产生不良影响，也会丧失权威机构的公信力。

三、大数据赋能舆情研判的应对对策

（一）强化数据管理，精准把控舆情数据库建设

紧密围绕新时期发展规律来看，很多舆情都是结合话题内容呈现，且具备一定的空间发展规律。以实践探究为例，很多研究人员结合近十年来网络舆情话题演替设定探究内容，会明确其发展过程具备一定的发展规律。如，很多研究人员结合数据资料发现，社会民生话题、公民应有的权利、按劳分配制度不科学或是涉及民族自豪感、国家利益等人们关注的焦点话题^[11]。结合时空分布情况来看，很多话题中心多是春夏衔接或是年末这两个节点。由此看出，对于传统舆情案例的梳理状况而言，可以凝练出针对性治理举措。基于此，要想升华舆情应用的效能，要建构更为完备的数据资源库，不断提升这一案例应用效能。

围绕案例库建设情况来看，舆情案例库主要包含舆情传播溯源整合、舆情时空数据整合、话题热度和数据整合等内容。借助此类数据的分析与凝练，深度挖掘过往数据传播过程中无法捕捉的变量值，并在上述基础上建构数列模型、仿真模型等方法，继而直观呈现舆情传播的节点，能够精准对接和控制舆论主体，探究其不同阶段的发展规律，着重围绕相关事件的衔接性，提升这一舆情话题存在的风险度和价值诉求。

大数据舆情的呈现规律由多种表达形式，也依托于海量数据信息进行。只有着重把握这一媒介传播的数据，方可熟练的展示其发展态势。现行的数据资源多掌握在企业 and 政府相关部门手中，舆情监管机构也会因为技术壁垒和数据资源受限，不能顺利

呈现各类数据信息，无法及时控制舆情发展态势^[12-13]。要想打破这一困局，需要不同部门间营造数据意识，正确看待舆情数据信息的价值和深度内涵，同时要建构数据库断层，创设更为科学的数据舆情治理体系，着重强调不同主体间的协调一致。主要有以下路径实施：创设集散式监管平台系统，注重对不同类型数据的整合协调，将政府数据、企业数据和科研机构数据互惠互通，实现不同主体间的共建共享，有助于提升舆情研判的效能。

（二）创设多元方法，加强研究模式的融合与应用

大数据分析技术可为舆情治理工作提供更充足的优化路径和研究模式，使得人们对于舆情发展一般规律有更深刻的理解。舆情治理涉及多个领域，且在实践过程中要融合不同学科的内容，如政治学、新闻学、管理学等等，继而营造以舆情数据为核心，融合多元社会场景为基准的框架建构。对于大数据舆情分析和实践来看，要及时创新认知理念和模式，将不同类型的视角融汇贯通，逐步提升这一数据的应用效能。如同拉泽尔等人强调，区别于对大数据的关注度，要想掌握数据领域的深度变革，首先要懂得运用各类数据信息，以此凝练对社会各个领域的认知。现行的大数据舆情监管要结合海量数据分析进行，为了解社会民意存在一定的局限性，要注重强化与完善。

以往的问卷调查多是使用科学抽样的形式检测数据可信度，数据代表性中包含一定的优势，且对社会数据信息完成补充说明。此外，设置调查问卷内容时还可获取不同类型的主体，选取数据分析模式可完成数据汇总，判定不同主体间对于舆情的关注度^[14]。针对数据舆情的整合也可称之为一种社会现象，是对社会背景、社会发展动态的诠释，也可折射出背后的思想风潮，具备阶段性和多元特征。舆情治理工作也是针对社会心态的整合，应当紧密结合人类发展实际需求进行。有关人员使用大数据完成数据分析时，不仅要对此一舆情数据做出统筹，还要深度探究其掩藏在背后的认知理念和核心要素。以致于在后续的实践工作中，要运用设备完成数据文本分析，明确其存在的语言特性，主动融合民众意志、语言分析、和修辞解释等内容，探究舆情数据背后掩藏的文化内涵，明确这一话题的针对性和认知建构，更能把握这一主题的心态和诉求。

表1 网络舆情研判指标体系

网络 舆情 研判 指标 体系	事件源 状态	参与人	直接当事人	影响力
				诉求目的
				所处社会状态
		相关参与者	影响力	
			诉求目的	
			所处社会状态	
	事件 影响力	事件等级	安全等级	
			事故等级	
		地点敏感程度		
		时间敏感程度		
发展趋势	引发衍生事件可能			
	扩大可能性			

网络舆情研判指标体系	事件源状态	事件现状	网页浏览量	
			网上跟帖量	
			网上转载量	
			网站发帖量	
	媒体介入状态	网上传播状态	报道的网站数量	
			发布的文章数量	
		网络媒体质量	主流权威媒体	真实程度
				全面程度
		非主流媒体	非主流媒体	真实程度
				全面程度
高校师生参与状态	师生态度	情感倾向分析		
		认知情况分析		
	主要分布地域分析			

(三) 融合人才培养，着力对舆情工作者的培训

网络舆情源头、潜在风险评定、社会网络生态、制度环境等内容分析，都离不开人力的操作。要想做出科学的判定，需要建构更为缜密的思维模式，将舆情理论、数据资源、科学判定等内容有效衔接。针对舆情数据的分析人员在其判定过程中也会有这

不可忽视的作用，从2013年底伊始，舆情分析人员要主动参与以国家为主导的培训活动，并获取相应的《网络舆情分析师职业培训合格证》，但是此项内容并非强制执行。结合舆情分析人员的工作特性，可以有效培养其个人能力和综合素养水平，助力其全面健康发展，在培训过程中也要遵循差异性原则。部分研究者指出，舆情研判人员队伍建设过程中，不仅要鼓励博学多才，还要着重术业专攻^[15]。总而言之，在精湛的实践能力和敏锐的观察视角下，舆情工作者方可做出正确评定。

四、结语

综上所述，围绕大数据背景下，网络中无时无刻均是人们在生活和工作中的活跃，此外，网络信息数据、各类传播途径和传播速度都产生巨大变革。政府有关部门要设定更为全面的监管体系，着重加强对社会内部舆情信息的监管力度，主要包含对各类舆情数据分析、整合和归纳指引等多个方面。面对可能出现的网络舆情或是危机事件做出科学研判，继而选取适宜的处理办法，有助于维护社会秩序的稳定性和不断提升网络安全效能。

参考文献

[1] 巢乃鹏, 吴兴桐, 黄文森, 等. 计算传播学研究现状与前沿议题 [J]. 全球传媒学刊, 2022(1): 19-40.
 [2] 冯艳茹. 基于 Python 的网络爬虫系统的设计与实现 [J]. 电脑与信息技术, 2021(6): 47-50.
 [3] 钟佳娃, 刘巍, 王思丽等. 文本情感分析方法及应用综述 [J]. 数据分析与知识发现, 2021(6): 1-13.
 [4] 荣婷. 机器人水军的演进、问题与治理研究 [J]. 中国出版, 2021(13): 72-75.
 [5] 李彪. 舆论学教程 [M]. 北京: 中国人民大学出版社, 2020: 119.
 [6] 曾润喜. 网络舆情治理的关键是“治未病” [J]. 中国传媒科技, 2018(12): 1.
 [7] 臧海平. 舆情监测和分析的难点及对策 [J]. 青年记者, 2019(19): 17-18.
 [8] 姚翼源. 人工智能时代政府网络舆情治理的逻辑、困局与策略 [J]. 西南民族大学学报(人文社科版), 2020(3).
 [9] 彭波, 张权. 中国互联网治理模式的形成及嬗变(1994—2019) [J]. 新闻与传播研究, 2020(8).
 [10] 王仕勇. 大数据时代的社会舆情治理: 何以可能与何以可为 [J]. 重庆社会科学, 2021(12).
 [11] 张志安, 聂鑫. 互联网平台社会语境下网络内容治理机制研究 [J]. 中国编辑, 2022(5).
 [12] 郑晨予, 范红. 从社会传染到社会扩散: 社交机器人的社会扩散传播机制研究 [J]. 新闻界, 2020(3).
 [13] 荣婷. 机器人水军的演进、问题与治理研究 [J]. 中国出版, 2021(13).
 [14] 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).
 [15] 夏德元, 刘博. 智媒时代网络内容生态治理与编辑的职业使命 [J]. 中国编辑, 2022(5).