

Analysis of Theme Characteristics and Emotional Tendency in Weibo Discourse of People with Autism Spectrum Disorder

Lin Meizhen¹, Ma Bosen², Zhang Yongwei³

(1. International College, Fujian Agriculture and Forestry University, Fuzhou, 350002; 2. School of International Studies, Zhejiang University, Hangzhou, 310058; 3. Institute of Linguistics, Chinese Academy of Social Sciences, Beijing, 100000)

Abstract: Weibo serves as an important platform for the general public to obtain information about people with autism spectrum disorder. This study focused on 33 popular Weibo trends related to autism during the period of 2020-2022, and employed a combination of LDA topic modeling and sentiment analysis to explore the attitudes and foci of Weibo users towards individuals with autism. The findings showed that Weibo users focused on five major topics: social care for individuals with autism, medical knowledge dissemination, educational and rehabilitation, parents of children with autism, and autistic savants. In terms of sentiment analysis, Weibo users exhibited overwhelmingly positive attitudes towards individuals with autism, with very few instances of neutral or negative sentiments. The image of the autism population is mainly portrayed as positive and optimistic. The findings provide insights into understanding the mechanism of autism information dissemination on social media and practical implications to the online health communication of autism on Weibo platforms.

Keywords: people with autism spectrum disorder; Weibo discourse; text mining; sentiment analysis; topic modeling.

自闭症人群微博话语中的主题特征及情感倾向分析¹

林美珍¹, 马博森², 张永伟³

(1. 福建农林大学国际学院 福建福州 350002; 2. 浙江大学外国语学院 浙江杭州 310058; 3. 中国社会科学院语言研究所 北京 100000)

摘要: 微博是社会大众获取自闭症人群信息的重要场域。本研究以 2020-2022 年新冠疫情期间与“自闭症”相关的 33 个微博热搜为研究对象, 采用 LDA 主题建模和情感分析相结合的方法, 探究微博社交媒体对自闭症人群的态度及关注焦点。研究发现, 微博用户关注主题涉及自闭症人群的社会关爱、医学科普、教育康复、自闭症家长以及自闭症天才等 5 大主题。从情感分析倾向上看, 微博用户对自闭症人群的态度呈现出一边倒的积极情感倾向, 鲜有出现中立或消极态度。被建构的自闭症人群形象主要以积极、正面为主。研究结果有助于加深对社交媒体上自闭症信息传播机制的理解, 并为自闭症的网路健康传播提供启示。

关键词: 自闭症人群; 微博话语; 文本挖掘; 情感分析; 主题建模

一、引言

自闭症(又称孤独症)谱系障碍(Autism Spectrum Disorder (ASD),以下简称自闭症)作为一种神经发育障碍,通常起病于儿童早期,以社会交往和沟通障碍,兴趣范围狭窄以及重复刻板行为为主要临床特征¹。近年来,自闭症谱系障碍发病率呈上升态势,仅 2012-2018 年,美国自闭症谱系障碍发生率的增幅超过了 56.8%²。自闭症在世界范围内影响巨大,全球至少有 7800 万人罹患自闭症,其中大多数没有获得适当的医疗、教育和社会支持³。在中国,自 20 世纪 80 年代南京脑科医院陶国泰先生首次报告 4 例婴儿自闭症以来⁴,目前我国的自闭症群体规模已经达到 1000 万,并以每年 10 余万的速度递增⁵。自闭症患病率的不断攀升给社会和家庭带来沉重的经济和精神压力,是影响人口健康的重大问题之一。而自闭

¹

本文系为国家社科重点项目:“汉语自闭症儿童和正常儿童的多模态指称策略对比研究”(项目编号:17AYY009)的研究成果之一。

¹ American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders (5th ed.)[C]. Washington, D.C.: Author.

² Centers for Disease Control and Prevention, 2021. Autism spectrum disorder: What is ASD? <https://www.cdc.gov/ncbddd/autism/facts.html>

³ Lord, C, Charman, T, Havdahl, A, et al., 2022. The Lancet Commission on the future of care and clinical research in autism[J]. Lancet, 399(10321): 271-334.

⁴ Tao, K. T, 1982. Diagnosis and classification of infantile autism[J]. Chinese Journal of Neurology and Psychiatry, (15):104-107.

⁵ 五彩鹿自闭症研究院, 2017. 中国自闭症教育康复行业发展状况报告(II) [C]. 北京:北京师范大学出版社.

症人群作为社会弱势群体中的一部分，经媒体报道影响了社会大众对其认知、评价和话语建构。因此，正确、有效地传播和建构自闭症人群媒体话语，讲好自闭症人群故事，展现真实、可信、多元的自闭症人群形象，不仅关涉社会大众对自闭症人群形象的建构和认知，也影响社会大众对自闭症人群的态度和行为。

报纸、电视、电影等传统媒体对自闭症人群的媒介呈现研究已经取得了一些有影响力的成果。学者们针对不同的媒体类型，采用传播学的内容分析法，考察了自闭症人群的媒体形象、媒体框架变迁以及自闭症议题的报道特征与报道策略²。其中围绕自闭症人群的媒体形象研究相对集中，且自闭症人群的媒介呈现往往刻板、消极。如 Jones & Harwood 研究表明，传统纸质媒体通过选择、加工和再传播等手段建构出危险的、不可控制的、不被爱的、不能治愈的自闭症人群媒体负面形象⁶，这不仅影响了社会大众对自闭症人群的认识、态度和行为，而且使社会大众对这类人群产生更多的误解、偏见和歧视。然而，随着社交媒体时代的到来，有关自闭症人群媒介呈现可能完全有别于传统媒体，但目前我们对其缺乏必要的关注。Mittmann 等人对 2010-2022 年间 PubMed 和 PsychInfo 数据库中收录的 31 篇相关文献进行梳理后发现，仅有 4 篇文献涉及社交媒体（如 YouTube, Twitter 等）上的自闭症人群话语研究⁷。与国外研究相似，国内研究也多采用传统的纸质媒体语料⁸⁻⁹，采用社交媒体如微博、快手、抖音等短视频媒体语料较少。尽管近期出现了两项涉及微博¹⁰和快手短视频¹¹的代表性研究，但相关研究甚少，且尚未发现采用 LDA（Latent Dirichlet Allocation）主题建模和情感分析方法来探讨自闭症人群的相关研究。

“人工智能导向的自然语言处理技术、大数据文本分析技术的迅猛发展给话语分析带来了挑战与发展空间。深度学习技术、情感分析、词向量技术、主题建模等信息技术正在进

⁶ Jones, S. C, & V. Harwood, 2009. Representations of autism in Australian print media[J]. *Disability & Society*, 24(1): 5-18.

⁷ Mittmann, G, Schrank, B. & V. Steiner-Hofbauer, 2023. Portrayal of autism in mainstream media – a scoping review about representation, stigmatization and effects on consumers in non-fiction and fiction media[J]. *Current Psychology*, <https://doi.org/10.1007/s12144-023-04959-6>.

⁸ 杨叶, 2014. 自闭症群体媒体形象建构研究[J]. *青年记者*, (34): 16-17.

⁹ 李莹, 2017. 中国大陆与台湾地区自闭症患者媒介形象研究——以《羊城晚报》、《联合晚报》为例[D]. 长春: 吉林大学硕士论文.

¹⁰ 黄月琴, 黄宪成, 2021. “转发”行为的扩散与新媒体赋权——基于微博自闭症议题的社会网络分析[J]. *新闻记者*, (5): 36-47.

¹¹ 章沫嘉, 黄月琴, 2021. 作为健康传播平台的短视频: 基于快手自闭症内容的探索性分析[J]. *传播创新研究*, (1): 89-109.

¹² 金碧希, 卫乃兴, 2023. 话语研究的语料库路径: 方法、挑战与前景[J]. *外语与外语教学*, (1): 1-11.

入语言研究”¹²。鉴于文本挖掘技术领域下的主题建模能归纳性地挖掘出海量文本语料背后潜藏的主题结构和词语搭配的深层关系³，且能弥补话语分析的主观性、缺乏系统性和操作性等问题¹³⁻¹⁴；而情感分析可用于探究文本态度¹⁵，且能更精细地展现节点词共现语境的态度意义¹⁶。基于此，本文采用 LDA 主题建模和情感分析相结合的方法，以 2020-2022 年新冠疫情期间与“自闭症”相关的 33 个微博热门搜索（以下简称热搜）词条下的评论部分为研究对象，探究以微博为代表的社交媒体对自闭症人群的态度及关注焦点，揭示微博话语背后隐含的意识形态和社会文化成因，旨在为自闭症的网路健康传播提供借鉴和参考。

二、研究方法

1、数据采集

本研究数据来源于新浪微博（以下简称微博）。微博自 2009 年上线后已逐渐成为国内用户活跃量最大的社交媒体平台之一。截至 2023 年 3 月 31 日，微博月活跃用户数已达到 5.93 亿，日活跃用户数为 2.55 亿，是具有用户群体庞大以及高传播影响力的重要社交媒体平台，也是社会大众获取自闭症人群信息的重要场域。正如知乎热榜、微信 24 小时、百度实时热搜等其他热搜板块一样，微博平台上也设置了热搜板块。微博用户依据平台热度榜单搜索特定话题获取信息。因此，微博热搜蕴含大量用户行为、用户关注热点以及情感态度等信息，具有传播学中所说的议程设置功能。“上热搜”已成为社会大众判断话题关注度的重要依据¹⁷。基于此，我们从微博“热搜引擎数据库”（<https://m.weibo.cn/>）爬取了 2020 年 1 月 1 日至 2022 年 12 月 31 日所有包含“自闭症”和“孤独症”为关键词的热搜，采集的元数据包括用户名称、用户文本、转发数、评论数以及发布时间等。共检索到相关热搜 43 条，其中有 10 条热搜在移动端上的评论被隐藏或被过滤，最终确定 33 条热搜为本研究的有效分析样本；随后，我们采用 Python 技术爬取每条热搜下核心报道的微博用户评论，经筛选、去重、去噪等处理后，共得到 21451 条有效数据。采用微博用户评论数据的原因在于，微博热搜评论行为蕴含着传播主体对传播内容的观点、态度和意见。

¹³ 刘文宇, 胡颖, 2020. 基于文本挖掘的非传统文本批评话语研究[J]. 天津外国语大学学报, (4): 29-41.

¹⁴ Jacobs, T. & R. Tschötschel, 2019. Topic models meet discourse analysis: a quantitative tool for a qualitative approach[J]. International Journal of Social Research Methodology, 22(5): 469-485.

¹⁵ Liu, D. & L. Lei. 2018. The appeal to political sentiment: An analysis of Donald Trump's and Hillary Clinton's speech themes and discourse strategies in the 2016 US presidential election [J]. Discourse, Context & Media, (25): 143-152.

¹⁶ 李晶洁, 胡奕阳, 陶然, 2022. 基于情感倾向分析的语义韵强度算法探析[J]. 外国语, (5): 65-74.

¹⁷ 刘胜枝, 2020. 微博热搜的价值、问题与完善[J]. 人民论坛, (31): 100-102.

2、数据分析

首先,对 33 条微博热搜评论进行 jieba 中文分词处理,使用主流的三大中文停用词库(哈工大、川大、百度)对分词后的博文进行停用词(stop words)处理,然后在 Python 中实现基于 LDA 模型的文本主题聚类,挖掘相关的主题特征词。经由参数调适,抽取主题分布及关键词,以明确微博用户所关注的内容焦点;其次,采用 HarvestText 情感分析技术对每条微博热搜评论进行情感倾向分析,分析结果采用 Matplotlib 生成可视化图形,以明确微博用户对自闭症人群的态度;最后,基于以上量化分析结果,对比、剖析和提炼微博用户对于自闭症人群的评论态度与关注焦点,并在此基础上总结微博话语建构的自闭症人群形象。

三、 结果与讨论

1、分布趋势

为了探究微博热搜评论话语产生的重要节点,以及社交媒体对孤独症人群的关注趋势,本研究统计了 2020-2022 年自闭症人群微博热搜评论数量(见图 1)。

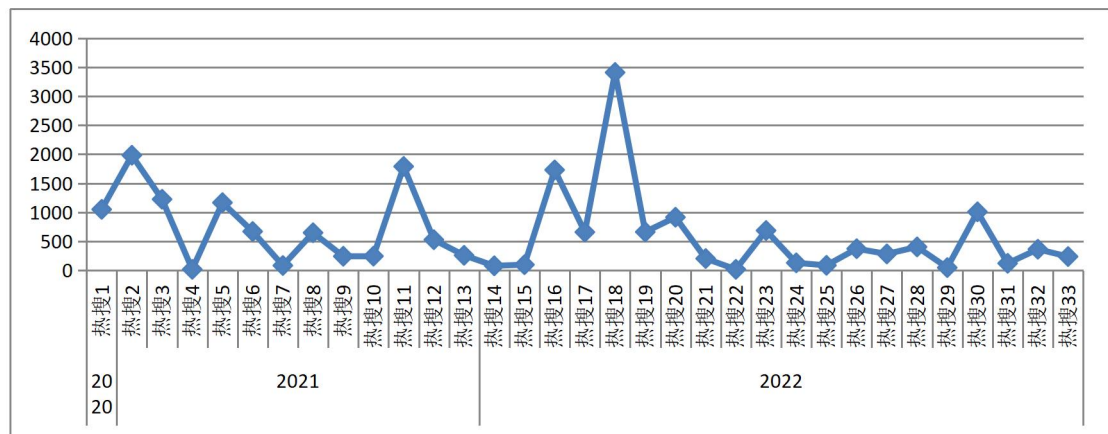


图 1 自闭症人群微博热搜评论数量分布情况(2020-2022)

如图 1 所示,2020-2022 年间,自闭症人群微博热搜评论数量呈现波动上升趋势,并伴有周期性上下波动,各年度热搜评论数量呈现骤然升降或时现断点的分布态势。从热搜数和评论数分布来看,2020 年仅有 1 条热搜关注自闭症人群,评论数为 1051 条(占总数 5%)。究其原因,可能源于 2020 年是全球范围内受 COVID-19 大流行影响最为严重的一年。媒体资源和公众关注度主要集中在与新冠疫情相关的新闻报道上。加之,媒体通常更倾向于报道与争议、突发事件或社会热点相关的话题。与 COVID-19 突发公共卫生事件相比,2020 年针对自闭症人群的报道可能没有足够的新闻价值或争议性,故出现媒体关注度明显不足的现象。

统计数据也显示,从2021年开始,关注自闭症人群的微博热搜数量不但增加,且热搜下的评论数量也明显增多,其中,2021年共出现12条热搜,评论数为8867条,占总数的41%;而2022年微博热搜数为20条,评论数为11533,占总数的54%。这说明,在疫情常态化发展的背景下,有关自闭症人群在社交媒体上的话语愈来愈受到关注。同时,相关研究也表明¹⁸,疫情对特殊人群的影响会更大⁴,因为,疫情使其生存状态面临更大的挑战和困境,例如社交隔离、日常活动的变化,以及线上康复训练的适应性问题等,这些因素共同引发了公众对自闭症人群的关注和关爱。值得一提的是,通过对语料进一步分析后发现,微博热搜评论数量在“世界自闭症关注日”的前后一周(3月26日-4月9日)出现“扎堆”现象,共出现10564条,占本年度评论总数的92%。其中,4月2日的“世界自闭症关注日”就出现了7条热搜,微博评论数达到了6028条,约占本年度一半有余(52%)。这可能是每年一度的“世界自闭症关注日”,对于孤独症患者及其家庭来说,这是他们能够获得社会关注和资源支持的“高光时刻”。因此,与传统主流媒体报道一样,社交媒体报道也存在较鲜明的“应景”特征和“节日化”报道趋向,缺乏主动的持续性关注议题¹⁹⁻²⁰。

2、主题建模分析

主题建模是一种文本分析技术,用于从一系列给定的文档集合中提取文本的话题或者主题。主题建模挖掘出的主题可以帮助人们理解海量文本背后隐藏的语义²¹。对高频主题的分析能揭示微博用户所关注的焦点及其建构的自闭症人群形象。我们综合考量所分析数据的实际情况后,确定主题词数量为5,进行主题挖掘。根据主题词提取结果,33个微博热搜,最终可归纳为5类主题,每个主题下列出了该主题文本中出现频次最高的10个关键词(见表1)。为验证主题词数设定是否为最优识别结果,我们还使用了词云软件制作了博文评论词云图(见图2)。

¹⁸麻敏洁,敖勇前,申承林,王庭照,2020.新型冠状病毒肺炎疫情影响下特殊儿童正负性情绪及居家干预调查[J].中国特殊教育,(4):3-12.

¹⁹Bie B. & L. Tang, 2015. Representation of autism in leading newspapers in China: A content analysis[J]. Health Communication, 30: 884-893.

²⁰林美珍,马博森,2020.基于语料库的自闭症人群媒体话语建构特征研究[J].语料库语言学,(2):13-27.

²¹张培晶,宋蕾,2012.基于LDA的微博文本主题建模方法研究述评[J].图书情报工作,(24):120-126.

表 1 自闭症人群微博热搜评论的关键词分布及主题类别

主题	关键词	主题类别
1	关注、社会、关爱、爱意、包容、理解、接纳、守护、马嘉祺、少年时代团	社会关爱
2	妈妈、父亲、爸爸、父母、家庭、陪伴、照顾、悲伤、可怜，苦涩	自闭症家长
3	星星、孩子、儿童、少年、患者、音乐、天才、天赋、可爱、善良	自闭症天才
4	医生、治疗、科普、医院、疾病、治愈、研究、发育、障碍、诊断	医学科普
5	康复、干预、治疗、机构、教育、学校、幼儿园、特教、老师、方法	教育康复



图 2 自闭症人群微博热搜评论词云图

从表 1 中博文主题聚类的关键词结果及可视化词云图 2 来看，微博用户所关注的主题可以概括为以下 5 类。（1）社会关爱（针对自闭症患者的救助行动、志愿者服务、捐赠活动等），如关注、关爱、包容、接纳、守护等关键词；（2）自闭症家长（自闭症患者的主要养育者），如妈妈、父亲、父母、陪伴、照顾等关键词；（3）自闭症天才（具有某些领域非凡天赋及才能者）如儿童、少年、音乐、天才、天赋等关键词；（4）医学科普（自闭症的病因、症状、诊断、治疗以及相关科学研究前沿进展等），如治疗、科普、医院、疾病、研究等关键词；（5）教育康复（自闭症的教育方法、康复技术、康复资源和支持等），如康复、干预、治疗、机构、教育等关键词。基于主题分析结果我们发现，自闭症人群在微博话语中被建构为被关爱者、家庭的负担、天资异禀者以及亟需医疗干预的病人形象。对语料进一步分析后发现，社会关爱类主题是微博用户的关注焦点，而医学科普类和教育康复类主

题涉及较少。⁵这与 Kang 考察电视新闻中自闭症议题的报道的研究结果相似²²，即传统媒体对自闭症干预措施、科普信息等关注较少。而本研究的结果可能源于所采用的语料为新冠疫情期间的报道。媒体和公众的关注焦点更集中在与疫情相关的问题上，如病毒的传播途径、症状、防护措施等。此外，从信息传递的优先级来看，疫情期间这 2 类主题可能被认为优先级较低，故微博用户关注度较低。限于篇幅，下文仅讨论前 3 类主题。

如上所述，社会关爱类主题（共有 12 条热搜，占总数的 36%）是微博用户的关注焦点。微博用户倾向于通过平台分享和传播关爱信息，并呼吁社会大众对这个特殊群体给予更多的关注和支持，因为疫情期间的社交隔离、封锁和限制措施可能使得自闭症人群的日常生活受到更大的影响和打击，他们可能面临着更加严重的社交障碍和心理压力。主题 1 中的关键词显示，“关爱”（859 次）和“关注”（441 次）高频出现，体现了微博用户呼吁关注和关爱自闭症群体的倾向，尤其在自闭症患病率逐年上升的背景下，呼吁社会的关注和关爱显得尤为重要。例如，主题 1 中的“马嘉祺”和“少年时代团”关键词来源于热搜 16 “#时代少年团呼吁关爱孤独症儿童#”。时代少年团作为公众关注度较高且为青少年偶像团体，通过呼吁“一起关注并守护星星的孩子”。他们的行动和言论引发明星粉丝团体的转发（28,213 万次）和点赞（5.2 万次），不但增加了自闭症信息的传播热度和广度，而且引发了公众对于自闭症的关注，提高了社会大众对于自闭症的认知度，从而减少对自闭症人群的误解和歧视。在某种程度上，社会观念态度的改变也有助于自闭症家庭获得更多的关注和支持，这对于创建包容有爱的社会支持体系有所裨益。然而，需要注意的是，明星公益慈善活动虽然在自闭症的宣传普及方面发挥着重要作用，但我们也必须思考：这些明星公益行为是否仅仅是商业策略，用于铺垫明星的道路或塑造明星形象？粉丝的转发和点赞行为是出于对自闭症人群的真正关注和关心，还是仅仅对明星本身的追随和关注？又或者，转发行为是否只是一种群体仪式，旨在从中获取情感和道德资源，以加强明星与粉丝之间的联系和群体凝聚力？对于这些问题的回答，还需要进一步的质性材料分析²³。

表 1 显示，在新冠疫情大流行期间，自闭症家长（共有 7 条热搜，占总数的 21%）也成为了微博用户重点关注的对象。关键词如妈妈（422 次）、父亲（117 次）、爸爸（212 次）、父母（224 次）、家长（184 次）等频繁出现。这可能源于关注自闭症家长和家庭可以了解到他们在这个特殊时期面临的困境和需要，以便提供更好的支持和帮助。同时，微博作为一

²² KANG S, 2013. Coverage of autism spectrum disorder in the US television news: An analysis of framing [J]. *Disability & Society*, 28(2): 245-259.

²³ 黄月琴, 黄宪成, 2021. “转发”行为的扩散与新媒体赋权——基于微博自闭症议题的社会网络分析 [J]. *新闻记者*, (5): 36-47.

个社交媒体平台，提供了信息共享和互助的渠道。疫情期间，自闭症家长可能面临着更大的压力和挑战，他们通过微博分享经验、获取支持和寻求帮助。²⁴微博作为一个非常重要的赋能平台，为孤独症家长提供信息分享、喘息服务、获得共情感受等自闭症家庭迫切需要的服务内容²⁴。语料分析还发现，与传统媒体一致，微博社交媒体也倾向于报道自闭症母亲（如热搜 8 “#妈妈在校打工 8 年陪护自闭症双胞胎#”）。这主要是因为自闭症患者的康复治疗周期漫长，自闭症母亲在照顾自闭症子女的过程中承受着巨大的经济和心理压力，同时也面临着更多的挑战和困境。这种经历更容易引起社会大众的共情和同情，因此媒体倾向于将其作为“卖点”进行传播。尤其在新媒体传播语境下，信息的传播和分享通常是基于用户的兴趣和关注度。值得一提的是，语料还显示，微博平台也倾向于报道自闭症父亲的故事，如热搜 11“#孤独症患儿父亲的 63 封家书#”和热搜 22“#父亲照顾自闭症女儿六年拒要二胎#”，这与传统媒体倾向报道自闭症母亲的研究结果不太一致²⁵。微博社交媒体采取平衡地报道自闭症父亲的故事有助于打破性别角色刻板印象，从而使公众意识到父亲在自闭症子女的抚养和陪伴中扮演的积极角色。此外，Bie & Tang 研究也表明，中国传统主流媒体常将家庭(家长)责任置于新闻报道的焦点²⁶。本研究语料分析也印证了这一结果。需要注意的是，微博用户倾向于采用柔性叙事方式讲述自闭症患者家庭故事，这虽能激发起公众心理层面上的共情机制，强化其情感认同，“但这可能导致公众认为自闭症是家庭的责任和负担，而忽视了社会对自闭症患者和家庭的支持和理解的重要性”²⁶。因此，媒体报道时还应该强调社会支持的重要性，减少对家庭的过度强调，以推动社会的包容和支持。

表 1 也显示，自闭症天才主题（共有 5 条热搜，占总数的 15%）位列第三，关键词如音乐（138 次）、天才（120 次）等多次被提及。这可能是由于少数自闭症患者在音乐、绘画等领域展示出非凡的天赋和才能，如热搜 30 “#4 岁孤独症男孩的绝对音感#：盲弹曲子准确率接近 100%”、热搜 12 “#3 岁自闭症男孩自学钢琴#：不识乐谱会弹几十首曲子”以及热搜 7 “#自闭症男孩学钢琴 15 年登上国际舞台#：钢琴老师陪伴自闭症男孩子 10 年学音乐”。此外，自闭症天才的故事通常涉及克服困难、追求梦想和取得成功的励志元素，这些故事能够引起公众的好奇和兴趣，激发人们的情感共鸣，故微博平台倾向于报道自闭症天才的故事，

²⁴ 中国孤独症家庭状况社会调查项目组, 2023. 中国孤独症及神经发育障碍人群家庭现状、需求及支持资源情况调查报告[C]. 北京: 北京大学医学出版社.

²⁵ Baroutsisa, A, Eckert, C, Newmanc, S. & D. Adamsd, 2023. How is autism portrayed in news media? A content analysis of Australian newspapers articles from 2016–2018[J]. *Disability & Society*, 38(6): 963–986

²⁶ Bie B. & L. Tang, 2015. Representation of autism in leading newspapers in China: A content analysis[J]. *Health Communication*, 30: 884–893.

以吸引更多用户的关注和分享。²⁷Bie & Tang 研究也发现，中国主流报刊媒体倾向于将自闭症患者描绘成具有特殊才能的天才。这种描绘可能导致公众对自闭症产生误解，认为所有自闭症患者都具有特殊能力，而忽视了自闭症患者的广泛个体差异。实际上在自闭症患者群体中，多数都存在语言迟缓、智力功能低下，生活难以自理等问题，而拥有特殊天赋的人仅占 10%²⁸。故需要说明的是，“无论是影视作品塑造的天才形象，还是大众媒体宣传的“来自星星的孩子”，都可能是一种对自闭症患者的过度美化。虽然这样做的初衷是为了提高社会对患者的包容和支持程度，但是过度美化的结果却容易使人们的关注焦点发生偏转。人们通常会更多地记住经过艺术加工之后的患者形象，而忽略了真实的患者模样，以及他们真正面临的痛苦与折磨”²⁹。

3、情感倾向分析

情感分析也被称为意见挖掘（Opinion Mining）或者倾向性分析，旨在识别和提取文本中的情感倾向或情感态度，以了解文本作者的情感、意见、态度或情绪。本研究的情感分析工具采用情感词典方法进行情感分析。通过提供少量标准的褒贬义词语（即“种子词”），从语料中自动学习其他词语的情感倾向，形成情感词典。对文本语料中情感词的加总平均则用于判断文本语料的整体情感倾向。本文选择 HarvestText 工具内置的通用情感词典作为情感“种子词”。按照通常做法，本文将情感分析指标分为消极、中立、积极三大类，分别用-1、0、1 表示。根据程序设置，情感综合指数大于 0 为积极情绪，代表肯定、赞美的褒扬态度；等于 0 为中立情绪，代表没有明显的态度倾向；而小于 0 则为消极情绪，代表批评、否定的反对态度。另外，正值的指数越靠近 1 表示越积极，负值的指数越近-1 表示越消极。图 3 展示了分析结果。

²⁷ Bie B. & L. Tang, 2015. Representation of autism in leading newspapers in China: A content analysis[J]. Health Communication, 30: 884–893.

²⁸ Huws J. C. & R. Jones, 2010. Missing voices: representations of autism in British newspapers, 1999–2008[J]. British Journal of Learning Disabilities, (39): 98–104.

²⁹ 李昂, 2023. 警惕对自闭症的污名化和过度美化[J]. 中国科技财富, (4) :45–46.

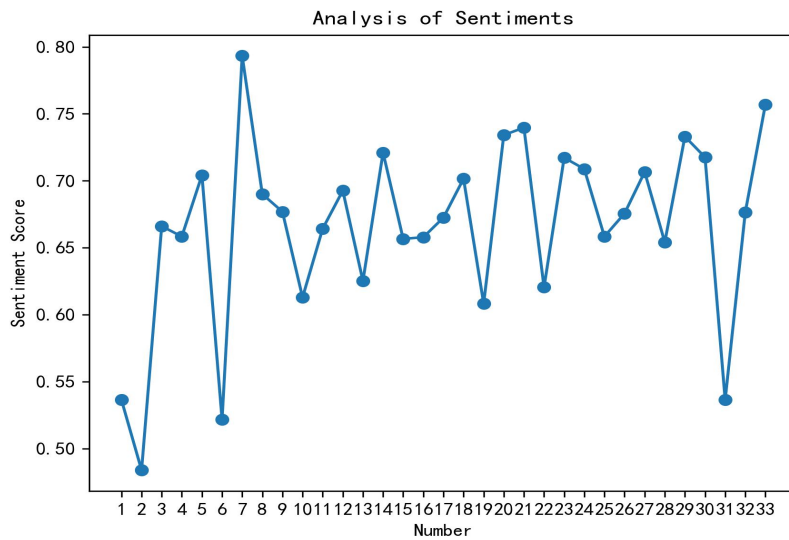


图 3 孤独症人群微博热搜评论博文情感值分布情况

从图 3 我们不难发现，32 条微博热搜的情感综合指数均在 0.5 以上，热搜 7 的情感综合指数最高（0.793365），仅有 1 条热搜（热搜 2，0.483892）的情感综合指数在 0.5 以下，说明微博用户对自闭症人群的态度呈现出一边倒的积极情感倾向，鲜有出现中立、消极态度。⁸换句话说，“用户反应呈现“正面”“积极”“关怀”的情感特征，表现了极强的共情性”³⁰。被建构的自闭症人群形象主要以积极、正面为主。究其原因，可能是因为微博作为一个相对匿名的社交媒体平台，用户可以自由发表观点和表达情感，而不受传统媒体的编辑和审查限制。这种匿名性和自由度使得微博成为了一个传达包容和支持声音的渠道，从而促使用户更倾向于积极地对待自闭症人群。

由于微博热搜评论部分是社会大众表达支持和积极情感的场域，且微博热搜评论信息的情感极性往往与其所要表达的主题密不可分，故再次对主题建模分析结果进行分析后发现，微博社交媒体对自闭症人群体现了正面、积极的态度，并以积极、正面的方式传递关爱自闭症人群。如呼吁社会“用接纳代替漠然，用良善驱散冰冷，让来自“星星的孩子”不再孤单”（热搜 28 “#你好，明天#”）、“多给他们一份了解、包容和关爱”（热搜 20 “#大龄孤独者如何融入社会#”）等。此外，涉及以自闭症音乐天才为代表的热门话题时，微博用户倾向于表达对自闭症孩子才艺特长的称赞和祝福，如“上帝为你关上了一扇门，却又为你打开了另一扇窗”、“希望孩子可以通过音乐走出自己的道路”（热搜 30 “#4 岁孤独症男孩

³⁰章沫嘉, 黄月琴, 2021. 作为健康传播平台的短视频: 基于快手自闭症内容的探索性分析[J]. 传播创新研究, (1): 89-109.

的绝对音感#：盲弹曲子准确率接近 100%”) 等。由此可见，微博用户对自闭症人群的评价是正面的，且态度呈现出积极情感倾向，究其缘由可能是微博热搜存在平台算法和内容过滤机制，倾向于筛选和推荐积极情感的评论，吸引更多的用户参与讨论、表达积极情感，以此维护积极的网络氛围和用户体验。加之，在微博新媒体传播语境下，每一位微博用户都可以是信息的生产者、转发者或者信息的接收者。而微博的内容在一定程度上也影响了接收者的回复、转发和评论。此外，社交媒体上的积极情感的评论也可能更容易引发他人的关注、转发和点赞。以点赞量为例，热搜 16 “#时代少年团呼吁关爱孤独症儿童#、热搜 11 “#父亲照顾自闭症女儿六年拒要二胎#、热搜 32 “#陌生人带走孤独症男孩同学上前阻拦#：其实是特教老师” 点赞量位居前三甲，分别为 5.2 万次、3.3 万次和 2.7 万次。说明微博用户更倾向于表达支持、鼓励和积极的情感，以符合社会期望和获得他人的认同，从而形成积极的互动循环。

四、 结语

微博社交媒体作为大众获取自闭症信息的重要途径，在传播自闭症知识、培养公众认知方面发挥着关键作用。本研究采用 LDA 主题建模和情感分析的方法，分析了以微博为代表的社交媒体对自闭症人群的关注焦点及其态度。LDA 主题建模结果显示，微博用户关注主题涉及自闭症人群的社会关爱、医学科普、教育康复、自闭症家长以及自闭症天才 5 大主题。自闭症人群在微博话语中被塑造为被关爱者、家庭的负担、天资异禀者以及亟需医疗干预的病人形象。情感倾向分析表明，微博用户对自闭症人群持有一边倒的积极情感。中立或消极态度的表达很少。被建构的自闭症人群形象主要以积极、正面为主。这一发现有助于重新审视和解读社交媒体形塑的自闭症人群形象，为未来优化自闭症信息传播的效果以及制定自闭症的网络健康传播策略提供有益的启示。

2016 年 10 月出台的《健康中国 2030》规划纲要强调“各级各类媒体应加大健康科学知识宣传力度，利用新媒体拓展健康教育”。在当前自闭症成为全球面临的严重公共卫生议题的背景下，如何消弭对自闭人群的刻板印象，讲好自闭症人群故事，以及在社交媒体传播形态下，如何提升自闭症人群故事在社交空间传播的广度和深度，以推动社会对自闭症人群态度转变，⁹需要我们“坚持科学、客观、严肃、认真的态度，尽力塑造客观、全面的特殊人群的媒体形象，提高报道特殊人群媒体形象的专业性，引导公众的正确认知”³¹。

³¹ 李慧冉, 马江霞, 李珍珍, 王鹏, 咎飞, 2019. 自闭症人群媒体形象分析——基于某新闻网的报道[J]. 现代特殊教育, (4):74-80.

作者简介：

林美珍，女，福建农林大学国际学院，讲师。研究方向：话语分析；语料库语言学。

马博森，男，浙江大学外国语学院，教授，博士生导师。研究方向：话语分析；语料库语言学。

张永伟，男，中国社会科学院语言研究所副研究员。研究方向：语料库语言学；计算语言学。