

公众与电商矛盾聚焦何处？

——基于LDA主题挖掘的2023年S市领导留言板公众诉求分析

朱 佳

上海工程技术大学管理学院，上海

收稿日期：2024年5月22日；录用日期：2024年6月4日；发布日期：2024年8月6日

摘 要

本研究采用LDA主题分析法，对2023年S市领导留言板上的公众留言进行了主题挖掘和诉求分析。通过分析留言内容，有效地识别出公众关注的核心问题和关键主题。研究结果显示，公众的诉求主要集中在电商商品、电商平台投诉与索赔、社会保障、社区治理等七个主题。其中，电商相关主题占据了2个。通过定量和定性的分析方法，本文详细阐述了主题的具体内容和公众的具体诉求，提供了电商与政府部门在优化服务过程和政策制定中的参考依据。

关键词

公众诉求，留言内容，市领导留言板，LDA主题分析法，电商

Where Do the Contradictions between the Public and E-Commerce Focus?

—Analysis of Public Demands of S City Leaders' Message Boards in 2023 Based on LDA Theme Mining

Jia Zhu

School of Management, Shanghai University of Engineering Science, Shanghai

Received: May 22nd, 2024; accepted: Jun. 4th, 2024; published: Aug. 6th, 2024

Abstract

In this study, LDA topic analysis was used to explore the topic and analyze the appeal of public messages on the message board of S City leaders in 2023. By analyzing the message content, we

can effectively identify the core issues and key themes of public concern. The research results show that the public's demands are mainly concentrated in seven themes, such as e-commerce goods, e-commerce platform complaints and claims, social security, and community governance. Among them, e-commerce related topics occupy 2. Through quantitative and qualitative analysis methods, this paper elaborated the specific content of the topic and the specific demands of the public, and provided a reference basis for e-commerce and government departments in the process of service optimization and policy formulation.

Keywords

Public Demands, Message Content, The Message Board of City Leaders, LDA Theme Mining, E-Commerce

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,随着网络的普及和技术的不断进步,电子商务行业呈现出蓬勃发展的趋势,同时,也浮现出不少问题。梁雯等人指出电子商务和物流快递业的相互推进也带来大量退货、产品垃圾、废旧物资循环利用率低下和环境污染等负面问题[1]。王夏阳和张斌对电商促销期产品缺货、电商企业临时取消订单等问题进行了深入研究,研究发现,电商战略性缺货更有利于电商企业[2]。上海社会科学院经济研究所课题组发现,对于小额跨境电子商务的趋势,我国现行管理体制、政策、法规及现有商业环境等条件还远远无法满足其发展需求,偷税漏税、检疫检测和市场监管三大问题突出[3]。刘小军和张滨提出,我国与“一带一路”沿线国家和地区跨境电商物流协作过程中仍存在诸多问题,如物流费用高、运输时间长、通关效率低、物流通道基础设施建设滞后等[4]。谢天成和施祖麟认为,当前,农村电子商务存在盲目、无序性,规模小,要素支撑薄弱等突出问题和困难[5]。

不难看出,学者们关注到了电商的一些现实问题,但少有研究从公众角度关注电商带来的负面影响,其中一个原因便是公众诉求很难收集。

如今,政务平台正逐渐成为公众与政府互动的主流方式之一,数字政府建设也成为政府发展的重点之一。2022年10月,党的第二十次全国代表大会在北京胜利召开。二十大报告指出的治理现代化等精神和工作要求对加强网上政府和数字政府建设极具指导意义。同年,国务院印发的《关于加强数字政府建设的指导意见》,就主动顺应经济社会数字化转型趋势,充分释放数字化发展红利,全面开创数字政府建设新局面作出部署,标志着我国数字政府建设步入新的发展阶段。

数字政府的发展为公众诉求的表达与研究提供了新的视界。

2. 文献综述

在学界,数字时代的政社互动也成了学者们的研究重点。数字政府的发展不仅推动了政府服务的现代化和高效化,更为公众诉求的表达与研究开辟了崭新的视界。通过数字化手段,政府能够更加便捷地收集、整理和分析公众的意见与需求,从而为政策制定提供更加科学、精准的依据。同时,数字政府的建设也促进了政府与公众之间的沟通与互动,使得公众的声音能够被更有效地传达和反馈。这不仅有助于提升政府的透明度和公信力,也为公众参与社会治理提供了新的途径和平台。

人民网设立的领导留言板因不受政府控制而引发了学者们的关注。目前,许多学者对领导留言板展开了分析。

从领导留言板建设的角度来看,段艺琳提出,领导留言板作为国内唯一的全国性领导干部留言板,集中体现了人民网作为中央重点新闻网站的权威性、大众化和公信力,其“社情民意的集散地,亲民爱民的回音壁”的宗旨正不断吸引着越来越多领导和网民的目光,带动着一大批同类平台的发展,吹起了“从留言板上获取社情民意”的一股新风[6]。李章程认为,人民网地方政府领导留言板呈现出如下特点:低成本互动,高效率办事;回应规格高,制度化趋势;凸显民生导向,民意满意度逐年上升;瞄准大数据发展前沿,强化数据理政;线上线下联动协同,公共治理迈向“标本兼治”;凸显建设性,在规则框架内实现良性互动[7]。侯莎莎提出,人民网“地方领导留言板”取得的突出成绩,显示了领导与网民良性互动的正面效果,也为通过互联网开展群众工作提供了诸多启示。各级领导干部要主动适应互联网时代的新要求,充分认识通过互联网开展群众工作的重要意义,建立健全网络舆情研判机制,积极解决群众通过网络反映的实际问题,采取多种措施保障群众的知情权,提高引导网络舆论的能力,切实做好互联网时代的群众工作[8]。

从领导留言板中的政社互动角度来看,李锋和孟天广收集了人民网地方政府留言板 2008 年至 2014 年期间的 21 万条帖子,并进行了分析。研究结果显示,公众更倾向于使用与工作、居住地等相关的集体认同表述,而不太愿意采用个体化的言辞或更广泛的集体认同来表达诉求[9]。郑石明、兰雨潇和黎枫通过对新冠肺炎疫情期间领导留言板的数据进行分析后发现,网络公共舆论与政府回应之间的互动表现呈现出从“民意先发、政府回应”的模式向“政府主导、政民高度互动”的趋势转变[10]。王国华等学者对领导留言板进行了研究,从政府回应的角度发现网络问政制度化水平与地方领导是否重视影响显著,提出强化问政制度建设,另一方面领导要重视网络问政[11]。王磊和易扬对疫情期间的领导留言板进行了研究,研究发现,网络问政平台忠实反映了疫情期间公众情绪变化,为政府了解社情民意充当了“传感器”,数字政府回应力高的地区,在疫情期间舆情更倾向于积极,问政平台能够扮演危机“缓冲器”的角色[12]。

不难看出,研究多数关注于网络问政制度建设的重要性与留言板中的政民互动的现象,对于公众诉求内容的深入挖掘尚有缺陷。因此本文选择对领导留言板中公众诉求内容进行深入研究,以求发现电商与公众之间的矛盾所在。

3. 数据来源与研究方法

3.1. 数据来源

领导留言板是由人民网于 2016 年创办的全国性网络问政平台。自 2006 年创办以来,数以万计的民众在留言板上通过留言的方式向各级领导积极表达自己的诉求,各省市也相继出台了留言办理相关的工作规定。此外,领导留言板的数据相对独立、透明度高且容易获取。

本文选取了 2023 年人民网领导留言板中 s 市的数据来进行研究,通过使用基于 Python 的爬虫技术对留言板的内容进行数据采集,数据内容包括用户昵称、留言类型、留言时间、留言领域、留言对象、留言内容、满意情况等字段。数据需要进行预处理,首先去除了重复评论、无效符号等,随后采用了 Python 中的 jieba 分词工具,停用词选用哈工大停用词词库,在进行分词与停用词处理后,本文共爬取了 4972 条有效公众诉求。

3.2. 研究方法

隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)是一种强大的主题建模工具,被广泛应用于自然语言处理领域。它以非监督学习的方式,将文档集中的每篇文档表示为概率分布的形式,从而揭示了文

本数据中的潜在语义结构[13]。在 LDA 模型中, 文档被看作是由多个主题组成的混合, 每个主题又由一组词汇构成。这种模型不仅能够发现文档中隐藏的主题, 还能够为每个词汇赋予不同主题的概率分布。通过对文档的主题分布和词汇的主题分布进行统计分析, LDA 模型可以有效地识别出文本数据中的关键主题, 并用于文本挖掘、情感分析、主题分类等任务中。LDA 模型的应用范围非常广泛, 不仅可以处理大规模文档集和语料库, 还可以用于分析社交媒体数据、新闻报道、学术论文等各种类型的文本数据。

LDA 被称为三层叶贝斯概率模型, 其中包含文档层、主题层和词层, 形成了一种层级结构。它利用文档中词汇的共现关系来对词汇进行主题类聚, 从而得到了“文档-主题”和“主题-词”的分布矩阵[14]。由于其广泛的应用和简单的操作性, LDA 被广泛用于识别用户评论中的主题和关注热点。因此, 本文选择使用 LDA 模型来识别领导留言板公众评论中潜在的主题信息。LDA 模型的应用可以帮助发现公众关注的话题和需求, 为政府回应改进回应效率与质量提供有益的参考。

4. 研究结果

将 4972 条公众留言进行主题一致性与主题困惑度评价, 评价结果见图 1。LDA 主题困惑度是一种用于评估 LDA 模型拟合效果的指标。困惑度越低, 表示模型在预测下一个词时的不确定性越小, 也就是说模型的拟合效果越好。LDA 主题一致性是一种用于评估 LDA 模型生成主题质量的指标。高主题一致性得分表示该主题中的词汇更加相关和连贯。通过评估 LDA 模型中各个主题的困惑度与一致性, 研究者可以更准确地确定最优主题数和模型参数, 从而提高主题提取的效果和质量[15]。

如图 1 所示, 当主题数为 7 时, 主题一致性评价最高, 与主题数为 8 相比困惑度差值小, 主题困惑度相对较低, 为了更好地进行主题识别, 本文将主题数定为 7, 聚类效果更佳。

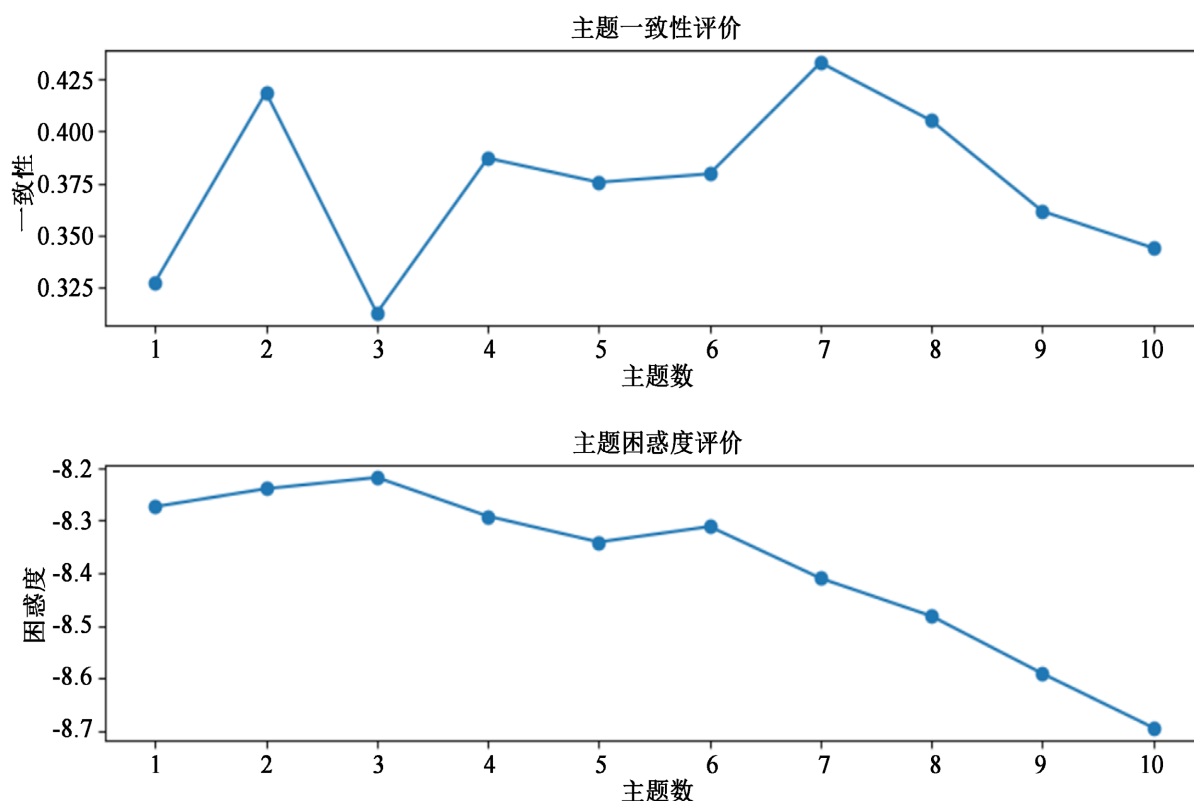


Figure 1. LDA topic coherence and perplexity evaluation
图 1. LDA 主题一致性与主题困惑度评价

由于数据量较大，本文主要摘取了前十条数据在正文作为展示。从表 1 可以看出公众诉求文档的主题概率分布，不同主题对应文档的隶属概率越高，则将该文档归类为该主题下。

Table 1. Topic probability distribution of public appeal documents
表 1. 公众诉求文档的主题概率分布

导入的语料	最大概率主题	主题 1	主题 2	主题 3	主题 4	主题 5	主题 6	主题 7
……口头辞退我，并恶意克扣拖欠工资……	2	0.227348492	0.758195817	0	0	0	0	0
……影响北面窗户外面的空气流通……	6	0	0	0	0	0.084906735	0.91051662	0
……和之前公示的线路存在差异……	6	0	0	0	0	0.131836623	0.673363507	0.182887778
……商家一直没有发货……	1	0.681736946	0	0	0.306562304	0	0	0
……拖欠三个月的工资……	2	0.397767991	0.575128019	0	0	0	0	0
……欺诈消费者，通过官方投诉，包庇商家……	4	0.152857199	0	0	0.773233652	0	0	0.066160582
……工资以公司亏本为由无故扣一半工资……	2	0	0.749929368	0	0.22340107	0	0	0
……巨大噪音惊醒……	6	0	0	0	0	0.286533087	0.655659914	0.053463012
……图书馆闭馆又已经是半个多月……	6	0.010552194	0.010553655	0.01056217	0.010556035	0.010569016	0.936664462	0.010542481
……发现噪音问题严重……	6	0	0	0	0	0	0.733933926	0.258632004

根据与网站议题的匹配结果，可以发现分类主题与留言板议题存在着一定的重合性，这进一步证实了模型分类的可靠性较高。与留言板上的 14 个议题相比，主题模型将其减少到了 7 个，但这些主题具有更高的凝聚性和相关性。

这一发现表明，主题模型能够更有效地捕捉到评论内容中的主要议题，并将其归类为更加具体和相关的主题类别。这种减少主题数量的趋势并不意味着信息的丢失，相反，它反映了模型在提取主题时更加精准和有效的能力。

因此，基于主题模型的分类结果不仅简化了对评论内容的理解，还提供了更加清晰和有针对性的分析基础，有助于进一步挖掘评论数据中的深层信息和潜在关联。

从主题分类以及主题特征词可以看出，公众诉求主要集中在七个主题，即电商商品问题、社会保障问题、社区治理问题、电商平台投诉与索赔问题、开发商问题、城市发展问题和教育问题，见表 2。其中，电商相关主题占据了 2 个，从主题一和主题四可以看出，在数字化浪潮席卷而来的今天，公众作为网购大军的一员，对于商品质量、售后服务、投诉与索赔机制等方面的关注程度日益上升。

Table 2. Topic classification and feature words
表 2. 主题分类以及主题特征词

主题关键词和权重	主题 1	主题 2	主题 3	主题 4	主题 5	主题 6	主题 7
关键词 1	商家 (0.023922)	公司 (0.020705)	小区 (0.018627)	平台 (0.011403)	开发商 (0.010582)	影响 (0.005086)	消费者 (0.003112)
关键词 2	平台 (0.022729)	工资 (0.016679)	居民 (0.010552)	消费者 (0.009666)	领导 (0.006656)	部门 (0.003637)	课程 (0.002961)
关键词 3	退款 (0.011276)	工作 (0.008746)	业主 (0.008860)	购买 (0.006217)	退款 (0.005405)	生活 (0.003630)	时间 (0.002891)
关键词 4	客服 (0.009012)	拖欠 (0.006493)	问题 (0.007826)	问题 (0.005884)	要求 (0.004853)	相关 (0.003584)	解决 (0.002811)
关键词 5	要求 (0.007493)	离职 (0.006310)	物业 (0.006703)	订单 (0.002920)	项目 (0.004512)	施工 (0.003475)	发现 (0.002692)
关键词 6	退货 (0.007360)	社保 (0.005585)	解决 (0.004886)	不能 (0.002887)	宣传 (0.004434)	目前 (0.003468)	销售 (0.002672)
关键词 7	购买 (0.006671)	支付 (0.004326)	严重 (0.004364)	赔偿 (0.002878)	购买 (0.003947)	规划 (0.002897)	严重 (0.002667)
关键词 8	商品 (0.006664)	合同 (0.003579)	车辆 (0.003998)	严重 (0.002798)	问题 (0.003612)	道路 (0.002761)	进行 (0.002400)
关键词 9	消费者 (0.006376)	拖欠工资 (0.003569)	车位 (0.003860)	投诉 (0.002717)	业主 (0.002385)	情况 (0.002845)	服务 (0.002085)
关键词 10	发货 (0.003762)	劳动合同 (0.003457)	停车 (0.003361)	诉求 (0.002708)	合同 (0.002007)	居住 (0.002563)	沟通 (0.002078)

主题一表现出公众作为消费者，在网购过程中对于电商商品的品质、发货速度以及退换货流程等方面十分关心。消费者对商品的品质有着高度的关注。他们期待购买到的商品能够与描述一致，质量可靠。其次，发货速度也是消费者十分关心的问题。他们希望能够尽快收到自己购买的商品，因此对于快递公司的服务质量和配送速度有着一定要求。最后，退换货流程的顺畅性对于消费者来说也至关重要。在网购过程中，难免会遇到商品质量不符合预期或存在瑕疵的情况。消费者希望能够通过与商家和平台的沟通，快速解决问题，获得合理的补偿或替换商品。

主题四与主题一有着诸多相似之处，但公众的焦点更多地放在了电商平台的投诉与索赔机制的完善与执行情况上。在这个日益成熟的网购时代，消费者对于维护自身权益的意识和要求也在不断提升。在网购过程中，难免会遇到商品质量问题、服务不达标等种种困扰。面对这些问题，他们不是默默忍受，而是更加期望能够有一个健全的投诉与索赔机制，为他们提供维权的通道和保障。一个完善的投诉与索赔机制，对于消费者而言，是维护自身权益的重要屏障。它不仅能够为消费者提供一个公平、透明的维权平台，还能够确保消费者在提出投诉时能够得到及时、有效的回应和处理。当消费者的权益受到侵害时，这样的机制能够为他们提供合理的补偿或解决方案，让他们感受到公平与正义。因此，公众对电商

平台的投诉与索赔机制寄予了期望。他们希望电商平台能够认真对待消费者的投诉，对涉及商品质量、服务不达标等问题进行深入调查，并根据调查结果给予消费者合理的解决方案。公众期望电商平台能够积极听取消费者的意见和建议，不断完善和优化投诉与索赔机制。他们希望平台能够建立起一套更加人性化、高效化的处理流程，为消费者提供更加便捷、高效的维权服务。

4972 条公众留言中，有近 30% 的公众留言主题是围绕与电商的矛盾而展开的，不难看出，电商问题已成为影响公众生活的重要因素之一。为了电子商务的健康发展与公众生活的和谐，解决电商与公众之间的矛盾已日渐凸显出其重要性。

5. 结论与建议

电商，作为现代生活的重要组成部分，其健康发展直接关系到公众的生活质量和社会的和谐稳定。电商问题之所以成为公众关注的焦点，一方面是因为电商平台的普及和便利为公众生活带来了极大的便利，另一方面则是电商在发展过程中暴露出的一系列问题，如 LDA 主题分析中所体现出的商品质量、售后服务、投诉与索赔机制等，这些都直接或间接地影响着公众的消费体验和权益保障。

商品质量问题是电商与公众矛盾的主要方面之一。消费者在网购时往往难以直接判断商品的质量，只能依赖于商家的描述和其他消费者的评价。然而，有些商家为了追求利润，可能会夸大商品质量或隐瞒瑕疵，导致消费者收到的商品与预期不符。这不仅损害了消费者的权益，也影响了电商平台的声誉。

售后服务问题同样不容忽视。在网购过程中，消费者可能会遇到各种问题，如物流延误、商品损坏等。此时，一个及时、有效的售后服务能够为消费者解决问题，提供必要的帮助。然而，有些商家在售后服务方面表现不佳，甚至存在推诿、敷衍的现象，这使得消费者在维权过程中感到无助和困惑。

投诉与索赔机制的完善与否直接关系到消费者权益的保障。一个健全的投诉与索赔机制能够为消费者提供维权的通道和保障，确保他们在遇到问题时能够得到及时、有效的解决。然而，目前一些电商平台的投诉与索赔机制还存在不完善、执行不力等问题，这使得消费者在维权过程中面临诸多困难。

为了有效解决电商与公众之间日益凸显的矛盾，进一步推动电子商务的稳健和可持续发展，可以从电商平台、政府和公众三个维度进行深入探讨和积极实践。

首先，电商平台在维护市场秩序和消费者权益方面具有不可推卸的责任。平台需要建立更加严格的商家入驻审核机制，确保只有信誉良好、经营合规的商家才能在平台上进行销售。同时，对于已入驻的商家，平台应实施动态监控，定期检查其商品质量，并督促其提供优质的售后服务。当发现商家存在违规行为时，平台应迅速采取措施，包括但不限于警告、罚款、甚至取消其销售资格，以确保市场的公平竞争和消费者的切身利益。

其次，政府在电商行业的规范和发展中扮演着举足轻重的角色。相关部门需要密切关注市场动态，及时发现并纠正行业中的不正之风。此外，政府还应根据电商行业的特殊性和发展趋势，制定或修订相关法律法规，明确各方责任和义务，为电商行业的有序发展提供坚实的法制保障。

最后，公众作为电商活动的重要参与者，也应积极提升自身的消费意识和维权能力。在选择商品和服务时，要保持理性，不被虚假宣传所迷惑，注重商品的质量和性价比。同时，当自身权益受到侵害时，要勇于维权，通过正规渠道进行投诉和申诉，以实际行动维护电商市场的良好秩序。

综上所述，解决电商与公众之间的矛盾，促进电子商务的健康发展，需要电商平台、政府和社会各界的共同努力。只有这样，才能构建一个更加和谐、健康、可持续发展的电商环境。

参考文献

- [1] 梁雯, 陈广强, 王欣. 电子商务发展中的逆向物流问题——电商企业与政府、消费者的不同利益关系博弈分析[J].

- 产经评论, 2017, 8(1): 67-79.
- [2] 王夏阳, 张斌. 消费者选择行为下的电商战略性缺货问题研究[J]. 管理科学学报, 2019, 22(10): 9-23.
- [3] 上海社会科学院经济研究所课题组, 石良平, 汤蕴懿. 中国跨境电子商务发展及政府监管问题研究——以小额跨境网购为例[J]. 上海经济研究, 2014(9): 3-18.
- [4] 刘小军, 张滨. 我国与“一带一路”沿线国家跨境电商物流的协作发展[J]. 中国流通经济, 2016, 30(5): 115-120.
- [5] 谢天成, 施祖麟. 农村电子商务发展现状、存在问题与对策[J]. 现代经济探讨, 2016(11): 40-44.
- [6] 段艺琳. 选择: 网络民意与政府回应路径——以人民网《地方领导留言板》为例[J]. 东南传播, 2010(1): 31-33.
- [7] 李章程. 网络时代公民和政府的信息沟通分析[J]. 电子政务, 2017(4): 23-36.
- [8] 侯莎莎. 关于通过互联网开展群众工作的思考——以人民网“地方领导留言板”为例[J]. 求实, 2012(4): 79-82.
- [9] 李锋, 孟天广. 策略性政治互动: 网民政治话语运用与政府回应模式[J]. 武汉大学学报(人文科学版), 2016, 69(5): 119-129.
- [10] 郑石明, 兰雨潇, 黎枫. 网络公共舆论与政府回应的互动逻辑——基于新冠肺炎疫情期间“领导留言板”的数据分析[J]. 公共管理学报, 2021, 18(3): 24-37+169.
- [11] 王国华, 罗泉, 方付建. 网络问政平台运行绩效影响因素研究——基于人民网地方领导留言板的分析[J]. 管理现代化, 2011(5): 36-38.
- [12] 王磊, 易扬. 公共卫生危机中的数字政府回应如何纾解网络负面舆情——基于人民网“领导留言板”回复情况的调查[J]. 公共管理学报, 2022, 19(4): 65-78+169.
- [13] Blei, D.M., Ng, A. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *The Journal of Machine Learning Research*, 3, 993-1022.
- [14] 谭春辉, 陈晓琪, 梁远亮, 等. 隐私泄露事件中社交媒体围观者情感分析[J]. 情报科学, 2023, 41(3): 8-18.
- [15] 李梦颖. 基于主题模型的学习方法研究及其在用户画像上的应用[D]: [硕士学位论文]. 北京: 北京交通大学, 2020.