

• 情报业务与情报服务 •

双重感知视角下的社交媒体类 APP 隐私 政策内容分析及优化策略研究

朱 光^{1 2} 卞淑莹¹ 杨健清¹

(1. 南京信息工程大学管理工程学院, 江苏 南京 210044;

2. 南京信息工程大学自贸区研究院, 江苏 南京 210044)

摘 要: [目的/意义] 探寻社交媒体类 APP 隐私政策的完善与优化策略, 为用户个人信息的合规采集与利用提供参考。[方法/过程] 基于技术接受模型(Technology Acceptance Model, TAM) 从双重感知视角对 30 款社交媒体类 APP 的隐私政策进行可读性分析和主题识别。对标国内外相关法律制度, 在分析感知易用性和感知有用性的基础上提出优化建议。[结果/结论] 研究发现, 现有社交媒体类 APP 隐私政策在命名方式、导语设置、小标题数量等可读性指标上表现较好, 但存在内容敷衍、文本篇幅不妥、术语占比较高等问题。政策主题涵盖用户权利、隐私政策概述、运营商行为、数据存储与共享、法律保护和监管与运营义务 6 个类别, 仍需在数据共享与跨境流动层面进一步完善内容。未来可从内容完整性、条款针对性和文本可读性 3 个方面进行完善与优化。

关键词: 社交媒体; APP; 隐私政策; 技术接受模型; 感知易用性; 感知有用性; 文本分析

DOI: 10.3969/j.issn.1008-0821.2022.10.004

(中图分类号) G251 (文献标识码) A (文章编号) 1008-0821 (2022) 10-0036-12

Research on Content Analysis and Optimization Strategy of Privacy Policy of Social Media APP from Dual Perception Perspective

Zhu Guang^{1 2} Bian Shuying¹ Yang Jianqing¹

(1. School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. Institute of Free Trade Zone, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract [Purpose/Significance] This paper investigated the improving and optimization strategy of privacy policy of social media apps, to provide suggestions for the collection and usage of personal information. [Method/Process] Based on the Technology Acceptance Model (TAM), this paper conducted the readability analysis and topic identification on 30 privacy policies of social media apps from a dual perception perspective. By comparing with the regulations domestic and abroad, the optimization strategy was proposed based on the analysis of perceived ease of use and perceived usefulness. [Result/Conclusion] The research results show that current privacy policies performed better in naming methods, introduction settings and the number of subtitles from the perspective of readability analysis, but has some limitations in perfunctory content, inappropriate text length, and a high proportion of terms. The identified topics of privacy policies included user rights, privacy policy overview, operator behavior, data storage and sharing, legal protection and regulation, and operator

收稿日期: 2022-04-02

基金项目: 国家社会科学基金项目“泛在智慧环境下多源主体隐私感知及协同保护机制研究”(项目编号: 19CTQ019)。

作者简介: 朱光(1986-), 男, 副教授, 博士, 研究方向: 信息隐私, 大数据分析。卞淑莹(1998-), 女, 硕士研究生, 研究方向: 信息隐私, 信息资源管理。杨健清(2000-), 女, 硕士研究生, 研究方向: 信息资源管理。

voluntary, while still necessary to be further improved in the field of data sharing and cross-border flow. In future, the optimization strategy of privacy policy can be implemented from the perspective of content integrity, clause pertinence and text readability.

Key words: privacy policy; Technology Acceptance Model; perceived ease of use; perceived usefulness; text analysis

随着无线通信技术的发展以及移动终端设备的广泛应用,微博、抖音、Instagram、Twitter 等社交媒体类 APP 为用户提供了网络交际和信息分享的渠道,成为大众日常生活中必不可少的一部分。然而,社交媒体 APP 为提供精准的个性化服务和广告营销,需获取用户手机的诸多权限(如录音、摄像、地理位置)并长期采集用户的使用行为,导致用户不知晓也未授权的个人数据被采集和使用等问题,增加了隐私泄露的风险^[1]。

为防止 APP 过度索取无关权限、恶意访问用户隐私数据等行为,各国(地区)都出台了相关法律法规,对个人信息使用的途径、范围、目的进行了明确规定。2013 年,我国发布的《电信和互联网用户个人信息保护规定》指出,电信业务经营者、互联网信息服务提供者应制定隐私政策,其内容应当包括收集、使用信息的目的、方式和范围,查询、更正信息的渠道以及拒绝提供信息的后果等事项^[2]。2018 年,欧盟出台的《通用数据保护条例》(General Data Protection Regulation, GDPR)在赋予用户处理自身数据更多权限的同时,也提升了企业对用户数据收集、存储和使用的标准^[3]。2020 年,网信办发布了《常见类型移动互联网应用程序(APP)必要个人信息范围(征求意见稿)》,对各类互联网平台和 APP 的个人信息收集范围做出界定,以保障公民个人信息安全^[4]。在此背景下,众多社交媒体类 APP 纷纷制定了个人信息保护政策、个人信息使用说明、隐私保护协议、隐私保护指引等隐私政策^[5-6]。

然而,现实情境下的隐私政策仍存在主题缺失、内容隐晦、结构混乱等问题,用户为了 APP 的正常使用通常不会仔细阅读隐私政策,并不真正了解哪些个人信息会被搜集与使用^[7]。2021 年 5 月,网信办发现有 33 款 APP 在出台隐私政策的同时,仍会违规收集个人隐私信息^[8]。2022 年 1 月,工信部指出 30 余款社交媒体类 APP 的隐私政策存

在内容表述不清、晦涩难懂等问题^[9]。因此,如何完善隐私政策的规范性、全面性和可读性成为业界和学界越发关注的问题。

基于上述背景,本文以社交媒体类 APP 隐私政策为研究对象,引入技术接受模型(Technology Acceptance Model, TAM)从感知易用性和感知有用性两个维度研究社交媒体类 APP 隐私政策的文本内容并提出优化建议。首先从感知易用性角度,对隐私政策的文本可读性进行测度,探寻其对用户阅读意愿的影响。进一步从感知有用性角度,运用 LDA(Latent Dirichlet Allocation)模型识别隐私政策主题,归纳成类别条目与国内外相关法律对比,找出存在的问题,提出优化建议。

1 国内外研究现状

隐私政策是指互联网环境下各类网站、软件以及 APP 就如何收集和使用个人信息所做出的说明与解释,其制定与完善对规范个人信息利用有着重要的意义^[10]。然而,皮尤研究中心的调查显示,只有约 9%的用户表示在使用 APP 或软件之前会阅读隐私政策^[11]。针对隐私政策的阅读意愿问题,Steinfeld N^[12]发现,当隐私政策全文显示时,用户会选择花费较多时间阅读全文;而当政策以折叠选项的方式呈现时,用户并不倾向于点开链接自愿阅读。Singh R I 等^[13]从隐私政策的外观呈现方面,发现字体大小、排版格式、背景颜色等因素会影响用户阅读意愿。Vail M W 等^[14]从文本结构维度,发现层次化的隐私政策内容比其他表现形式更能提升用户的阅读意愿。朱侯等^[15]从个人信息搜集、信息安全、用户权利、联系与申诉途径等方面对社交媒体隐私政策内容的可读性特征进行了统计,验证了可读性对用户阅读意愿会产生正向影响的结论。朱光等^[16]基于技术接受模型和自我效能理论,从用户主观感知和隐私政策的内容特征两个维度分析主观规范、信息质量、信息过载、自我效能、感知易用性、感知有用性、经验等因素对社交媒体隐私政

策阅读意愿的影响。张玥等^[17]基于认知负荷理论,采用虚拟实验和多因素方差分析测评了用户在阅读不同类型隐私政策时的主观认知负荷以及客观成绩,从而进一步研究信息表征对移动医疗 APP 隐私政策阅读意愿和阅读效果的影响。

为提升用户的阅读意愿,研究学者运用内容分析法和比较分析法,分析不同类型隐私政策的文本特征,比较主题差异,提出优化方案。袁红等^[18]以政务 APP 的隐私政策文本为研究对象,运用内容分析法从外部特征和文本内容两个维度分析其存在的问题。储节旺等^[19]以美国政府开放数据的相关政策文本为数据来源,运用 NVivo 文本分析工具,对政策文本进行内容多层编码分析,梳理美国政府开放数据隐私保护框架。何培育等^[20]以微信、淘宝等 10 款常用的 APP 为研究对象,对其隐私政策内容进行定性梳理与归纳,并提出完善建议。严炜炜等^[21]以在线学术社区的隐私政策框架内容为研究对象,从信息管理、技术安全、用户权利、责任意识、基础功能和高级功能等维度提出了完善思路和优化路径。Martiny K 等^[22]运用文本挖掘技术分析上百个隐私政策的摘要内容,帮助用户了解其结构与主题。Kaur J 等^[23]收集 2 000 多例不同类型的网站、软件和 APP 的隐私政策,运用 LDA 模型对政策文本进行主题识别,结果显示,所在地的法律法规对隐私政策的术语和关键词具有显著影响。Strahilevitz L J 等^[24]从隐私政策的语言描述和句法结构层面,分析了法官和消费者在法律层面对于隐私政策理解的差异性。

针对隐私政策的规范性、友好性、全面性等测评指标,赵杨等^[25]以医疗健康 APP 的隐私政策文本为研究对象,综合运用卷积神经网络、循环神经网络和长短期记忆人工神经网络进行隐私政策合规性的测评。张艳丰等^[26]以 10 款短视频 APP 为研究对象,从政策表征、信息收集与存储、信息使用与分享、用户权利、隐私保护这 5 个维度构建隐私政策的评价指标体系,进一步利用灰色加权关联分析法对隐私政策的阅读感知性进行关联度计算。姚胜译等^[27]以电子商务类 APP 的隐私政策文本为研究对象,运用层次分析法,从内容可读性和交互友好性两个维度构建了多层级的隐私政策的友好度评价指标体系。周林兴等^[28]以我国 78 个档案网站为

研究对象,对其是否设有隐私政策以及隐私政策的形式、内容进行调研。调研结果显示,我国档案网站的隐私政策规范性不足,内容有待进一步健全。Chang Y 等^[29]提出了隐私边界评价模型,从信息公平等 5 个维度评估用户对隐私政策的感知有用性和感知易用性。Aïmeur E 等^[30]从隐私政策的个性化结构和文本内容两个维度,分析了网站隐私政策与用户信任之间的关系。

整体来看,现有研究大多关注隐私政策的阅读行为、内容分析及评价指标等方面,少有研究结合呈现形式和主题类别两个层面对隐私政策进行内容分析,也缺少与国内外相关法律制度的对比分析。基于此,本文引入 TAM 模型综合考虑隐私政策的感知易用性和感知有用性,运用文本可读性分析和 LDA 主题识别,从呈现形式和内容主题两方面研究社交媒体隐私政策的优化问题。

2 研究设计

TAM 模型用来研究用户接受信息系统或 IT 技术的影响因素,主要从感知有用性和感知易用性两个维度分析行为态度、使用意愿和实际行为三者之间的关系^[31]。基于此,一方面,若社交媒体类 APP 的隐私政策内容全面、主题规范,可以满足用户的隐私保护、心理感知及社交匿名等需求,则用户的感知有用性较强,隐私政策发挥的效果更好;另一方面,若隐私政策内容逻辑清晰,简明易懂,则用户感知易用性较强,可以引导用户深入阅读,让其了解其个人信息采集和使用的准则。

基于上述分析,本文从感知有用性和感知易用性两个维度对隐私政策内容进行分析。一方面,利用 LDA 模型识别隐私政策主题,并与国内外知名法规进行对比以分析其感知有用性;另一方面,设计政策可读性指标对感知易用性进行测评。在此基础上,提出社交媒体类 APP 的隐私政策优化方案。

2.1 数据采集与预处理

本文基于小米应用商店排行榜(2022 年 2 月 28 日),选择累计下载量在 1 亿次以上的 33 款社交媒体类 APP(包括微博、微信、知乎、抖音等)作为研究样本,采用网络爬虫与人工下载相结合的方式获取隐私政策文本。剔除文本内容相似和未能正确采集的部分,最终得到 30 款社交媒体类 APP 隐私政策文本,如表 1 所示,文本总字数约 46 万字。

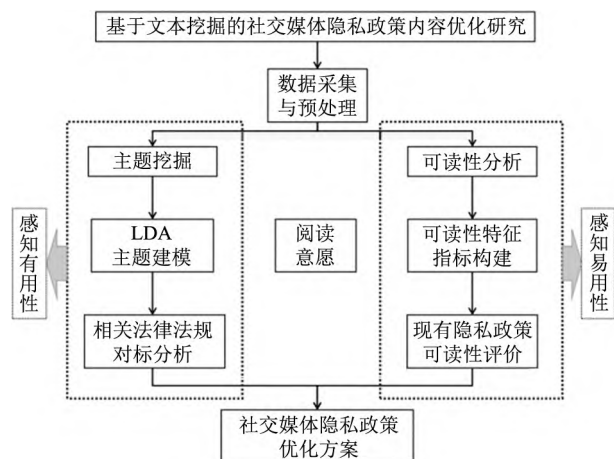


图 1 研究流程

为便于后续 LDA 主题识别和文本可读性分析，需要对采集文本进行预处理，步骤如下：

1) 数据清洗。对隐私政策文本中的标点、无意义数字和符号进行清洗过滤，如标题序号“—”(1)“1.1”等。

2) 文本分词。采用 Jieba 分词包对隐私政策文本进行分词处理，并依据《中华人民共和国个人信息保护法》《中华人民共和国网络安全法》以及《信息安全技术 个人信息安全规范》(GB/T35273-2020)中的术语解释构建词典，对分词结果进行停用词处理。

表 1 30 款社交媒体类 APP 隐私政策概况

序号	应用名称	政策概况	更新/生效日期	序号	应用名称	政策概况	更新/生效日期
1	微信	共 9 章 59 条	2022-02-17	16	秒拍	共 9 章 26 条	2020-06-08
2	QQ	共 10 章 42 条	2021-11-01	17	哔哩哔哩	共 13 章 52 条	2021-09-30
3	Soul	共 10 章 25 条	2022-02-17	18	腾讯视频	共 11 章 43 条	2022-02-18
4	陌陌	共 9 章 20 条	2022-02-22	19	爱奇艺	共 10 章 23 条	2021-12-27
5	探探	共 9 章 26 条	2022-02-11	20	优酷视频	共 9 章 22 条	2022-02-23
6	钉钉	共 9 章 30 条	2021-11-18	21	芒果 TV	共 11 章 40 条	2021-12-14
7	脉脉	共 11 章 28 条	2021-12-29	22	平安健康	共 8 章 22 条	2022-02-22
8	企业微信	共 10 章 35 条	2022-01-10	23	微博	共 9 章 30 条	2022-02-09
9	拼多多	共 10 章 35 条	2021-12-31	24	小红书	共 11 章 22 条	2021-12-17
10	淘特	共 9 章 30 条	2022-01-26	25	知乎	共 11 章 24 条	2021-10-29
11	京喜	共 8 章 30 条	2021-11-30	26	百度贴吧	共 9 章 24 条	2021-08-11
12	抖音	共 8 章 32 条	2022-03-01	27	QQ 空间	共 10 章 19 条	2021-11-14
13	快手	共 10 章 37 条	2019-11-08	28	皮皮虾	共 10 章 44 条	2022-01-21
14	腾讯微视	共 10 章 40 条	2021-11-01	29	最右	共 11 章 28 条	2021-07-20
15	好看视频	共 9 章 28 条	2021-01-04	30	王者营地	共 10 章 29 条	2022-02-23

3) 特征提取。对分词后的隐私政策文本进行 TF 和 TF-IDF 特征提取，选取相对较优的特征作为 LDA 主题模型的输入参数。

4) 文本向量化处理。采用 Gensim 包中封装的 Doc2bow 方法进行文本向量化处理，过滤文本中出现频率过高或过低的词语，以保证 LDA 主题识别的效果。

2.2 隐私政策的可读性测度

在隐私政策文本中，突出重点的标题设置、条理清晰的段落划分以及便于阅读的整体排版都将大幅提升用户的阅读体验，从而提升用户对隐私政策的阅读意愿^[32]。结合前人关于文本可读性的研究

以及隐私政策文本结构的独特性^[33]，本文选取命名方式、小标题数量、导语、文本篇幅、术语密度、平均句长 6 个指标对社交媒体类 APP 隐私政策的可读性进行测评，以探究隐私政策的感知易用性情况。指标描述如表 2 所示。

表 2 隐私政策文本可读性测评指标

指标名称	指标描述	文献参考
命名方式	隐私政策的标题措辞	[27]
小标题数量	隐私政策全文小标题总数	[38]
导语	隐私政策文本开头是否有引导性语言	

表 2 (续)

指标名称	指标描述	文献参考
文本篇幅	隐私政策全文总字数	[23]
术语密度	专业术语在隐私政策全文词汇中的占比	
平均句长	隐私政策全文句子平均长度	[25]

其中,术语密度的计算公式如下:

$$\text{术语密度} = \frac{\text{术语总数}}{\text{词语总数}} \times 100\% \quad (1)$$

平均句长的计算公式如下:

$$\text{平均句长} = \frac{\text{文本篇幅}}{\text{句子总数}} \quad (2)$$

2.3 隐私政策主题识别

与传统文本分析方法相比,LDA 主题模型克服了文本矩阵稀疏、语义缺失等诸多缺陷,常用来识别各类文本中隐含的主题状况^[34]。基于此,本文利用 LDA 主题模型对社交媒体类 APP 的隐私政策文本进行主题识别,并将结果作为与国内外知名法规进行主题对比的依据,以探寻现有社交媒体隐私政策文本的感知有用性。

LDA 模型认为文档到主题服从多项式分布,主题到关键词也服从多项式分布,其核心结构是由词、主题和文档组成的 3 层贝叶斯模型^[35],如图 2 所示。

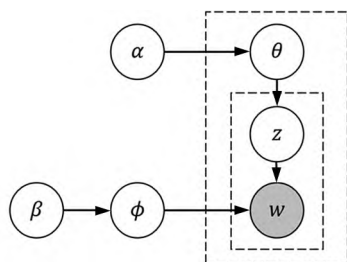


图 2 LDA 主题模型结构

其中, θ 表示主题分布, ϕ 表示关键词分布, α 是主题分布 θ 的 Dirichlet 分布参数, β 是关键词分布 ϕ 的先验分布参数, w 表示模型最终生成的关键词, z 表示模型生成的主题。LDA 主题模型的联合概率分布如式 (3) 所示^[36]:

$$p(w|\theta) = \sum_{z=1}^k p(w|z) p(z|\theta) \quad (3)$$

3 结果与讨论

3.1 关键词统计

对 30 款社交媒体类 APP 的隐私政策文本进行分词,排序前 20 的关键词词频统计如表 3 所示,得到 4 499 个词汇共出现 106 178 次。“信息”“个人信息”“服务”“提供”“收集”等关键词出现的频率最高,反映出现有 APP 隐私政策大多针对用户个人信息的收集进行规定和解释。此外,“第三方”“设备”“授权”“设置”“权限”等高频词说明了 APP 对用户移动设备通话、录音、摄像、存储等权限的获取,若隐私政策含糊其词或用户没有仔细阅读,极有可能造成未经授权的隐私泄露。

表 3 词频统计结果

序号	关键词	词频	序号	关键词	词频
1	信 息	4 414	11	保 护	923
2	个人 信息	3 229	12	隐 私	886
3	服 务	2 640	13	第 三 方	873
4	提 供	2 124	14	设 备	854
5	收 集	1 419	15	授 权	829
6	功 能	1 409	16	设 置	826
7	相 关	1 309	17	同 意	782
8	用 户	1 049	18	内 容	782
9	产 品	1 003	19	方 式	704
10	政 策	972	20	权 限	702

3.2 可读性分析

为分析社交媒体类 APP 隐私政策的感知易用性,在关键词统计的基础上,对政策文本的可读性进行测评,部分结果如表 4 所示。具体分析如下:

1) 命名方式。共有 14 款社交媒体类 APP 直接以“隐私政策”作为标题,占比达 46.7%;7 款社交媒体类 APP 的隐私政策标题为“隐私保护指引”,占比 23.3%,主要为腾讯系应用;3 款社交媒体类 APP 的隐私政策相关内容以“个人信息保护政策”为标题,占比 10%;此外还有部分隐私政策的命名为“隐私权政策”(3 款)、“个人信息保护指引”(1 款)、“隐私权保护政策”(1 款)、“用户隐私政策”(1 款)。整体来看,有 26 款社交媒体类 APP 的隐私政策在标题表述中直接提及了“隐私”字样,占比超过 86.7%。由于在标题

中突出了“隐私保护”“隐私策略”“个人信息保护”等关键字样,用户能够较为直观地定位阅读入口,并初步了解隐私政策的作用。

2) 小标题数量。小标题可以体现隐私政策文本的层次结构,但小标题数量过多也会增加用户的阅读负担,降低感知易用性^[37]。从小标题数量上看,社交媒体类 APP 隐私政策的小标题数量均在 10 个左右,其中最少的 8 个,包括“淘特”“京喜”“平安健康”“优酷”“抖音”“钉钉”6 款应用,最多的应用是“哔哩哔哩”,共有 13 个小标题。

3) 导语。导语的设置便于用户第一时间获得隐私政策的总体信息,本文选取的 30 款社交媒体类 APP 隐私政策均在开头设置了导语,但在导语的具体内容上存在很大差异,“微信”“QQ”等腾讯系应用在导语中对隐私政策的全文内容进行了概括,方便用户及时获取所关注的相关信息,而“秒拍”等 APP 只是简单地进行了目录和框架介绍。

4) 文本篇幅。隐私政策的文本篇幅如果过短则难以传达完整的信息,过于冗长又会增加用户的阅读难度。不同社交媒体类 APP 隐私政策的篇幅差异较大,篇幅最短的是“QQ 空间”,政策全文共 6 560 字;篇幅最长的是“芒果 TV”,隐私政策全文达到 26 755 字。总体上看,大多数社交媒体类

APP 隐私政策的文本篇幅都超过 1 万字(93.3%),这与朱侯等关于社交媒体隐私政策篇幅集中在 4 000~8 000 字的结论差别较大^[38]。该分析结果表明,随着个人信息保护法律法规的不断完善,以及公众隐私保护意识的提高,社交媒体运营商在制定隐私政策条款时内容更加具体,完整性较之前有了较大提升。

5) 术语密度。30 款社交媒体类 APP 隐私政策的术语密度均集中在 4%~10% 之间,其中密度较高的是“QQ 空间”(9.596%)和“陌陌”(9.474%)。“QQ 空间”由于文本篇幅较短,导致术语密度高,而“陌陌”的文本篇幅在 1 万字以上,专业术语密度明显偏多,会显著降低用户的感知易用性。“抖音”(4.024%)和“钉钉”(4.439%)的术语密度较低,而下载量和使用度最高的“微信”的术语密度为 5.385%。

6) 平均句长。相关研究表明,汉语单句的最佳长度应控制在 12 字以内,超过此长度则会带来阅读上的困难^[39]。本文选取的 30 款社交媒体类 APP 隐私政策的平均句长在 5~11 字,没有出现明显长难句的情形。值得注意的是,“陌陌”的术语密度(9.474%)和平均句长(10.290)两项均较高,其隐私政策可读性较低。

表 4 社交媒体类 APP 隐私政策可读性分析结果(部分)

指标名称	微 博	微 信	知 乎	快 手	百度贴吧
命名方式	个人信息保护政策	隐私保护指引	个人信息保护指引	隐私权保护政策	隐私政策
小标题数量	9	9	11	10	9
导 语	有	有	有	有	有
文本篇幅(字数)	12 699	15 774	16 732	15 023	13 913
术语密度(%)	7.728	5.385	6.905	7.412	6.730
平均句长	6.556	6.870	8.416	5.814	7.231

整体来看,30 款社交媒体类 APP 的隐私政策在命名方式上均做到直观明了,便于用户快速定位,也设置了导语和一定数量的小标题,使得隐私政策的重点突出、结构分明。然而,不同 APP 的隐私政策可读性还存在较大差异,部分隐私政策存在导语内容敷衍、文本篇幅不妥、术语占比较高等

问题,导致了政策文本可读性和用户阅读意愿的下降。

3.3 主题识别

3.3.1 停用词去除

从表 3 看出,“信息”“个人信息”“服务”“提供”“收集”等高频词属于隐私政策文本中的常见

词汇, 广泛存在于隐私政策全文各处。若直接提取其向量特征作为 LDA 模型的输入参数, 易导致出现大量的重复、无效词汇, 影响主题识别效果。因此, 为尽量减少常见高频词对 LDA 主题识别的干扰, 本文首先采用去除局部停用词的方法, 将隐私政策文本中的常见高频词去除。

3.3.2 向量特征提取与主题数确定

LDA 主题模型是一种无监督学习模型, 本文选用主题一致性 (Coherence) 得分作为衡量模型优劣和确定主题数的依据, 主题一致性得分越高, 表示该主题模型效果越好。设定模型迭代次数为 200, 分别基于 TF 和 TF-IDF 特征绘制的主题一致性得分分布图, 结果如图 3 所示, 其中横坐标表示主题数, 纵坐标表示主题一致性得分。可以看出, 将 TF-IDF 特征作为 LDA 模型的输入参数, 其识别效果在不同主题数下均好于 TF 特征。因此, 本文基于 TF-IDF 特征进行 LDA 主题识别, 并根据主题一致性得分的极大值点设定最优主题数为 13。

3.3.3 可视化与结果分析

基于最优主题数, 利用可视化工具包 pyLDAvis 对所构建的 LDA 主题模型进行可视化建模, 结果如图 4 所示。图 4 左侧以二维坐标的形式展示了不

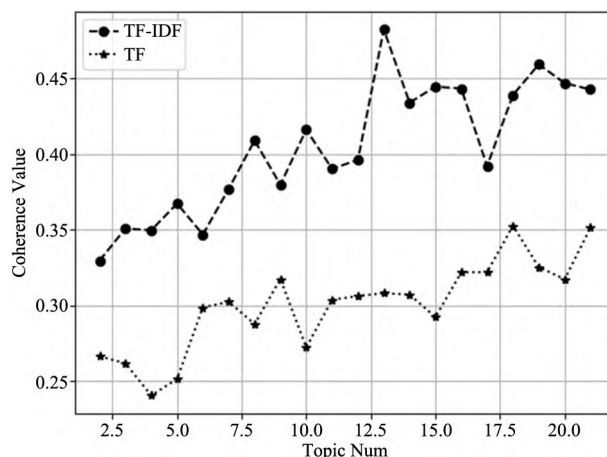


图 3 TF/TF-IDF 方法对 LDA 建模影响

同主题之间的相关程度, 圆圈大小表示包含该主题的文本在整篇文档之中的占比。右侧条形图是展示当前主题下前 30 个关键词的频率分布情况, 其中蓝色条块体现为整篇文档中的关键词频率, 红色条块体现为当前主题中的关键词频率。主题识别及关键词分布结果如表 5 所示。结合图 4 的可视化结果, 本文将 13 个政策主题进行归纳与合并, 整理得到 6 个类别, 分别是隐私政策概述、用户权利、运营商义务、运营商行为、数据存储与共享、法律保护 and 监管。

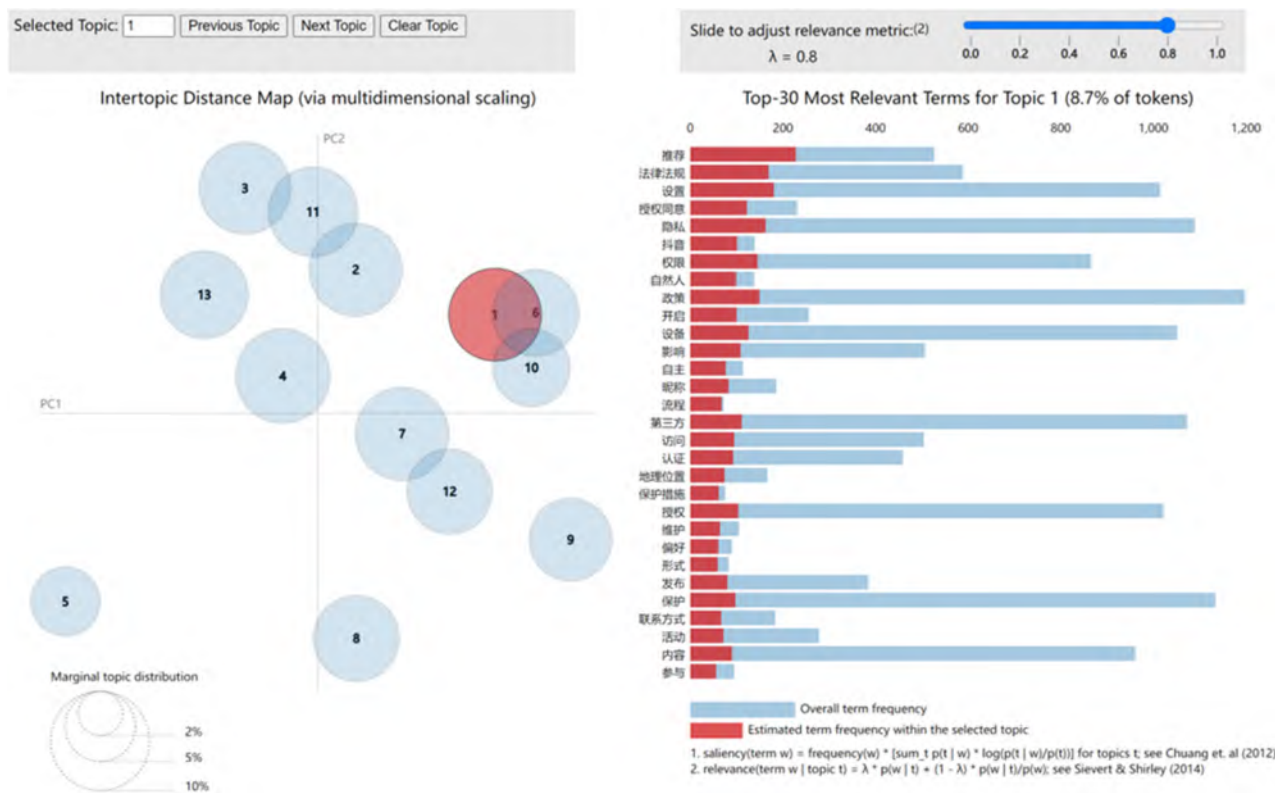


图 4 LDA 模型的可视化建模

1) 类别 1——用户权利。在图 4 中,主题 1 与主题 6、主题 10 之间存在一定的重叠。从关键词分布来看,主题 1 的主要内容是对用户权利主体进行界定,偏向于对用户权利的概述性说明。主题 6 的关键词包括“麦克风”“相册”“照片”等,主要是对 APP 所需权限的说明与解释,体现用户知情权。主题 10 的关键词体现了用户对 APP 的拒绝权利,如“注销”“取消”“拒绝”等词,对应用户拒绝权。因此,主题 1、主题 6 和主题 10 可合并概括为用户权利。

2) 类别 2——隐私政策概述。从图 4 可以看出,主题 2 与主题 3、主题 11 存在一定的重叠。从关键词分布来看,上述 3 个主题都属于隐私政策的概括性表述,但在具体内容上又各有侧重。主题 2 对要获取的数据进行了简单概述,关键词包括“请求”“业务功能”“手机号码”等。主题 3 侧重于争端出现与解决方面,关键词包括“违反”“争议”“地址”等。主题 11 对未来可能发生的变更与修改进行了概述性说明,关键词包括“修改”“发生变化”“通知”等。各主题虽然侧重点不同,但在具体内容上均呈现概述形式,较少涉及细节,因此,将主题 2、主题 3 和主题 11 合并以隐私政策概述作为概括。

3) 类别 3——运营商行为。从可视化结果上看,主题 5 较独立,单独处于一个区间,关键词包

括“人脸”“指纹”“身份证”“匿名化”等,均是指向性非常明确的细节词语,与用户敏感数据,尤其是生物特征数据紧密关联。从图 4 上看,主题 7 和主题 12 存在一定重叠,主题 7 的内容偏向于网络服务、Cookie、登录账号等,主题 12 的关键词偏向个性化服务(访问、视频、分享等),二者都属于运营商主体下的行为活动,而主题 5 对敏感数据的访问和使用处理同样也可以归至运营商行为。因此,将主题 5、主题 7 和主题 12 合并概括为运营商行为。

4) 类别 4——数据存储与共享。从可视化结果来看,主题 8 和主题 9 同处于一个区间。主题 8 的关键词包括“境内”“备份”“跨境”等,属于数据共享的主题内容,偏向对数据跨境流动做出规定。主题 9 的关键词“第三方”“合作方”“保存”“存储”等,更多涉及数据在运营商与第三方之间的共享。因此,将主题 8 和主题 9 合并为数据存储与共享。

5) 类别 5——法律保护和监管。主题 13 的关键词多涉及法律和制度层面,包括“法律法规”“国家”“遵守”等,因此,将其归至法律保护和监管类别。

6) 类别 6——运营商义务。主题 4 的关键词包括“记录”“查询”“用途”等,与运营商应承担的义务相关,因此将其归至运营商义务类别。

表 5 社交媒体类 APP 隐私政策的主题识别结果及关键词分布

编号	主题标签	关键词分布
1	用户权利 (主体界定)	法律法规 设置 授权同意 隐私 权限 自然人 政策 设备 自主 昵称 流程 第三方 访问 认证 保护措施 授权 维护 偏好 形式 保护
2	隐私政策概述 (信息获取)	账号 App 合作伙伴 获取 SDK 请求 注册 语音 识别 业务功能 同意 姓名 方式 网络 手机号码 开启 编辑 科技 推送 消息
3	隐私政策概述 (争端与解决)	平台 情况 政策 共享 关联 我方 违反 协议 授权 公司 委托 客服 情形 提出 地址 部门 行业 定义 过程 争议
4	运营商义务	第三方 指引 路径 义务 设备 情形 查询 需求 同意 技术 权利 用途 视频 记录 类型 隐私 管理 判断 存储 联系方式
5	运营商行为 (敏感数据处理)	管理 敏感 操作 人脸 提交 委托 选择 方式 名称 仅会 开通 订阅 文档 匿名化 对外 接口 联系人 同意 指纹 个人用户 身份证
6	用户权利 (知情权)	合作 权限 手机 实名 授权 验证 健康 存储 数据 广告 登录 沟通 相册 照片 客户端 麦克风 观看 变更 机制 意见 搜索 改进 删除
7	运营商行为 (Cookie 使用)	内容 Cookie 接收 登录 保护 展示 申请 方式 账号 公开 授权 推送 浏览 告知 音视频 互动 特定 第三方 信息处理 服务提供商 相机

表 5 (续)

编号	主题标签	关键词分布
8	数据存储与共享 (跨境流动)	头像 账号 保护 公司 必需 删除 注销 告知 事先 优化 填写 境内 发生 添加 未成年人 隐私 备份 跨境 主动 共享 保障
9	数据存储与共享 (第三方)	第三方 内容 登录 合作方 发布 视频 选择 保障 停止 统计 保存 关系 图 片 变化 身份 去标识化 公众 权限 删除 互动 存储 数据
10	用户权利 (拒绝权)	权限 隐私 设备 目的 隐私权 个人敏感信息 注销 文件 特别 取消 位置 状 态 数据 面部 弹窗 公告 提示 随时 号码 版本 原则 拒绝
11	隐私政策概述 (变化告知)	账户 政策 合法权益 风险 技术 授权 同意 保护 关注 密码 单独 确认 修 改 通知 应用程序 业务功能 分析 披露 发生变化 客户端 存储
12	运营商行为 (个性化服务)	访问 视频 分享 运行 中心 拒绝 展示 同意 明示同意 征得 位置 存储 目 的 个性化 商品 查阅 操作 关闭 性别 浏览 技术 姓名
13	法律保护与监管	身份 记录 授权 安全事件 共享 国家 征得 期间 广告 删除 法律法规 匹配 解决 安全性 方式 商品 业务 遵守 签署 生物 性别 唯一

3.4 主题对比分析

为分析社交媒体类 APP 隐私政策在主题上的全面性和规范性, 本文将政策文本的主题识别结果与国内外知名法律法规进行对比, 以期找出现有隐私政策的疏漏与不足之处。对比分析的对象选取《中华人民共和国个人信息保护法》(下称《个人信息保护法》)和欧盟颁布的《通用数据保护条例》(General Data Protection Regulation, GDPR)。

3.4.1 与《个人信息保护法》的对比结果

《个人信息保护法》于 2021 年 11 月 1 日正式施行, 对处理个人信息主体的权责、信息跨境流动规则、个人信息保护与监管等方面都做出了顶层设计, 具体内容主要包括信息处理规则、信息跨境流动规则、用户权利、运营商义务、部门监管和法律责任 6 个纲目。本文选取的 30 款社交媒体类 APP 的隐私政策主题与《个人信息保护法》内容纲目的对比结果如图 5 所示。可以看出, 《个人信息保护法》所规定的相关内容在隐私政策中均有所涉及, 但在内容完整性和规范性上还有所差异, 具体分析如下:

1) 从具体内容上来看, 《个人信息保护法》对个人信息处理和使用的多个情境都进行了具体说明和规则限制, 而本文识别的隐私政策主题关于这部分内容没有详细的阐述, 部分政策文本也没有提及有关同意撤回、紧急情况下信息使用等方面的信息处理规则。

2) 在信息跨境流动规则方面, 《个人信息保

护法》对个人隐私数据跨境流动的安全评估及信息保护均做出了明确的规定, 并列举了不得提供信息的情境, 而现有隐私政策主题中在信息跨境流动方面仍属于概述型描述, 缺少详细说明。

3) 在个人用户权利界定方面, 现有隐私政策对用户权利的界定和描述大多还停留在概述性文字上, 但《个人信息保护法》对用户知情权、拒绝权等众多权利都有着详细的定义和描述, 并且就个人向信息使用者和信息处理者提出合理要求的各种情况也有详细阐述, 上述内容在现有隐私政策中还处于严重欠缺状态, 这也是导致社交媒体隐私泄露和监管不力的根本原因。

4) 在信息处理者的义务方面, 现有社交媒体类 APP 的隐私政策依然存在内容完整性不足的问题。比如《个人信息保护法》对个人信息保护评估的需求和内容均做出了明确的阐释, 但大部分的隐私政策却鲜有体现。此外, 《个人信息保护法》对个人信息泄露、篡改以及丢失时所需要采取的补救措施做出了明确要求, 并规定必须落实到责任人或责任部门, 而现有隐私政策在此方面大多为概述性文字, 没有针对性的操作解释。

5) 在法律责任与处罚方面, 《个人信息保护法》的内容纲目与 30 款社交媒体类 APP 的隐私政策主题大体相同, 考虑到法律法规和隐私政策的发布主体不同, 这里不做详细对比分析。

3.4.2 与 GDPR 的对比结果

GDPR 对用户数据收集、存储、保护和使用的

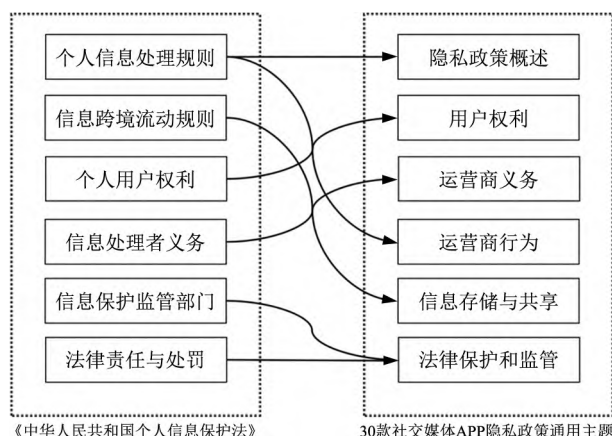


图 5 社交媒体类 APP 的隐私政策主题与《个人信息保护法》内容纲目的对比

标准进行了界定，赋予了用户对自身数据更大的操作权限，其内容纲目包括条例原则、数据处理权利、数据处理者、数据分享与跨境流动、监管机构体系、责任与处罚措施等内容。本文选取的 30 款社交媒体类 APP 的隐私政策主题与 GDPR 内容纲目的对比结果如图 6 所示，可以看出 GDPR 对生命周期视角下数据处理的全流程进行了清晰而详细的规定，对应主题上现有隐私政策虽有涉及，但在内容阐述和表现形式上仍有较大差异，具体分析如下：

1) 在数据主体权利方面，GDPR 详细规定了如何行使数据主体权利的步骤和模式，并具体阐述了访问权、更正权、收集权、擦除权、限制处理权、数据携带权等各项权利的详细内容与适用情境。在隐私政策主题中，对用户权利的具体划分远不如 GDPR 清晰、详细，这也导致了用户社交数据的使用失控及隐私侵犯。

2) 在数据处理者方面，GDPR 规定数据处理者所“操作”的数据范围并非控制者所拥有的全部数据，而是其中为满足使用目的所必须的部分数据，需满足最小特权原则。因此，在数据控制者未明确授权处理全部数据的前提下，数据处理者不得对任何逾越使用目的的其他数据进行任何处理。然而，现有社交媒体类 APP 的隐私政策对数据使用范围与时限甚少涉及，即使有所介绍，但用户不同意该规则便无法正常使用 APP，导致政策并无实际效果。

3) 在数据分享与跨境流动方面，用户身份、地理位置、网络习惯等社交媒体数据的严格管控和

合理利用对国家数据安全十分重要。GDPR 对数据主权、数据本地化及数据出入境均做出了明确的规定，并对特殊情形进行了补充说明，而现有社交媒体类 APP 的隐私政策在这一部分几乎是空白。

4) 在监管机构体系、责任与处罚措施方面，考虑到监管主体及适用法律的差异性，这里不做详细对比。

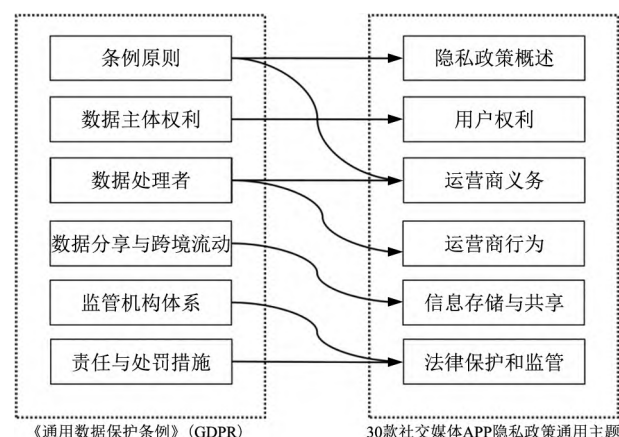


图 6 社交媒体类 APP 的隐私政策主题与 GDPR 内容纲目的对比

3.5 存在问题与优化策略

3.5.1 存在问题

通过对 30 款社交媒体类 APP 隐私政策的可读性分析与主题识别，可以发现当前隐私政策的制定总体上呈现不断完善趋势，但仍存在诸如完整性不足、针对性不强、可读性不高等问题。

1) 完整性不足。从 LDA 主题识别与相关法律法规的对比结果来看，现有社交媒体类 APP 隐私政策所涵盖的主题比较全面，但在具体的文字表述和内容解释上，仍存在完整性不足的情况。①用户权利沦为“查改删”。用户权利保障作为个人信息保护的重要组成部分，部分隐私政策只提及“查阅”“更正”“复制”“删除”的权利，而对用户知情权、决定权以及转移权等缺乏相关说明，在何种情形下可要求运营商删除个人信息也缺乏场景列举；②运营商的数据保护义务“空口无凭”。大多数社交媒体类 APP 隐私政策都会提及“保证用户的信息安全”，却鲜有提出针对性的解决方案和措施，内部管理制度和操作规程不透明，需要采用的安全技术措施也缺少详细说明，部分隐私政策甚至对可能发生的个人信息泄露等安全事件无预案说

明; ③信息存储和共享机制不够透明。信息存储的类型、地点、期限缺少详细说明, 哪些用户数据会被第三方共享、哪些数据涉及跨境流动缺少具体列举, 隐私政策中的信息存储共享相关内容总体还停留在文字性概述。

2) 针对性不强。从主题识别结果来看, 不同类型社交媒体类 APP 的隐私政策主题大同小异, 缺乏针对性。比如聊天社交类 APP(如“微信”)、社区社交类 APP(如“豆瓣”)、自媒体平台类 APP(如“哔哩哔哩”)缺乏基于自身平台特点的差异化表述, 隐私政策的主题内容和文本表述相似度较高, 部分段落、语句甚至出现明显雷同。

3) 可读性不高。隐私政策可读性不高会导致其感知易用性大大降低, 并影响用户的阅读意愿。部分社交媒体 APP(如“芒果 TV”“哔哩哔哩”)隐私政策的篇幅明显过长, 需要耗用户大量的阅读时间。此外, 部分 APP(如“秒拍”)隐私政策的导语部分流于形式, 未能起到帮助用户快速掌握全文要点以及辅助阅读的效果。还有部分 APP(如“陌陌”)隐私政策的专业术语过多, 给用户理解政策内容带来较大难度。

3.5.2 优化策略

针对上述问题, 本文从内容完整性、条款针对性和文本可读性 3 个方面提出优化建议。

1) 对标法律法规, 改进内容完整性。我国在个人信息保护方面已陆续制定了相关法律制度, 立法层面的不断完善为社交媒体类 APP 的隐私政策改进提供了权威参照。①用户权利层面。结合《个人信息保护法》与 GDPR 的对比结果, 现有隐私政策首先要在用户权利层面予以明确和充实, 隐私政策中应将法律规定的用户在个人信息活动中所拥有的各项权利均予以明确说明, 并同时给出情境列举, 使用户通过阅读隐私政策, 不仅能够全面了解自身有哪些权利, 更能够清晰认识到自身在何种情况下可使用这些权利对个人信息进行保护; ②运营商义务层面。运营商义务和用户权利应具有对等篇幅, 需要在隐私政策文本中体现“权责对等”以及运营商履行义务的详细步骤, 比如社交媒体运营方的内部管理制度和操作规程、发生信息安全问题时的应急方案、以及因个人信息保护出现问题时

责任承担与处罚赔偿等方面的具体措施。同时考虑到文本可读性和逻辑结构的要求, 在适当位置应用小标题突出标注; ③数据共享与跨境流动层面。随着互联网、大数据、人工智能和实体经济的深度融合, 数据作为基础性战略资源的重要性地位日益凸显, 监管层对个人身份、地理位置、网络行为等信息的采集、利用以及跨境流动等方面已重点关注, 相关法律和规定也在这方面做出了详细规定和要求, 社交媒体类 APP 的隐私政策有必要对这一部分加以详细标注, 在内容上明确阐述个人信息的存储和使用规则, 以及有可能发生的信息共享和跨境流动情形, 以保障用户对自身信息流动的知情权。

此外, 考虑到社交媒体用户分布的全球性, 隐私政策除了对标国内法律法规外, 也需要借鉴多个国家(地区)的个人信息保护条款, 进一步完善其内容的规范性和全面性。

2) 立足应用特点, 加强条款针对性。社交媒体类 APP 隐私政策的优化与完善要避免“泛泛而谈”的情形, 参考国内外相关法律制度并不表示文本主题一定要趋同。社交媒体类 APP 根据其用途, 可进一步划分为聊天类、博客类、问答社区类、社交电商类和短视频社交类等诸多领域, 隐私政策应结合其应用领域加强条款针对性, 力求重点突出。例如, 针对聊天类 APP, 聊天记录的存储与信息安全有必要重点说明; 针对短视频社交 APP, 隐私条款需要对人物图像等生物信息的安全保护进行针对性说明。社交电商类 APP 对支付账户、用户资产等隐私数据的利用与保护应着重阐述。

3) 创新外在形式, 提高文本可读性。隐私政策文本可读性的提升涉及多方面内容, 主要包括: ①隐私政策在篇幅上要做到合理控制, 对无关内容(如产品介绍、功能特点等)应做到应减尽减, 将篇幅留给隐私政策的主体内容, 尽可能减少文字冗余; ②避免导语内容流于形式。导语设置应立足于帮助用户快速领悟全文要点, 并辅助用户快速查找自身关注信息, 因此导语长度不宜过长, 应做到简洁明确、重点突出; ③尽量减少专业术语的使用, 对确实有必要使用术语的段落, 应在易于查找的部分标注, 并予以详细的名词解释, 最大程度降低用户的阅读难度。

4 结 语

本文基于 TAM 模型,从感知易用性和感知有用性两个维度对 30 款社交媒体类 APP 的隐私政策进行文本分析。针对隐私政策的感知易用性,从命名方式、小标题数量、文本篇幅等方面对隐私政策文本的可读性进行测评。针对隐私政策的感知有用性,运用 LDA 主题模型识别出用户权利、隐私政策概述、运营商行为、数据存储与共享、法律保护、监管与运营商义务 6 个类别的主题,并与国内外法律法规进行对比。研究结果发现,现有社交媒体类 APP 隐私政策存在完整性不足、针对性不强、可读性不高等问题,需要在用户权利、运营商业务、数据流动与跨境共享层面进一步完善内容,创新外在呈现形式,并根据 APP 类型的细分定位制定差异化的隐私政策。

本文选取下载量靠前的 30 款社交媒体类 APP 隐私政策文本作为研究对象,并没有涵盖使用人数较少的冷门社交媒体平台,在未来的研究中可以扩大样本范围。在研究方法层面,未来可以结合情境实验法、实证分析法对隐私政策的阅读意愿及规范性评价进行交叉比较研究。

参 考 文 献

- [1] Ma X, Qin Y, Chen Z, et al. Perceived Ephemerality, Privacy Calculus, and the Privacy Settings of an Ephemeral Social Media Site [J]. Computers in Human Behavior, 2021, 124: 106928.
- [2] 中华人民共和国国家互联网信息办公室. 电信和互联网用户个人信息保护规定 [EB/OL]. http://www.cac.gov.cn/2012-07/29/c_133142088.htm, 2022-03-02.
- [3] Gdpr Data Protection Statement [EB/OL]. <https://www.fendahl.com/gdpr-data-protection-statement>, 2022-03-04.
- [4] 中华人民共和国国家互联网信息办公室. 常见类型移动互联网应用程序(App) 必要个人信息范围(征求意见稿) [EB/OL]. http://www.cac.gov.cn/2020-12/01/c_1608389002456595.htm, 2022-03-04.
- [5] 中华人民共和国国家互联网信息办公室. 国家互联网信息办公室关于《网络数据安全条例(征求意见稿) 》公开征求意见的通知 [EB/OL]. http://www.cac.gov.cn/2021-11/14/c_1638501991577898.htm, 2022-03-04.
- [6] Kumar R, Kumar P, Kumar V. Design and Implementation of Privacy and Security System in Social Media [J]. International Journal of Advanced Networking and Applications, 2022, 13 (4): 5081-5088.

- [7] 张艳丰, 邱怡. 我国移动阅读应用个人信息保护政策合规性测度研究 [J]. 图书情报工作, 2021, 65 (22): 35-43.
- [8] 中华人民共和国国家互联网信息办公室. 关于征求《App 违法违规收集使用个人信息行为认定方法(征求意见稿) 》意见的通知 [EB/OL]. http://www.cac.gov.cn/2019-05/06/c_1124455735.htm, 2022-03-04.
- [9] 搜狐. 重要提醒! 工信部通报 35 款 APP: 涉嫌收集个人信息 [EB/OL]. https://www.sohu.com/a/433599022_758182, 2022-03-06.
- [10] 杜荷花. 国外政府数据开放平台隐私保护政策的考察与借鉴 [J]. 图书馆建设, 2020, (3): 67-75.
- [11] Pew Research Center. Privacy in the Digital Age [EB/OL]. <https://www.pewinternet.org/2015/06/03/privacy-in-the-digital-age>, 2022-03-13.
- [12] Steinfeld N. "I Agree to the Terms and Conditions": (How) Do Users Read Privacy Policies Online? An Eye-tracking Experiment [J]. Computers in Human Behavior, 2016, 55: 992-1000.
- [13] Singh R I, Sumeeth M, Miller J. A User-centric Evaluation of the Readability of Privacy Policies in Popular Web Sites [J]. Information Systems Frontiers, 2011, 13 (4): 501-514.
- [14] Vail M W, Earp J B, Antón A I. An Empirical Study of Consumer Perceptions and Comprehension of Web Site Privacy Policies [J]. IEEE Transactions on Engineering Management, 2008, 55 (3): 442-454.
- [15] 朱侯, 张明鑫. 移动 APP 用户隐私信息设置行为影响因素及其组态效应研究 [J]. 情报科学, 2021, 39 (7): 54-62.
- [16] 朱光, 李凤景, 沈雨萌, 等. 社交媒体隐私政策的阅读意愿研究——基于 TAM 模型与自我效能理论视角 [J]. 现代情报, 2022, 42 (1): 150-166.
- [17] 张玥, 王坚, 余姝, 等. 信息表征对移动医疗 APP 隐私政策阅读效果的影响研究——基于认知负荷理论 [J]. 图书情报工作, 2021, 65 (11): 3-13.
- [18] 袁红, 张苏娇. "互联网+政务服务" 环境下政务 APP 隐私政策的内容分析 [J]. 现代情报, 2022, 42 (3): 121-132.
- [19] 储节旺, 丁辉. 美国政府开放数据个人隐私保护政策及对我国的启示——基于 52 个政策文本的内容分析 [J]. 图书情报工作, 2021, 65 (8): 140-150.
- [20] 何培育, 王潇睿. 智能手机用户隐私安全保障机制研究——基于第三方应用程序“隐私条款”的分析 [J]. 情报理论与实践, 2018, 41 (10): 40-46.
- [21] 严炜炜, 谢顺欣, 曹灿瑜. 在线学术社区隐私保护比较研究: 保护维度、功能设置和优化路径 [J/OL]. 图书馆论坛: 1-12 [2022-03-30]. <http://kns.cnki.net/kcms/detail/44.1306.G2.20220121.1700.010.html>.
- [22] Martiny K, Briesemeister L, Denker G, et al. Protecting Privacy during a Pandemic Outbreak [C] //ICISSP, 2021: 308-318.

(下转第 70 页)

- [12] 王玉林, 钟敏. 数据原权利人的权利与限制 [J]. 情报理论与实践, 2017, 40 (1): 36-40.
- [13] 王东方. 涉数据权利的区分及其保护的应然路径 [J]. 图书馆建设, 2022, (3): 44-51.
- [14] Dosis A, Sand-Zantman W. The Ownership of Data [J]. The Journal of Law, Economics, and Organization, 2022.
- [15] 周环, 幸强国, 唐泳. 基于政策文本计算的数据开放与隐私保护政策协同度研究 [J]. 图书馆论坛, 2021, 41 (11): 118-127.
- [16] 张涛, 马海群. 基于政策文本计算的开放数据与数据安全政策协同研究 [J]. 情报理论与实践, 2020, 43 (6): 149-155, 141.
- [17] 洪伟达, 马海群. 我国开放政府数据政策的演变和协同研究——基于 2012—2020 年政策文本的分析 [J]. 情报杂志, 2021, 40 (10): 139-147, 138.
- [18] 张玲玲. 基于内容分析的政府数据开放政策协同研究 [J]. 图书馆研究与工作, 2020, (3): 42-48.
- [19] 毛子骏, 郑方, 黄膺旭. 政策协同视阈下的政府数据开放研究 [J]. 电子政务, 2018, (9): 14-23.
- [20] 洪伟达, 马海群. 我国开放政府数据政策协同机理研究 [J]. 情报科学, 2020, 38 (5): 126-131.
- [21] 姜鑫, 王德庄. 开放科学数据与个人数据保护的政策协同研究——基于政策文本内容分析视角 [J]. 情报理论与实践, 2019, 42 (12): 49-54, 93.
- [22] 徐蕾, 李庆, 肖相泽. 基于扎根理论的大数据政策共词网络研究 [J]. 现代情报, 2018, 38 (6): 157-164.
- [23] 唐要家. 数据产权的经济分析 [J]. 社会科学辑刊, 2021, (1): 98-106, 209.
- [24] 万方. 大数据背景下个人信息保护的新路径——被遗忘权的更新兼论信息技术革新 [J]. 图书情报知识, 2016, (6): 107-113.
- [25] 闫境华, 石先梅. 数据生产要素化与数据确权的政治经济学分析 [J]. 内蒙古社会科学, 2021, 42 (5): 113-120.
- [26] 文禹衡, 贺亚峰. 数据产权市场主体的认知调查与矫正 [J]. 图书馆论坛, 2022, 42 (3): 31-42.

(责任编辑: 陈 媛)

(上接第 47 页)

- [23] Kaur J, Dara R A, Obimbo C, et al. A Comprehensive Keyword Analysis of Online Privacy Policies [J]. Information Security Journal: A Global Perspective, 2018, 27 (5-6): 260-275.
- [24] Strahilevitz L J, Kugler M B. Is Privacy Policy Language Irrelevant to Consumers? [J]. The Journal of Legal Studies, 2016, 45 (S2): S69-S95.
- [25] 赵杨, 严周周, 沈棋琦, 等. 基于机器学习的医疗健康 APP 隐私政策合规性研究 [J/OL]. 数据分析与知识发现: 1-19 [2022-03-30]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20211126.1717.006.html>.
- [26] 张艳丰, 邱怡. 硬规则下我国移动阅读 APP 隐私政策合规性研究 [J]. 现代情报, 2022, 42 (1): 167-176.
- [27] 姚胜译, 吴丹. APP 隐私政策用户友好度评价研究 [J]. 信息资源管理学报, 2021, 11 (1): 30-39, 58.
- [28] 周林兴, 徐承来. 信息安全视域下国内档案网站隐私政策合规性研究 [J]. 档案学研究, 2022, (1): 85-91.
- [29] Chang Y, Wong S F, Libaque-Saenz C F, et al. The Role of Privacy Policy on Consumers' Perceived Privacy [J]. Government Information Quarterly, 2018, 35 (3): 445-459.
- [30] Aïmeur E, Lawani O, Dalkir K. When Changing the Look of Privacy Policies Affect User Trust: An Experimental Study [J]. Computers in Human Behavior, 2016, 58: 368-379.
- [31] Yoon C. Extending the TAM for Green IT: A Normative Perspective [J]. Computers in Human Behavior, 2018, 83: 129-139.
- [32] Sur S, Goswami K. Digital Privacy: Case Study Analysis on WhatsApp Privacy Policy Changes [J]. International Journal of Applied Science and Engineering, 2021, 9 (2): 157-167.
- [33] Elkoumy G, Fahrenkrog-Petersen S A, Sani M F, et al. Privacy and Confidentiality in Process Mining: Threats and Research Challenges [J]. ACM Transactions on Management Information System, 2021, 13 (1): 1-17.
- [34] Kim E H J, Jeong Y K, Kim Y H, et al. Exploring Scientific Trajectories of a Large-scale Dataset Using Topic-integrated Path Extraction [J]. Journal of Informetrics, 2022, 16 (1): 101242.
- [35] Hagen L. Content Analysis of E-petitions with Topic Modeling: How to Train and Evaluate LDA Models? [J]. Information Processing & Management, 2018, 54 (6): 1292-1307.
- [36] Chen Y, Zhang H, Liu R, et al. Experimental Explorations on Short Text Topic Mining Between LDA and NMF Based Schemes [J]. Knowledge-Based Systems, 2019, 163: 1-13.
- [37] Martinc M, Pollak S, Robnik-Šikonja M. Supervised and Unsupervised Neural Approaches to Text Readability [J]. Computational Linguistics, 2021, 47 (1): 141-179.
- [38] 朱侯, 张明鑫, 路永和. 社交媒体用户隐私政策阅读意愿实证研究 [J]. 情报学报, 2018, 37 (4): 362-371.
- [39] 王军, 李子舰, 刘潇蔓. 不同文本长度的体验型产品在线评论时间序列研究——以电影评论为例 [J]. 图书情报工作, 2019, 63 (16): 103-111.

(责任编辑: 陈 媛)