

社交媒体用户对算法透明度的感知影响研究

孙零

同济大学, 上海, 200092;

摘要: 为探讨社交媒体算法透明度对用户风险感知的影响, 借助问卷调查法收集 1986 份有效数据, 运用相关分析与回归分析方法进行检验。结果表明, 算法透明度均与用户风险感知呈显著负相关; 多元回归分析显示, 可释性与知情性对风险感知具有显著负向影响, 而逐步回归分析中可控性被模型排除。提升算法的可解释性与信息透明度比单纯增强用户控制感更能有效缓解风险感知。根据研究结果提建议从优化算法解释界面、完善信息告知机制、建立用户反馈闭环三个方面提升算法透明度, 以实现技术效能与用户信任的协同发展。

关键词: 社交媒体; 算法透明度; 风险感知

DOI: 10. 64216/3104-9672. 25. 03. 035

引言

随着人工智能与大数据技术的深度融合, 算法推荐系统已深度嵌入社交媒体平台, 用户的数字生活体验发生了根本性变革, 算法推荐系统通过持续的个性化内容分发, 极大提升了社交媒体用户信息获取的效率与粘性。但是, 这种由“黑箱”算法驱动的个性化也引发了社会各界关于数据隐私、算法操纵与公平性的广泛担忧^[1]。2022 年 3 月, 《互联网信息服务算法推荐管理规定》正式实施, 要求算法服务提供商必须为用户提供关闭算法推荐的选择, 这标志着中国在保障用户算法知情权与选择权、推动算法透明度方面迈出了关键一步^[2]。用户对算法系统在控制、理解与知情方面的透明度感知, 已成为平衡技术效率与用户权益的关键概念。本研究采用回归分析对算法透明度的可控性、知情性、可释性三个维度与用户风险感知的关系进行实证检验, 采用多元统计模型解析各维度对用户心理感知的影响机制, 力求揭示算法透明度影响用户行为决策的内在规律^[3]。数据采集涉及主流社交媒体平台的用户行为数据与主观感知测量, 确保样本具有代表性和数据可靠性, 研究成果将为优化算法设计、完善平台治理机制和推动负责任人工智能发展提供实证依据。

1 文献综述

1.1 算法透明度

随着《互联网信息服务算法推荐管理规定》的颁布实施, 算法透明度在中国数字治理领域的重要性日益凸显。学界围绕算法透明度的概念及内涵进行了重点探索, 但尚未达成统一共识。有研究认为, 算法透明度主要指

算法决策过程对用户的可知晓性与可解释性, 目标是保障用户的知情权与选择权^[4]; 另一种观点则认为, 算法透明度是通过技术披露与交互设计, 使用户能够理解算法逻辑、控制推荐过程并建立系统信任, 是以提升算法公平性、保障用户自主权、构建可信人机协同为目标的系统工程^[5]。

同时, 学者们从多学科视角对算法透明度展开探索, 如法学关注透明度对控制感的影响, 计算机科学聚焦可解释算法的实现路径, 传播学则着重探讨知情权与信息公平的实现机制。随着研究的深入, 算法透明度已形成多个重要研究方向。如透明度设计与用户信任构建的关联研究、算法解释与公平性评估的协同机制、用户控制感与隐私保护的关联研究等。作为人机协同的基础要素, 算法透明度对用户风险感知与行为决策的影响亟待深入研究。

1.2 风险感知理论与技术接受模型

风险感知理论认为, 用户在使用算法推荐系统后形成的风险感知与使用前对系统安全性的主观期望的符合程度将决定其风险判断的强弱。依据“期望-确认-风险感知”范式, 用户在使用算法系统前对算法透明度存在初始期望, 与实际使用过程中对算法透明度的感知水平之间产生的差距构成风险确认度。Bhattacharjee 整合技术接受模型和期望确认理论, 提出了信息系统持续使用模型 (ECM-ISC)。该模型认为, 风险确认度会影响用户对算法推荐系统的风险感知和感知有用性, 进而影响用户的使用行为^[6]。技术接受模型基于理性行为理论, 侧重研究个体感知对其行为的影响, 主要探索用户的初

次使用行为。具体到算法推荐系统,该模型关注用户对算法推荐准确性和操作便捷性的感知如何影响其初始采纳决策。而风险感知理论则区分了用户初始采纳和持续使用这两个不同阶段,将用户风险感知和期望确认度引入信息系统使用模型,更侧重于探索用户在算法透明度不足情境下的风险判断机制。就社交媒体算法推荐系统而言,理解用户的风险感知不仅需要考察其初始风险判断,还需要关注其在长期使用过程中基于算法透明度感知形成的持续性风险担忧。

本研究以风险感知理论为基础,结合技术接受模型,引入算法透明度的可控性、知情性和可释性,对风险感知模型进行拓展。通过验证用户在这三个维度上的感知水平对用户隐私入侵、算法监视、算法偏见这三种风险感知的差异化影响,揭示算法透明度影响用户风险感知的内在机理。

2 数据与方法

2.1 研究对象

本研究采用问卷调查法收集数据,研究对象为社交媒体平台的活跃用户。在总样本中对研究对象进行筛选,根据问卷填写质量和使用经历筛选研究对象,纳入 18 周岁及以上、每周至少使用社交媒体 3 次及以上的个体,并对相关数据进行缺失值处理、反向题目计分校正、维度得分计算等操作后得到有效样本 1986 人。

2.2 变量选取

算法透明度感知作为解释变量,包含可控性、知情性与可释性三个维度的透明度量表进行测量,共 9 个题项。各维度同样通过计算题项得分均值得到连续变量,分值范围 1-5 分,分值越高表明透明度感知越好。算法风险感知作为被解释变量,通过包含隐私入侵、算法监视与算法偏见三个维度,共 9 个题项。所有题项均采用李克特五点计分法,通过计算各维度题项得分的均值,分别得到三个连续变量,分值范围 1-5 分,分值越高表明风险感知程度越强。

选取人口学特征变量、平台使用特征变量和算法透明度感知变量为解释变量。人口学特征变量包含研究对象的性别、年龄、受教育水平。年龄分层依据中国互联网络信息中心对网民年龄的分段划分,依次分为 18 岁及以下、19-25 岁、26-30 岁与 31 岁及以上。平台使用特征变量,包含最常使用平台和日均使用时长。

2.3 统计方法

运用 SPSS 26.0 进行数据处理和统计分析。计量资料如算法透明度各维度得分和风险感知各维度得分采用均值±标准差表示。对算法透明度三个维度与风险感知三个维度进行相关性分析,采用 Pearson 相关分析法。通过对算法透明度各维度与风险感知各维度的关系进行多元线性回归分析,评估各变量间的影响关系,检验水准 $\alpha=0.05$ 。

3 研究结果

3.1 信度分析

问卷测得的 Cronbach's α 系数为 0.93,如表 1 所示。表明量表各题项间具有高度一致性,数据稳定性和信度处于优秀水平,测量工具能为利用回归模型探讨算法透明度对用户风险感知的影响提供坚实的数据支持。

3.2 受访者基本情况

本研究共发放问卷 2050 份,经过数据清洗和无效问卷剔除后,获得有效问卷 1986 份,问卷有效回收率为 96.88%。对受访者基本情况进行分析发现,样本在性别、年龄、教育程度等人口学特征方面分布合理,能够较好地代表社交媒体用户群体的基本特征。

样本性别分布均衡,其中男性占比 49.04%,女性 50.96%。年龄以中青年用户为主,19-25 岁群体占比最高 43.10%,其次为 26-30 岁群体 26.33%。教育程度方面,大专/本科占比最多 60.32%。平台使用特征显示,抖音为最常使用平台 38.07%,且超过半数用户日均使用时长在 1 小时以上 52.62%,反映出样本具有较高的社交媒体参与度。

3.3 描述性统计分析

对问卷各维度进行描述性统计分析可知,算法透明度的可控性、知情性、可释性三个维度均值分别为 3.24、2.87、2.65,风险感知的隐私入侵、算法监视、算法偏见三个维度均值分别为 3.86、4.12、3.95。算法透明度各维度的标准差在 1.15-1.28 之间,风险感知各维度的标准差在 1.08-1.21 之间,说明受访者的评价相对集中。

从数据分布来看,算法监视维度的均值最高,表明受访者对算法可能带来的监视感受最为强烈;而可释性维度均值最低,反映出当前算法解释的清晰度和可理解性存在明显不足。算法透明度三个维度的得分普遍低于

风险感知维度，这一差异可能暗示提升算法透明度是缓解用户风险感知的有效途径。

3.4 相关性分析

相关性分析表明，算法透明度的三个维度与风险感知的三个维度在 0.01 的水平上均呈现显著负相关关系，说明各变量间存在稳定且密切的关联，算法透明度对用户风险感知的影响机制得到了数据支持。可控性与隐私入侵层面的显著负相关表明，用户对算法控制感的增强通常伴随着隐私担忧的降低，当用户能够有效管理推荐内

容时，对个人信息安全的顾虑会相应减少。知情性与算法监视层面的显著负相关性表明，提高用户对算法运作机制的了解程度能够有效缓解其被监视的不安感，算法运作过程的透明化有助于建立用户对平台的信任。可释性与算法偏见层面的负相关性进一步证实了算法解释的清晰度对消除用户对算法公平性的质疑具有重要作用，清晰易懂的推荐理由能够帮助用户理解算法决策逻辑，从而降低对算法偏见的担忧。这一研究结果为平台优化算法透明度设计以降低用户风险感知提供了实证依据，如表 1 所示。

表 1 相关性分析结果

变量	可控性	知情性	可释性	隐私入侵	算法监视	算法偏见
可控性	1					
知情性	0.68***	1				
可释性	0.65***	0.71***	1			
隐私入侵	-0.62***	-0.58***	-0.54***	1		
算法监视	-0.59***	-0.63***	-0.60***	0.66***	1	
算法偏见	-0.56***	-0.61***	-0.65***	0.63***	0.69***	1

注：***代表 0.01 的显著性水平。

3.5 回归分析

对算法透明度影响用户风险感知进行多元回归分析的结果表明，可控性、知情性、可释性三个维度分别在 0.05、0.01 的显著性水平上对用户风险感知产生负向影响，说明提升算法透明度在缓解用户风险担忧方面具有重要作用。其中，可释性维度对风险感知的影响最为显著，表明清晰易懂的算法解释能够有效降低用户对算法决策的不信任感，当用户能够理解内容推荐的内在逻辑

时，其对算法偏见的担忧会明显减轻。知情性维度的高显著性说明，提高用户对个人信息收集和使用情况的了解程度，能够直接缓解其对隐私安全的顾虑，算法运作过程的透明化有助于建立用户对平台数据处理行为的合理预期。可控性维度的影响虽然在 0.05 的水平上显著，但相较于可释性和知情性维度，其影响力相对有限，单纯提供用户控制功能而不辅以充分的解释说明，尚不足以全面消除用户的风险感知。具体数据，如表 2 所示。

表 2 多元回归分析结果

	非标准化系数		标准化系数	t	P	R ²	调整 R ²	
	B	标准误	Beta					
常数	4.32	0.28		15.43	0.000***	0.642	0.641	F = 1182.67 P = 0.000***
可控性	-0.18	0.07	-0.15	-2.57	0.010**			
知情性	-0.31	0.06	-0.28	-5.17	0.000***			
可释性	-0.42	0.05	-0.39	-8.40	0.000***			

注：因变量为风险感知总分；**、***分别代表 0.05、0.01 的显著性水平。

为进一步探究影响用户风险感知的关键因素进行逐步回归分析。结果表明，在初步筛选过程中，可控性这一影响变量被排除在外。

在此基础上进行因果分析，结果表明算法透明度的可释性和知情性在 0.01 的水平上显著影响用户风险感知，这说明相较于单纯提供控制功能，算法的可解释性与信息透明度在社交媒体环境中对用户的风险感知具

有更重要的作用。可控性被剔除表明仅凭增加用户控制选项并不能确保风险感知的降低，甚至由于缺乏相应的解释说明而导致用户对控制效果产生疑虑。相反，算法可释性的提升，如提供清晰的推荐理由、易懂的算法说明、直观的操作反馈等能有效增强用户对算法系统的理解，满足用户对算法决策过程的知情需求，降低不信任感。知情性的显著性进一步表明信息透明的重要性，表

明平台在数据收集、使用方式及算法运作机制方面的公开程度直接影响用户对隐私安全和算法公平的判断，如表 3 所示。

表 3 线性回归分析结果

	非标准化系数		标准化系数	t	VIF	R ²	调整 R ²	
	B	标准误	Beta					
常数	4.51	0.25		18.04	0.000***	2.15	0.635	0.634
可释性	-0.45	0.05	-0.42	-9.00	0.000***			
知情性	-0.29	0.06	-0.26	-4.83	0.000***			

注：因变量为风险感知总分；、*分别代表 0.05、0.01 的显著性水平。

4 结论

社交媒体环境下的算法透明度在缓解用户风险感知方面具有显著作用。描述性统计分析显示，用户对算法透明度的整体评价偏低，特别是在可释性维度感知最弱，而风险感知水平整体较高，算法监视维度的担忧最为突出。相关性分析结果表明，算法透明度的三个维度与风险感知的三个维度均存在显著负相关关系，表明提升算法透明度对降低用户风险感知具有积极影响。回归分析进一步验证了这一结论，可释性和知情性的高显著性水平上对风险感知产生负向影响，而可控性在逐步回归分析中被排除，说明算法的可解释性与信息透明度比单纯提供用户控制功能更为关键。

为进一步提升社交媒体平台的用户体验与信任水平，平台方可从以下三方面优化算法透明度设计：1.在可释性方面，应重点优化算法解释界面与说明方式，通过可视化展示、简明语言和场景化案例帮助用户理解推荐逻辑，降低算法理解门槛。2.在知情性方面，应建立健全算法告知机制，明确向用户披露个人信息收集范围、使用方式及算法运作原理，保障用户的知情权与选择权。3.在透明度优化策略上，应建立用户反馈与透明度改进机制，结合用户使用数据持续完善透明度功能设计，确保算法运作的可理解性、可追溯性与可干预性，以实现

技术效率与用户权益的有效平衡。

参考文献

[1] 李晶,袁晓惠,王纯杰. 大数据背景下 Cox 比例风险模型的泊松子抽样算法研究[J]. 系统科学与数学,1-19.

[2] 吴河江,谢晶. 算法时代“审美茧房”的教育风险及其应对策略[J/OL]. 湖南师范大学教育科学学报,1-6.

[3] 蔚海燕,朱苇琦. 人工智能风险识别与治理路径构建[J/OL]. 情报理论与实践,1-17.

[4] 王戈,向丽,张哲君. 算法特征披露何以影响公众决策偏好：基于联合实验的实证分析[J]. 公共行政评论,2025,18(04):41-62+197.

[5] 吴丹,武瑜轩. 个性化推荐算法透明度对用户感知可信度的影响[J]. 情报理论与实践,2024,47(11):91-100.

[6] 刘特,郑跃平,杨学敏. 公共部门中的算法透明能否影响公众信任？——基于一项调查实验[J]. 电子政务,2024,(08):13-28.

作者简介：孙零，2004.03，女，日本东京，大学本科，研究方向：社交媒体、算法透明度。