arXiv:2601.01094v1 [cs.HC] 3 Jan 2026

SoulSeek ：探索社交线索在基于大语言模型的信息寻求中的应用

YUBO SHU，复旦大学，中国PENG ZHANG，复旦大学，中国MENG WU，联想集团，中国YAN CHEN，弗吉尼亚理工大学，美国HAOXUAN ZHOU，复旦大学，中国GUANMING LIU，复旦大学，中国YU ZHANG，联想集团，中国LIUXIN ZHANG，联想集团，中国QIANYING WANG，联想集团，中国TUN LU，复旦大学，中国NING GU，复旦大学，中国

社交线索传达他人的存在、行为或身份，在人类信息寻求中发挥着关键作用，帮助个体判断相关性与可信度。然而，现有的基于大语言模型的搜索系统主要依赖语义特征，与自然信息寻求背后的社会化认知存在不一致。为弥合这一差距，我们探究将社交线索融入基于大语言模型的搜索后，如何影响用户的感知、体验和行为。聚焦于开始采用基于大语言模型的搜索的社交媒体平台，我们结合设计研讨会、原型系统（SoulSeek）的实现、被试间研究以及混合方法分析，考察结果层面和过程层面的发现。研讨会为原型系统的线索整合设计提供了依据。研究表明，社交线索能够改善用户感知的结果和体验，促进反思性信息行为，并揭示当前基于大语言模型的搜索的局限性。我们提出了相关设计启示，强调更好的社会知识理解、个性化线索设置以及可控交互。

附加关键词与短语：社交线索、社交内容平台、基于大语言模型的搜索、信息寻求

1 引言

信息寻求通常被视为这样一个过程：当个人遇到知识缺口时，会搜索并使用信息。在信息时代，信息寻求已变得无处不在[5], ，信息系统的快速发展大幅提升了信息的可获取性，同时也加剧了信息过载[49]。因此，个人越来越依赖多种线索来筛选信息并理解信息

作者联系信息：Yubo Shu，复旦大学，中国上海，ybshu20@fudan.edu.cn；Peng Zhang，复旦大学，中国上海，zhangpeng_@fudan.edu.cn；MENG WU，联想集团，中国北京，wumeng5@lenovo.com；YAN CHEN，弗吉尼亚理工大学，美国弗吉尼亚州布莱克斯堡，ych@vt.edu；HAOXUAN ZHOU，复旦大学，中国上海，24210240428@m.fudan.edu.cn；GUANMING LIU，复旦大学，中国上海，gmliu24@m.fudan.edu.cn；YU ZHANG，联想集团，中国北京，zhangyu29@lenovo.com；LIUXIN ZHANG，联想集团，中国北京，zhanglx2@lenovo.com；QIANYING WANG，联想集团，中国北京，wangqya@lenovo.com；Tun Lu，复旦大学，中国上海，lutun@fudan.edu.cn；Ning Gu，复旦大学，中国上海，ninggu@fudan.edu.cn。

允许免费为个人或课堂使用而制作本作品全部或部分内容的数字或纸质副本，但不得出于营利或商业目的制作或分发副本，且副本首页必须附上本声明和完整引文。必须尊重非作者拥有的本作品组件的版权。注明出处后允许摘录。如需以其他方式复制、再版、发布至服务器或向列表重新分发，须事先获得明确许可和/或支付费用。请通过 permissions@acm.org 申请许可。© 2026 版权归所有者/作者所有。出版权已授权给 ACM。稿件已提交至 ACM

[2, 14, 61]。这些线索既包括信息线索（内容中的语义特征）[35]，也包括超越语义的**社交线索** [62]。社交线索能够反映信息交互中他人的存在、行为或身份，例如信息发布者的身份、内容的作者风格，或其他读者的反馈。社交线索在人类信息寻求中发挥着关键作用。从用户角度来看，社交线索是情境因素，能够帮助人们更有效地评估信息的可信度和相关性，尤其是在信息过载和存在不确定性的情况下，从而影响注意力分配和决策倾向 [25, 56]。从系统设计角度来看，信息系统可以提取并建模社交线索，以优化检索和推荐流程，提高信息呈现的准确性和系统响应的效率 [24]。

信息，尤其是社交媒体上的用户生成内容（UGC），通常包含丰富的社交线索[55, 59, 63]。近年来，**社交媒体平台**逐渐成为信息寻求的重要渠道[37]。例如，用户会在 Facebook 上寻求本地活动的信息 [20],，或在 TikTok 上搜索心理健康问题的应对策略 [36]。社交媒体不仅提供丰富的信息线索，还提供易于获取且多样化的社交线索，例如发帖者的公开个人资料、关注者数量以及评论区中的互动。用户可以利用这些线索支持信息寻求活动，并帮助筛选和判断多样化的帖子。随着人工智能技术的迅速发展，许多主流社交媒体平台（例如 X、微博和小红书）已开始整合**基于大语言模型的搜索系统**，以协助用户进行信息寻求 [18]。基于大语言模型的搜索通常遵循“检索—理解—生成”的核心流程 [13]。如图1所示，该流程首先根据用户查询从海量帖子中进行初步检索；随后，大语言模型对检索内容进行筛选、理解和整合；最后，生成自然语言响应。基于大语言模型的搜索系统能够理解自然语言查询、匹配多源信息，并自动整理和生成综合答案，从而显著提升信息寻求的效率[17]。

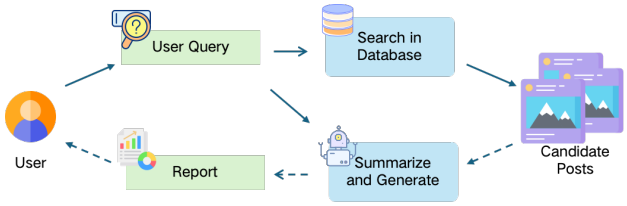


图 1. 基于大语言模型的搜索系统的核心流程。

由于社交线索的整合与利用有限，现有的基于大语言模型的搜索系统仍主要依赖语义线索，这**与人类在自然信息寻求过程中的认知社会化并不相符**。也就是说，其局限性主要体现在：系统通常只根据语义计算或基于规则的匹配（例如关键词匹配）来筛选和总结信息，[58], 而没有充分考虑用户对语义之外社交线索的需求。例如，用户搜索帖子时，可能希望系统优先呈现评论区中获得正面反馈或讨论热度较高的内容，因为这反映了社会认同或群体共识。然而，主流的基于大语言模型的搜索工具目前尚不支持这种基于社交线索的检索意图，从而限制了它们在社会化信息寻求场景中的有效性。

尽管上述不一致已受到关注，但作为一种新兴的信息检索范式，**社交线索在基于大语言模型的搜索中的整合及其机制仍缺乏系统研究**，目前尚不清楚。稿件已提交至 ACM

用户在搜索前如何理解并期望结果中呈现社交线索，在搜索过程中如何根据社交线索控制和筛选结果，以及在搜索后如何解读和利用带有社交线索的结果，这些问题仍不明确。为弥补这一研究空白，我们开展了一项探索性研究，探讨以下研究问题：**研究问题1：在基于大语言模型的信息寻求情境中，用户期望使用哪些社交线索，以及应如何使用这些社交线索？研究问题2：整合社交线索会如何影响用户的信息寻求结果和过程？**

为开展本研究，我们选择社交媒体提供的基于大语言模型的搜索工具作为研究背景。社交媒体平台不仅已成为人们寻求信息的重要渠道，也提供了丰富且易于获取的社交线索。此外，近年来，多个主流平台已推出或正在试验基于大语言模型的搜索，进一步体现了这一趋势的广泛性和现实意义。基于这些因素，社交内容平台为研究社交线索与基于大语言模型的搜索相结合提供了具有代表性的背景。我们采用探索性研究方法，核心流程包括原型部署和实证研究。针对研究问题1，我们组织研讨会，收集用户对社交线索的需求与期望，并据此设计和实现了支持社交线索整合的基于大语言模型的搜索原型系统 *SoulSeek*。针对研究问题2，我们招募用户在真实的信息寻求情境中使用该原型，并通过被试间实验记录他们的操作和交互过程。最后，我们采用结合定性与定量分析的混合方法，从结果和过程两个层面系统地得出研究发现。本研究结果揭示了社交线索在基于大语言模型的搜索中的多重作用：整合社交线索显著改善了用户的感知与体验，提升了结果的有用性和可信度，使搜索过程更具方向性且更可控，并促使用户反思自身的信息行为。同时，研究发现，模型在理解社会化知识和兼顾多线索匹配方面仍存在局限。本文进一步指出，用户主动运用社交线索有助于恢复并增强其在信息寻求中的能动性，并据此为未来的社交线索感知型大语言模型搜索系统提出设计启示，包括增强对社会知识的理解、支持个性化线索配置，以及实现透明且可控的人机协作搜索过程。总之，本文的贡献包括：

- 本研究是首批探究基于大语言模型的搜索中社交线索机制的研究之一。
- 我们通过设计研讨会识别用户需求，并开发了原型系统 *SoulSeek*，以支持社交线索集成。

- 通过实证研究，我们揭示了社交线索如何塑造用户的感知、经历和行为。
- 我们为未来的基于大语言模型的搜索提出设计启示，指导如何将社交线索融入设计。

2 相关工作

2.1 信息寻求

信息寻求是一种基本的人类认知与社会过程，个体在其中识别信息需求，并主动收集、筛选和整合信息，以支持理解与决策[26, 43, 60]。在学术研究或调查工作等复杂情境中，这一过程具有高度迭代性和非线性，涉及不确定性下的持续交互以及多源信息的整合 [42]。

关于信息寻求的理论已从注重效率的范式演变为注重体验的范式。信息觅食理论强调在信息环境中进行策略性的成本—收益权衡 [44],，而意义建构理论则着重于用户解读和组织信息时所经历的情境化循环 [11]。信息漫游者这一概念进一步将信息寻求视为一种探索式且富有反思性的实践 [12]。这些理论共同揭示了信息寻求具有动态性和体验性。在此基础上，人机交互研究考察了不同平台和情境中的信息寻求。研究表明，社会语境如何塑造用户的健康信息行为 [10], 社群线索如何在超本地搜索中促进归属感[20], 以及个人信息生态系统如何支持反复验证与反思 [36]。在这些研究中，用户并非被动的信息接收者，而是信息路径与意义的积极塑造者 [3, 14]。因此，认识到用户能动性已成为设计更个性化、更具适应性的信息系统的核心。

本研究聚焦于社交内容平台，以此拓展相关文献。这类平台包含丰富的社交线索，例如作者身份、反馈和对话语气。这些线索会影响用户的判断与意义建构，因此，对于理解信息寻求的社会化特性以及设计更具情境感知能力的智能系统而言，它们至关重要 [23, 36]。

2.2 信息寻求中的社交线索

社交线索是传递社会或人际信息，并影响人们如何解读他人及回应他人的信号 [41, 53, 62]。在面对面交流中，社交线索包括注视、表情和语气等非语言行为；在社交平台上，则表现为点赞、评论、作者身份和互动程度。Feine 等人[15] 提出了一个包含四种类型的社交线索分类体系——语言、视觉、听觉和隐性线索，涵盖从语言到时机等显性与隐性特征。

社交线索在信息寻求中发挥着关键作用。Winter 等人 [62] 指出，信息选择不仅取决于内容，也受到用户社会动机的影响，而人们往往会无意识地处理受欢迎程度等线索。从计算机媒介传播的视角来看，Tanis 和 Postmes [54] 发现，线索的可用性会影响印象形成；Huang 等人 [21] 和 Liu 等人 [32] 则揭示，用户会根据不同情境和媒介灵活调整或切换线索，以传达意图并提升可访问性。此外，Yang 等人 [65] 表明，用户在阅读社交媒体内容时会产生钦佩、焦虑等社会情绪反应，而基于传感器的测量能够准确检测这些反应。总的来说，这些研究凸显了社交线索对注意力、信任和互动模式的广泛影响。近期研究还将社交线索融入智能系统。Li 等人 [29] 证明，将社会关系与视觉特征结合能够提升检索性能。Klein 的荟萃分析 [24] 发现，类人线索能够适度提升信任和满意度，但过度拟人化可能引发恐怖谷效应 [9]。Reinkemeier 等人 [48] 进一步指出，线索的作用高度依赖情境，会随用户期望和任务设置而变化。

总而言之，先前的研究已经确立了社交线索在信息寻求中的重要性，但大多数研究关注的是社交线索的呈现与感知，而非用户如何主动运用和配置这些线索，尤其是在新兴的基于大语言模型的交互中。

2.3 基于大语言模型的搜索系统

大语言模型（LLM）和检索增强生成（RAG）的快速发展催生了将检索与自然语言生成相结合的新型搜索方式 [1, 13, 27]。用户以自然语言表达信息需求，大语言模型解析查询、检索相关知识，并生成整合后的回答。提交至 ACM 的稿件

这类系统提供灵活、交互式且迭代式的搜索体验，使用户能够在多轮对话中不断完善目标。尽管前景广阔，当前基于大语言模型的搜索系统仍存在明显局限。在知识层面，检索增强生成方法擅长事实检索，但在情境理解方面表现不佳。Wang 等人 [58] 指出，忽略群体互动和情境背景往往会导致输出偏离需求。在交互层面，检索与生成过程仍不透明，限制了用户的理解和控制；为提升透明性和可解释性，人们提出了 RAGTrace 等可视化工具 [7] 。

近期研究因此强调用户能动性和反思能力。尽管对话式搜索简化了交互，但也可能削弱用户主动探索的意愿。Schroeder 等人 [50]观察到，自动摘要可能会使用户从主动寻求者转变为被动接收者。Mei 等人 [33] 倡导混合主动式协作，即系统与用户共同制定搜索目标。同样，Neshaei 等人 [39] 认为，系统应促进反思，而非取代用户思考。

总之，先前的工作从两个方向推进了基于大语言模型的搜索设计：提升透明性并恢复用户主动性。基于这一趋势，本文探讨如何整合社交线索，以进一步支持用户在社交内容平台上自然且可控地进行信息寻求。

3 设计研讨会与实施

3.1 设计研讨会

为探究研究问题1并为解决研究问题2的系统设计提供依据，我们开展了一场两阶段设计研讨会，考察用户在社交内容平台上使用基于大语言模型的搜索时，如何理解和使用社交线索，以及对社交线索有何期待。此次研讨会旨在确定用户在信息寻求过程中依赖哪些社交线索，以及应如何将这些线索整合到基于大语言模型的搜索工具的交互与生成过程中。

我们选择小红书（小红书）[1] 作为研究背景，原因如下。小红书是中国最具影响力的社交媒体平台之一，月活跃用户超过3.5亿，平均每日使用时长超过74分钟。该平台支持多样的使用场景，用户也经常在其中进行信息寻求活动。根据其2024年年度报告，约70%的活跃用户会使用平台内的搜索功能来寻求信息，涉及旅行、消费、学习和职业规划等领域。因此，小红书已成为研究信息寻求行为的重要环境。

我们总共招募了 10 位活跃用户（年龄为 25–35 岁，性别比例为 5:5，且均至少拥有学士学位）。所有参与者使用该平台的时间均超过一年，并报告称每周使用搜索功能三次以上以获取实用信息。研讨会通过线上会议进行，每场约持续两小时。经知情同意，所有场次均进行了录音和记录，每位参与者获得了 10 美元的报酬。

我们遵循参与式设计范式 [38]，采用务实且迭代的两阶段研讨会结构。**第一阶段：** 我们请参与者回忆最近一次在小红书上通过搜索解决问题的经历（例如旅行、消费或求职），并说明影响其内容选择的主要因素。随后，我们提供了示例线索卡片（例如“作者头像”“点赞数”“评论氛围”“位置标签”）。参与者补充了遗漏的线索，并写下他们认为重要的要素。各组通过卡片归类和标注展开讨论，形成高层级类别，并对每种线索的重要性进行投票。各组对卡片进行归类和标注，通过讨论形成更高层级的类别，并对每种线索的重要性进行投票。最终形成了

[1] https://www.xiaohongshu.com/explore

多样的社交线索实例，以及为后续系统设计奠定基础的结构化分类体系。**第二阶段：**我们回顾了第一阶段的结果，并介绍了主要线索类型。参与者提供了若干基于大语言模型的搜索工作流场景草图，展示了典型流程（用户输入查询，人工智能返回结果并展示帖子层面的社交信息）。参与者思考系统应如何理解和使用这些社交线索，并使用便利贴和图示标注应在何处以及如何引入线索。随后，各组通过角色扮演搜索对话，测试所提机制是否合理。研究团队汇总了讨论和草图，并将共同创作的想法归纳为三个关键模块。

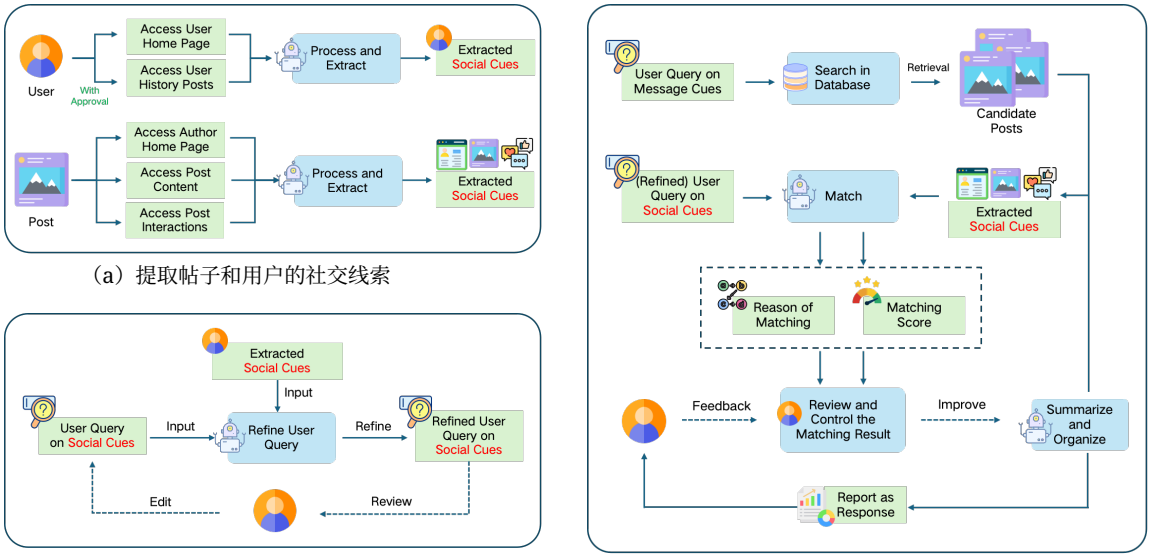
经参与者同意，我们对所有研讨会进行了录音，并使用语音识别工具将录音转写成文本。第一作者核实并修正了转写稿。我们进行了主题分析 [4]，并使用 MAXQDA 工具对编码进行系统化整理，考察不同编码之间的关系。两位核心研究人员分别阅读了所有反馈，以熟悉数据并识别初步概念。随后，我们通过迭代过程不断完善编码方案：通过开放式编码捕捉新出现的概念和模式；通过主轴编码将语义相近的初始编码合并为更高层次的主题，并剔除冗余或薄弱的条目；通过选择性编码，由研究团队共同裁定并提炼与两个核心主题相关的编码——“社交线索的类型及其在不同阶段的使用”以及“整合方法与设计预期”。这些步骤确保了逻辑一致性，并使分析与研究目标保持一致。

总体而言，用户明确表示，他们希望在社交内容平台的对话式搜索中主动利用社交线索。他们还指出了当前系统在对这些线索进行建模和响应方面的不足。如表1所示，搜索过程分为三个阶段——搜索前、搜索中和搜索后；社交线索则来自三个来源：发布者（来源方）、内容本身（自身）和受众反馈（接收方）。在搜索前阶段，用户会在查询中指定基于线索的约束条件，并结合自己在平台上的社交线索（例如关注关系、互动历史）对其加以补充。在搜索中阶段，用户会检查候选帖子的多维社交线索，验证系统的匹配结果，并提供即时反馈，帮助校准线索识别和排序。在搜索后阶段，用户会根据呈现出的社交特征完善查询，探查潜在偏好，并优化后续搜索。

表1. 按线索来源和信息寻求阶段分类的、借助社交线索进行信息寻求的使用预期。

线索来源	**搜索前**	**搜索中**	**搜索后**
发布者（来自）	根据以下内容定义查询约束条件：发布者特征；结合用户的个人社交线索	识别并匹配发布者侧社交线索（例如个人资料简介、关注者数量、标签）	基于匹配结果，分析跨不同发布者
内容本身（自身）	基于以下内容定义查询约束条件 内容特征定义查询约束；并以……进行补充 用户的个人社交线索	提取帖子级特征（例如位置、风格、时间），并进行匹配	基于匹配结果，分析跨帖子级特征关联帖子
受众反馈 (To)	根据以下内容定义查询约束条件 受众反馈特征；并结合用户个人的社交线索	分析互动特征（例如，点赞、收藏、评论）并进行匹配	基于匹配结果，分析跨帖关联，探究受众指标与评论之间的内容

在第二次研讨会中，我们明确了用户期望如何在基于大语言模型的搜索工作流中引入和使用社交线索。我们将这些期望概括为图 2 所示的三个关键模块：(1) 社交线索



（a）提取帖子和用户的社交线索

（b）根据社交线索优化用户查询

（c）基于社交线索进行匹配

图 2. 将社交线索整合到基于大语言模型的搜索系统中的核心工作流程设计。

提取。系统首先从用户的个人资料和发帖历史中提取个人社交线索。同时，系统还从检索到的帖子中提取多源社交信息，包括作者资料属性、帖子属性和评论区互动信号，从而形成社交线索的多层级表示。(2) 社交线索查询优化。在用户提供初始线索条件后，系统利用用户的社交历史（例如此前的帖子、互动偏好）来优化并完善查询，使线索规范更加准确且个性化。(3) 社交线索感知的匹配与生成。系统将用户的内容查询与线索条件相结合，从平台数据库中检索候选帖子。随后，大语言模型提取并匹配帖子中的社交线索，生成带有理由的匹配分数，并支持用户进行查看和控制。最后，系统根据用户确认的结果，为信息寻求生成综合答案和报告。

3.2 最终设计与实现

为验证在真实场景中整合社交线索的有效性，我们根据研讨会的发现，设计并实现了一个交互式原型 *SoulSeek*，并将其部署在社交内容平台环境中。交互流程包含三个阶段：

在搜索前阶段，系统在标准搜索框旁添加了一个社交线索输入面板（图 3-1-a），让用户在输入内容查询时指定线索偏好和约束条件（图 3-1-b）。例如，搜索“新加坡自由行指南”时，用户可以补充“偏好真实用户体验；避免商业推广”。系统会解析这些线索条件，并构建一个多维社交线索模式，涵盖发布者、内容本身以及受众反馈。在获得用户授权后，系统还会从历史行为中提取个人社交线索（图 3-1-c）。随后，用户可以通过与大语言模型的对话界面进一步完善条件（图 3-1-d）。点击界面右上角的搜索图标，即可进入搜索中阶段。



图 3. *SoulSeek* 中搜索前用户界面。



图 4. *SoulSeek* 中搜索中用户界面。



图 5. *SoulSeek* 搜索后界面。

在搜索中阶段，系统使用大语言模型从候选帖子中提取并匹配社交线索，并生成带有解释的匹配结果。左侧栏列出检索到的帖子；点击帖子封面即可打开内容平台查看详情（图 4-2-a）。中间栏展示从每个帖子中提取的社交线索，涵盖发布者、内容本身和受众反馈（图 4-2-b）。右侧栏根据用户的线索条件展示模型的匹配结果及理由（图 4-2-c）。我们采用 1–5 分的评分方案：5 分表示与用户条件高度相关，1 分表示毫不相关。在我们的实现中，得分高于 3 分的项目被视为符合用户的社交线索约束。系统还提供“同意匹配”切换按钮，方便用户快速反馈和控制。如果模型标记的匹配结果不符合预期，用户可以标记“不同意”。系统会记录这些反馈，并在随后的摘要阶段将其反馈给大语言模型，以改进匹配逻辑和解释，从而增强用户对搜索过程的控制。点击界面右下角的“生成报告”按钮，即可进入搜索后阶段。

在搜索后阶段，系统汇总结果并生成摘要报告（图 5-3-a）。报告突出显示与用户需求高度匹配的内容，并清晰标注其来源，帮助用户了解结果的来源和背景。用户可以继续与模型互动，探索结果中社交线索之间的关联（图 5-3-b）。例如，用户可以要求大语言模型综合评论区中不同的观点并提炼共识，从而对信息形成更高层次的理解。

该系统使用通义千问-Max 作为大语言模型。它属于通义千问系列，这是领先的开源大语言模型产品线，兼具高性能与本地部署能力，能够在数据隐私环境中稳定运行。通义千问-Max 继承了在通用理解和生成方面的强大能力。在 LMArena 文本排行榜[2]上，其排名第三，超过了 GPT-5-Chat。鉴于社交线索与个人身份之间存在紧密联系，我们采用通义千问系列，以确保其适用于优先考虑数据安全和用户隐私的受控环境，

[2]https://lmarena.ai/leaderboard/text

与隐私保护和边缘部署人工智能系统这一新兴趋势保持一致 [57]。该原型基于 Coze 框架构建，前端支持搜索交互和社交线索可视化。后端使用大语言模型进行社交线索提取、查询优化和结果生成。系统在访问个人信息时严格遵循用户授权政策。如果用户未提供个人资料链接，系统不会自动提取社交线索来优化查询。用户仍可手动指定线索条件。在搜索过程中，系统仅访问平台上公开可见的内容，这与正常使用相一致。所有实验数据在存储时都会进行匿名化处理，以保护参与者隐私和数据安全。

4 实证研究与发现

4.1 研究方案

4.1.1 被试间设计。 基于前期研讨会和原型系统，我们开展了一项采用被试间设计的实证研究，以考察社交线索在信息寻求中的作用。本研究的核心目标是回答以下研究问题：社交线索的整合如何影响用户的信息寻求结果和过程？为此，我们从以下两个方面探讨这一问题：(1) 结果层面效应——用户对搜索结果的感知发生变化，例如感知有用性和可信度——对应于我们在 F1（§4.2）中的发现；以及 (2) 过程层面效应——用户的信息搜寻行为和体验发生变化，包括其方向感、意外发现、用户控制和使用意愿——对应于 F2（§4.3）和 F3（§4.4）中的发现。

我们招募了20名活跃的小红书用户（男性10名、女性10名），所有参与者均与此前研讨会的参与者相互独立，并且至少拥有学士学位。每位参与者使用该平台的时间均超过一年。此外，所有参与者都熟悉并曾使用过小红书搜索和基于大语言模型的人工智能搜索工具，表明他们习惯于在该平台上开展信息寻求活动。详细的人口统计信息见表2。本实验采用被试间设计，将参与者随机分配至实验组（E）或对照组（C）。两组均使用我们开发的系统完成相同的任务。实验组使用的系统支持社交线索的输入与呈现，而对照组使用的系统不具备这一功能。除社交线索外，两个版本的系统在界面和功能上完全相同，从而确保结果具有可比性。

考虑到参与者的任务负荷，每次实验时长限制为一小时。每位参与者根据自身的信息需求提出两个搜索查询，随后使用系统进行搜索、浏览检索内容并阅读生成的报告。为确保任务具有代表性和开放性，研究人员采用了启发式任务生成方法。研究人员提供了九个预定义的主题类别，包括：实用指南、职业发展、产品／服务比较、健康与心理学、学术与知识学习、生活经历分享、消费与口碑、旅行与地点探索，以及社会问题与价值判断。参与者可以根据个人兴趣选择主题，并自由生成具体查询。为保持任务的复杂性和探索深度，过于简单或事实性的提问（例如“今天天气怎么样？”或“谁获得了2025年诺贝尔物理学奖？”）将被排除。实验结束后，研究人员采用结合问卷和访谈的混合方法进行评估。参与者使用五点李克特量表评价其在信息寻求过程中的感知和交互体验。

表2. 对照组和实验组参与者的人口统计信息。

ID	Age	使用小红书（年）	小红书的搜索频率（次）	基于大语言模型的搜索（次数）	性别	教育	组
C1	27	3	>3/周	>3/周	女性	学士	对照组
C2	26	4	>3/周	>3/天	Male	硕士	对照组
C3	25	2	>3/天	>3/周	女性	学士	对照组
C4	27	5	>3/天	>3/天	Male	学士	对照组
C5	23	3	>3/周	>3/月	女性	学士	对照组
C6	30	6	>3/天	>3/周	女性	硕士	对照组
C7	28	1	>3/月	>3/天	Male	硕士	对照组
C8	23	4	>3/天	>3/月	Male	学士	对照组
C9	25	4	>3/天	>3/月	Male	学士	对照组
C10	31	2	>3/天	>3/月	女性	硕士	对照组
E1	24	6	>3/周	>3/月	女性	学士	实验性
E2	27	3	>3/周	>3/天	Male	硕士	实验性
E3	23	2	>3/天	>3/周	女性	学士	实验性
E4	29	5	>3/天	>3/周	Male	硕士	实验性
E5	33	4	>3/天	>3/月	女性	硕士	实验性
E6	24	7	>3/天	>3/周	女性	学士	实验性
E7	24	2	>3/周	>3/天	Male	学士	实验性
E8	27	3	>3/天	>3/天	Male	学士	实验性
E9	26	2	>3/天	>3/月	女性	学士	实验性
E10	26	3	>3/天	>3/周	Male	硕士	实验性

4.1.2 数据收集与分析。我们采用了混合方法，整合定性与定量分析。我们的目标不仅是考察社交线索是否带来了更好或更差的结果，还希望了解用户如何看待这些线索，以及这些变化为何发生。考虑到不同参与者角色在动机和体验方面存在显著差异，仅靠定量测量不足以捕捉用户行为背后的细微机制。受先前研究的启发 [46, 47, 65], 因此，我们在数据收集与分析中结合了定性见解与定量评估。

对于定量分析，我们重点考察了参与者对信息寻求结果感知质量的评估及其体验。具体而言，在感知层面，我们考察了感知有用性、可信度 [22, 45, 64]。在交互体验层面，我们重点关注意外发现、使用意愿、方向感和用户控制 [12, 45, 60]。详细的定量维度汇总于表3：

我们计算了各项指标的均值和方差，将其数据分布可视化，并进行了显著性检验，以评估实验组与对照组之间的差异。

对于定性数据，我们结合出声思考法和半结构化访谈，收集了用户的认知与行为反应。参与者在评分时被鼓励说出自己的想法，使我们能够记录他们的推理过程。此外，我们还观察并记录了用户在使用系统进行信息寻求时的互动行为。我们采用归纳式主题分析方法分析收集到的数据 [4], ，旨在识别社交线索的融入如何影响参与者的信息寻求过程。例如，当一位参与者提到“模型无法区分常见家用食材，真令人沮丧”时，我们将此类表述编码为“难以准确解读社交线索条件”和“对社交线索的误解引发负面情绪”。这一开放式编码过程确保分析始终

表 3. 用户感知与体验的评估维度。

维度	描述
感知有用性	系统提供的信息能有效帮助我完成当前任务，或回答我的问题。
感知可信度	我相信该系统呈现的搜索结果真实可靠。
使用意愿	综合考虑交互过程和获取的信息，我愿意继续使用该系统进行信息寻求。
用户控制	我觉得自己能够主动控制搜索和筛选的方向，不受限制，也不会被系统主导。
方向感	在整个信息寻求过程中，我始终清楚自己在寻找什么，也不会感到迷失或困惑（例如，忘记最初的目标，或不确定接下来该做什么在搜索中）。
意外发现	使用该系统时，我能够发现意想不到但有价值的新信息。

以参与者自己的语言和反思为依据，同时涵盖客观观察和主观体验。

初始编码后，研究人员共同将相关编码归纳为更广泛、更抽象的主题。例如，“通过评论区中的共识发现新信息”和“内容之间的空间关联激发进一步探索”被合并为“信息意外发现”这一主题。同样，“模型对复杂社会关系的理解有限”“社会常识知识不足”和“对区域文化的理解有限”等编码，则被归入“模型的社会知识表征有限”这一主题。这个迭代过程经过多轮交叉比较，以确保每个主题都能准确反映参与者的经历，以及他们与系统互动时所体现的细微差异。

最后，通过整合定性分析、定量分析和交叉验证，我们确定了四个主要主题，揭示了将社交线索融入基于大语言模型的搜索所蕴含的潜在价值与局限性：(1) 提升用户对信息需求满足情况的感知；(2) 改善用户体验；(3) 增强用户反思；(4) 在复杂社会线索查询中面临用户对齐方面的挑战。

表4. E组与C组各项指标的比较。Δ 表示均值差异（E–C）。

类别	维度	C_{Mean}	C_{SD}	E_{Mean}	E_{SD}	Δ	p值
感知	可信度	3.50	0.83	4.55	0.60	+1.05	0.0002
	有用性	2.75	0.85	3.85	1.09	+1.10	0.0011
体验	用户控制	2.60	0.88	4.20	0.52	+1.60	<0.0001
	方向感	3.05	0.94	4.50	0.61	+1.45	<0.0001
	意外发现	2.20	0.83	4.00	0.86	+1.80	<0.0001
	使用意愿	2.55	0.83	4.20	0.62	+1.65	<0.0001

4.2 F1：提升用户对信息需求满足的感知

表 4 展示了用户评估的详细对比，图 6 展示了评分分布。实验组报告的感知有用性（+1.10）和可信度（+1.05）均高于对照组。与有用性评分相比，可信度评分的分布更为集中，这表明参与者对结果可靠性的看法更加一致，而不同个体对有用性的感知差异更大。我们的定性分析进一步解释了这些改善背后的原因。

4.2.1 提升感知有用性。在搜索过程中，用户综合运用各种社交线索来判断和完善搜索标准，从而让大语言模型更有效地筛选内容，并生成更符合自身需求的结果。参与者通常会考虑多个维度，包括内容风格、评论和作者身份。E7 描述了如何利用风格线索推断内容的可靠性：“我会根据内容风格来判断。例如，搜索权威信息时，我可以排除那些使用‘宝子们’（亲爱的朋友们）或‘家人们’（家人）等过于亲昵称呼的帖子。”。E3 进一步强调，评论作为集体意见具有诊断价值，而不仅仅是互动指标：“评论比点赞数显得更真实，因为人们写评论需要投入更多精力。系统可以有效利用这些观点来提升回复质量。” E8 提到，了解发帖者的社会身份或背景，有助于他们更好地判断帖子的可信度和适用性：“对于开放式问题，我真正想了解的是相关人员的身份……我可以筛选出与自己情况相似的人分享的经历。” 相比之下，对照组的参与者也表示希望基于社交线索进行筛选，但他们使用的系统无法充分支持这一功能。C8 表示，当检索结果中出现模型未能过滤掉的标题党或营销类帖子时，他们感到很沮丧：“……像‘突发新闻！薪资暴涨！’这样的标题，很可能是夸大宣传的帖子。我真正想要的是来自个人的真实经历。” C5 还强调了搜索实用指南时评估评论反馈的重要性，并表示他们“希望隐藏那些有很多负面评论的帖子，但系统不支持……”

在结果筛选阶段，用户普遍认为，整合社交线索有助于他们更高效地排除不相关或未预期的内容。他们指出，大语言模型通常能够正确解读他们的线索条件，并据此筛选结果。例如，E3 提到，看到系统明确解释线索与结果的对应关系后，他们对模型的理解能力更有信心：“看到解释后，我觉得它理解得很正确。”

此外，当大语言模型根据筛选后的帖子生成最终结果报告时，几位参与者指出，社交线索不仅影响了他们对结果的理解，也影响了系统组织信息的方式。他们发现，在表达社交线索后，系统能够更好地按照他们的个人兴趣组织内容并确定优先级，从而生成更易于理解且更切合需求的报告。E1 描述说，纳入线索后，系统的组织方式明显更加清晰：“表达社交线索后，系统在组织报告时更清楚地体现了我的要求。”

4.2.2 提升感知可信度。社交线索的整合显著提升了用户对搜索结果的感知可信度。参与者表示，当系统根据社交线索条件匹配结果时，会明确说明每项匹配背后的依据。这种解释性反馈帮助用户理解系统的判断过程，并评估其推理是否符合自己的预期，从而增强了他们对结果的信任。几位参与者提到，这种透明性清楚地说明了为什么某些帖子会被纳入或排除。E1 观察到，在查看模型的解释后，它准确地反映了

他们的意图和筛选逻辑："审阅模型的解释后，我发现它准确地转述了我的要求。"其他人则赞赏系统提供了分级解释，而非简单的分数，同时给出了数值和文本理由。E8 指出，这种做法增强了他们对模型判断的信心："……解释合情合理时，它帮助我对结果建立了信任。"

除了通过可解释的线索匹配建立信任外，参与者还提到，社会参照——看到反映自身社会条件或身份的结果——也增强了他们对信息可信度和适用性的信心。他们认为，检索到的帖子与自身情况更加契合，从而为判断信息可靠性提供了明确的社会锚点。E1 指出，来自社会背景相似用户的真实个人经历尤其值得信任："真实生活中的个人经历体现得更加清楚。一群位置、目标或身份与我相似的人，能够提供真正有帮助的经历和建议。" E3 强调，相比机构来源，同伴分享的共同经历更有分量："作为消费者，我往往更信任其他消费者，而不是企业。如果我看到其他人也在寻找类似信息或分享他们的经历，我就更有可能信任这些内容。" E4 补充说，意识到他人总结过相关内容或亲身使用过，会提升结果的可信度："感觉他们确实用过它，或者提供了一些总结，所以从这些总结中进行选择似乎更可靠。" 类似地，E7 提到，来自背景相似用户的结果增强了可信感："如果某个身份与我相似的人——也是计算机科学专业的学生——正在搜索类似的招聘帖子，那么来自同伴的信息就更值得信任。"

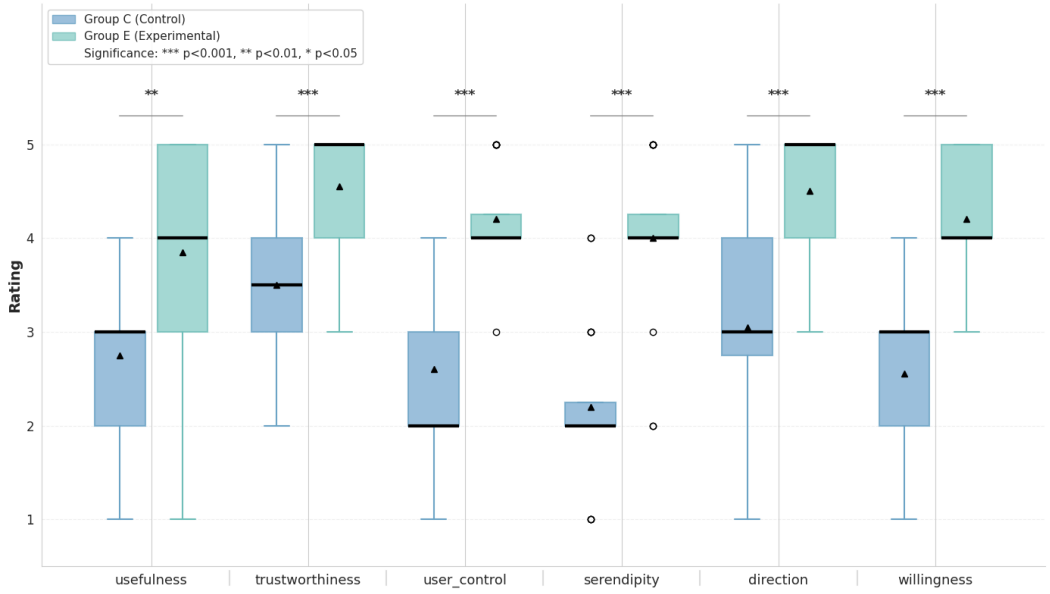


图6. E组与C组在各评估维度上的用户评分分布。

4.3 F2：用户体验提升

社交线索的融入不仅提升了用户对自身信息需求得到更好满足的感知，也改善了他们在信息寻求过程中的主观体验。如表 4 和图 6所示，论文已提交至 ACM

实验组在所有体验维度上的评分均持续较高，包括用户控制（+1.60）、方向感（+1.45）、意外发现（+1.80）和使用意愿（+1.65）。根据箱线图分布，在这四个维度中，对照组的下四分位数（25%）均低于 3 分，表明整体体验偏负面；而实验组的所有下四分位数均高于 3 分。我们的定性分析进一步阐释了这些提升背后的具体模式。

4.3.1 增强用户控制。 在信息寻求过程中，实验组参与者展现出更强的控制能力，他们能够根据自身特点和偏好灵活调整社交线索，从而获得更符合自身需求的结果。与传统平台被动的推荐机制不同，该系统允许用户通过互动直接介入模型的决策过程，使他们能够在搜索早期参与内容筛选和生成。正如 E1 所述，能够查看并回应模型的中间推理，让他们得以塑造结果，而不是被动等待：“我可以表示同意或不同意，模型会据此进行调整。”

这种预先干预使用户从被动的审阅者转变为搜索过程中的主动参与者。他们不再等待系统输出，而是通过个性化反馈引导模型的匹配逻辑，从而获得更强的控制感。E3 表示，这种体验与自己平时在小红书使用原有的基于大语言模型的搜索时截然不同：“我通常会先看整体概述，然后浏览单篇帖子。而在这里，情况正好相反——我先阅读单篇帖子，再总结自己认同的内容。我觉得控制感更强了，因为我可以在整个过程中查看并调整，而最终报告也反映了我自己的选择。”相比之下，C10 指出了传统基于大语言模型的搜索的局限，特别是缺乏精细化控制：“我自己搜索帖子时，可以刷新或重新选择标签来调整看到的内容。但使用基于模型的搜索时，它经常会检索出一大批帖子——其中很多都不是我想要的——即使重新开始，那些不想要的帖子也可能再次出现。”

通过与模型持续交互，一些用户经历了从将其视为机械式摘要工具，到认识到其是能够进行联合推理和共同创作的智能协作者的转变。E2 描述了这一变化：“起初，我只把它看作摘要工具，但后来我意识到，我可以掌控它。现在，它更像是一位能够理解我的意图并据此整理信息的协作者。”这种认知转变体现了用户在信息寻求中的控制权重构：用户不再只是搜索系统的操作员，而是与模型共同塑造结果的共同创作者，从而获得更具掌控力和参与性的搜索体验。

4.3.2 增强意外发现。 尽管我们最初引入社交线索的目的是帮助用户筛选和解读信息，但有趣的是，我们发现，社交线索也为用户探索新信息提供了机会——在某些情况下，甚至成为信息寻求本身的目标。通过展示作者的地理位置等线索，系统帮助用户注意到此前被忽视的关联，并作出意外却颇有价值的发现。例如，E7 搜索了“如何制作馄饨汤”，并认为所显示的位置促使其从地域角度重新解读食谱：“如果没有列出这些位置，我不会从那个角度去思考。南方馄饨汤通常会使用更多猪油，而北方则更注重芝麻油。”

通过利用社交线索，系统帮助用户发现了他们此前未曾注意到的细节或关系，从而带来了意想不到却很有价值的发现。例如，一位参与者搜索“如何制作馄饨汤”时，注意到系统显示了表明作者地理位置的社交线索。这促使他从地理角度重新解读食谱，并有趣地发现这道菜在南北方的差异。*E7*：“如果系统没有列出这些位置，我就不会从那个角度去思考

了。这些信息很有价值——南方的馄饨汤往往会用更多猪油，而北方则更常用芝麻油。”

参与者指出，社交线索帮助他们超越了传统的语义检索范式，使他们能够通过社会语境重新组织信息空间。E8 解释说，这一转变为探索提供了更大的灵活性：“以前，无论是基于嵌入还是基于关键词，搜索始终围绕内容本身展开。有了社交线索，我可以放宽对内容的约束，并利用这些线索为探索界定一个更有意义的边界。”

参与者观察到，社交线索帮助他们发现了超越单篇帖子的共识模式，揭示出更广泛的社群动态。评论和反馈的结构不再被视为补充信息，而被看作集体意见和社会情绪的重要指标。E10 指出，“评论中反映的共识实际上可以提供极具价值的信息。”从这个意义上说，社交线索带来的新颖性不仅在于新信息出乎意料，还在于对社群动态的重新诠释。通过观察他人的回应、共鸣或分歧，用户形成了更高层次的理解——从获取单篇信息转向对社群认知和社会结构的更加整体性的感知。

4.3.3 增强方向感。参与者将社交线索描述为导航辅助工具，能在信息探索过程中提供稳定的逻辑脉络。与单纯依赖关键词或语义匹配相比，社交线索帮助他们在复杂的信息空间中建立更清晰的参照点。E1 将社交线索称为“把一切联系起来的逻辑脉络”，而 E10 则描述为“沿着两条平行线搜索——一条关注内容，另一条关注社交线索”。正如 E2 所指出的，社交线索还能让他们在整个搜索过程中更明确、更清晰地追踪自己的意图，“不断提醒我在找什么”。这种双轨过程使用户能够同时围绕内容和社交维度进行搜索与推理，并在结果生成前形成更清晰的方向感。

对于目标导向搜索任务，这种方向感尤为明显。社交线索发挥着注意力支架的作用，帮助用户高效过滤无关或分散注意力的内容（例如广告），并保持认知专注。E3 解释说，这些线索在学术探索过程中强化了预期情境：“有广告时，系统会把它们标记为广告，所以我自然会忽略它们。而且，当系统显示相关人员的信息——比如他们所属的大学或实验室——这些信息能帮助我始终处于学术搜索情境中。”正如 E9 所指出的，这类搜索提醒也取决于任务：“如果我只是在随便浏览，就不需要……但当我有明确目标时，这种匹配过程对于保持方向感就至关重要。”

评估生成结果时，用户会依据社交线索来判断报告的可信度和丰富程度。社交属性的差异——例如关注者数量、评论层级和用户类型——帮助他们快速判断聚合内容的真实性和有用性。E1发现，这些线索揭示了真实参与的模式：“我想阅读真实的个人经历，而且我注意到，很多真正的用户关注者其实很少。”E6很赞赏这类信息集中展示的方式，因为这样就不必逐一查看每位用户的个人资料：“在原来的系统中，我必须点进每位用户的个人资料，才能查看他们的领域或关注者数量。这里，所有这些详细信息都集中显示在一处。”相比之下，C9提到：“查看作者的个人资料时，我会被过往帖子中的图片分散注意力。”这表明用户需要帮助，才能专注于搜索。

4.3.4 提升用户的使用意愿。 在基于大语言模型的信息检索系统中，使用社交线索的用户通常表现出更高的系统使用意愿。尽管引入社交线索会在表达和使用过程中带来一定的认知负荷，但参与者认为这是可以接受的，甚至值得付出。E1 指出，“作为一种社会性存在，我觉得这完全是本能的，”并进一步解释说：“在社交内容平台上搜索时，

反正我自然会处理这类线索，而系统对这些线索之间关系的可视化让这种付出变得值得。”。同样，E2 表示：“总体而言，这感觉毫不费力，几乎是出于本能。事实上，在平台上，我经常会自然而然地使用社交线索。”

此外，参与者喜欢使用自然语言来表达社交线索，重视其灵活性以及模型处理模糊或定义宽泛描述的能力。E4 解释道：“由于我可以直接用自然语言进行搜索，约束条件就少了。以前我必须想出精准的关键词，还要担心拼写错误；而现在，系统会提供解释，我可以根据需要对其进行完善或更正。”

一些参与者指出，即使在提交查询之前，社交线索也会形成一种“心理匹配模式”，帮助他们迅速判断系统是否理解了自己的意图，从而减轻验证负担。E2 解释道：“社交线索的匹配通常比内容匹配更明显——你可以立刻看出系统是否理解了我的意思。”这种预期效应使用户能够更快地关注关键帖子，减少反复验证所耗费的时间和精力。

4.4 F3：增强用户反思

通过观察用户互动并分析访谈，我们发现，社交线索也促使用户反思自身的信息寻求行为。用户开始重新评估信息来源的有效性，并在表达和利用社交线索的过程中，持续完善自己的信息需求和评估标准。

4.4.1 信息来源的有效性。在借助社交线索进行信息寻求时，一些参与者开始反思该平台本身是否适合满足自己的信息需求。例如，E5 搜索公务员考试策略，并设定了严格的社交线索——将目标限定为认证公务员或资深同行发布的帖子，且内容须正式、详尽、不含广告。当许多检索到的帖子被模型标记为不符合要求时，该参与者意识到平台的生态与任务目标并不匹配。同样，E4 用一个比喻描述了这种体验：“这就像钓鱼——社交线索帮助我感知水的深浅和流向，让我知道想找的鱼是否在这里。如果不在，我就知道该换个地方了。”

4.4.2 需求表达。参与者指出，思考在查询中纳入哪些社交线索，促使他们重新审视并完善问题定义。E2 解释说，明确对内容、作者和反馈的社会约束，有助于更清楚地理解信息预期：“思考如何约束内容和作者，让我反思自己对信息的实际预期。”

几位参与者谈到，社交线索促使他们重新考量查询中不同条件的优先级。有些人意识到需要区分必要条件和可选条件。E9 指出，严格设置线索会减少匹配的帖子数量，这让他发现，即使内容只部分匹配，也仍然有用：“应用社交线索后，我发现严格匹配的帖子变少了，但有些不匹配的帖子仍然很有价值。这让我思考，是否应该放宽某些条件，或者告诉模型哪些线索不是硬性要求。”这一过程也鼓励用户反思自己更深层的信息目标。正如 E8 所说：“一开始，我要求寻找从业界转入学术界的一手经历，但后来发现，二手叙述同样有帮助。这让我意识到，我真正想要的是可以与自身情况进行比较的信息。”

4.5 F4：在复杂社交线索查询中协调用户所面临的挑战

4.5.1 社会知识不足。大语言模型对社会知识的理解仍然有限，尤其是在识别社群动态和社会关系方面。E2 指出，这一缺口可能导致对细微的群体情境产生误解：“大语言模型缺乏对某些社群动态的理解。例如，在一个偶像团体中，存在不同的粉丝子社群——有些粉丝喜欢成员 X 和 Y，却不喜欢 Z。如果模型忽略这一点，就可能推荐用户不想要的内容。”

当社交线索涉及常识或情境知识时，模型的理解能力会受到更大限制。它往往难以将用户的自然语言条件转化为可执行的筛选标准。E6 展示了这一局限：“我搜索只用常见家用调味料制作的馄饨汤食谱时，模型似乎不确定‘常见’是什么意思，还加入了一些大多数人平时并不会备有的食材。”

模型对地理和文化情境的理解不准确，进一步限制了社交线索匹配的有效性。E10 指出，模型经常误解情境中的细微差别：“模型把一篇帖子标记为旅游指南，但它实际上是一次从昆明到玉溪的本地周末短途旅行——这两个地方离得很近。它仍然无法区分这类细微的地理差异。”同样，E1 观察到，尽管模型能够正确识别某些社交线索，却难以理解这些线索应如何体现在内容中：“当我要求一篇帖子包含‘足够详细的信息’时，这个标准只存在于我的脑海中。模型需要从我的反馈和行为中学习，才能更好地符合我的预期。”

4.5.2 社交线索过度处理。一些参与者观察到，模型有时会过分强调某些社交线索，导致匹配结果带有偏差或过于严格。即使某一线索维度得到了正确解读，结合来源、内容和评论特征的多条件查询也常常会触发过度筛选，降低召回率。E8指出，这种严格筛选并无必要：“召回率无故降低了。我们通常优先考虑召回率，而不是精确率，因为我们的目标是检索相关条目。帖子本身没有问题，但一条评论误导了模型，导致帖子被筛掉了。”

此外，在生成结果时，模型倾向于放大占主导地位的社交线索，例如来自专业创作者的线索，同时忽视普通用户发布的内容。这种偏差有可能降低观点多样性，随着时间推移导致摘要趋于同质化。E7表达了对这种失衡的担忧：“当前的结果掩盖了社交线索的多样性。虽然我之后可以指示模型纳入特定线索，但我更希望能在生成报告时主动设定这一点——例如，确保观点多样，而不是让所有人都重复相同的看法。”

5 讨论

5.1 将社交线索作为透镜：理解用户对自主性的需求

在当前基于大语言模型的搜索中，系统能够自动总结大量信息并生成回答，似乎可以减轻用户的认知负荷。然而，这种高度自动化的功能往往会削弱用户的能动性[7, 50], 在信息寻求中，能动性可以理解为用户主动发起行动，并使这些行动与自身价值观和身份认同相一致的程度[3]。在自动整合与生成过程中，信息来源和推理逻辑往往被掩盖，造成黑箱效应，使用户难以理解、介入或控制这一过程。这种不透明性扰乱了用户原有的信息处理认知路径；在解读模型生成的摘要时，还会带来新的认知负荷，并使用户从主动探索者转变为被动接受者。然而，人类信息寻求本质上是迭代且渐进的——一篇提交至 ACM 的稿件

假设形成、探索和评估的渐进循环——其中每一步都必须处于寻求者的认知承受能力范围内 [60]。

研究问题1（用户期望如何运用社交线索）的研究发现表明，用户希望通过配置作者资料、内容风格和评论语境等社交线索来参与信息寻求过程，从而介入查询条件、调整匹配结果，并探索替代结果或报告。研究问题2（社交线索如何发挥作用）的研究发现表明，用户借助社交线索，重新获得了在筛选和理解检索到的信息时失去的能动性。除了这种能动性的恢复，我们还观察到用户能动性在三个方面出现了新的形式。

用户能动性在澄清信息需求中的作用。当用户通过社交线索来表达查询时，例如考虑想看谁发布的帖子、哪种类型的帖子合适，以及读者如何回应，他们便能更清晰地表述消息层面的查询，并揭示其信息需求的本质。换言之，通过社交线索表达搜索意图，为反思性澄清提供了机会，使用户能够更好地理解自己真正寻求的内容。

用户在维持信息寻求过程中的能动性。社交线索的整合为用户在整个搜索过程中提供了更清晰的方向感。从定量结果来看，与对照组相比，社会线索条件下的用户平均方向感评分高出 1.45 分。从定性结果来看，参与者解释说，与社交线索相匹配起到了持续提醒或锚定的作用，帮助他们在信息寻求过程中保持注意力、专注于任务，并维持认知连续性。

用户在更高维度上探索信息的能动性。在大语言模型文本处理能力的支持下，用户能够跨越多个层次的社交信息开展高阶意义建构。借助社交线索，他们构建多维信息图式，例如识别特定作者社群中的共识、综合多个评论串中的见解，以及比较不同社会语境下的内容。这些做法表明，社交线索如何将用户能动性从个体解读拓展到集体层面的探索与综合。

这些新发现的能动性形式可能源于这样一种变化：在与基于大语言模型的系统互动时，用户对社交线索的隐性认知加工 [62] 变得愈发显性且具有反思性，使他们能够更加清楚地认识到线索本身及其相互关系。

5.2 基于大语言模型的信息寻求中社交线索的优势与不足

将社交线索融入基于大语言模型的搜索，能够在整个信息寻求过程中提升用户感知到的需求满足、体验质量和反思能力。在提升用户感知方面，社交线索能够提高结果的可解释性和可信度，使用户理解结果生成的逻辑，从而增强他们对模型输出的信任和接受度。这与“社交线索促进社会认知和信任”的观点一致 [15]，也体现了社会化检索增强生成的效果 [58]，，即系统可以利用过往交互或群体反馈，提高推荐的相关性和透明性。在提升用户体验方面，社交线索能够增强用户的主动控制感和探索感，使用户从被动的结果接收者转变为主动的过程参与者[28, 30]。这体现了人机交互领域的混合主动机制 [33]，，即系统与用户通过分享社交线索共同塑造检索目标。同样，[20] 和 [10] 表明，在线环境中的社交线索能够激发用户的社会动机和归属感，增强他们持续参与社群的意愿。我们的实验也证明，社交线索能够触发这种社会驱动的探索，形成

基于大语言模型的搜索中的信息漫游者风格 [12]。在促进用户反思方面，社交线索促使用户核实信息来源和表达需求，从而提高搜索质量，并为培养信息素养创造机会。这与以下观点一致： [39], 系统应促使用户反思，而不是取代他们的思考过程。用户可以借助社交线索识别信息背后的社会结构和价值取向，这也呼应了个人信息生态系统中提出的循环式验证—整合—反思动态 [36]。

带有社交线索的基于大语言模型的搜索仍面临社会知识储备和多维度线索平衡方面的关键弱点。首先，大语言模型在理解复杂关系、文化情境和社会常识方面的能力有限，导致社交线索匹配出现偏差。这一问题与大语言模型对群体社会语境理解不足的情况相一致 [58], ，表明模型的知识建模仍停留在浅层语义匹配阶段，缺乏对互动动态的捕捉。其次，在处理多个条件时，模型往往过度强调某一维度，导致结果同质化，削弱了社交线索本应带来的多样性。具体而言，在生成过程中，大语言模型倾向于放大最显著的信号，而忽视不同线索之间的平衡。由于这些社交线索涵盖身份、群体反馈和互动行为等多种类型，模型可能会受到统计概率最高的特征引导（例如，搜索结果中某一职业占多数），从而引入对特定线索的偏差。信息检索和推荐系统研究中也观察到了类似的偏差[31, 52]。在信息检索中，模型往往偏好高频词或处于显著位置的特征 [51], ，而推荐算法则容易产生流行度偏差和曝光放大 [6]。因此，未来的模型设计应在社交线索整合阶段引入明确的权重分配或多目标优化策略，以缓解显著线索放大和多样性损失等问题。

5.3 面向未来信息寻求系统的设计启示

结合上述发现与既有研究，未来基于大语言模型的信息寻求系统应在模型层面和系统层面进行双重改进。

在模型层面，应致力于加强社会知识的整合与多样化语境推理。实证研究揭示了大语言模型在理解社会知识方面的不足 [58], ，例如在社会规范任务中表现不佳（如在 SocKET 等基准测试中仅达到中等准确率[8]）。近期关于基于大语言模型的语音助手的研究也强调了这一差距 [46], ，指出应加强基础模型的训练，使其能够更好地捕捉和推理社会语境。未来的模型应在训练过程中引入社会知识图谱和跨模态社会数据，以增强对隐含社会语义的理解。例如，通过图注入技术[40], ，模型可以从社交网络数据中学习隐含关系，进而更好地理解对话中的细微线索，如作者身份或整体互动氛围。此外，跨模态训练可以通过自监督学习整合文本、图像和音频数据（例如来自社交媒体帖子的内容），以实现模态对齐，最终更可靠地识别跨越多种输入形式的社交线索 [16, 19]。此外，“社交线索分类体系”这一概念可作为对社会语境建模的指导框架，涵盖个体、群体和文化三个特征表征。个体表征编码语言风格和人格线索；群体表征识别社会角色和互动模式；文化表征则反映社群规范和语用惯例。在推理过程中，模型应动态平衡

这些表示方式可以避免对某些显著信号过拟合，从而增强生成内容的上下文敏感性和社会连贯性。

在系统层面，未来的搜索界面应支持可协商的线索配置和透明的反馈机制。受混合主动式设计原则的启发， [33, 34, 39], 系统可以允许用户个性化设置并反复优化其社交线索偏好，随着时间推移形成自适应线索模板。此外，基于可解释检索生成的理念， [7], 系统应通过线索卡片或线索轨迹图呈现模型推理过程，展示引用的社交线索及其相关权重，让用户能够点击并干预匹配逻辑，从而促进用户理解，并及时介入，进一步提升协同搜索体验。此外，系统应提供查全率—查准率平衡控件，例如集成式滑块控件，使用户能够根据任务目标（探索式、验证式或比较式）切换模式，从而动态优化搜索模式。这种设计符合信息觅食理论中的成本效益平衡[43] ，也与近期强调用户能动性和探索式参与的研究相呼应 [3]。

6 局限性

尽管本研究揭示了社交线索在基于大语言模型的搜索中的实际影响，但研究范围和实验条件仍存在若干局限。（1）本研究在小红书上开展，其独特的文化和用户群体可能会影响研究结果。虽然该平台提供了丰富且真实的社交线索，但研究发现未必能推广到其他平台（如 Reddit、知乎、Twitter）。此外，参与者均具备数字素养且是活跃用户，他们可能比普通用户更熟悉基于大语言模型的搜索和社交线索，这可能影响研究结果的普适性。（2）实验聚焦于开放式信息寻求任务，未涵盖事实检索或多轮协作场景。社交线索的有效性可能因任务目标而异（如决策支持与知识获取）。未来研究应考察社交线索在不同任务类型中的适用性。（3）该系统在一个中文内容平台上进行测试，其中线索提取和解读依赖特定的语言及文化惯例。由于不同语言和文化中的社交表达与规范各不相同，该系统的跨文化适应能力仍需进一步验证。

7 结论

本研究系统地考察了社交线索在基于大语言模型的信息寻求中发挥作用的方式。通过设计工作坊以及使用原型系统 *SoulSeek* 开展的实证研究，我们揭示了用户如何在基于大语言模型的搜索过程中主动利用并配置社交线索，以及这些线索如何在不同阶段（搜索前、搜索中和搜索后）影响用户的判断、方向感和意义建构。结果表明，社交线索不仅在信息筛选和理解方面发挥支持作用，还能促进用户主动参与和认知反思，使信息寻求过程从被动接收转变为积极协作。整合社交线索显著增强了用户对结果的信任、方向感和信息意外发现。然而，研究也发现，在复杂的社会语境下，大语言模型在具备充分的社会知识、解读跨文化情境以及平衡多维条件方面仍存在局限，这可能导致线索过度拟合或语义偏差。未来研究可以在模型层面增强模型对社会和文化语境的表征与推理能力，并在系统设计层面探索如何平衡透明性、可控性与自动化，从而支持更具社会意识的人机协同参与式搜索体验。

致谢

本研究由 CAAI-Lenovo 蓝天科研基金资助。Tun Lu 还任职于上海市数据科学重点实验室、复旦大学老龄研究院、教育部国家发展与智能治理实验室，以及复旦大学上海智能电子系统研究所。

参考文献

[1] Akari Asai、Zeqiu Wu、Yizhong Wang、Avirup Sil 和 Hannaneh Hajishirzi。2024年。Self-rag：通过自我反思学习检索、生成和批判。（2024年）。[2] Dan Bennett、Oussama Metatla、Anne Roudaut 和 Elisa D Mekler。2023年。人机交互如何理解人的能动性与自主性？载于《2023年CHI计算系统中的人为因素会议论文集》。1–18。[3] Daniel Bennett、Oussama Metatla、Anne Roudaut 和 Elisa D. Mekler。2023年。人机交互如何理解人的能动性与自主性？载于《2023年ACM CHI计算系统中的人为因素会议论文集》。doi:10.1145/3544548.3580651[4] Virginia Braun 和 Victoria Clarke。2006年。在心理学中运用主题分析。《心理学定性研究》3，2（2006年），77–101。[5] Donald O Case 和 Lisa M Given。2016年。寻找信息：关于信息寻求、信息需求与行为的研究综述。（2016年）。[6] Jiawei Chen、Hande Dong、Xiang Wang、Fuli Feng、Meng Wang 和 Xiangnan He。2023年。推荐系统中的偏差与去偏：综述及未来方向。《ACM信息系统汇刊》41，3（2023年），1–39。[7] Sizhe Cheng、Jiaping Li、Huanchen Wang 和 Yuxin Ma。2025年。RAGTrace：理解并优化检索增强生成中的检索-生成动态。载于《第38届ACM用户界面软件与技术研讨会论文集》。1–20。[8] Minje Choi、Jiaxin Pei、Sagar Kumar、Chang Shu 和 David Jurgens。2023年。大语言模型理解社会知识吗？使用Socket基准评估大语言模型的社会性。arXiv预印本 arXiv:2305.14938（2023年）。[9] Leon Ciechanowski、Aleksandra Przegalinska、Mikolaj Magnuski 和 Peter Gloor。2019年。恐怖谷的阴影之中：一项关于人与聊天机器人交互的实验研究。《未来一代计算机系统》92（2019年），539–548。[10] Munmun De Choudhury、Meredith Ringel Morris 和 Ryen W White。2014年。在线寻求和分享健康信息：比较搜索引擎与社交媒体。载于《SIGCHI计算机系统中的人为因素会议论文集》。1365–1376。[11] Brenda Dervin 等。1983年。意义建构研究概述：迄今为止的概念、方法与成果。（1983年）。[12] Marian Dörk、Sheelagh Carpendale 和 Carey Williamson。2011年。信息漫游者：重新审视信息寻求。载于《SIGCHI计算机系统中的人为因素会议论文集》。1215–1224。[13] Wenqi Fan、Yujuan Ding、Liangbo Ning、Shijie Wang、Hengyun Li、Dawei Yin、Tat-Seng Chua 和 Qing Li。2024年。大语言模型与检索增强生成：面向检索增强型大语言模型的综述。载于《第30届ACM知识发现与数据挖掘会议论文集》。6491–6501。[14] Xianzhe Fan、Qing Xiao、Xuhui Zhou、Jiaxin Pei、Maarten Sap、Zhicong Lu 和 Hong Shen。2025年。用户驱动的价值对齐：理解用户对人工智能伴侣中偏见和歧视性言论的看法及应对策略。载于《2025年CHI计算系统中的人为因素会议论文集》。1–19。[15] Jasper Feine、Ulrich Gnewuch、Stefan Morana 和 Alexander Maedche。2019年。对话代理的社交线索分类体系。《国际人机交互研究期刊》132（2019年），138–161。[16] Shan Gao、Kaixian Yu、Yue Yang、Sheng Yu、Chenglong Shi、Xueqin Wang、Niansheng Tang 和 Hongtu Zhu。2025年。利用大语言模型构建知识图谱，探索心理健康。《自然通讯》16，1（2025年），7526。[17] Weiwei Gao、Kexin Du、Yujia Luo、Weinan Shi、Chun Yu 和 Yuanchun Shi。2024 年。Easyask：一款面向老年人、支持语音和触控输入的应用内情境化教程搜索助手。《 ACM 交互、移动、可穿戴与泛在技术汇刊》8，3（2024年），1–27。[18] Yunfan Gao、Yun Xiong、Xinyu Gao、Kangxiang Jia、Jinliu Pan、Yuxi Bi、Yixin Dai、Jiawei Sun、Haofen Wang 和 Haofen Wang。2023年。面向大语言模型的检索增强生成：综述。arXiv预印本 arXiv:2312.10997 2，1（2023年）。[19] Philipp Hallgarten、Verena Jasmin Hallitschke、Enkelejda Kasneci、Michael Beigl 和 Tobias Grosse-Puppendahl。2025年。RouteLLM：具备原生路由上下文理解能力、支持上下文感知推理的大语言模型。《ACM交互、移动、可穿戴与泛在技术汇刊》9，3（2025年），1–34。[20] Yuheng Hu、Shelly D Farnham 和 Andrés Monroy-Hernández。2013年。Whoo.ly：利用社交媒体促进超本地化社区的信息寻求。载于《SIGCHI计算机系统中的人为因素会议论文集》。3481–3490。[21] Wendy Huang、Ravi Teja Chadalavada、Roberta Paradeda、Kerstin Fischer 和 Omar Mubin。2024年。向服务机器人传达人类意图：理解面对面对话中社交线索的使用。载于《2024年ACM/IEEE人机交互国际会议论文集》。[22] Yucheng Jiang、Yijia Shao、Dekun Ma、Sina Semnani 和 Monica Lam。2024年。走进未知的未知：通过参与语言模型代理对话促进投入式人类学习。载于《2024年自然语言处理经验方法会议论文集》。9917–9955。

[23] Kyung-Sun Kim 和 Sei-Ching Joanna Sin。2015年。探究旅游协作信息寻求中的移动社交媒体用户行为。载于《2015年ACM计算机系统中的人为因素会议（CHI）论文集》。doi:10.1145/2702123.2702508[24] Stefanie Helene Klein。2025年。类人社交线索对基于文本的对话代理所引发社交反应的影响：一项荟萃分析。《人文与社会科学通讯》12，1（2025年），1–16。[25] Ekaterina M Kosova、Ivan V Spasenov 和 Elena S Gorbunova。2025年。新闻界面中的社交线索：建立初始在线信任的关键。《人机交互》（2025年），1–26。[26] Carol C Kuhlthau。1991年。搜索过程探析：从用户视角看信息寻求。《美国信息科学学会期刊》42，5（1991年），361–371。[27] Patrick Lewis、Ethan Perez、Aleksandra Piktus、Fabio Petroni、Vladimir Karpukhin、Naman Goyal、Heinrich Küttler、Mike Lewis、Wen-tau Yih、Tim Rocktäschel 等。2020年。面向知识密集型自然语言处理任务的检索增强生成。《神经信息处理系统进展》33（2020年），9459–9474。[28] Yaqiong Li、Peng Zhang、Hansu Gu、Tun Lu、Siyuan Qiao、Yubo Shu、Yiyang Shao 和 Ning Gu。2025年。DeMod：一种用于有害内容审查的综合工具，具备可解释检测和个性化修改功能。《ACM人机交互汇刊》9，2（2025年），1–24。[29] Zhi Li、Yanwei Li、Yuntao Zhang 和 Jian Li。2024年。通过多模态线索进行视频中的社交情境感知人物搜索。载于《IEEE/CVF计算机视觉与模式识别会议（CVPR）论文集》。[30] Baoxi Liu、Peng Zhang、Yubo Shu、Zhengqing Guan、Tun Lu、Hansu Gu 和 Ning Gu。2022年。构建用于社交媒体文本内容审查的个性化模型。《ACM人机交互汇刊》6，CSCW2（2022年），1–31。[31] Jiahao Liu、Dongsheng Li、Hansu Gu、Peng Zhang、Tun Lu、Li Shang 和 Ning Gu。2025年。利用公平采样实现无偏协同过滤。载于《第48届国际ACM SIGIR信息检索研究与发展会议论文集》。2555–2559。[32] Ming Liu、Ze Ma、Shuo Zhao、Qinglin Wu、Xue Zhang 和 Jiliang Tang。2024年。SocialCueSwitch：通过多种感官呈现社交线索，实现可定制的无障碍体验。arXiv预印本 arXiv:2403.05840（2024年）。[33] Yu Mei、Yuanxi Wang、Shiyi Wang、Qingyang Wan、Zhuojun Li、Chun Yu、Weinan Shi 和 Yuanchun Shi。2025年。InterQuest：用于对话式搜索中动态用户兴趣建模的混合主动式框架。载于《第38届ACM用户界面软件与技术研讨会论文集》。1–23。[34] Christian Meurisch、Cristina A. Mihale-Wilson、Adrian Hawlitschek、Florian Giger、Florian Müller、Oliver Hinz 和 Max Mühlhäuser。2020年。探索用户对主动式人工智能系统的期望。《ACM交互、移动、可穿戴与泛在技术》4，4，文章146（2020年12月），22页。doi:10.1145/3432193[35] Joan Meyers-Levy 和 Brian Sternthal。1991年。使用信息线索和作出判断方面的性别差异。《营销研究期刊》28，1（1991年），84–96。[36] Ashlee Milton、Juan F Maestre、Abhishek Roy、Rebecca Umbach 和 Stevie Chancellor。2024年。循环式信息寻求：用户如何利用个人信息生态系统寻找心理健康信息。载于《2024年CHI计算系统中的人为因素会议论文集》。1–16。[37] Lena Mulansky、Rüdiger Pryss、Caroline Cohrdes、Harald Baumeister 和 Felix Beierle。2022年。新冠疫情期间社交媒体应用使用与PHQ-9抑郁评分的关系。载于《2022年ACM普适与泛在计算国际联合会议及2022年ACM可穿戴计算国际研讨会附属论文集》。188–192。[38] Michael J Muller 和 Sarah Kuhn。1993年。参与式设计。《ACM通讯》36，6（1993年），24–28。[39] Seyed Parsa Neshaei、Matea Tashkovska、Paola Mejia-Domenzain、Thiemo Wambsganss 和 Tanja Käser。2025年。以用户为中心的反思性写作辅助：利用检索增强生成增强个性化支持。载于《CHI计算机系统中的人为因素会议扩展摘要》。1–8。[40] Shirui Pan、Linhao Luo、Yufei Wang、Chen Chen、Jiapu Wang 和 Xindong Wu。2024年。统一大语言模型与知识图谱：路线图。《IEEE知识与数据工程汇刊》36，7（2024年），3580–3599。[41] Gordon Pennycook、Jamie McPhetres、Bertrand Bago 和 David G. Rand。2023年。社交媒体情境会干扰真伪辨别。《科学进展》9，24（2023年），eabo6169。doi:10.1126/sciadv.abo6169[42] Peter Pirolli。2009年。十的幂次：在多个尺度上对复杂信息寻求系统建模。《计算机》42，3（2009年），33–40。[43] Peter Pirolli 和 Stuart Card。1995年。信息获取环境中的信息觅食。载于《SIGCHI计算机系统中的人为因素会议论文集》。51–58。[44] Peter Pirolli 和 Stuart Card。1999年。信息觅食。《心理学评论》106，4（1999年），643。[45] Chandra Prabha、Lynn Silipigni Connaway、Lawrence Olszewski 和 Lillie R Jenkins。2007年。多少才算足够？信息需求满意化。《文献学期刊》63，1（2007年），74–89。[46] Zhigu Qian、Jiaojiao Fu 和 Yangfan Zhou。2025年。探索面向中国老年人的语音助手设计中的文化与代际动态。《ACM交互、移动、可穿戴与泛在技术汇刊》9，1（2025年），1–18。[47] Pragnya Ramjee、Bhuvan Sachdeva、Satvik Golechha、Shreyas Kulkarni、Geeta Fulari、Kaushik Murali 和 Mohit Jain。2025年。CataractBot：一款面向白内障患者、由大语言模型驱动且有专家参与的聊天机器人。《ACM交互、移动、可穿戴与泛在技术汇刊》9，2（2025年），1–31。[48] Fabian Reinkemeier、Philipp Spreer 和 Waldemar Toporowski。2021年。语音商务中的语音助手：社交线索对信任和满意度的影响。载于《国际经济信息学会议》。Springer，130–135。稿件提交至 ACM

[49] Peter Gordon Roetzel。2019。信息时代的信息过载：一项综述，回顾了工商管理、商业心理学及相关学科的文献，并采用文献计量方法和框架开发方法。《商业研究》12，2（2019），479–522。[50] Hope Schroeder、Doug Beeferman、Maya Detwiller、Dimitra Dimitrakopoulou 和 Deb Roy。2025。Forage：理解社区对话中的基于检索增强生成的意义建构。载于《CHI计算系统中的人为因素会议扩展摘要集》。1–12。[51] Nikhil Sharma、Q Vera Liao 和 Ziang Xiao。2024。生成式回音室？大语言模型驱动的搜索系统对多元信息寻求的影响。载于《2024年CHI计算系统中的人为因素会议论文集》。1–17。[52] Yubo Shu、Haonan Zhang、Hansu Gu、Peng Zhang、Tun Lu、Dongsheng Li 和 Ning Gu。2024。RAH! 推荐系统–助手–人类：一种由大语言模型智能体支持的以人为中心的推荐框架。《IEEE计算社会系统汇刊》11，5（2024），6759–6770。[53] Yubo Shu、Peng Zhang、Hansu Gu、Wenjia Liu、Xinyue Yu、Tun Lu 和 Ning Gu。2024。理解危机响应中协作桌的使用。载于《CCF计算机支持的协同工作与社会计算会议论文集》。Springer，197–212。[54] Martin Tanis 和 Tom Postmes。2003。计算机媒介传播中的社交线索与印象形成。《传播学杂志》53，4（2003），676–693。[55] Miao Tian、Zhiwen Yu、Zhu Wang 和 Bin Guo。2015。结合社交媒体与基于位置的服务进行店铺类型推荐。载于《2015年ACM普适与泛在计算国际联合会议附属论文集暨2015年ACM可穿戴计算国际研讨会论文集》。161–164。[56] Cecilie S Traberg、Trisha Harjani、Jon Roozenbeek 和 Sander Van der Linden。2024。社交线索与来源效应对错误信息易感性的说服效果。《科学报告》14，1（2024），4205。[57] Shreshth Tuli、Giuliano Casale 和 Manuel Roveri。2025。SELA：智能边缘大语言模型智能体，优化人工智能助手的响应权衡。《ACM交互、移动、可穿戴与泛在技术汇刊》9，3（2025），1–18。[58] Ruotong Wang、Xinyi Zhou、Lin Qiu、Joseph Chee Chang、Jonathan Bragg 和 Amy X Zhang。2025。Social-RAG：从群体互动中检索信息，为人工智能生成提供社会依据。载于《2025年CHI计算系统中的人为因素会议论文集》。1–25。[59] Napong Wanichayapong、Monsak Socharoentum 和 Wasan Pattara-Atikom。2015。改进事故统计数据：社交媒体的作用。载于《2015年ACM普适与泛在计算国际联合会议附属论文集暨2015年ACM可穿戴计算国际研讨会论文集》。1153–1158。[60] Tom D Wilson。1999。信息行为研究中的模型。《文献学期刊》55，3（1999），249–270。[61] Tom D Wilson。2006。重新审视用户研究与信息需求。《文献学期刊》62，6（2006），680–684。[62] Stephan Winter、Miriam J Metzger 和 Andrew J Flanagin。2016。选择性使用新闻线索：社交媒体环境中信息选择的多重动机视角。《传播学杂志》66，2（2016），208–228。doi:10.1111/jcom.12241[63] Xuhai Xu、Bingsheng Yao、Yuanzhe Dong、Saadia Gabriel、Hong Yu、James Hendler、Marzyeh Ghassemi、Anind K Dey 和 Dakuo Wang。2024。Mental-LLM：利用大语言模型通过在线文本数据预测心理健康。《ACM交互、移动、可穿戴与泛在技术汇刊》8，1（2024），1–32。[64] Takehiro Yamamoto、Yusuke Yamamoto 和 Sumio Fujita。2018。探索人们对网页搜索中审慎信息寻求的态度与行为。载于《第27届ACM信息与知识管理国际会议论文集》。963–972。[65] Bufang Yang、Yunqi Guo、Lilin Xu、Zhenyu Yan、Hongkai Chen、Guoliang Xing 和 Xiaofan Jiang。2025。Socialmind：一种基于大语言模型的主动式增强现实社交辅助系统，具备类人感知能力，可支持现场实时互动。《ACM交互、移动、可穿戴与泛在技术汇刊》9，1（2025），1–30。