

生成式人工智能对话式搜索对个人用户信息搜寻行为的影响研究

蒙新梦

四川大学公共管理学院, 四川 成都

收稿日期: 2024年5月31日; 录用日期: 2024年7月16日; 发布日期: 2024年7月26日

摘要

[目的/意义] 文章旨在研究在生成式人工智能背景下, 对话式搜索如何影响用户的信息搜寻行为模式。[方法/过程] 通过文献理论研究方法进行论证, 提取关键概念与成果, 阐明对话式搜索在技术应用和产品实践上的发展动向, 阐述对话式搜索对传统模式的影响。[结果/结论] 研究发现, 对话式搜索重塑了用户信息搜寻的结果获取模式、人机交互方式及语言表达模式。它通过主动生成答案满足个性化需求, 重塑了从被动接收到主导定制的信息过程。对话式搜索被验证为一种重要的信息搜寻模式, 已成为未来信息搜寻发展的新动向。它将引领信息服务产业走向用户中心和知识服务的新阶段。信息机构需要重视AI应用, 跟进技术变革以应对挑战。

关键词

生成式人工智能, 对话式搜索, 信息搜寻行为

The Impact of Generative Artificial Intelligence Conversational Search on Individual Users' Information-Seeking Behavior

Xinmeng Meng

School of Public Administration, Sichuan University, Chengdu Sichuan

Received: May 31st, 2024; accepted: Jul. 16th, 2024; published: Jul. 26th, 2024

Abstract

This study aims to investigate the impact of conversational search on users' information-seeking

behavior patterns in the context of generative artificial intelligence. Through literature review and theoretical research, key concepts and findings are extracted to elucidate the development trends of conversational search in technological applications and product practices, and to expound its influence on traditional modes. The study reveals that conversational search reshapes users' patterns of information-seeking, human-computer interaction methods, and language expression modes. It satisfies personalized needs by generating answers and reshapes the information process from passive reception to active customization. Conversational search is validated as an important information-seeking mode and has become a new direction for future information search. It will lead the information service industry towards user-centric and knowledge-based services. Information organizations need to prioritize AI applications and keep up with technological transformations to meet the challenges.

Keywords

Generative Artificial Intelligence, Conversational Search, Information-Seeking Behavior

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,生成式人工智能领域取得长足进步,代表性成果不断问世。2022年11月,OpenAI面向公众发布ChatGPT模型,这标志着用户生成内容模式向人工智能辅助内容生成模式的重要转变[1]。随后,一批顶尖科技企业如谷歌、百度等也推出Apprentice、文心一言等大规模预训练语言模型。

与此同时,信息检索领域研究者积极探索人工智能技术在对话式检索中的应用前景。微软、谷歌先后公布相关计划,旨在打造支持对话式交互的新一代搜索产品。作为目前自然语言处理能力最强大的技术范式,生成式人工智能为对话式搜索服务提供铺垫,预计将成为未来主流的智能信息服务模式[2][3]。

对终端用户而言,相较传统检索服务,对话式搜索不仅可以提供更多元的选择空间,其重新定义信息需求表达方式及检索结果呈现机制也将深远影响用户的信息采集行为。本研究将以生成式人工智能为背景,探讨对话式搜索如何重塑用户的信息搜寻模式,研究成果或有助于信息服务机构规划未来的业务布局和资源运营策略。

2. 生成式人工智能在信息检索服务领域的应用前景

生成式人工智能并非一个全新的概念,其理论基础可以追溯到机器学习领域一些经典模型的设计,例如条件随机场模型与生成对抗网络等。但直到最近几年,随着大数据驱动下模型规模和计算能力的飞速增长,生成式人工智能才真正携带强大生产力得以广泛应用。

信息检索服务业是最早受到生成式人工智能技术影响和推动的行业之一。作为一个信息密集型的产业,它受益于大数据和人工智能技术日新月异的发展。随着网络信息量的指数增长,用户搜索语言和需求模式日益复杂多样,传统检索模式已难满足个性化体验。而生成式人工智能支持的对话式交互和内容生成模式,为信息服务企业提供了全新的思路。

经过近年来,生成式人工智能在理论基础和实际应用两个层面都取得长足进展。在技术实现水平和产品应用规模方面都成熟许多。这为信息检索这一信息密集型行业,提供了潜在的技术革新引擎。它已

经成为信息检索产业持续改进和创新发展的动力，给行业带来深远影响。未来，在人工智能和大数据得到进一步成熟的基础上，生成式人工智能在信息检索服务中的应用前景将更加广阔。

2.1. 生成式人工智能技术发展概况

2.1.1. 生成式模型与生成式人工智能

生成模型及其在 AI 应用中的演进路径一直受到学术界广泛关注。作为机器学习经典范式之一，生成模型与判别模型在原理设计上存在重要差异：前者面向样本空间进行概率分布建模，计算联合概率分布，然后通过贝叶斯公式得出条件概率。它注重建立样本特征的内在联系。而后者基于决策函数划分不同类别边界，依赖条件概率进行分类预测任务。这两种模型各自解决了机器学习的不同子任务。但随着深度学习算法与计算资源的大幅进步，生成模型在自然语言、图像以及其他多模态内容中的应用表现日臻成熟。特别是通过无监督学习大规模语料，构建深层神经语言模型已成为目前预测性能最强的生成 AI 范式[4] [5]。

基于此，学术界开始将其定性为完全不同于传统计算的“生成式 AI”范式。它具有自动产出多元内容的能力，为许多应用场景提供了前所未有的解决方案[4] [5]。

2.1.2. 内容生产模式在互联网发展背景下的演进

随着 Web 技术的迭代，内容生产模式也在不断演进。Web1.0 时代，基于早期信息门户，专业机构产出的内容(PGC)占主导地位。进入 Web2.0 阶段，互联网平台架构的成熟。以社交媒体为代表的 UGC 模式广泛采用，深刻影响人们的工作生产模式。ChatGPT 问世被视为 Web3.0 时代的里程碑事件。在此之前，人工智能辅助绘画、电影制作以及 NFT 等新兴应用的盛行，预示着基于 AI 自动生成内容(AIGC)正成为下一阶段关键资源形式。

综上，随着互联网技术和应用模式的深入发展，内容生产已从初期的专业机构为主，逐步向开放共享和技术赋能的方向迈进。ChatGPT 标志着 AIGC 产业链逐步成熟[1] [4]。

2.1.3. 生成式人工智能模型从预训练到联网迭代的历程

OpenAI 自 2018 年开始研发 ChatGPT 项目，经历 1.0~4.0 多个版本升级[6] [7]。其中大众熟知的 3.5 版采用 Transformer 架构，通过大规模无监督语料训练构建了强大的语言理解能力。值得注意的是，ChatGPT 4.0 版实现了联网访问功能。这意味着生成模型可以跳出单纯依赖预训练数据集的限制，直接利用互联网上实时更新的内容进行训练。它标志着生成式 AI 从起初的静态模型走向动态迭代过程。模型可以根据最新数据进行在线优化，最大限度提升在复杂互联网环境下的应用价值。这被视为生成性 AI 研究一个重要阶段性成果[4] [8]。

2.2. 生成式人工智能在信息搜寻中的应用潜能与机遇

作为一种新兴的生成式人工智能技术范式，在信息检索领域的应用中充满了巨大的机遇与潜能。这主要表现在：它具有通过大规模无监督学习深度训练词汇知识的强大数据处理能力，有效应对不断增长的信息量；不仅能进行内容匹配检索，还可以根据上下文智能生成个性化信息，从而丰富了信息表达维度；支持跨模态资源的深度整合，可以实现文本、图像、结构数据在统一平台上的融合挖掘；采用对话式交互体现出强大的个性化交互能力，有力提升了用户体验满意度。基于这些独特的技术优势，生成式人工智能无可避免地将成为信息检索技术研发的重要方向和趋势。它能够很好满足个性化搜索在内容多样性和交互形态上的需求，给信息服务行业带来系统升级的契机。目前，包括谷歌、微软在内的多家公司已经在这一前沿领域开展了积极的技术布局与产品探索[4] [8]。

2.2.1. 信息搜寻行为概述

信息行为是一个广泛的概念，它指代用户在整个信息需求产生和满足过程中采取的一系列行为[9]。作为信息行为研究中重要的一个分支，信息搜寻行为指用户在感知到信息需求后所进行的主动获取与处理信息的一系列行为[10]。具体来说，它包括但不限于[11]：根据信息需求主动进行内容检索获取；对检索结果进行过滤选择，匹配自己的实际需求；为了提高利用效率，对搜集到的信息进行归类整理；基于个性特征对信息展现进行定制，以提高用户体验满意度；信息服务机构方面不断优化提供方式，丰富信息表达形式等内容。这些行为构成了信息搜寻行为的内在属性及其持续发展演变的动力。随着用户需求个性化程度不断深化，信息搜寻也将随之呈现出更多元化的发展态势。深入研究这一复杂过程，对优化信息服务与指导机构决策具有重要意义[3]。

2.2.2. 传统搜索引擎面临的挑战

随着互联网深入应用，信息技术日新月异的发展正在深刻改变用户信息需求的特征。随着大数据时代的到来，用户信息查询不再限于简单的关键词检索，而是一个个性化和多层次的过程[12]。这对传统搜索引擎提出了诸多挑战，例如：用户查询方式从词汇扩展到任务导向，内容形式呈现出日益复杂的趋势；海量信息资源的增长提高了对高效智能处理的要求；跨模态资源整合已经成为新浪潮，单一文本检索模式难满足使用场景的融合需求；基于静态列表结果的传统模式欠缺主动回答能力等。此外，多语种并存环境和个性化服务水平亦需不断提升。总体来说，基于关键词检索的传统模式难以适应信息需求快速演变的复杂特征。这为构建面向用户个性化服务的新一代智能检索机制奠定重要基础。

2.2.3. 对话式搜索引擎浪潮来临

ChatGPT 问世标志着以生成式内容为导向的对话式信息检索技术万众关注[12]。与传统搜索模式不同，对话式搜索引擎具备以下优点：它可以解决数据规模问题，实现内容创造；支持跨模态整合检索；提供自然互动交互体验[4]，这为用户个性化需求提供了全新表达渠道。众多对话式搜索引擎，代表产品如 Andi 和 Phinch 等，已经对传统搜索引擎带来了巨大的市场压力。许多传统搜索引擎龙头企业也在积极推动传统搜索引擎，比如微软推出的对话式搜索引擎新必应(New Bing) [13]，谷歌宣布将 AIGC 的相关技术与目前搜索引擎融合[14]等等。商业实践表明对话式搜索已成主流趋势。它优化用户体验，重新定义信息检索范式，引领产业升级[3]。

3. 对话式搜索引擎重塑信息搜寻行为

与传统检索模式不同，对话式搜索引擎深远影响用户信息搜寻各个环节：它通过 AI 主导智能生成，为用户直接提供答案，而非单一列表信息；采用自然语言对话式交互取代传统表单查询，使用户充分发挥主导性；支持跨平台深度融合，助推资源深度开放与合作应用；理解用户语境和错误输入，使搜索语言个性化自由度更强。总之，对话引擎以生成式 AI 为基础，于信息发布获取各环节进行重构。这将深远影响用户未来的信息行为特征。

3.1. 信息搜寻结果的演进

3.1.1. 从检索到生成：信息搜寻结果的迭代

传统搜索引擎通过信息检索技术，实现数据库内既有信息的匹配检索，但其本身无法产生新的内容。相比之下，对话式搜索引擎利用了深度学习语言模型，它不仅可以实现信息的检索，更重要的是可以在用户交互基础上进行智能生成，为用户提供全新内容。尽管两者在工作过程上存在差异，但我们不能忽略它们的根本目标都是解决用户的信息需求。传统搜索模式通过分类与选择实现信息提取，而对话搜索

则利用 AI 模型的总结归纳能力将知识点生成出来呈现给用户。实际上，无论是选择式检索还是智能生成，只要能满足用户信息查询需求，其本质上都可以视为一种信息搜寻行为。而对话搜索相比传统模式而言，通过不断迭代创造性地提炼和生成信息，其实已经实现了信息搜寻结果的一个新阶段。综上，我们可以认为对话搜索虽使用了不同的工作机制，但仍属于一种信息搜寻服务，它推进了信息搜寻结果的演化进程。

3.1.2. 从结果到答案：信息搜寻结果的精准化

传统信息检索模式的一个问题在于，它提供的往往不是用户期待中的直接答案，而是一些与需求相关但需要后续处理的信息结果。满足用户实质需求往往需要额外步骤，例如格式转换、内容总结和提炼等。相比之下，对话式搜索引擎通过与用户的高度交互，可以更快捕捉到核心问题点，直接生成满足需求的精准答案。这与信息服务一直追求的理想状态是一致的——能够直接提供目标用户想得到的结果，而非需要额外处理就能得到答案。相比传统模式，对话搜索在这一点上表现出色，很好地实现了从结果向答案的精准演变。这不仅体现了交互体验的提升，也更好满足了用户个性化信息需求。它推进了信息检索服务向更人性化智能化方向的发展。

3.2. 信息交互的革新：从“一对多”到“一对一对多”

随着互联网的兴起和普及，人们信息获取模式发生重大变化。从过去以“一对一”交互为主导，逐步转向以“一对多”交互(见图 1)为主。个体能够通过互联网接触来自全球各地海量信息源。这一时期被称为信息大爆炸时代，每个人面临的信息量无比庞大，可能增长了几何级数。

但是，海量信息同时也带来新的挑战。个体难以消化和处理如此庞大的数据量，也很难在其中高效找到需要的信息。这使得每个人的信息查找需求达到前所未有的高峰。面对信息爆炸，个体很难单独完成大数据分析计算。这在一定程度上提高了人们参与和处理信息的门槛，也降低了信息服务的个性效能。

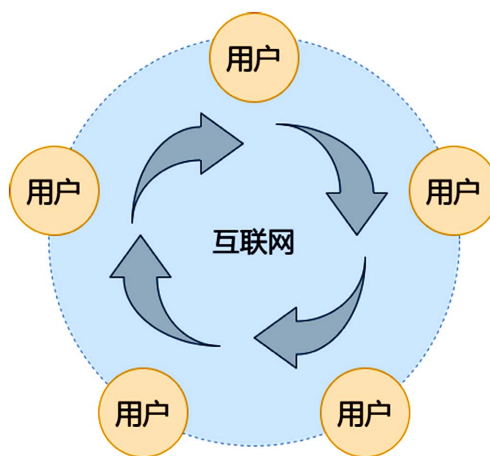


Figure 1. “One-to-many” information interaction model

图 1. “一对多”信息交互模式

然而，对话式搜索引擎作为一种代表性的生成式人工智能，其出现深刻改变了用户信息交互的模式。用户不再需要直接面对海量互联网络数据进行信息交互选择，也无需亲自处理大规模信息资源。通过同对话搜索引擎进行高效的“一对一”人机对话，即可完成信息查询任务提交。我们可以视对话搜索引擎为一种智能化的“信息助手”，它代表用户利益完成与互联网于数据库间的后台“一对多”信息交互与处理工作。这实现了一个前所未有的“一对一对多”交互新模式(见图 2)：用户作为主体，可以通过自然

互动与“信息助理”对话，实现对信息海洋中任意资源的访问呈现，而无需自行处理大数据任务。这大幅降低了个人检索门槛，极大提升了信息访问效率。

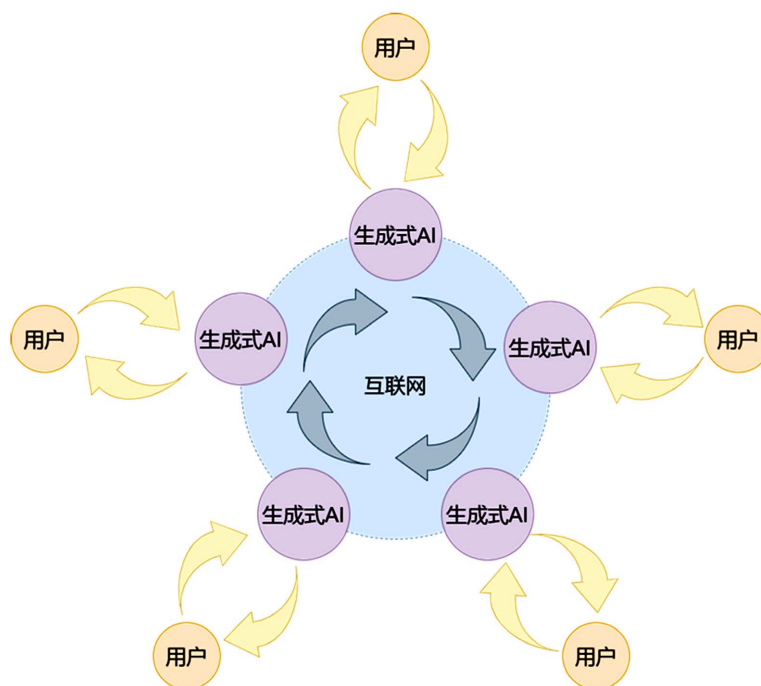


Figure 2. “One-to-one-to-many” information interaction model
图 2. “一对一对多”信息交互模式

虽然个人通过编程已可开展部分大数据处理，但就整体效率和门槛来看，自行完成海量信息资源的处理与挖掘仍需面临一定难度和成本。幸而，基于“一对一对多”的新交互模式极大降低了这个门槛。未来只要持有智能设备，每个人都将可以通过与生成式 AI 的自然对话，轻松完成大规模信息的处理分析和个性化搜寻。这意味着在信息大爆炸后，个人首次真正掌握了面对海量信息资源的能力。以生成式人工智能为基础，人们像开启宝库大门一般，得以轻松访问信息深海。随着 AI 语言模型不断优化，它将给每个人提供充分地消化信息的“钥匙”。预计未来将迎来一个信息集约化新纪元，个人自由调动知识的同时，信息将呈现前所未有的开放共享局面。对于个人来说，信息大塌缩时代即将到来。

3.3. 信息共享的新途径：生成式人工智能间接交互

ChatGPT 推出联网功能后，一些人关注它是否会带来信息安全问题，目前已经出现了部分以生成式人工智能为媒介得信息泄露。不可否认这些事件是负面的，然而它们也表明对话式搜索引擎为信息共享生态带来革新。通过对生成模型的交互，用户可以直接或间接获取对方输入的信息，这为信息交流提供另一渠道。这种交流过程大多无意识和间接，但它使得用户在自行搜索信息的同时，也能共享自己的知识给他人。这实质上改变了信息共享的传统模式。它不再限于主动上传和下载，而是用户在日常询问对话中，通过数据对话而进行隐性共享。这为开放知识社区提供机会，但同时也需重视安全管理，以免信息外泄。整体来说，这反映出对话 AI 如何重塑人与人之间的互联互通新范式。

3.4. 检索语言的灵活与个性化

传统搜索由于技术限制，检索语言的要求比较严格，容错能力相对不足，导致用户输入词不当可能

难以得到满意答案。但对话式搜索不同，它通过强大的语言理解和人机交互能力，可以更好地扣捉用户真实需求，同时在必要时进行智能纠错。这使得检索词汇的表达获得空间，从严格词到主题式表达都能理解，大幅提升了检索语言的灵活性与个性化程度。随着 AI 发展，未来检索模式很可能全面转向主题检索主导。这不仅扩展了用户参与度，也极大促进信息获取效率。总体来说，这丰富了人机交互体验，为用户提供更人性化智能服务。

4. 信息服务机构在新环境下的信息服务智能与路径选择

正如分析中提到的，作为信息服务机构，积极应对新的技术变革显有必要。生成式 AI 不仅代表着信息产业潮流，更标志着知识传播模式将迈向一个全新阶段。不适应这一变革将面临较大挑战。因此，信息机构需要主动顺应时代，积极拥抱这一技术革新。利用自身丰富的信息资源优势，可以考虑建设专属对话搜索应用，实现信息与服务的无缝对接。通过打造智能自然语言接口，提供精准的查询服务，这不仅可以提升用户体验满意度，也将成为机构服务赋能的有效途径。只有主动学习运用新技术，信息服务才能真正迈入智能化新纪元。这将使机构服务走在时代前列，并且有效提升社会影响力。总体来说，积极应对就是信息机构应有的姿态。

4.1. 保留原始信息资源

保留完整的原始信息资源与生成互动式搜索结果的价值平衡是一个不容忽视的问题。不同于传统搜索模型下提供链接到原始信息源的方式，对话式搜索引擎会产生全新的自动生成内容，这些内容经过总结和提取，可以很好地满足用户的即时信息需求。然而，在信息提取过程中，原始信息资源中某些细节或隐藏性信息可能不易于呈现，从而难以包含在生成结果中。此外，对信息进行自动分类和概念表达的过程必然会与原始信息的真实性和完整性产生一定偏差。尽管提取性搜索结果的方便性能满足大部分场景下的效率要求，但原始完整信息资源中未被主动提取但同样重要的信息可能因此丧失。考虑到信息在不同阅读与研究场景下可能面临的多样需求，信息服务机构，特别是承担长期信息保存与传承职责的图书馆、档案馆等，还应高度重视原始信息资源本身的保存与开放，以兼顾信息利用的长远价值。

4.2. 控制和评估生成内容质量

面对基于生成式 AI 技术自动生成情报性内容的趋势，如何有效地评估和控制其质量成为信息服务领域一个重要而富挑战性的课题。目前，这一新型计算机辅助情报获取模式对应的算法和技术体系尚处在不断探索与优化的阶段，其自动生成结果的主观性和客观性尚需长足的提升。即便相关技术日益成熟，我们也不宜盲目地信任和采纳机器生成的信息，特别是在重要的决策支持或学术研究等场景下。信息服务机构应夯实自己对信息质量的独立判断能力，对自动生成内容进行深入质疑和多方验证，开发合理的评价体系追踪和反馈生成结果的优缺点。同时加强与人工智能开发单位的通力合作，共同探讨如何在技术手段上进行必要补救，以确保数字情报在服务社会过程中的可信任度。只有通过持续不断的监测与改进，该新型信息获取模式才能成为公信力较强的有效辅助，而非盲目依赖。

5. 总结

总结之，随着生成式 AI 技术不断走向成熟，以对话方式提供个性化信息服务的新型交互模式正成为未来信息获取与应用的必然趋势。用户行为模式的重构也在多个层面产生深远影响，从传统的信息检索模式向主动参与式的信息生成与共享模式转变，这将极大推动数字信息资源的开放应用。对信息服务机

构来说,要迎接信息时代不断深化的变革,积极适应这一趋势变化是十分重要的。需要重构内部工作流程,完善技术体系,同时加强对信息质量的把控。与此同时,还应主动参与到相关技术的研发与优化当中,为信息服务在新模式下持续改进提供支持。只有掌握研发趋势和用户需求演变规律,不断优化自身能力与服务模式,信息机构才能在信息社会高速迭代的大环境下屹立不懈,继续发挥核心作用。要充分做好各项准备,以迎接信息界面的深层次变革,为用户提供更智能化的服务体验,这将是未来各项工作的重中之重。

参考文献

- [1] 蒋俏蕾,陈宗海,张雅迪. 可供性视角下的人工智能生成内容[J]. 青年记者, 2023(20): 84-86.
- [2] 蔡子凡,蔚海燕. 人工智能生成内容(AIGC)的演进历程及其图书馆智慧服务应用场景[J]. 图书馆杂志, 2023, 42(4): 34-43+135-136.
- [3] 卢新元,张进澳,雷晓鹏. 人工智能生成内容环境下用户信息行为研究——以对话式搜索引擎为例[J/OL]. 情报理论与实践, 2023, 46(12): 84-92. <http://kns.cnki.net/kcms/detail/11.1762.G3.20231107.1443.006.html>, 2023-12-03.
- [4] 李白杨,白云,詹希旎,等. 人工智能生成内容(AIGC)的技术特征与形态演进[J]. 图书情报知识, 2023, 40(1): 66-74.
- [5] Wang, Y.H., Herron, L. and Tiwary, P. (2022) From Data to Noise to Data for Mixing Physics across Temperatures with Generative Artificial Intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, **119**, e2203656119. <https://doi.org/10.1073/pnas.2203656119>
- [6] ChatGPT: Optimizing Language Models for Dialogue. <https://openai.com/blog/chatgpt/>
- [7] Lund, B.D. and Wang, T. (2023) Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries. *Library Hi Tech News*. <https://doi.org/10.1108/LHTN-01-2023-0009>
- [8] 李书宁,刘一鸣. ChatGPT 类智能对话工具兴起对图书馆行业的机遇与挑战[J]. 图书馆论坛, 2023, 43(5): 104-110.
- [9] 乔欢. 信息行为学[M]. 北京: 北京师范大学出版社, 2010.
- [10] 岳剑波. 信息管理基础[M]. 北京: 清华大学出版社, 1999.
- [11] 于瑶,钟隽誉. 国内研究生群体学术信息查寻行为研究综述[J]. 情报科学, 2022, 40(8): 177-184.
- [12] 赵鑫,窦志成,文继荣. 大语言模型时代下的信息检索研究发展趋势[J]. 中国科学基金, 2023, 37(5): 786-792.
- [13] Forbes (2023) Google Stock Loses \$57 Billion amid Microsoft's AI "Lead"—And Reports It Could Be Replaced by Bing on Some Smartphones. <https://www.forbesafrica.com/technology/2023/04/17/google-stock-loses-57-billion-amid-microsofts-ai-lead-and-reports-it-could-be-replaced-by-bing-on-some-smartphones/>
- [14] Bloomberg (2023) Google "Racing" to Build New AI Tools for Search, NYT Says. <https://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=163114590&lang=zh-cn&site=ehost-live>