

人工智能时代高校图书馆知识服务体系的重塑研究^{*}

肖荣荣 郭晨晨

(东北林业大学, 黑龙江 哈尔滨 150040)

[摘要]从知识服务的底层逻辑思考人工智能时代知识服务的根本性变革,对改造和发挥情报系统传统资源能力,针对用户复杂问题及其解决方案的认知型决策型知识服务能力,知识服务支持进行了深入的研究和探讨。主要从数据支撑、服务重构、创新赋能、能力重塑4个方面详细讨论知识服务体系重塑,并提出当前面临的困境及思考。

[关键词]知识服务 人工智能 ChatGPT

[分类号]G252.7

1 引言

2022年11月30日,OpenAI发布了新一代聊天机器人程序ChatGPT,在短短两个月内其用户数量突破了一亿,引发了广泛关注。ChatGPT是一种基于互联网数据训练的深度学习模型,具备语言理解和文本生成能力,通过持续的迭代改进,该模型逐渐具备了对生成答案进行评判的能力,是一项革命性的生成式人工智能技术^[1]。除了作为聊天机器人,ChatGPT还具备智能化内容创作能力,可以根据对话的上下文环境生成特定内容,这一创新的人工智能技术对教育、科学研究、医疗健康和生物工程等领域产生了重要影响,并促使潜在应用研究快速增加。

ChatGPT的突破为图书馆发展带来了新的机遇,根据《新媒体联盟地平线报告:2017年图书馆版》,人工智能(AI)被列为图书馆发展的六大技术方向之一,其中包括内容智能分析能力的提升和服务效能的改善。目前,已有高校图书馆开始运用ChatGPT提供知识服务和情报咨询。这些聊天机器人利用自然语言处理和机器学习等技术,能够与用户进行实时互动和交流,回答有关图书馆资源、检索技巧、研究方法等方面的问题,并根据用户需求快速获取和分析大量图书馆数据,提供个性化的知识解答。人工智能聊天机器人的引入为高校图书馆的知识服务和情报服务带来了新的机遇。

中国科学院文献情报中心张智雄等人在《ChatGPT对文献情报工作的影响》一文中阐述了人工智能技术迅速

发展对文献情报工作的启示,以及ChatGPT对文献情报工作多个方面的影响,从数据组织方式、知识服务模式、情报分析方法、文献使用方式、文献情报队伍建设要求以及文献情报工作重点6个方面分析了ChatGPT对文献情报领域的影响,并基于文献情报工作的特点,提出人工智能时代文献情报领域发展的9条建议^[2]。吴建中提出图书馆可以借力人工智能,提升内容智能分析能力,探索提升图书馆服务效能的研究与实践^[3]。茆意宏指出,智能识别(图像识别、语音识别、其他感知技术)、智能处理(组织、分析)、智能服务(检索、推荐、咨询)是人工智能对图书馆的主要影响方向,特别是在数字资源处理与服务方面将带来更大的变革^[4]。张晓林深入探讨了ChatGPT对知识服务带来的挑战和应对方法,从知识服务的底层逻辑对人工智能时代知识服务的根本性变革,改造和发挥情报系统传统资源能力,针对用户复杂问题及其解决方案的认知型决策型知识服务能力,知识服务支持进行了深入的分析和探讨^[5]。笔者以此为切入点,重新思考人工智能时代知识服务的未来走向及服务体系重塑路径。

2 人工智能时代知识服务体系的重塑

以ChatGPT为代表的AI革命对知识服务带来的巨大挑战使高校图书馆不得不认真思考AI时代知识服务体系的重塑问题,以下将从数据支撑、服务重构、创新赋能、能力重塑4个方面进行详细讨论。

^{*}本文系2022年黑龙江省高校图工委课题“基于智能体验的高校智慧图书馆建设路径研究”(项目编号:2021-077-A)的阶段性研究成果之一。

2.1 数据支撑:构建专业领域语料库,建立并组织垂直领域信息资源

ChatGPT的语料库来源于各种公开的互联网资源,如书籍、文章、维基百科、网页、论坛、社交媒体等,这些文本资源经过深度清洗和处理,确保了数据的质量和准确性。ChatGPT通过大规模的预训练来学习语言模型,并努力让模型能够对输入的问题作出有意义和准确的回答,但不会从特定的作者或来源处引用信息,它所产生的回答基于训练过程中接触到的广泛语料库的总体知识和理解,无法提供特定来源的实时或专业领域的最新信息,而是根据先前学习到的信息进行组织。因此,ChatGPT回答的质量和可信度可能受到训练数据和模型的限制。语料是AI获取知识的源泉,高价值语料工作是一切AI的基础^[6],构建该专业垂直领域信息资源的语料库极为重要。

构建一个知识库,整合领域知识、决策模型和规则。知识库中包括专业数据库资源、专家知识、经验数据、规则库等,与智能服务相结合,形成一个综合的决策支持系统。知识库的更新和管理可以通过专家反馈、领域知识更新等方式进行。首先,在构建高校图书馆专业领域语料库之前,需要明确目标和范围,确定好要收集的专业领域以及相关学科的范围;同时,需要考虑语料库的规模和深度,以及所涵盖的材料类型。其次,通过多种途径进行数据收集和抽取,包括订阅和购买专业期刊、数据库、学术搜索引擎等,还可以与相关学院、研究机构合作,获取他们的研究成果和数据。针对文献类资源,可以利用文献检索工具进行全文或摘要的抽取,确保语料库的广度和质量;在数据收集和抽取过程中,要注重数据的整理和标注。为了提高语料库的可利用性,应该对数据进行分类、归类和标记,以便后续的检索和利用。可以使用自动化的方法进行初步的整理与分类,再结合人工的校对来提高准确性和完整性。再次,建立适当的信息组织与存储模型,可以使用知识图谱、本体论和语义网络等技术手段来组织和表示垂直领域的信息,使用户能够更加方便地浏览、搜索和发现信息资源。此外,还可以利用标签、关键词和元数据等方式来描述和索引语料库中的文献和资源。为了提高垂直领域信息资源的可用性和互操作性,可以将语料库与其他已有的知识库和信息资源进行链接和集成,通过建立统一的标准和接口,实现各种资源之间的相互补充和共享,这样不仅能够增强垂直领域信息资源的价值,也能够促进学术交流和合作。最后,需要注意语料库的持续维护和更新,专业领域的知识和信息在不断发展和变化,因此语料库也需要及时更新和补

充。同时,要对语料库进行专业的质量控制和评估,确保信息的准确可靠。

2.2 服务重构:传统知识服务的AI智能化改造

ChatGPT能够通过自然语言处理和生成,实现智能问答、文档摘要、自动翻译等功能,提供更加高效和精准的知识服务。这将改变传统的人工检索和信息提供方式,使知识情报服务更加自动化、智能化,大大减少人工工作量。

2.2.1 个性化咨询和智能检索

传统情报知识服务系统的检索功能通常依赖于关键词匹配,而ChatGPT等AI技术可以根据用户的需求和背景提供个性化咨询服务,并通过自然语言理解和生成实现智能、灵活的检索,使传统的一对多咨询模式变得更加精准和高效。用户不再受限于特定的关键词或查询语法,可以用自然语言描述所需信息,更轻松找到知识。

2.2.2 实时情报和跨学科知识整合

ChatGPT等AI技术可以通过与多个数据源实时连接,提供即时的情报服务。传统情报知识服务系统需要人工整理和归纳大量信息,而ChatGPT可以将这一过程自动化,快速提供最新的信息和动态更新,帮助用户及时了解最新的发展动态。传统的学科知识服务往往由单个领域专家提供,而ChatGPT等AI技术可以整合多个学科领域的知识,提供跨学科的综合解决方案。通过对多个领域的知识进行深度学习和推理,ChatGPT可以为用户提供更全面、综合的知识服务。

2.2.3 个性化教育培训和多模态知识支持

AI技术可以根据用户的学习需求和进度,提供个性化的教育培训服务。通过与用户的对话交互,AI技术可以了解用户的学习水平和目标,并根据用户的反馈和表现调整教学内容和方式,提供有针对性的学习支持。此外,AI技术还可以支持多种模态的知识表达和呈现,如图像、语音、视频等,使知识更加直观和易于理解。

2.2.4 海量数据分析以及自动文献挖掘

传统情报知识服务系统往往无法有效分析和处理海量的数据,而ChatGPT等AI技术可以通过强大的计算能力和机器学习算法实现对大数据的快速挖掘和分析。这将帮助用户更好地理解 and 利用庞大的数据资源,发现隐藏在数据背后的价值点。传统的科技文献挖掘通常需要人工阅读和筛选大量文献,而ChatGPT等AI技术可以通过机器学习和自然语言处理,自动分析和提取科技文献中的关键信息。张智雄在2023年第六届未来智慧图书馆发展论坛中介绍了ChatGPT在国家重点研发计划项目“科

技文献内容深度挖掘及智能分析关键技术和软件”中的应用,这将大大提高文献挖掘的效率和准确性,帮助用户快速获取所需的科技知识。

通过以上4个方面的改造,传统知识服务得以智能化,同时提升了个性化咨询、检索和教育培训的效果,实现实时情报、跨学科知识整合、多模态知识支持、大数据分析及自助文献挖掘,为用户提供更智能、便捷和全面的知识服务体验。

2.3 创新赋能:AI创新助力解决多维复杂知识服务课题

随着AI技术的发展和应用范围的扩大,越来越多的科学领域开始利用AI方法和技术解决科学问题,推动科学研究的进展。利用AI方法可以加速科学研究的过程,改善科学实验和观测的效率,发现新的科学知识和洞见,并提供对复杂系统和现象的解释和预测。通过将AI与科学研究相结合,可以提高科学研究的效率 and 创新能力,推动科学的发展。同时,AI还可以通过积极的用户反馈和需求分析,了解用户提出的复杂问题的背景、具体需求和期望解决方案,与用户进行深入的对话和交流,以确保对问题的准确理解,并获取更多关键信息,对于复杂问题通过多次迭代和确认,不断优化回答以满足其需求和期望。

2.3.1 改善科学实验和观测的效率、加速科学研究的过程

AI可以处理大规模数据集,帮助科研人员从中发现模式、关联和趋势。借助AI的巨大技术优势,可将数学计算和科学模型相结合,高效处理海量数据,使传统的科研过程变得自动化、规模化、并行化和平台化,从而解决原来传统科学研究范式无法解决的问题,帮助科学家突破科研瓶颈。AI可以帮助科研人员在大量文献中搜索、过滤和整理信息,提供全面和准确的文献综述。此外,利用自然语言处理和知识图谱技术,还可以构建领域知识图谱,帮助科研人员理解领域内的概念和关系。

2.3.2 发现新的科学知识和洞见

AI可以帮助科研人员从大量数据中挖掘潜在的模式和关联。通过机器学习和数据挖掘技术,AI能够识别出科学领域中的新现象、规律或趋势,这些新的洞察为科研人员提供了观察和探索的起点,激发了新的研究方向和问题。通过处理来自不同学科领域的大量数据和信息,AI能够融合多个学科的知识 and 数据,发现不同学科之间的联系和相互影响,促进跨学科的创新。同时AI还为科研人员提供了许多创新工具和技术,例如基于深度学习的图像识别和分析软件、药物虚拟筛选工具、自然语言处理的文献综述系统等。这些工具和技术可以帮助科研人员更好地处理和理解复杂的数据,从而推动知识创新的

进程。

2.3.3 提供对复杂系统和现象的解释和预测

利用自然语言处理、文本挖掘和机器学习等技术,对复杂问题进行知识分析和解决方案生成,将知识库中的信息与用户问题进行匹配和推理,生成针对性的解决方案,并提供详细的解释和理由以增强可信度,整合专家知识和AI技术,建立专家系统和决策支持系统。通过与领域专家的合作和知识的融合,提供更加深入和全面的解决方案,同时保留专家的判断和经验。

利用机器学习和反馈机制,不断优化和提升知识服务能力。通过分析用户的使用数据和反馈意见,修正和更新知识模型,提高决策的准确性和针对性。要具备针对用户复杂问题及其解决方案的认知型决策型知识服务能力,需要不断学习和改进,从用户需求和问题本身出发,整合专家知识与技术手段,提供个性化、深入和全面的知识服务。

2.4 能力重塑:重塑基于用户需求的感知型知识服务能力

重塑基于用户需求的感知型知识服务能力对图书馆服务提升具有重要意义。它可以提高用户满意度、优化服务质量、增加用户参与度,支持学术研究与教学,帮助图书馆提供更精细化、个性化、高质量的知识服务,提升用户体验和服务质量,实现图书馆的可持续发展,进而推动整个图书馆行业的创新与进步。

2.4.1 技术培训与教育

提供AI技术培训,组织培训课程,使馆员了解AI技术的基本原理、应用场景和潜在价值,掌握AI在知识服务中的具体应用方法。强化AI素养,培养馆员对AI技术的认知和理解,使其能够主动了解和关注AI技术的发展动态,并将其运用到实际的知识服务工作中。

2.4.2 培养数据分析与判断能力

培养馆员的数据分析和判断能力,使其能够对大量数据进行有效整合和分析,培养数据驱动的决策思维,从中发现用户需求和行为特征,为知识服务提供有针对性的改进和优化策略。鼓励馆员结合用户反馈和AI输出进行综合研判,准确把握用户需求,拓展和精细化知识服务内容,提高满意度和质量。

2.4.3 基于用户体验的知识服务设计

鼓励馆员与用户保持密切联系,定期进行用户满意度调研和需求分析,了解用户对AI技术应用的期望和需求,为设计和改进知识服务提供指导。借助AI技术,培养馆员提供个性化的知识服务能力,如通过智能推荐系统和智能问答系统提供精准、个性化的图书推荐和问题解答服务。

3 面临的困境与思考

3.1 语义理解和上下文理解的困境

ChatGPT等AI在语义理解和上下文理解方面仍然面临挑战,存在对问题理解不准确,或者无法理解复杂的上下文信息的问题。这可能导致回答的不准确或无法满足用户的需求,需要不断改进模型来提高其语义理解和上下文理解的能力,也需要优化方法以提高回答的质量,比如提供更详细的足够明确和具体的问题、在问题中添加上下文、对模型的回答进行主动约束、进行迭代反馈等,避免模棱两可的问题。对于这类问题,最好明确要求ChatGPT给出客观事实或提供多个可能的观点。ChatGPT生成的回答有时会存在错误或不准确的情况,在使用回答之前,最好对其进行人工检查和必要的编辑,以确保回答的准确性。

ChatGPT虽然强大,但仍然存在一定的局限性。持续的反馈和优化可以帮助改善回答的质量,但在关键任务或重要决策中,最好仍然依靠人类专业知识和判断。与用户、专业人士和领域专家建立密切的合作关系,通过不断收集反馈和经验教训,改进和优化ChatGPT的性能和回答质量。同时,持续关注最新的技术进展和研究成果,公开透明的沟通可以增加用户对服务的信任度。

3.2 数据隐私和安全性的挑战

ChatGPT在提供知识服务时需要访问和处理用户的数据。这可能带来数据隐私和安全性的问题,如数据泄露、滥用等。在使用ChatGPT提供知识情报服务时,需要保护用户的数据隐私,并采取相应的安全措施。确保ChatGPT处理用户数据的合法性和隐私保护,遵守相关的隐私和数据保护规定。建立适当的数据存储壁垒,保障用户数据的安全性和保密性。

3.3 数据偏见和不公平

AI模型的训练数据往往存在偏见和不公平性,这会导致模型在处理特定群体或敏感问题时出现偏差或歧视。缺乏多样性和代表性的训练数据,以及对数据收集和过程的不透明性,是导致这种困境的主要原因之一。应获取具有多样性和代表性的训练数据,避免数据中的偏见。可以通过广泛涵盖不同群体、地域和文化背景的数据来提高模型的包容性和公正性。在数据预处理和模型训练过程中,采取措施检测和纠正潜在偏见。这可能涉及对训练数据进行分析、使用公平性指标来评估模型输出,以及应用公平性增强技术来平衡模型的决策。

3.4 信息可信度与错误传播

AI模型在生成文本时可能受到虚假信息或误导性内

容的影响,这可能导致错误的信息传播。模型的输出缺乏审查和验证机制,容易被滥用或误解,对社会产生负面影响。加强对AI模型输出的审核和验证,确保生成文本的内容准确可信。引入领域专家或人工审核,以审查和过滤误导性或虚假信息,需要向用户提供关于AI技术及其应用的教育,增强信息素养和媒体素养,使用户具备判断虚假信息和误导性内容的能力。

3.5 法律与伦理问题

AI的广泛应用引发了一系列法律和伦理问题,例如责任和道德、知识产权和数据使用权问题等。法律法规的滞后和对新问题适用的不确定性,使解决这些问题变得复杂而具有挑战性。

4 结语

随着AI技术的不断发展,高校图书馆将迎来更多的机遇与挑战。应保持敏锐的洞察力,紧跟时代步伐,灵活应用AI技术,推动高校图书馆知识服务体系的深度重塑。同时,要注重人文关怀,保持人类智慧的核心地位,通过合理的规划和管理,确保AI与人类智慧的有机结合,为用户提供丰富多样、人性化的知识服务。

参考文献:

- [1] 喻国明,苏健威.生成式人工智能浪潮下的传播革命与媒介生态——从ChatGPT到全面智能化时代的未来[J].新疆师范大学学报:哲学社会科学版,2023(5):81-90.
- [2] 张智雄,等.ChatGPT对文献情报工作的影响[J].数据分析与知识发现,2023(3):36-42.
- [3] 吴建中.再议图书馆发展的十个热门话题[J].中国图书馆学报,2017(4):4-17.
- [4] 茆意宏.人工智能重塑图书馆[J].大学图书馆学报,2018(2):11-17.
- [5] 张晓林.从猿到人:探索知识服务的凤凰涅槃之路[J].数据分析与知识发现,2023(3):1-4.
- [6] 刘文艳.面向专题领域的垂直信息服务模式研究[J].信息技术与信息化,2007(1):93-95.

肖荣荣 女,1980年生。硕士,副研究馆员。研究方向:智慧图书馆与知识服务。

郭晨晨 女,1985年生。硕士,馆员。研究方向:数据分析与数据研究。

(收稿日期:2023-12-31;责编:刘清扬。)