

Chat GPT 在档案知识服务中的价值及风险应对

许宏轩 杨 铮

摘 要 在档案知识服务领域,Chat GPT 的融合应用给档案馆服务带来无限生机。从 Chat GPT 对档案知识服务的应用价值出发,探究其对档案知识服务的有利影响与价值的实现,在此基础上发掘 Chat GPT 在档案知识服务工作中潜在的风险及问题,并根据潜在的问题提出应对的策略与办法。

关键词 Chat GPT 档案 档案知识服务 人工智能

中图分类号 G273.5 文献标识码 A 收稿日期 2023-09-18

★作者简介:许宏轩,广西民族大学管理学院,在读硕士研究生,研究方向为图书馆学;杨铮,广西民族大学管理学院,在读硕士研究生,研究方向为档案学。

Value and Risk Response of Chat GPT in Archival Knowledge Service

Xu Hongxuan, Yang Zheng

Abstract In the field of archival knowledge service, the integrated application of Chat GPT brings infinite vitality to archival service. This paper starts with the application value of Chat GPT in archival knowledge service, explores its beneficial impact on archival knowledge service and the realization of its value, and on this basis, excavates the potential risks and problems of Chat GPT in archival knowledge service work, and proposes countermeasures and methods according to the potential problems.

Keyword Chat GPT; archives; archival knowledge service; artificial intelligence

2023 年,Chat GPT 横空出世,标志着人工智能技术研究进入了全新的阶段。Chat GPT 的出现也引发了社会上诸多热议,各大媒体对其争相报道。Chat GPT 的应用像是一把“双刃剑”:一方面,Chat GPT 技术的出现提高了工作效率,优化了工作方式,极大地解放了社会生产力;另一方面,Chat GPT 在现实中可能产生一系列的法律风险、人员就业、安全隐私及伦理道德等问题,存在巨大风险与隐患。近年来,随着档案知识服务理念提出,传统的档案工作开始转型,Chat GPT 的出现为档案知识服务提供了极大便利。如何更好地将档案知识服务工作与 Chat GPT 结合,发掘其隐藏的价值的同时规避潜在的风险,是本研究的核心内容。

一、相关概念解析

1.Chat GPT。Chat GPT 是 Chat Generative Pre-trained Transformer 的简称,由美国公司 Open AI 发布的一款聊天机器人程序,本质上是人工智能技术驱动的自然语言处理工具。Chat GPT 使用基于 GPT-3.5 架构的大型语言模型并通过强化学习进行训练^[1]。Chat GPT 目前推广的最火热的功能即问答系统,以人机交

互方式为主进行一问一答交互。此外,Chat GPT 还具有语言翻译、自然语言及文本生成、智能服务等多个应用场景。

2.档案知识服务。随着时代的发展与技术的进步,档案信息服务正从单纯的提供文本数据转向知识层面。档案知识服务的理念最早由特里·库克在《电子文件与纸质文件观念:后保管及后现代主义社会信息与档案管理面临的一场革命》中提出。他指出,档案工作者身份应当从保管员向知识提供者转变。目前,认可度较高的档案知识服务的概念是徐拥军等人提出的:“档案知识服务是以档案资源为基础,以用户需求为中心,运用现代信息技术对数字档案信息进行加工,集中和整合显性知识资源,挖掘和提炼隐性知识资源,实现知识的转化和增值,达到为用户提供所需知识的目的。”^[2]提供档案知识服务的关键在于,如何从海量的档案资源中快速提取信息并转化为知识提供给用户利用。为实现高效服务,学者将机器学习、数据挖掘、知识库等新兴技术应用于档案知识服务中,如 ANDER-SON 等通过机器学习技术挖掘了有关控制论发展的历史档案材料,以帮助学者

更好地理解控制论的发展历程,并使得科学家和工程师们可以重用和获取相关数据^[9]。

3.Chat GPT 相关特征。

(1)技术特征。Chat GPT 主要包括 Transformer 模型架构、大规模语料库训练、自动回复系统、Fine-tuning 等核心技术特征。①Transformer 模型架构。Transformer 是一种广泛应用于自然语言处理领域的模型架构,由编码器和解码器两部分构成。Chat GPT 的实现是以 Transformer 模型架构为基础,通过多层自注意力机制和前向神经网络捕捉获取长距离的依赖关系。②大规模语料库训练。大规模语料库训练是指利用海量的文本数据进行机器学习或深度学习模型训练的过程。Chat GPT 通过大规模语料库对语言知识和规律进行训练,其主要流程包括数据准备、特征提取、模型训练、模型评估及优化等。③自动回复系统。自动回复系统是一种人工智能技术,通过计算机程序自动处理用户输入的文本内容,并根据用户提问给予相应回复。Chat GPT 通过自动回复系统模仿人类的语言语法及思维,根据用户输入的问题生成人性化的回答。④Fine-tuning。Fine-tuning 作为一种机器学习技术,主要应用于在已有训练好的模型基础上实施进一步的训练,来适应新的任务与场景。Chat GPT 通过 Fine-tuning 技术进行微调,来适应不同应用下的场景和任务,其主要步骤包括选择预训练的语言模型、准备训练数据、对模型进行微调及评估与调优。

(2)功能特征。Chat GPT 将得到广泛应用,并对人们的生产和生活产生深远影响,展现出前所未有的应用能力。目前,Chat GPT 的应用能力主要涵盖以下几个方面:①自然语言处理功能。Chat GPT 具有先进的自然语言处理技术,可以通过用户输入的问题迅速高效地生成相关回答。此外,其自然语言处理功能支持多种语言,包括英文、中文、日语、韩语等语种,可以根据其语言习惯做到真正的语义理解。②智能化问答。Chat GPT 拥有海量知识储备,其数据来源包括网页数据、书籍文本、百科内容等,并涵盖音乐、艺术、科学知识等多个领域的专业知识,能够满足用户不同的需求。③辅助决策。Chat GPT 能够对用户的日常生活和工作起到简单高效的决策辅助作用,如在工作中帮助用户编辑文案、设计代码、编辑修图,在日常生活中帮助用户完成餐厅预订、规划出行和医疗建议等。④个性化推荐。Chat GPT 能够根据用户的历史记录数据,分析用户的信息需求与个人喜好,进行个性化

的信息推送与建议。例如,在教学活动中,Chat GPT 可以根据不同学生的兴趣爱好为其提供感兴趣的知识和相关教育资源,满足不同个体的需要。

二、应用价值

1.提高工作效率,提升用户满意度。Chat GPT 在档案知识服务领域的应用助力档案馆工作人员解放双手,提高工作效率。在以往的档案知识服务中,由工作人员与用户进行电话、邮件或面对面沟通交流。对于用户来说,要付出的时间成本较多,便利程度较低,如最终经过咨询未能达到用户预期,便会降低用户的满意度;而对档案馆工作人员来说,传统的服务方式不仅加大了工作量,同时也因用户咨询时间的长短不同而无法保障工作效率。“用户至上”是服务行业、机构一直推崇并遵循的原则,而 Chat GPT 能恰到好处地满足这一要求。Chat GPT 包含智能客服应用场景,可 24 小时在线提供问答服务,既节约了用户的时间成本,也便于用户随时随地进行咨询和提问,在提供给用户所需的精准信息后,会使用户满意度不断提升。另外,此应用也减少了档案馆工作人员的工作量,从而提高工作效率。

2.拓宽服务范围,提升服务质量。档案开发与利用一直是档案界推行的服务方式,但在用户体验过程中还是更加偏向档案存取的方式。档案虽不断向数字化方向发展,但档案信息的数据挖掘、语义关联、智能问答等智能应用技术在档案知识服务过程中应用较少^[4],而 Chat GPT 正是改善这一薄弱环节的桥梁。Chat GPT 的语义理解分析、知识问答系统等都无疑加强了档案信息与用户之间的联系,使档案不再以存取为重心,从而拓宽了服务范围,使用户能够多方面地了解档案、应用档案。Chat GPT 的助力更是提升了服务质量,将更有效、更高质量的档案信息呈现给用户。

3.促进信息多元化获取。Chat GPT 包含信息分类应用场景,可对档案信息进行智能分类,利用计算机技术对档案信息资源进行加工,形成一种更易于网络用户接受的网络传播模式,提高了档案信息资源的利用率^[5]。此外,Chat GPT 在档案知识服务中的应用促进了其知识获取途径由检索式向生成式转变^[6]。档案馆可通过 Chat GPT 智能化的信息检索与生成能力,通过简单的对话或关键词输入,帮助用户在海量的数据中获取所需要的信息,剔除不相关的信息,避免了在不同数据库中反复检索的麻烦。

三、潜在风险

1.信息质量良莠不齐,数据存在偏差。Chat GPT 所提供的内容是根据所收集数据和算法机制生成的,这也导致所输出的信息质量良莠不齐。这些低质量信息包括存在事实性错误的知识,以及违反伦理道德、社会价值观的信息。虽然已经禁止回答关于宗教、犯罪、性别歧视及地域歧视等内容的问题,但研究实验表明,大部分 Chat GPT 使用者都经历过接收低质量信息的情况。尽管现阶段 Open AI 对于 Chat GPT 的这种错误进行积极纠正,但仍无法消除数据偏差问题^[7]。信息质量良莠不齐,用户获取的数据存在偏差,会给档案知识服务带来消极影响。这些弊端主要体现在如下几方面:一是档案具有真实记录性,如 Chat GPT 所生成的内容不实或存在偏差,导致档案知识服务失去了根本意义。二是档案知识的价值涵盖政治、经济、文化、教育及生态等多个领域,低质量的档案知识信息会影响社会的发展,盲目相信偏差数据会造成严重的经济损失。

2.隐私存在泄露风险,数据不够安全。自从实施档案开放利用以来,一直存在用户数字隐私保护问题。而 Chat GPT 作为新时代人工智能的产物,虽然其收集数据与分析数据的能力突出,但也存在数据不够安全、隐私泄露的风险。其存在的风险问题主要包括如下几方面:一是在档案馆应用各种场景与用户进行数据交换的过程中,用户提交的信息或数据有可能被 Chat GPT 收录并且输出给其他用户,从而产生个人隐私泄露风险。二是 Chat GPT 所生成内容的来源是否真实可靠、是否合法,仍有待商榷。三是 Chat GPT 会对用户访问信息及个人信息进行记录,包括姓名、电话号码、家庭住址、查询数据、日志数据等。这些个人隐私信息易被不法分子恶意传播,从而造成隐私泄露,给用户带来损失。

3.输出内容难溯源,数据缺乏透明度。Chat GPT 根据用户需求所输出的内容不同于以往用搜索引擎所检索出的信息。检索模式下获取的信息来源明确,多源自于期刊、报纸、数据库等,具有权威性与真实性,且检索到的为原始未加工的信息。Chat GPT 所输出的内容来源广泛,难以追溯。由于算法黑箱效应,其输出的内容是经历算法机制和神经网络加工处理过的信息,但是其处理过程与方式并未公开,数据缺乏透明度,权威性与可靠性难以令人信任^[8]。因此 Chat GPT 在知识服务过程中会带来不小的潜在风险。

4.缺乏人性化。作为人工智能的最新产物,Chat GPT 虽然具备强大的逻辑运算能力与广阔的知识储备,甚至可以模仿人类的风格进行对话,但其本质仍是冰冷的工具,并不具备人类的情感、价值观与自我意识,无法站在人类的视角和观点来考虑和解决问题,故而 Chat GPT 在档案知识服务工作中始终无法取代档案管理人员,其更多是作为一种工具来弥补或优化知识服务工作。一味地依赖 Chat GPT 的算法功能,由 Chat GPT 而不是档案管理人员来主导档案知识服务,会导致知识服务工作越发程序化格式化,缺乏人性化,无法很好地满足用户对于信息的需求。此外,过度依赖 Chat GPT 还会导致档案馆员的工作能力下降,逐渐成为 Chat GPT 的附庸。

四、问题应对

1.提升管理者辨别信息真伪的能力。虽然应用 Chat GPT 有很多积极效应,但不可完全依赖冰冷的技术。Chat GPT 目前仍在不断更新换代,仍有需要完善的地方,而在应用中也确实存在一些漏洞。Chat GPT 的数据库系统会自动筛选数据,但在与用户进行交互的过程中提供的信息是否完全正确、正向还有待商榷。这也给档案部门工作人员带来了挑战。第一,档案部门工作人员在使用过程中需要发现智能技术的漏洞,辨别 Chat GPT 所提供的信息是否准确,是否具有正向含义,而不是虚假信息和误导信息。第二,档案部门技术人员在使用 Chat GPT 进行档案知识服务时,应尽量避免系统的随机性,进行多次获取,进而过滤低质量信息,选择最适合的数据信息,来强化信息筛选功能。第三,要健全信息风险与失误纠察机制,确保用户所获取的信息是符合法律要求的高质量信息^[9]。

2.完善数据管理,加强隐私保护。任何事物都具有双重性,Chat GPT 虽然在很大程度上给人们带来了便利,但也在一定程度上“监视”着人们的隐私,尤其在档案利用方面。从开发者角度来说,AI 技术的发展需要进一步加强对于用户的隐私保护。从应用 Chat GPT 系统的档案馆角度来说,在注重工作便利的同时,也要加强档案数据库的建设,完善数据管理系统,将档案资源按照保密程度分级管理。《生成式人工智能服务管理办法(征求意见稿)》第四条(五)提道:“尊重他人合法利益,防止伤害他人身心健康,损害肖像权、名誉权和个人隐私,侵犯知识产权。禁止非法获取、披露、利用个人信息和隐私、商业秘密。”^[10]在智能技术不断发展的当下,在利用 Chat GPT 进

行档案知识服务的基础上,应进一步降低 Chat GPT 的数据安全与用户个人隐私泄露风险,强化数据管理与监控,保护用户个人隐私不被侵犯。

3.健全信息获取与利用方面的法律法规。现阶段虽然出台了关于人工智能的一系列法律法规,但因为 Chat GPT 功能的特殊性,现有的政策法规无法很好地规范 Chat GPT 的功能与用途。因此,需要健全现有的法律法规,确保 Chat GPT 使用的合规合理合法。Chat GPT 在档案知识服务工作中的法律规范应包括两个方面:一是建立完善的信息获取制度^[9]。确保 Chat GPT 的信息来源公开透明,获取方式正当合法。档案馆的机密档案与核心数据对 Chat GPT 设置访问权限,防止数据被不法分子利用,给国家社会造成危害。二是健全信息利用制度。在利用 Chat GPT 进行档案知识服务的过程中,应通过立法来确保对于信息的合法使用,补齐非法数据利用方面的法律短板,保障 Chat GPT 在不违反法律制度的前提下最大限度地推动档案知识服务的高速发展。

4.贯彻以人为本的核心服务理念。“以人为本”是服务型机构的核心准则。在档案馆知识服务中,要以满足用户需求为核心,帮助用户获取所需知识与信息,有针对性地满足不同用户群体的多元化需求,利用 Chat GPT 多种应用场景为用户提供个性化服务。贯彻以人为本的服务理念,需要从两方面入手:其一是始终以档案馆员为档案馆的核心与关键。在当下技术不断进步的时代,档案馆员应增强专业自信,努力提升专业素养,学会合理利用新兴技术,掌握并合理运用 Chat GPT 的技术与功能,利用其完成复杂烦琐的重复性工作,提升工作效率^[11]。其二,档案馆员应转变工作观念,更多地用自身的专业经验与思维等人工智能无法具备的功能,从用户的情感与需求角度出发帮助其解决实际问题。人工智能技术与档案馆员的存在不应该是对立的,而应该是互补的。档案馆员应合理利用 Chat GPT 等人工智能技术来更好地开展知识服务工作,从而实现 1+1>2 的效果。

参考文献

- [1]张海,刘畅,王东波,等.Chat GPT 用户使用意愿影响因素研究[J].情报理论与实践,2023,46(4):15-22.
- [2]徐拥军,周艳华,李刚.基于知识服务的档案管理模式的理论探索[J].档案学通讯,2011(2):24-28.
- [3]ANDERSON I B, PROMI C, HAMILTON I K, et al. The cybernetics thought collective project: Using computational methods to reveal intellectual context in archival material[C]/2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017: 2213-2218.
- [4]钱德凤.高校档案知识服务模式创新研究[J].浙江档案,2022(11):56-58.
- [5]谢维焱.档案信息多元化服务模式研究[J].科技创新与应用,2016(3):275.
- [6]游俊哲.Chat GPT 类生成式人工智能在科研场景中的应用风险与控制措施[J].情报理论与实践,2023,46(6):24-32.
- [7]OpenAI Used keyanal Workers on Less Than \$21 Per Hour. Exclusive Trime ([EB/OL]. [2023-03-05]. https://time.com/6247678/openai-chatgpt-kenya-workers/.
- [8]周旭.机遇与挑战:Chat GPT 普及背景下图书馆的应对分析[J].图书馆,2023(6):34-41+48.
- [9]国家互联网信息办公室.《生成式人工智能服务管理办法(征求意见稿)》公开征求意见的通知 [EB/OL]. [2023-04-11]. http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm
- [10]蔡士林,杨磊.Chat GPT 智能机器人应用的风险与协同治理研究[J].情报理论与实践,2023,46(5):14-22.
- [11]施志唐.Chat GPT 对图书馆的影响与应对措施[J].图书情报导刊,2023,8(4):1-7.