



Research on the Enhancement Paths of Scientific Researchers' Information Literacy under the Background of AIGC Technology

Hao Jianjun, Xu Hui* Hemuqi

(Library of Inner Mongolia University of Technology, Hohhot, Inner Mongolia 010051, China)
51179734@qq.com.

Abstract: With the rapid development of artificial intelligence technology, generative intelligent technologies such as ChatGPT are profoundly transforming the paradigm of scientific research. As an essential support institution for scientific research, libraries need to adjust and improve their service models for scientific researchers, offering more personalized and diversified services to meet their needs. Scientific researchers may sometimes lose focus and divulge confidential research information when using ChatGPT. Therefore, when utilizing ChatGPT and similar technologies, scientific researchers should enhance their information literacy to prevent risks. This paper explores multidimensional paths for enhancing scientific researchers' information literacy and puts forward corresponding suggestions.

Keywords: AIGC Technology, Information Literacy, Scientific Research Information, Knowledge Management, Risk Prevention

In the era of big data, the scale and types of data have proliferated, requiring scientific researchers to expend significant time and effort in processing vast amounts of information. AIGC (Artificial Intelligence Generated Content), represented by ChatGPT, addresses this issue effectively. AIGC can assist scientific researchers in tasks such as programming, literature reading, video and image generation, as well as precise analysis and processing of textual data, producing natural and fluent texts. [1] AIGC significantly improves researchers' efficiency, enabling them to focus more on scientific research and innovation. Currently, few scholars have systematically studied the impact of ChatGPT technology on scientific researchers from a library perspective. This paper delves into this topic from various angles, providing valuable insights for related research.

1 Overview of ChatGPT

ChatGPT is an advanced natural language processing tool driven by intelligent AI technology developed by OpenAI. It relies on a neural network model based on a deep learning framework, adopting the Transformer architecture. ChatGPT's efficient and precise dialogue response capabilities stem from its pre-training on large-scale text data, enabling it to generate coherent and logical responses, providing users with a smooth and natural communication experience.

ChatGPT's origins can be traced back to 2018, when OpenAI first introduced GPT (Generative Pre-trained Transformer), a language model based on recurrent neural networks, marking a new era in natural language processing. Subsequently, OpenAI continuously optimized and upgraded GPT, releasing GPT2 and GPT3, [2] offering more accurate, engaging, and practical dialogue experiences. ChatGPT, as OpenAI's latest model, underwent further pre-training and optimization based on previous versions, achieving an even higher level of natural language processing capabilities.

© The Author(s) 2024

C. Bai et al. (eds.), *Proceedings of 2023 China Science and Technology Information Resource Management and Service Annual Conference (COINFO2023)*, Advances in Economics, Business and Management Research 293,

https://doi.org/10.2991/978-94-6463-498-3_12

With the advancement of deep learning technology, ChatGPT's natural language processing abilities have continued to strengthen. Through extensive training, it can understand and generate human language, providing targeted responses through dialogue and intelligent retrieval. Nowadays, as AI technology progresses, corpora are constantly refined, personal digital literacy improves, and the efficiency of human-computer interaction is enhanced. ChatGPT is widely applied in fields such as intelligent consulting, education, and intelligent retrieval.

Scholars and research institutions have shown keen interest in exploring ChatGPT's value in natural language processing, AI technology development, and related applications. Researchers delve into the model's architecture, refine training methods, enhance ChatGPT's performance, and expand its application scope. This research not only promotes the researchers' own intellectual progress but also sparks new ideas, driving continuous innovation in generative AI technology and the emergence of new versions.

2 Characteristics of ChatGPT Utilization in Libraries

ChatGPT is an efficient, intelligent, and personalized language processing tool. By empowering libraries, ChatGPT is embedded into search interfaces and document databases, assisting users in generating comprehensive search queries, thereby enhancing literature recall rates. Through analyzing documents and presenting them in mind map formats, it facilitates users' understanding and saves time spent on searching and analyzing literature, thus improving literature precision.

2.1 Intelligence

Leveraging natural language learning and adaptive capabilities, ChatGPT possesses an initial level of intelligence akin to humans. It automatically answers user queries, comprehends, and generates natural language texts, mathematical formulas, etc., enhancing the efficiency and quality of library services.

2.2 Personalized Services

ChatGPT analyzes and mines users' historical behavior data, enabling it to provide precise book recommendations and research information services tailored to individual needs, achieving personalized and diversified services.

2.3 Efficiency

ChatGPT promptly responds to user requests with corresponding answers and suggestions, enhancing the speed and efficiency of library services. Users are advised to ask questions accurately or train ChatGPT through multiple queries for optimal results.

2.4 Discipline-Specific Services

ChatGPT meets users' knowledge demands round-the-clock and across time zones. Drawing from its vast knowledge base, it responds to queries from users of diverse disciplinary backgrounds, inspiring research ideas among scholars.

2.5 Scientific Management

Through deep data mining, analysis, and correlation, ChatGPT facilitates data integration and visualization. Libraries can leverage these insights for scientific management, including database procurement strategies, literature development strategies, shelf management, space management, reader management, and lending management.

3 Research on ChatGPT

Since its inception, numerous scholars have explored ChatGPT from various angles, including intelligent robotics, language models, applications, and ethical and privacy considerations.

3.1 Research on Intelligent Robotics

Intelligent robots possess perception, thinking, and action capabilities, enabling them to make decisions and execute tasks based on environmental changes. Research in this field encompasses machine learning, computer vision, and natural language processing. Tianyu Wu et al. have outlined ChatGPT's history, current status, and future developments, summarizing its core technologies.^[3] Qian Li et al. have analyzed ChatGPT's corpus, algorithms, and models, detailing technical aspects and component algorithms.^[4] Chu Xulong et al. used CiteSpace software to summarize the research status, hotspots, and trends in China's intelligent robotics field over the past 20 years.^[5]

3.2 Language Models

As a large language model based on Transformer, ChatGPT's research focuses on improving its architecture, including depth, width, and parameter count, to enhance performance. Researchers also design more effective pre-training tasks and utilize unsupervised learning to extract language patterns and rules from massive datasets. Zhang Kejun et al. have improved and optimized classic CBOW and skip-gram models.^[6] Zhao Lina et al. have studied multi-channel knowledge base push services using Python.^[7] Cai Jiacheng et al. have reviewed the basics and current research status of depth estimation.^[8] Li Tuohang discovered the root cause of low-quality absent keywords generated by deep models and proposed a joint model.^[9] Shi Tongyue and Wang Zhongqing introduced pre-trained language models based on Transformer, their "pre-training + fine-tuning" paradigm, related models and progress, as well as improvement methods.^[10]

3.3 Application Scenarios

ChatGPT has widespread applications, including online consultation, virtual assistants, academic research, and intelligent question-answering systems. Scholars have studied these scenarios and proposed solutions and improvements. Yang Jiezhuma and Wang Zhihui have reflected on the impact of AI on academic knowledge production, advocating for a paradigm shift in sports research.^[11] Wu Xiaolin and Xing Yifei have examined the opportunities and challenges AI poses to graduate education, emphasizing personalized training, scenario creation, and ideological guidance to mitigate risks and

promote creative teaching reforms. ^[12] Zhao Yang et al. have analyzed the opportunities and challenges AIGC brings to smart libraries, proposing a transformation framework, pathways, and challenges to advance AIGC's application in smart library construction. ^[13] Zheng Shilin et al. have analyzed ChatGPT's development opportunities and challenges across multiple domains. ^[14]

3.4 Ethics and Privacy Protection

With the continuous development of artificial intelligence technology, various issues have gradually emerged, including concerns over data privacy, data pollution, algorithmic bias, and human job displacement. Pu Qingping and Xiang Wang believe that ChatGPT has the potential to spark the Fourth Industrial Revolution, yet it also poses risks and hidden dangers, necessitating the strengthening of talent development, emphasizing practical application, prioritizing legislation, and reinforcing ideological guidance. ^[15] Meanwhile, Ling Xiaoxiong and others, from the perspectives of science and technology ethics as well as academic ethics, have conducted empirical research on ChatGPT's ability to think and respond, subjecting it to ethical scrutiny. ^[16] Sun Weiping argues that traditional human-machine relationships are being disrupted by AI technology, necessitating the reconstruction of a new type of human-machine relationship. ^[17] Zhang Xiaoheng acknowledges that ChatGPT has positive effects on human life but also stresses that the risks posed by the technology cannot be overlooked, advocating for measures to address these issues. ^[18]

4 The Impact of ChatGPT on Scientific Research Information Management

Generative artificial intelligence provides researchers with a novel tool for information acquisition and application. With its intelligent interactive mode, the acquisition and utilization of scientific research information have become more convenient and efficient.

4.1 Convenience in Information Acquisition

Researchers can engage in dialogues with ChatGPT to obtain a vast amount of scientific and technological information. By posing questions related to technology, research, projects, and other aspects to ChatGPT, researchers can receive detailed and accurate answers within seconds. If unsatisfied, researchers can repeatedly train ChatGPT until they obtain the desired answers or inspirations.

4.2 Knowledge Management

Libraries form thematic and disciplinary databases through re-recognition, re-discovery, and re-aggregation of resources. ChatGPT builds its own knowledge base on these databases. It autonomously understands and analyzes problems, connects them with solutions from its knowledge base, and presents relevant knowledge to researchers in textual form.

4.3 Overcoming Technical Bottlenecks

ChatGPT's extensive knowledge base and potential for application in various scenarios can provide technical support and solutions for researchers. In areas such as program editing, computation, data

collection, and processing, ChatGPT can assist researchers in completing tasks, thereby reducing the workload of technical personnel.

4.4 Information Service Capability

ChatGPT's round-the-clock availability, repeatability, and automation capabilities enhance information management efficiency, enabling timely and efficient responses to questions from team members, supervisory departments, and collaborating organizations.

4.5 Intelligence Information Service

Through its autonomous sensing capabilities, ChatGPT collects, understands, and analyzes scientific and technological information, forming a complete knowledge system and chain, which it recommends to pre-customized researchers. Based on this intelligence, researchers can grasp the demands of the technology market, promptly identify research topics, adjust research directions, collect relevant literature, and carry out scientific research. Simultaneously, researchers can collaborate with libraries to promptly push their research achievements into the market.

5. Transformation of Library Service Paradigms for Researchers

With the introduction of ChatGPT technology, libraries have witnessed changes in their service subjects, content, and modes. The emergence of intelligent machines has enabled seamless, round-the-clock intelligent knowledge services both online and offline.

5.1 Resource Search

Generative AI technologies, represented by ChatGPT, offer significant advantages over traditional search methods in terms of resource search efficiency, recall rate, and precision. When librarians use ChatGPT for resource searches, they input precise keywords or questions that convey their ideas. ChatGPT then quickly retrieves relevant literature based on the context, assisting researchers in analyzing the literature and presenting the results. Researchers can also quickly retrieve the materials they need through the library's online intelligent robot. The advancement of intelligent technology has undoubtedly transformed the paradigm of scientific research, allowing researchers to focus more on academic pursuits rather than spending considerable time searching, organizing, and analyzing information.

5.2 Personalized Recommendations

ChatGPT leverages user interaction data and customized information to recommend the latest literature, research trends, and achievements in specific fields. This enables researchers to promptly access the information they require. By embedding ChatGPT into library databases, libraries can better understand researchers' needs and preferences, providing detailed and precise personalized recommendations across various disciplines and research areas. The collaboration between librarians and ChatGPT transforms the paradigm of discipline-specific services, offering customized, timely, and accurate support for researchers' scientific endeavors. Furthermore, with the market information provided by

ChatGPT, libraries can facilitate the introduction of research achievements into the market, accelerating the speed and efficiency of technological commercialization, promoting industrial innovation and upgrading, and driving societal progress, thereby embodying the value of libraries.

5.3 Reference and Consultation Services

Libraries re-evaluate and reorganize their resources to form discipline-specific resources. The integration of ChatGPT technology with retrieval technologies and disciplinary resources leads to the creation of intelligent retrieval systems. Researchers frequently rely on library resources throughout the entire lifecycle of academic research. Now, libraries provide both online and offline intelligent retrieval services, assisting librarians in providing reference and consultation services. The intelligent system aids researchers in conducting deep analyses of information, summarizing abstracts of articles, and saving literature while generating citation formats.

6 Issues with ChatGPT

ChatGPT, being a large language model, excels in generating fluent language but lacks human-like thinking and judgment. As we navigate the world of ChatGPT, we may sometimes find ourselves lost, inadvertently divulging research secrets, or receiving false and inappropriate content. Therefore, researchers must remain vigilant and adopt necessary security measures to protect their research when using ChatGPT.

6.1 Losing Oneself

Researchers using ChatGPT may become overly reliant on it, losing their own thinking and judgment abilities, as well as their capacity to analyze problems and solve them. They may mistakenly consider ChatGPT's opinions as correct and use them as research materials or arguments. ChatGPT sometimes generates a vast amount of responses, leading to information overload, distortion, obsolescence, and anxiety among researchers. Handling, discerning, and digesting this information can be challenging, resulting in mental confusion, poor concentration, and difficult or erroneous decision-making. Thus, researchers must maintain independent thinking and judgment, using ChatGPT mindfully and within limits.

6.2 Leakage of Research Secrets

Since ChatGPT's servers are located abroad, researchers who have not adequately upgraded their digital literacy may inadvertently disclose research secrets while posing questions, causing severe consequences and losses for individuals, organizations, and even nations. Researchers must refrain from providing sensitive information or personal identifiers when using ChatGPT.

6.3 Inappropriate Content

ChatGPT may generate inappropriate or unsuitable content, necessitating researchers to verify, monitor, and filter information to ensure its accuracy and legality.

7 Improvement Strategies

Researchers are advised to use ChatGPT through official channels or leverage generative AI software developed by libraries to assist in research.

7.1 Interactive Preprocessing of Data

Before formal use, engage in a conversation with ChatGPT to let it understand you and your research field. Gradually expand the interaction, and ChatGPT will generate contextually relevant content tailored to your needs.

7.2 Knowledge Graph-Guided ChatGPT for User-Specific Knowledge Generation

Create knowledge graphs from large models and databases, integrating ChatGPT with APIs to form ChatGPT neurons, enhancing its knowledge base and analytical abilities.

7.3 Multimodal Data Training and Sensory Capability Development

Scientific research often presents findings through various modalities such as papers, images, physical objects, or videos. As research data may come from multimodal sources, ChatGPT's corpus must adapt to real-time sensing and collection of such data. Both unsupervised and supervised training should be employed to expand ChatGPT's knowledge base and sensory capabilities.

7.4 Real-Time Expansion of ChatGPT's Database

Given the diversity and real-time nature of research data, ChatGPT's auxiliary functions require continuous enrichment or training of its corpus to improve adaptability and accuracy.

7.5 Collection, Analysis, and Improvement Based on Feedback and Invalid Interactions

ChatGPT is currently still in its low-intelligence stage, far from achieving general-purpose intelligence. Sometimes, it may provide incorrect information or be unable to answer questions posed by researchers, leading to dissatisfaction among researchers who find that the information provided does not meet their needs. Recognizing the issues and limitations of ChatGPT, libraries should promptly take action to improve and optimize its knowledge base. This ongoing effort aims to enhance the quality and relevance of ChatGPT's responses, particularly in addressing the specialized inquiries of researchers.

8 The Path for Libraries to Enhance Researchers' Information Literacy

8.1 Guiding Researchers in Basic Information Knowledge

Through media like WeChat official accounts and short video platforms, libraries should educate researchers on computer science, AI, and information retrieval, covering topics like data structures, algorithms, programming languages, and search techniques. In the AIGC-driven intelligent era, knowledge updates rapidly, necessitating continuous learning for researchers. Libraries can organize

academic exchanges, lectures, and other events to encourage researchers to write papers, attend lectures, and maintain a learning mindset.

8.2 Developing AIGC Information Literacy Education

Libraries should research AIGC information literacy, exploring educational paths and methods that integrate literacy education with AIGC, particularly focusing on search construction, keyword selection, and content verification. Through experimentation, libraries should establish a suitable AIGC information literacy education model and timely conduct training and seminars. Researchers will gradually comprehend AIGC technology's principles, algorithms, and application scenarios through learning and research.

8.3 Conducting Multi-dimensional Training and Promotion of Data Literacy

Data, as a production factor, plays a crucial role in the field of scientific research. Researchers interact with AIGC, inputting or generating vast amounts of data, which involves data security education, data cleaning, data analysis and mining techniques, as well as visualization technologies. The education of data literacy falls within the scope of library education and can be integrated with information literacy education and grid information education. By doing so, libraries can provide researchers with data security education and information technology education. Additionally, libraries can utilize their platforms to conduct promotional activities and immersive experiential education on data literacy.

9 Conclusion

Since the advent of generative AI technologies represented by ChatGPT, people have actively explored the integration of technology and scenarios, continuously driving technological improvement and innovation through application. As a field for information technology practice and an educational sector for information technology literacy, libraries have the responsibility and obligation to promptly absorb, transform, and upgrade these technologies to provide comprehensive and efficient information services for scientific research. To make better use of ChatGPT, researchers need to strengthen their information literacy education and data security education. Libraries can offer relevant training courses and lectures to provide literacy education for researchers^[19].

References

1. Wang Shuyi, Zhang Qingwei: ChatGPT's Opportunities and Challenges for Researchers Library Tribune 43(3),109-118 (2023)
- 2 Blue Ocean Thoughts Sharing:What are GPT Series Products and Their Development History? , http://www.360doc.com/content/12/0121/07/71784469_1075004560.shtml,2023/4/4
- 3 Tianyu Wu,,Shizhu He,,Jingping Liu,,Siqi Sun,,Kang Liu,,Qing-Long Han & Yang Tang :A Brief Overview of ChatGPT:The History,Status Quo and Potential Future Development.IEEE/CAA Journal of Automatica Sinica10(5),1122-1136(2023)

- 4 Qian Li, Liu Yi, Zhang Zhixiong, et al: An Analysis on the Basic Technologies of ChatGPT. *Data Analysis and Knowledge Discovery* 7(3),6-15(2023)
- 5 Chu Xulong, Shi Dongmei, Liu Jinchang: The Hot Spot of Intelligent Robot Research in China in Recent 20 Years: Bibliometric Analysis Based on CiteSpace. *Science Technology and Engineering* 23(6),2477-2484(2023)
- 6 Zhang Kejun, Shi Taimeng, Li Weinan, Qian Rong: Word2Vec Optimization Strategy Based on an Improved Statistical Language Model. *Journal of Chinese Information Processing*33(7),11-19(2019)
- 7 Zhao Lina, Li Wei, Kang Wu, Zhang Kai: Python-Based Intelligent Robot Multi-Channel Knowledge Base Push Simulation. *Computer Simulation*37(3),328-332(2020)
- 8 Cai Jiacheng, Dong Fangmin, Sun Shuifa, et al.: Unsupervised Learning of Monocular Depth Estimation:A Survey. *Computer Science*51(02),117-134(2024)
- 9 Li Tuohang: Research on Automatic Keyphrase Technology inAcademic Corpus. Jilin University, (Accession Number: 10.27162/d.cnki.gjlin.2023.000003)(2023)
- 10 Shi Tongyue, Wang Zhongqing: Overview of Transformer-based Pre-trained Language Models for Natural Language Processing. *Information & Computer*34(10), 52-56 (2022)
- 11 Yang Jiezhuoma, Wang Zhihui: Seeking Change in Changes:Reflection on Sports Science in ChatGPT Era——A Qualitative Research on the Narration on “Academic Research on Sports in ChatGPT Era:Body as Method” . *Sports & Science* 44(2), 1-8 (2023)
- 12 Wu Xiaolin, Xing Yifei : Knowledge Replication or Innovation Stimulation?——Challenges and Opportunities of Artificial Intelligence(ChatGPT)for Social Science Graduate Education. *Journal of Guangxi Normal University: Philosophy and Social Sciences Edition*59(2), 66-74 (2023)
- 13 Zhao Yang, Zhang Xue, Fan Shengyue.: AIGC-driven Intelligent Library Transformation: Framework, Pathways and Challenges . *Information Studies: Theory & Application*,46(7), 9-16 (2023)
- 14 Zheng Shilin, Yao Shouyu, The economic and social impacts of ChatGPT next-generation AI technology development. *Review of Industrial Economics*(3), 5-21 (2023)
- 15 Pu Qingping, Xiang Wang: Opportunities and challenges aroused by ChatGPT as generative AI and strategy for response . *Journal of Chongqing University (Social Sciences Edition)*29(3), 102-114 (2023)
- 16 Ling Xiaoxiong, Wang Dingmin, Yuan Jian: Reflection on Technology Ethics and Academic Ethics in the Context of ChatGPT Craze. *Journal of Xinjiang Normal University (Philosophy and Social Sciences Edition)*44(4), 123-136 (2023)
- 17 Sun Weiping: The New Human-Machine Relationship in the Age of Intelligence and Its Construction. *Journal of Hubei University (Philosophy and Social Sciences)*50(3), 18-25, 168 (2023)

18 Zhang Xiaheng.: New Generation of Artificial Intelligence Technology (ChatGPT) and Its Impact and Transformation on Human Society. Review of Industrial Economics, (2023)

19 Lu Yuquan: Study on the service mode and construction path of university scientific research data in the era of big data. Nanjing University of Aeronautics and Astronautics, (2017)

Inner Mongolia Autonomous Region Higher Education Research Project.

《Research on the Construction of the Classic Reading Promotion System in Libraries in Ethnic Minority Areas from the Perspective of Cultural Inheritance》 Project Approval Number: NJSY21310

Hao Jianjun was born in July 1975, male, associate research librarian, specializing in library science, information science, and information literacy research.

Address: 49 Aimin Street, Xincheng District, Hohhot City, Inner Mongolia Autonomous Region, Postcode: 010051. Phone: 13789414762, Email: 393063246@qq.com.

Contact Person: Xu Hui, born in February 1979, female, associate research librarian, specializing in information management research. Phone: 13500693838, Email: 51179734@qq.com.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

