

Using Artificial Intelligence (AI) in Improving the Accuracy of Electronic Sources Indexing

استخدام الذكاء الاصطناعي في تحسين دقة فهرسة المصادر الالكترونية

منى حازم

Muna Hazim

Information Technology Management Technical College of
Management-Baghdad Middle Technical University Baghdad,
Iraq

0009-0002-1493-7622

mhy@mtu.edu.iq

الملخص

الاهداف

تبحث هذه الدراسة في تطوير تقنيات الذكاء الاصطناعي التي تعزز دقة وكفاءة فهرسة المصادر الإلكترونية، مما يُحسِّن استرجاع المعلومات وإدارتها في البيئات الرقمية. **مشكلة البحث:** تتبثق مشكلة البحث من نقاط الضعف والقيود المفروضة على الفهرسة الإلكترونية التقليدية، والتي كانت دائماً يدوياً أو شبه تلقائي. لقد كانت الفهرسة غير الدقيقة وغير المكتملة والبطيئة للموارد الرقمية مشكلة شائعة في هذه الورقة، حيث تحاول توضيح كيفية تحسين الدقة والاتساق والفعالية الكلية في فهرسة المصدر الإلكتروني من جودة الاسترجاع عن طريق الاستفادة من تقنيات الذكاء الاصطناعي. لذلك، يؤدي تحسين جودة الاسترجاع إلى إمكانية وصول أفضل وتجربة مستخدم جيدة في بيئات المعلومات الرقمية. **المنهجية:** منهج البحث النوعي: تم اختيار عينة طبقية لدراسة حالة تتعلق بكفاءة ودقة فهرسة الموارد الإلكترونية باستخدام الذكاء الاصطناعي. شملت الدراسة مقابلات مع 32 خبيراً (أمناء مكتبات، باحثين، ومطوري ذكاء اصطناعي).

جمع البيانات:

ركزت الدراسة على سبعة محاور لتقييم تأثير الذكاء الاصطناعي على دقة الفهرسة.

النتائج:

يُحسِّن الذكاء الاصطناعي بشكل كبير دقة فهرسة المصادر الإلكترونية. تدعم النتائج نظرية نظم المعلومات التي تؤكد أن التقدم التكنولوجي يمكن أن يعزز عمليات إدارة المعلومات في المؤسسات.

الاستنتاجات:

يلعب الذكاء الاصطناعي دوراً إيجابياً وفعالاً في تحسين دقة وكفاءة فهرسة المصادر الإلكترونية. وتعتبر الابتكارات التكنولوجية مثل الذكاء الاصطناعي أدوات قيمة لتطوير إدارة المعلومات الرقمية. **الكلمات الدالة:** الذكاء الاصطناعي، الدقة، المصادر الإلكترونية، الفهرسة، إطار نظم المعلومات.

Abstract

Objective: This study investigates and develops artificial intelligence (AI) techniques that enhance the accuracy and efficiency of indexing electronic sources, thereby improving information retrieval and management in digital environments.

Research Problem: It stems from the weaknesses and limitations of traditional electronic indexing, which has always been manual or semi-automatic. Inaccurate, incomplete, and slow indexing of digital resources has been a common issue in this paper, as it attempts to demonstrate how improved accuracy, consistency, and overall effectiveness in electronic source indexing can enhance retrieval quality by leveraging AI Technologies. Therefore, improved retrieval quality leads to better accessibility and a good user experience in digital information environments.

Methods: Qualitative research methodology.

A stratified sample was selected for a case study related to the efficiency and accuracy of indexing electronic resources using artificial intelligence. The study consisted of Interviews with 32 experts (librarians, researchers, and AI developers).

Data Collection: It focused on seven themes, assessing AI's impact on indexing accuracy.

Results: Artificial intelligence has significantly improved the accuracy of electronic resource indexing. Findings support the Information Systems (IS) theory that technological advancements can enhance organizational information management processes.

Conclusions: Artificial intelligence has made significant contributions to improving the accuracy and efficiency of electronic resource indexing. Technological innovations such as artificial intelligence are valuable tools for advancing digital information management.

Keywords: Accuracy, artificial intelligence, indexing, electronic sources, information systems framework.

Introduction

Artificial intelligence application has already had an essential impact on many things in the world, including the management and retrieval of information. According to Wu et al., indexing is one of the most crucial areas, which currently presumes manual sources of information ahead. As Fagbolla mentions, the indexing is the first step of being able to access and manage information that comes from the vast digital repositories. The actual manual way of integrating this information into faster means and more appropriate ways of learning require innovative solutions. As it has been mentioned by Wang et al., the AI technologies like natural language processing and machine learning can be applied in the actual indexing process. This in turn not only reduces time spent on wasteful manual operations but also reduces the number of mistakes. These technologies can act based on the detailed pattern analysis of all brittle digital content and classify the information, which can be done with the highest accuracy for proof. Artificial intelligence can be interpreted as software or computers developed by a human that can be able to perceive their environment, improve their experience, reason, and interpret data to perform a set of complex operations and tasks further making decisions or actions to reach a specific goal. The most developed AI is capable of executing human tasks, which require intelligence such as problem-solving, understanding of language, recognition of patterns and adapting to new situations, sometimes coming with unintentional autonomy. Many organizations, according to Oloruntoba, are already using the AI technologies to make electronic indexing resources more reliable. This way, the search process is sped up, and the right access to needed and accurate information enables the right people to reach informed decisions on time. This paper is intended to research and explain the theory of AI and its real-life applications. In order to discuss how it can help users, as well as further specify the level of accuracy it is capable of reaching, the paper will contain the analysis of the theory of AI and its elaboration in the medium.

Artificial intelligence plays a significant role in enhancing the accuracy of electronic resources. Schuett (2019) defines "artificial intelligence" as the specific ability of computers to solve problems, make decisions, be creative, and use technology autonomously. Dobrev (2012) defines artificial intelligence as a scientific field dedicated to developing computers and machines that can think, learn, and act like

humans, or process data beyond human comprehension. According to Collins (2024), the indexing of electronic resources is the systematic arrangement and classification of digital information materials to ensure proper organization and ease of use and location. These components represent organized entries that are usually associated with metadata such as keywords, titles, authors, abstracts, structures, destinations, itineraries, and even agents, which are useful pieces of information that can be used to find a location.

This research proposes an AI-based application integrating machine learning and natural language processing technologies to improve the indexing process. This method allows for the creation of more efficient and accurate information systems while addressing linguistic complexity and human error. Although AI technologies have been widely applied in various information-processing systems, a significant theoretical gap remains in understanding their influence on the indexing of electronic resources. Existing studies on indexing vocabularies mostly focus on traditional contexts or general understandings of AI, without systematically defining their specific applications in indexing.

Consequently, it is still unclear how modern AI technologies—such as natural language processing and machine learning algorithms—can be applied to enhance electronic resource indexing. No comprehensive framework has yet been established to demonstrate how AI can improve semantic disambiguation, context update speed, or relevance aggregation, thereby increasing the precision of search results in digital libraries and databases. This research aims to fill this gap by showing how AI-based approaches to electronic resource design and development can achieve maximum precision and effectiveness.

Problem Statement

The problem statement is that the accuracy of indexing electronic sources is still quite an issue. It occurs because the digital domain is quite vast and diverse, and conventional methods of semantic indexing and relevance to context cannot be fruitfully applied thereto. It seems that such a scenario may be developed with the help of AI, but the use of AI technologies for indexing electronic sources has not yet

been properly developed. This research attempts to answer the following question: how can AI be used to productive ends while improving the accuracy and efficacy of indexing electronic sources.

Research Objective and Question

The aim of the research is to develop and implement an artificial intelligence techniques language system, which may help to improve the accuracy and efficiency of indexing electronic sources. Taking into consideration the current state of digital library technology, as well as information retrieval and management, the research problem is as follows: “To what extent may AI assist in improving the indexing of electronic sources?”

Research Significance

The purpose of this research is to facilitate electronic source indexing, as this process faces many challenges that are difficult to overcome using manual methods with low precision and productivity, as well as traditional methods, which have become inefficient as the growing volume of digital content has largely surpassed their capacity. In reality, the challenge in this process arises from the extremely high semantic complexity of the content in question, which can exist in a virtually unlimited number of contexts, making it difficult to retrieve promptly. Improving knowledge organization through artificial intelligence and related technologies is especially important in this case to ensure higher retrieval precision for researchers. Such technologies help libraries and information sources meet the increasingly difficult challenge of enabling users to discover resources. This step is intended to contribute to the practical development of innovative processes in indexing and to meeting the challenges of the digital age.

Literature Review

Theoretical Framework

The theoretical framework of this research includes a literature review and several theories concerning the use of artificial intelligence to increase the accuracy of electronic source indexing. When discussing this theoretical research on artificial source indexing and AI, this study highlights the most recent breakthroughs in the

fields of machine learning, data structuring, and knowledge organization. According to Izzard (2024), AI-powered indexing has a fundamental advantage through the ability of machine learning algorithms to detect data regularities or patterns. These patterns allow for the use of advanced semantic-based search techniques, where intent is more important than mere word matching.

However, as Zhao (2021) states, these rules are not reflected in current state-of-the-art Transformer-based natural language processing architectures, as they can detect the intent and meaning encoded in text and subtext, which are generally ignored by rule-based methods. Additionally, Izzard (2024) uses these Transformer-based architectures to draw term-level meaning from large amounts of data, which, according to the author, is particularly effective in increasing the relevance of metadata.

Moreover, Kaur (2018) explains that this approach represents a complete paradigm shift from hash tables and static B-trees; learned indexes use machine learning models to predict data locations in database table structures. These models take advantage of distribution regularities and allow, in the long term, for searches with sub-logarithmic time complexity. The advantage of these models is that they reduce storage overhead and increase search efficiency in read-intensive scenarios. Additionally, Molina et al. (2024) emphasize that this theoretical framework draws on the cooperation between artificial intelligence and human intelligence, creating a manageable equilibrium between assisted automation and human input to address misapprehensions such as "hallucinations".

According to Gani et al. (2016), Model collapse is the term used for when AI-generated results are used to train AI. These can lower the data diversity and make the information less reliable. This is one of the indications as to why the system should be trained with high-quality, human complied data. To avoid model collapse, incremental learning and hybrid methods are suggested. Chen and Li's et al. research indicated that AI-generated metadata is far from that of humans in most cases since they usually deal with multi-lingual or unstructured content. Several theoretical insights and case studies can be found in the AutoSE project. Kasprzik (2024)

explains that ZBW's automatic indexing system combines the open source machine learning toolkit Annif with human verification.

Regarding Microsoft's hybrid approach, Vernadakis et al. (2011) note that by using query execution data to refine index recommendations, researchers have overcome optimizer estimation errors and call for empirical validation in AI-driven optimization. Ethical and bias mitigation frameworks could be developed to examine metadata for cultural or representational biases. The underlying theory for the potential trade-off between innovation and robustness in AI indexing suggests that learning and evolution are relevant to the development of systems that blend technological and human capabilities.

The theory most applicable to this study is information systems (IS). This is a framework for AI research proposed by Collins, Dennehy, Conboy, and Mikalef (2021). The connection of AI practice with information systems research was established in the article (Collins et al., 2021). The framework developed systematically connects AI practice with information systems research and provides a structured approach for understanding how AI technologies can be integrated to enhance tasks handling related to information processing, such as the electronic source indexing. Specifically, the framework discusses the role of AI in automating the indexing process and improving the accuracy of electronic source indexing. The framework categorizes AI applications and supports researchers in identifying appropriate AI technologies for a variety of information activities, from semantic analysis to metadata generation (Bawack et al., 2021).

Existing conceptual gaps in electronic source indexing are overly dependent on traditional algorithms that do not maximize the advantages of AI. Despite AI adoption in related disciplines, little is known about how its techniques can be newly and concretely conceptualized toward increasingly accurate indexing. The existing models hardly account for the dynamic and complex nature of electronic sources as results of their semantic nuances and contextual relevancy. This research is intended to bridge the conceptual gap by designing a new framework for AI methods integration toward increased precision and adaptability of electronic source indexing.

Previous Studies

The following previous studies aim to clarify the content of this top. The work by Jain and Mitra (2025) towards scaling artificial intelligence with application to agility, accuracy, and retrieval in a records management setting is reviewed in this study. This research used a multiple-case study to verify an actual application for scaling, accuracy, and retrieval speed with artificial intelligence in the systems of records management. The study collected primary data from healthcare, governance, and government organizations.

It is proven at an empirical level that artificial intelligence via ML & NLP is some pertinent techniques that add value to RMS by automating indexing, context-aware retrieval, and general processing times. In addition, it improved end-user satisfaction as their tasks were reduced and made easier without manual effort. Poor data quality and system integration and ethical challenges were found, and thus the need for customized implementation strategies. This, in turn, adds to the pool of existing research on the role of AI in data management while it informs companies moving towards an AI-based RMS on how to attain efficiency and enforce regulations.

Musa et al. (2021) conducted a study on artificial intelligence and machine learning in oncology: historical overview of papers indexed in the Web of Science database. Artificial Intelligence (AI) and Machine Learning (ML) are innovations that help with the diagnosis and treatment of cancer. This paper presents an overview of the applications of AI and ML in oncology research. Data has been taken from the Web of Science. Bibliometric and R packages and VOS viewer software helped in mapping and network analysis. Results show that 214 publications were written by 1161 authors and published in 133 journals from 1988 to 2021 there has been a trend of increasing research between these years. AI and ML in oncology research have drawn the attention of the authors and the readership. The top documents are 173 citations. It includes hot topics about common errors in diagnostic categories in clinical and the potential future opportunities for precision oncology using AI. Aneja S and Thompson RF from the USA are the authors of the highest productivity. *Frontiers in Oncology* remains the most productive journal. Leading the research effort on the topic is the United States, followed by Korea. The collaboration and

network between these countries in AL or ML in oncology research was documented. Therefore, AI and ML in oncology research should draw interest from the scientific community and readership. The trend of research has been steadily increasing on a global scale.

Hameed et al. (2017) carried out a study on the use of artificial intelligence (AI) techniques in water quality index to manage AI techniques in predicting water quality index. The river water quality management is one of the most important environmental problems. The water quality index (WQI) describes several water quality variables at a particular aquatic environment and time. Normally, WQI is most popularly computed using the conventional method which is time-consuming, and sometimes there are subindex base accidental errors, and the runs are long. Thus, there is a great need to provide a WQI model that is precise. In the last few years, artificial neural networks (ANNs) were tested for the same type of prediction applications for prediction applications and appeared quite well in capturing the nonlinearity pattern between the predictors and predict and. Two different ANN algorithms, that is radial basis function neural network (RBFNN) and back propagation neural network models, have been used to study and develop the relationship for simulation of WQI based on the water quality parameters in a tropical environment (Malaysia). The input variables are of two different architectures and are examined. In addition, a full analysis of the performance evaluation and sensitivity analysis of the variables was carried out. The obtained results depicted high-performing accuracy in both scenarios in the model of RBFNN. Furthermore, the proposed approach provides a simple way to calculate and predict WQI since the manual calculation of WQI needs lengthy computations, transformations, and different sub index formulae for each observation of the measured water quality variables, and is time-consuming.

Wu et al. (2015) carried out a study of Citeseerx: AI in a digital library search engine. CiteSeerX is a digital library search engine that gives access to more than five million scholarly documents, with almost a million users and millions of hits per day. We describe here the key AI technologies used for document classification and de-duplication, along with document and citation clustering, automatic metadata

extraction and indexing, and author disambiguation. These AI technologies were developed by members of the CiteSeerX group over the past 5-6 years. We describe the usage status, payoff, development challenges, main design concepts, and deployment and maintenance requirements. AI technologies are introduced in table and algorithm search, which are special search modes in CiteSeerX. While it would be difficult to re-engineer a system like CiteSeerX from the foundation up, many of these AI technologies are portable to other digital libraries and/or search engines.

Research Methodology

The methodology of this study includes research design, data collection procedures, research tool, and data analysis. This study applies a qualitative approach. It investigates the artificial intelligence in improving the accuracy of electronic sources indexing. According to Moser and Korstjens (2017), qualitative approach is the type of research that offers and clarifies deeper insights into real-world problems. Rather than collecting numbers data points or intervening treatments as in quantitative research, which help create a hypothesis that can further be explored with more detail but where a specific phenomenon should be measured, how many Qualitative research collects experiences, perceptions, and behavior on participants.

The population of this study includes a collection of case studies related to this topic, database, interview questions for a certain number of participants, while the sample of this study consists of 32 certain experts like librarians, researchers, and AI developers who have experience with AI-enhanced indexing systems. The participants are interviewed by the researcher about some main themes that relate to the content of this research (These main themes are stated in Appendix A).

The sampling technique used in the study is the Stratified Sampling Method. complexities, qualitative research investigates the unquantifiable meanings and complications of human experience, behavior, and social phenomena. It is applied in natural settings and uses verbal, not literal, methods such as in-depth interviews, observations, and focus groups. The particular significant descriptive knowledge is considered to be a description of the participants' aspects, social context, and an inductive interpretation of the patterns and patterns with their concepts. Qualitative analysis is versatile should be versatile, subjective, and reflective, recognizing that

reality is socially constructed and open to diverse interpretations. The aim is to have a sample that properly represents the population to enable subgroup research.

The interview themes are first judged by a jury of doctors and experts who have the same major to be sure that these themes cover the main objective of this research. Then, the researcher can meet the participants of this study by using social media applications such as WhatsApp, Zoom, or Telegram. Then, he can discuss these themes based on experiences in their workplaces. After that, he can summarize their responses and compare them to the literature reviews, previous studies, and theoreticians' comments and points of view.

Research Tools

The research tools represent literature reviews, a collection of case studies related to this topic, database interview, and themes for conducting discussion by certain experts like librarians, researchers, and AI developers who have experience with AI-enhanced indexing systems. The researcher constructed 7 themes that refer to the content of this topic. They cover a part of the research question. These 7 themes (**See appendix A at the end of this research.**) are on using artificial intelligence in improving the accuracy of electronic sources indexing.

Data Collections

The interview themes were distributed to the participants by using the social media particularly WhatsApp and Zoom. After discussing these themes, the researcher asked them to send their responses that are discussed and analyzed carefully. This is often referred to as a free interview and is typically applied in qualitative research. The present study also chooses individual interviews using the semi-structured interview method. Gelo et al. (2008) states that in this kind of interview the researcher prepares a set of themes and can change their sequence, delete some of them, or add others to them based on the flow of the interview and the information collected. Therefore, the researcher derives the meaning of what he hears and notes from the in-depth interviews. He then makes guesses in the light of those notes which later develop into some facts that he works on to confirm or deny through

the rest of the information in subsequent interviews and finally comes up with explanations and results.

Study Procedures and Data Analysis

This topic explores the artificial intelligence in enhancing the accuracy of electronic sources indexing. Thematic analysis proves to be the most appropriate subjective data analysis approach for this study. Riger and Sigurvinsdottir (2016) describe thematic analysis as a technique of evaluating qualitative data that involves reading the data and attempting to locate patterns within the meaning of the data to learn about topics.

Braun and Clarke (2012) believe thematic analysis can apply to qualitative data and works toward detecting patterns and topics. Therefore, the present study conducts a thematic analysis of the interview transcripts, and any other qualitative data that may be collected is read to get an understanding of the content and to identify initial patterns. This pertains to themes on artificial intelligence in enhancing the accuracy of indexing electronic sources. After identifying the themes and subthemes that should be discussed with the respondents, the researcher reviews their conversation and reflections. At this point, cross-checking to ensure that all the items listed under the themes are, in fact, themes, becomes possible. "Themes can be cross-supported" with evidence. This is how the present study moves one step further by 'writing up' the analysis, synthesizing findings through the use of vivid extracts from the data that eventually congeal to underline and expound upon the subject. Themes, therefore, can guide the investigation and present clear ideas on the subject. Eventually, a paper can be written. Riger and Sigurvinsdottir (2016) explain that a thematic analysis paper comprises an "introduction, approach, what comes out, and the output." In drafting the report, the researcher must be able to give all details necessary for the reader to be able to appraise the findings rich, complex, and intricate description. The flexibility of thematic analysis allows an opportunity for the researcher to describe the data in a very rich, very intricate, and very sophisticated way.

Research Results

The most important finding of the study is that AI has made a positive contribution to improving the accuracy of electronic resource indexing. Interviews

with experts such as librarians, researchers, and AI developers who have experience with AI-assisted indexing systems showed that more than 85% of respondents confirmed that AI has made a significant contribution to improving the accuracy of electronic resource indexing.

Due to the huge amount of data, non-text content (such as images or videos), and many context-dependent decisions, human indexers cannot maintain indexing capabilities, As it is done on an item-to-item basis, this results in an inability to organize and analyze entire datasets. Similarly, ontological databases are pursued for increased accuracy, though they are still highly labor-intensive. These tools require manual checklists that are difficult to develop and inflexible in application. This is a significant drawback compared to current parsing tools, which can sort any set of documents without additional input once deployed properly.

Data patterns and contextual relationships are learned by machine learning models and Transformer-based architectures to address these deficiencies. Machine learning index structures that predict where a record is located by (e.g., by replacing B-trees) will optimize performance in the future if data distributions are well defined. Transformers understand the semantics of terms and big language models and semantic search tools, such as Elastic search, have them. As a result, the index is analyzed in a dynamic and contextually informed way based on the meanings of words and not through a comparison of them.

AI can also take some of the tiresome maintenance and parameter-optimization work out of the equation without having to rely heavily on a fixed set of rules and terminology. It is possible for AI to process terabytes of data in hours, something that human indexers are unable to do. However, training and operations still require resources, and there are also setup costs to consider. AI can be trained continuously on changing datasets and updates indexes on an ongoing basis to ensure sustainability. The downside is that accuracy will vary. That said, AI is still a must-have technology for big data applications.

Here are some examples regarding the interview themes:

Discuss the limitations and typical error rates of manual or rule-based indexing approaches before AI implementation against the baseline accuracy of traditional indexing methods.

. Discuss the specific artificial intelligence technologies applied to indexing electronic sources.

. Examine measurable changes such as metrics (error reduction rates, recall, and precision scores) in index accuracy change after AI adoption.

. Assess AI's ability to manage ambiguous terms, context-dependent keywords, synonyms, and polysemy in electronic sources; in contrast with traditional methods.

Discussion

It should be noted that artificial intelligence significantly affects the improvement of the resource indexing accuracy. It is supported by the results of the study conducted by Musa et al. (2021)., as compared to which, AI-based systems of indexing show much higher accuracy, especially when it is expanded to the indexing and clasterization of non-textual information. Much in contrast, the traditional use manual interference or the application of static taxonomies.

One example of such a system is created for more than 5 million papers, using machine learning to generate metadata and perform document clustering in order to make a quick and effective of diverse scientific works called CiteSeerX. which applies machine learning to extract metadata and perform document clustering from more than 5 million scientific papers to efficiently index diverse scholarly content.

This is consistent with the findings of Amer and Elboghdadly (2024). Especially in the biomedical literature, fully automated indexing can lead to unacceptably low retrieval performance. However, probabilistic logic combined with a small amount of human control can improve retrieval accuracy to 95%. This result is consistent with the findings of Wagner et al. (2022). The huge benefits of scalability are huge: AI reduces the manual indexing workload for multimedia archives by 70% to 80%. According to Schellnack-Kelly and Modiba (2024), Microsoft's AI-based Index Tuner achieves just that, as it replaces the error-prone cost estimator with a machine learning classifier.

AI is therefore extremely useful in processing big data, although research by Oyighan et al. (2024) warns against over-reliance on AI. Automated listings of open-access books have proven to be quite weak in terms of human-computer interaction, highlighting the need for human-computer interaction. Research by Collins et al. (2021) in the context of information systems (IS) shows a significant correlation between AI and accuracy in indexing electronic resources. This supports information systems theory, which posits that technological advances can improve an organization's information management processes. However, this overall recommendation goes a step further and affirms that the integration of AI partners can maximize the effective capabilities and efficiency of information systems.

Conclusion

The article contributes to the current knowledge by demonstrating that using sound artificial intelligence techniques allows achieving a high degree of accuracy and efficiency in indexing electronic resources. Such benefits are likely to address the common deficiencies related to manual and rule-based mechanisms or fit-for-purpose approaches. Despite the reasonableness of traditional algorithms in very particular cases, it is inevitable to acknowledge the limited ability of archivists to apply indices on a larger scale and adapt them to emerging demands and needs, including new terms and data diversity. The importance and application of artificial intelligence techniques, and in particular algorithms, as well as the machine learning frameworks associated with the introduced model, are justified by their understanding of contextual relations, relevant index structures, and source-specific needs.

As a result, the use of artificial intelligence in indexing is proven to elevate the overall precision of indexing, the pace of the process, and to relieve a massive human burden associated with controlling massive digital resources as part of such activities. Whereas the benefits in precision can differ depending on the environment and usage, clear advantages in terms of speed and scalability prove the transformative nature of artificial intelligence tools and techniques in the context of information retrieval. Such acknowledgments are crucial for a better understanding of the problem, as well as further research proposals and initiatives aimed at elaborating and mainstreaming new artificial intelligence-based resources and solutions.

References

1. Amer, E., & Elboghdadly, T. (2024, November). The End of the Search Engine Era and the Rise of Generative AI: A Paradigm Shift in Information Retrieval. In 2024 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (pp. 374-379). IEEE.
2. Bawack, R. E., Fosso Wamba, S., & Carillo, K. D. A. (2021). A framework for understanding artificial intelligence research: insights from practice. *Journal of Enterprise Information Management*, 34(2), 645-678.
3. Braun, V. & Clarke, V. (2013) Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist*, 26(2), 120-123.
4. Chen, S., & Li, M. (2024). AI for Cataloging and Metadata Creation: Perspectives and Future Opportunities from Cataloging and Metadata Professionals. *Technical Services Quarterly*, 41(4), 317-332.
5. Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60, 102383.
6. Collins, A. M. (2024). *Bibliographic and Indexing Sources. Reference and Information Services: An Introduction*, 355.
7. Dobrev, D. (2012). A definition of artificial intelligence. arXiv preprint arXiv:1210.1568.
8. Fagbola, O. O. (2018). Indexing and Abstracting as Tools for Information Retrieval in Digital Libraries: A Review of Literature. *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications*, 905-927.
9. Gani, A., Siddiqa, A., Shamshirband, S., & Hanum, F. (2016). A survey on indexing techniques for big data: taxonomy and performance evaluation. *Knowledge and information systems*, 46, 241-284.
10. Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative psychological and behavioral science*, 42, 266-290

11. Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., & Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications*, 28, 893-905.
12. Intelligence, A. (2020). Artificial intelligence (AI). How Does AI Work.
13. Jain, V., & Mitra, A. (2025). Leveraging Artificial Intelligence for Enhanced Scalability, Accuracy, and Retrieval in Records Management Systems: A Case Study Approach. In *Artificial Intelligence in Records and Information Management* (pp. 367-392). IGI Global Scientific Publishing.
14. Izzard, T. (2024). Generative artificial intelligence (AI) and its performance at indexing tasks. *The Indexer*, 42(4), 383-400.
15. Kasprzik, A. (2024). The Automation of Subject Indexing at ZBW and the Role of Metadata in Times of Large Language Models. *Procedia Computer Science*, 249, 160-166.
16. Kaur, T. (2018). Learned Index Structures: Practical Implementations and Future Directions.
17. Molina, D. A., Kharlov, V., & Chen, J. S. (2024, August). Towards Effective Human-AI Collaboration in Decision-Making: A Comprehensive Review and Conceptual Framework. In *2024 Portland International Conference on Management of Engineering and Technology (PICMET)* (pp. 1-6). IEEE.
18. Moser, A., & Korstjens, I. (2017). Series: Practical guidance to qualitative research. Part 1: Introduction. *European Journal of General Practice*, 23(1), 271-273.
19. Musa, I. H., Zamit, I., Okeke, M., Akintunde, T., & Musa, T. H. (2021). Artificial intelligence and machine learning in oncology: historical overview of documents indexed in the web of science database. *Eurasian Journal of Medicine and Oncology*, 5(3), 239.
20. Oloruntoba, N. O. (2025). AI-Driven autonomous database management: Self-tuning, predictive query optimization, and intelligent indexing in enterprise it environments. *World Journal of Advanced Research and Reviews*, 25(2), 1558-1580.

21. Oyighan, D., Ukubeyinje, E. S., David-West, B. T., & Oladokun, B. D. (2024). The role of AI in transforming metadata management: Insights on challenges, opportunities, and emerging trends. *Asian Journal of Information Science and Technology*, 14(2), 20-26.
22. Riger, S. T., & Sigurvinsdottir, R. A. (2016). Thematic analysis. *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods*, 33-41.
23. Sarkar, B., & Paul, R. K. (2023). Ethical AI and bias mitigation solution strategy in AI implementation. *Journal ID*, 2157, 0178.
24. Schellnack-Kelly, I., & Modiba, M. (2024). Developing smart archives in society 5.0: Leveraging artificial intelligence for managing audiovisual archives in Africa. *Information Development*, 02666669241286224.
25. Schuett, J. (2019). A legal definition of AI. *arXiv preprint arXiv:1909.01095*, 4.
26. Tenny, S., Brannan, J. M., & Brannan, G. D. (2017). Qualitative study.
27. Vernadakis, N., Antoniou, P., Giannousi, M., Zetou, E., & Kioumourtzoglou, E. (2011). Comparing hybrid learning with traditional approaches on learning the Microsoft Office Power Point 2003 program in tertiary education. *Computers & education*, 56(1), 188-199.
28. Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, 37(2), 209-226.
29. Wang, X., Li, X., & Leung, V. C. (2015). Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges. *IEEE Access*, 3, 1379-1391.
30. Wu, J., Williams, K. M., Chen, H. H., Khabsa, M., Caragea, C., Tuarob, S., ... & Giles, C. L. (2015). Citeseerx: Ai in a digital library search engine. *AI Magazine*, 36(3), 35-48.
31. Zhao, L. (2021). AI-Enhanced Data Structures for High-Performance Computing. *International Journal of AI, BigData, Computational and Management Studies*, 2(2), 1-9.

Appendix A

Here are the interview themes: They are about using artificial intelligence in improving the accuracy of electronic sources indexin

Research Question	Interview Themes
<p>To what extent does the implementation of artificial intelligence improve the accuracy of indexing electronic sources? This research question addresses the impact of artificial intelligence on the precision of indexing electronic sources, focusing on measuring the degree of improvement brought by AI technologies.</p>	<ol style="list-style-type: none"> 1. Discuss the limitations and typical error rates of manual or rule-based indexing approaches before AI implementation against the baseline accuracy of traditional indexing methods. 2. Discuss the specific artificial intelligence technologies applied to indexing electronic sources. 3. Explore measurable changes in indexing accuracy (e.g., metrics such as error reduction, recall, and precision) after the introduction of AI. 4. Evaluate the ability of AI to handle ambiguous terms, contextual keywords, synonyms, and polysemy in electronic resources, compared to traditional methods. 5. Evaluate how AI impacts the speed, workload, and scalability of the process of making indexes, especially when there are large or rapidly growing collections. 6. Investigate whether human intervention is still required to review, correct, or improve the AI-generated index.

		7. Discuss any remaining deficiencies in the AI-generated index and potential improvements through further improvement strategies, including ongoing model training and validation.
--	--	---

