

# Bridging Islamic Knowledge and AI: Inquiring ChatGPT on Possible Categorizations for an Islamic Digital Library (full paper)

Amina El Ganadi<sup>1,4,5,†</sup>, Riccardo Amerigo Vigliermo<sup>1,4,\*,†</sup>, Luca Sala<sup>2,4,†</sup>,  
Matteo Vanzini<sup>3,4,†</sup>, Federico Ruozzi<sup>4,†</sup> and Sonia Bergamaschi<sup>4,†</sup>

<sup>1</sup>PhD Student, Digital Humanities

<sup>2</sup>PhD Student, ICT Doctorate school

<sup>3</sup>PhD Student, CDS Doctorate school

<sup>4</sup>University of Modena and Reggio Emilia, UNIMORE

<sup>5</sup>University of Palermo, UNIPA

## Abstract

This research evaluates the capabilities of ChatGPT in assisting with the categorization of an Islamic digital library exploiting incremental Machine Learning and Transfer Learning techniques. Noticeably, ChatGPT showcased a remarkable familiarity with Islamic knowledge, evident in its ability to classify subjects hierarchically based on their importance, from Qur'anic Studies to Modern Islamic Thought. The library aimed to cater to a diverse Arabic Islamic audience with collections sourced from varied digital donations. Despite ChatGPT's commendable proficiency, challenges arose. In light of ChatGPT's significant performance, several challenges arose, with interpretability, generalization, and the hallucination issue standing out as the most critical obstacles.

## Keywords

Libraries and Archives in CH, Digital Libraries and Religious Archives, ChatGPT, Islamic studies, Arabic script languages, Islamic knowledge classification, Islamic subjects

---

*2nd Italian Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2023, <https://ai4ch.di.unito.it/>), co-located with the 22nd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2023). 6-9 November 2023, Rome, Italy*

\*Corresponding author.

†These authors contributed equally.

✉ amina.elganadi@unimore.it (A. El Ganadi); r.a.vigliermo@unimore.it (R. A. Vigliermo); luca.sala@unimore.it (L. Sala); matteo.vanzini@unimore.it (M. Vanzini); federico.ruozzi@unimore.it (F. Ruozzi); sonia.bergamaschi@unimore.it (S. Bergamaschi)

ORCID 0000-0002-8196-2628 (A. El Ganadi); 0000-0001-9914-3295 (R. A. Vigliermo); 0000-0002-4833-8882 (L. Sala); 0000-0003-0471-1101 (M. Vanzini); 0000-0003-2729-5016 (F. Ruozzi); 0000-0001-8087-6587 (S. Bergamaschi)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

# 1. Introduction

## 1.1. A Brief Overview of the Digital Maktaba Project

In the present time, there is a pressing need for a global sharing of multicultural heritage, which is made simpler by technology that can develop semi-automatic solutions to improve document readability, comprehension, preservation, and advanced digital use with sophisticated consultation and search capabilities. In this context, the Digital Maktaba project, hereafter referred to as DM, is the result of a collaboration between the Foundation for Religious Studies (FSCIRE), a national infrastructure located in Bologna and Palermo, the innovative startup mim.FSCIRE, and the University of Modena and Reggio Emilia. This collaborative effort integrates the interdisciplinary expertise of diverse professionals from both the Humanities and ICT fields[1, 2, 3, 4]. DM is part of a broader research project known as ITSERR,<sup>1</sup>. Its objective is to improve the European Research Infrastructure RESILIENCE [5] in response to the demands of the scientific community in Religious Studies in terms of technology integration and capacity to increase innovation. The aim of DM is to provide an intelligent system that can extract valuable information from documents in various languages, encompassing rich metadata content. Additionally, it assists in the manual work typically associated with the cataloging process, presenting an innovative solution tailored for libraries specializing in religious studies, particularly those tasked with managing multilingual and multi-alphabetic cultural heritage documents.

## 1.2. ChatGPT and Islamic Texts: An Initial Stride Towards Connecting AI with the Categorization and Cataloging of Islamic Knowledge

A project such as DM, involving the aforementioned research perspectives, cannot ignore the recent progress made in OpenAI's generative language model, GPT. Consequently, we carried out an experiment using its AI chatbot to evaluate the capacity to generate potential text classifications for an Islamic digital library. This stimulated us to delve into the formulation of interesting queries regarding how Islamic texts could be effectively categorized by comparing the outcomes with other digital libraries in the Islamic domain (e.g., *Shamela*, *Quranic thought*, *Ketābxāneh-ye Diğitāl-e noor*, etc.). The experiment conducted with ChatGPT was not intended to establish a definitive categorization of Islamic texts for a digital library. Nevertheless, it serves as the inaugural attempt to combine a Large Language Model (LLM) with the classification of Islamic knowledge and digital libraries. In this preliminary study, specific linguistic and religious coordinates were used to assess ChatGPT's behavior and analyze its responses, which included proposed classifications for Islamic topics. It's important to note that this study should not be viewed as a basic exercise involving input-processing-output, nor the resolution of the centuries-old debates and scholarly studies related to the categorization of Islamic knowledge and bibliographic sciences; rather, it represents a pioneering step in exploring the intersection of AI and Islamic scholarship, offering valuable insights for future research and development in this field.

---

<sup>1</sup>The ITSERR project was launched in November 2022 and funded by the Italian Ministry of Research with NextGenerationEU program funds. It involves the University of Modena and Reggio Emilia, CNR, University of Palermo, University of Turin and the University of Naples "L'Orientale"

The paper is structured as follows: in Section 2, we will provide an overview of previous studies concerning Islamic Knowledge classification and related works with ChatGPT. Subsequently, we will delve into the details of our experiment involving the categorization of Islamic knowledge for digital libraries in Section 3, highlighting our preliminary findings and engaging in discussions pertaining to this inquiry. The conclusion and future perspectives will follow.

## 2. A State-of-the-Art Exploration: Knowledge Classification and ChatGPT in Libraries

To contextualize the proposed experiment on Islamic classification of knowledge, we must consider the most important studies produced by Muslim information scientists (2.1), to better introduce some related works connecting ChatGPT and libraries (2.2).

### 2.1. Islamic Knowledge Classification and Bibliographic Studies

The classification of Islamic knowledge has been investigated since the 8th A.D. century and it is still a discussed topic in Information science and Librarianship fields. Early Muslim scholars have already devised a subject-based classification when dealing, for example, with the *ḥadi* literature.<sup>2</sup> In more recent times, the field of Information management for Islamic literature has highlighted the urgency of establishing a sort of bibliographical control on one side and evolving new classification schemes on the other. Works such as 'Ishāq [6], Anees and Athar [7] have dealt with the former. The latter emerged as a response to the Western-originated library classification systems such as Dewey Decimal Classification (DDC), Library of the Congress Classification (LCC), Universal Decimal Classification (UDC) and Colon Classification (CC) which could not provide proper enumeration and suitable hierarchy to Islamic topics [8, 9, 10]. For this reason, more recently, librarians have proposed translations [11] or expansions of the standard classification systems: DDC [12, 13], LCC [14]. At the same time independent frameworks were proposed as an alternative [15, 16]. These studies are only an example of what has been produced on the Islamic libraries classification topic. What emerges is a pretty scattered frame where many attempts were proposed at a local level, but, as Sabzwari argued, a universal Islamic classification has not yet been subjected to full discussion or analysis [17].

### 2.2. Related Works Considering ChatGPT and Libraries

LLMs' application in library services is an emerging and auspicious research domain. As far as a review of the literature has shown, the topic is still in a preliminary phase. Literature on ChatGPT and libraries is currently limited, though in quick development. What emerged is a focus on the use of AI chatbots in the educational field on topics such as Islamic education [18] or Arabic Language in Higher Education [19], to cite a few. Scientific papers discussed possible implementations of ChatGPT in various librarian and editorial fields [20], as well as some aimed to determine its possible impact and implications on academic libraries [21, 22]. Additionally,

---

<sup>2</sup>Also Hadith, trans. "narrative", "talk", "speech". The term *ḥadi* indicates the tradition of words and deeds of the Prophet. As a proper Islamic literary genre, it includes also traditions on Companions of the Prophet and other kinds of historical reports.

numerous conversations have been posted in blogs related to library science, exploring the potential applications of ChatGPT and other AI tools within library settings [23, 24, 25]. To our knowledge, the work most closely related to the one presented here is a recent publication by Brzustowicz [26]. Through experiments with the cited AI chatbot, the author demonstrates its ability to generate accurate MARC records using RDA and other standards.<sup>3</sup> As evidenced by the current state of the art, there has not been an experiment that combines LLMs with Islamic topics classification and cataloging. Therefore, this paper can be regarded as the pioneering effort in this genre, aligning LLMs with the classification of Islamic knowledge within the digital library framework.

### 3. ChatGPT for an Islamic Digital Library Classification

This experimentation tries to assess the AI model's competence in a novel capacity—that of a knowledgeable librarian. Our focus was twofold: first, to ascertain its proficiency in Islamic studies nuances, and Library information science, and then to explore its potential in devising an effective categorization system for our extensive collection of Islamic literature. The foundation of our experiment rested on the collaboration between human expertise and AI capabilities. It began by formulating a comprehensive categorization structure for our digital library. This involved brainstorming and generating categories spanning a myriad of topics, including Qur'ānic studies, Hadith, Islamic law, Theology, History, and more. Each category was meticulously tailored to cater to diverse audiences while maintaining scholarly integrity. Once the initial categorization framework was established, ChatGPT was integrated into the process. The AI's proficiency in multiple languages, including Arabic, Persian, and Azerbaijani proved to be a useful asset in transcending linguistic barriers. The model's role extended beyond translation; it actively participated in generating categories and subcategories across languages. These AI-generated outputs underwent thorough review and refinement by human interventions. This iterative collaboration resulted in an initial attempt to create a cohesive categorization structure that tries to accommodate the linguistic diversity intrinsic to Islamic scholarship. The AI chatbot's capacity to identify connections across disparate categories within Islamic studies brought interdisciplinary perspectives to the forefront. Its suggestions facilitated discussions on conceptual overlaps, enabling the creation of cross-disciplinary categories that could eventually expand our library's knowledge ecosystem (e.g., Contemporary Issues > Bioethics in Islam). Moreover, ChatGPT's involvement streamlined the process, significantly reducing the time required for category creation. Its rapid idea generation, multilingual capabilities, and contextual understanding expedited the creation of categories that remained faithful to the essence of Islamic scholarship.

#### 3.1. Methodology

As anticipated, our methodology was rooted in unleashing ChatGPT's capabilities as a knowledgeable librarian for our Islamic digital library. We interacted with the model by posing a series

---

<sup>3</sup>MARC, MACHine-Readable Cataloging. It provides the mechanism by which computers exchange, use, and interpret bibliographic information. RDA, Resource Description and Access. It is a standard for descriptive cataloging providing instructions and guidelines on formulating bibliographic data.

of different queries strategically designed to mirror real-world librarianship challenges. More specifically, we asked the model to consider historical contexts, the spectrum of authors who have contributed to Islamic literature, and the interaction between different Islamic schools of thought. By analyzing its responses to these queries, we could gauge its ability to not only formulate categories but also its adeptness at weaving a cohesive framework that encapsulates the intricacies of Islamic scholarship. Each query represented a distinct facet of the categorization puzzle, allowing us to evaluate the model's versatility, and depth of comprehension as a virtual librarian. The proposed methodology traversed beyond mere technical evaluation, venturing into the realm of generative AI, and unfolding its role in scholarly pursuits. By entrusting ChatGPT to propose categories mirroring the intricate tapestry of Islamic knowledge, we aimed to test the boundaries of AI's potential while unraveling its role in the context of knowledge organization. This meticulously designed methodology stands as the foundational guide to our experiment – a roadmap that led us to unveil the fusion of AI and human interventions in the context of knowledge categorization for Islamic studies. For this preliminary experiment, we exploited ChatGPT 4 during the frame time that went from June 27th to July 4th 2023. Due to the limit of messages in this version (25 maximum at that time, now 50), in order to maintain the workflow, we decided to also use ChatGPT 3.5 as a backup. The total number of queries was 78 spread over 12 chats.<sup>4</sup>

### **3.2. Experiment Observations and Preliminary Results**

In the very preliminary phase, we decided to ask the model to catalog a book by looking at its title and author. We followed up by asking for information about other digital libraries and how they are structured from a classification perspective. These couple of preliminary chats were designed to make ChatGPT familiar with our goal from one side and to understand what level of knowledge the model had on this topic, on the other. After this introductory phase, we started with specific questions on possible cataloging schemes for a digital library of Islamic studies. As long as questions proceeded we gradually refined prompts such as the expected audience, languages involved, the extension of the collection, and so forth. We have selected three key points related to the building of a digital library in order to create our prompts: A) Digital library users; B) Collection characteristics; C) the Overall objective of the project. For each variation of the parameters, a corresponding query has been introduced to the GPT model that answered with a particular classification. Regarding the A and B points, specific attention is given to the language(s) of both users and collection. Those parameters were chosen based on the specific needs of our ongoing DM project. It should be noted that we are well aware that these are not the only factors concurring in building a digital library. However, in our specific case, they represent a common thread between our classification of Islamic knowledge needs and the implementation of Machine Learning (ML) and Transfer Learning (TL) features, with respect to other Islamic digital libraries. All suggestions the model produced during the chats gave us a corpus with well-organized classifications. As a final step, we compared those with the one already in use in the “La Pira” library. In order to do that, we downloaded and exploited

---

<sup>4</sup>See Github repository <https://github.com/projectmaktabah/Bridging-Islamic-Knowledge-and-AI>

the PDF reader plugins available for ChatGPT 4 (AskYourPDF <sup>5</sup> and Chat with PDF <sup>6</sup>).

### **General observations**

- The more questions are asked, the more the AI librarian increases its familiarity with the languages that are proposed to it.
- If we pose similar questions within the same chat, the model offers minor changes to its classifications.

### **A) Digital library users/audience**

- For a general audience, there's been an inclusion of introductory categories about Islam, e.g. "Introduction to Islam".
- When shifting from a general audience to an Arabic one, linguistic responses in both Arabic and English are given. Yet, in the English version, specific technical Islamic terms, such as "Hadith", are not translated. For an Arabic-speaking audience, there is also a subdivision into subcategories with a greater degree of specificity. Interestingly, even in this scenario, an introductory category to Islam is consistently referenced. This observation appears to stem from ChatGPT's interpretation, which may account for the consideration that not all Arabs are necessarily Muslims.
- As we further refine the user profile of the digital library to emphasize the Islamic religious aspect, the introductory category on Islam disappears while adjustments occur or new subcategories, like Muslim minorities, become evident.
- When specifying the audience both from a linguistic perspective (Arabic) and a religious perspective (Islamic), categories and subcategories are presented in Arabic (first language) and English (translation, when applicable). As anticipated, these categories and subcategories undergo further specialization.

### **B) Collection characteristics**

- As a variation, we introduced a prompt on the extensive multilingual (Arabic, Persian, Iranian Azerbaijani) La Pira Library's digital books dataset.
- Emphasizing the actual multilingualism of the collection at our disposal, clear differences in classification depth are observable. We have noticed the emergence of categories that both reflected multilingualism and a multicultural understanding. This classification, in general, tends to accommodate linguistic diversity at the expense of thematic specificity. For example, a process of translating categories into the three aforementioned languages was observed, leading to the creation of subcategories such as "Seerah <sup>7</sup> in Arabic", "Seerah in Persian", and "Seerah in Azerbaijani". Clearly, the latter two seem more like a hallucination (3.4) as they do not correspond to any existent classification.
- By specifying the distinction between printed text and manuscript, the classification only adds the subcategory "manuscripts" for every proposed category.
- Various digital formats were also taken into account. In this specific case, ChatGPT provided interesting suggestions regarding data management, user experience, digital format

---

<sup>5</sup><https://docs.askyourpdf.com/askyourpdf-docs>

<sup>6</sup><https://www.chatpdf.com/>

<sup>7</sup>Also *sira*, is an early Islamic literature genre. It literally means "way of going", "way of life" (similar to Sunna in this sense). The term has been widely used to indicate the accounts/biographies of the Prophet Muhammad.



AM

Let's assume that you are working at the building of a digital library for Islamic studies, which categories and subcategories would you choose ?



Building a digital library for Islamic studies is a massive endeavor, and the categories and subcategories chosen should reflect the breadth and depth of the discipline. Here's a suggested structure for the categories and subcategories:

1. **Quran Studies (دراسات القرآن)**
  - Interpretation/Tafsir (التفسير)
  - Recitation/Tajweed (التجويد)
  - Quranic Sciences ('Ulum al-Quran) (علوم القرآن)
2. **Hadith Studies (دراسات الحديث)**
  - Hadith Collections (مجموعات الحديث)
  - Hadith Sciences ('Ilm al-Hadith) (علم الحديث)
  - Biographical Evaluation ('Ilm al-Rijal) (علم الرجال)
3. **Islamic Law (القانون الإسلامي)**
  - Jurisprudence/Fiqh (الفقه)
  - Principles of Islamic Jurisprudence/Usul al-Fiqh (أصول الفقه)
  - Comparative Fiqh (الفقه المقارن)

**Figure 1:** Partial caption of ChatGPT's suggested category structure.

preservation, and especially Copyright and intellectual property.

### ***C) Overall objectives of the project***

- Indicating to the model that the ultimate goal of the project is to create a comprehensive digital library of Islamic studies, we have observed a sort of tension between accommodating multilingualism and the depth of specialization in Islamic subjects.
- By stating as the ultimate goal the creation of a digital library for Islamic studies that exploits incremental ML and TL to produce a semi-automated system for cataloging texts in non-Latin alphabets, ChatGPT resumed unexpectedly a very detailed classification by expanding sections such as "Hadith studies" or "Islamic Jurisprudence" (Figure 1).

ChatGPT's performance in its knowledgeable librarian role proved to be particularly impressive, unveiling a substantial grasp of Arabic language intricacies, Islamic studies, and the underpinnings of library science principles. Its contributions materialized in the form of some detailed and comprehensive classifications, demonstrating its adeptness at navigating the complexities of Islamic knowledge. The suggested categories demonstrated a level of comprehension of the AI model that went beyond superficial recognition. The model seemed to possess a significant grasp of the intricate dimensions of Islamic scholarship. Nonetheless, the challenge of interpretability arises because we cannot determine the extent of its knowledge due to the unpredictability of the patterns that underlie LLM's generative process.

- **Qur'ānic Studies:** Qur'ānic commentaries (tafsīr) and studies on Qur'ānic sciences.
- **Hadith Studies:** Hadith collections, commentaries and so on.
- **Fiqh (Jurisprudence):** ChatGPT showed knowledge when categorizing *fiqh* according to different schools of thought such as Hanafī, Malikī, Shafi'ī, and Hanbalī.
- **Islamic History and Civilization:** This category exhibited a sophisticated breakdown into significant historical periods, showing the AI model's knowledge of the chronological evolution of Islamic societies.
- **Modern Islamic Thought:** ChatGPT demonstrated relevance to contemporary discourses by suggesting this category, attesting to its grasp of ongoing intellectual trends.

Beyond the categorization structure, ChatGPT's prowess extended to proposing subcategories and cross-references. These additions optimized user navigation and accessibility, a testament to the AI's role in enhancing user experience. In essence, the model's logical and systematic arrangement of categories, coupled with its intuitive suggestions, not only showcased its technical competence but also highlighted its potential to bridge the gap between complex Islamic knowledge and user-friendly access. This phase of the experiment underscores GPT's capacity to contribute meaningfully to knowledge organization in a digital Islamic library.

### 3.3. Comparison Between AI-driven and Human-driven Classifications

The AI-driven categorization by ChatGPT was remarkable in its speed, contextual skills, and ability to deal with the vast complexity of Islamic literature. It was capable of producing a well-structured categorization system significantly quicker than traditional human-led methods. However, while the AI chatbot demonstrated a certain level of understanding of Arabic and Islamic studies, it lacked the experiential knowledge and contextual understanding that human librarians possess. This limitation became apparent in some instances where human intuition and deep cultural understanding are vital for making complex categorization decisions. Considering that it is extremely difficult to determine a definitive classification of Islamic knowledge, we aimed to compare the various classifications with those formulated by ChatGPT as a result of our queries (AI-driven). To do so, we started from the FSCIRE "La Pira" Library's topographic project, which is composed of a detailed classification with respect to the library's specific needs (Human-driven). Then, we selected from all the 36 classifications a range of 11 (that represent 31% of the total) the most promising classification proposals based on project needs such as comprehensiveness, specialization, multilingualism, and so forth. We put all the 12 classifications (that represent 33% of the total) in a PDF file<sup>8</sup> and then let Chat GPT read them through its plugins to choose the better proposals (enumerated from 1 to 12). The following is an extract from the chat:

- *Query:* Could you compare all the proposals starting from 1 to 12 and choose what you think could be the better to accomplish the task mentioned at the beginning?
- *ChatGPT response:* To build a digital library for an Islamic studies research center that caters to both specialized and non-specialized audiences, it's important to have a classification system that covers a wide range of topics in Islam, from the basics to more advanced theological and

<sup>8</sup>See the file "Classification Proposals" within the GitHub repository <https://github.com/projectmaktabah/Bridging-Islamic-Knowledge-and-AI>



philosophical studies. The library should also include resources on contemporary issues in Islam and the contributions of Islam to various fields such as science, art, and economics.

Considering these requirements, here's a comparison of all the proposals:

- **Proposals 1, 2, 3, 4, 7, and 10** are more focused on religious texts and theological studies. They cover topics such as Qur'ānic Studies, Hadith Studies, Seerah (Life of the Prophet), Fiqh (Islamic Jurisprudence), Aqidah (Islamic Theology), and Islamic Philosophy and Mysticism. These proposals would be more suitable for a specialized audience.
- **Proposals 5, 6, 9, and 12** are more comprehensive and cover a wide range of topics in Islam. They include sections on the basics of Islam, Qur'ānic and Hadith Studies, the biography of the Prophet, Islamic theology and jurisprudence, Islamic history and civilization, women and family in Islam, ethics and spirituality in Islam, Islamic art and culture, Islamic economics and finance, and science and Islam. These proposals also include sections on contemporary issues in Islam, which would be relevant also to non-specialized users.
- Given the need to cater to both specialized and non-specialized audiences and to cover a wide range of topics, **Proposals 5, 6, 9, and 12** would be the most suitable for building the digital library. They provide a broad overview of Islam, making them accessible to non-specialized users, while also including in-depth resources for a specialized audience.
- However, among these, **Proposal 5** seems to be the most comprehensive and well-structured. It covers all the essential areas of Islamic studies and also includes sections on contemporary issues and the contributions of Islam to various fields. Therefore, Proposal 5 would be the most suitable for building a digital library for an Islamic studies research center that caters to both specialized and non-specialized audiences.

### **3.4. Potential Obstacles in Generative AI Models: the Hallucination Case in our Inquiry**

During our experimentation with ChatGPT as an expert librarian for the categorization of our Islamic digital library, we encountered a phenomenon known as "hallucination" within generative AI models. This phenomenon refers to instances where AI generates information that seems plausible but is actually non-verifiable, incorrect, or fabricated [27]. Although ChatGPT has demonstrated a noticeable level of understanding and competence, it has not been immune to occasional cases of hallucination. These instances of hallucination emerged when the AI model generated information that seemed coherent and contextually relevant but lacked factual accuracy or historical authenticity. The phenomenon of hallucination underscores a crucial consideration when using AI models for tasks that require factual accuracy and scholarly precision. Although GPT's responses are generally reliable and well-informed, its ability to generate entirely new information may inadvertently result in the creation of inaccurate content. This emphasizes the importance of cross-referencing AI-provided information with authoritative sources and consulting subject-matter experts. As we reflect on our exploration, it becomes evident that while AI models like ChatGPT offer remarkable capabilities, they are not immune to limitations inherent in their design, training data, interpretability, and generalization challenges.

As such, the integration of AI into scholarly endeavors must be accompanied by careful monitoring, validation, and human oversight to ensure the accuracy and reliability of the

generated content. This raises broader questions about the ethical and epistemic implications of AI-generated content in fields that require deep expertise and nuanced understanding, such as Islamic studies. Some queries were misread or misinterpreted by ChatGPT resulting in a sloppy, incoherent, or incomplete answer. This has been a clear obstacle to our workflow that we addressed by rephrasing the prompts. This outcome, though expected, inevitably led us to intervene on some of them in order to overcome this issue and obtain an analyzable output.

#### 4. Conclusions and Future Perspectives

The assessment of ChatGPT's potential as a knowledgeable librarian for organizing an Islamic digital library provided valuable insights into the integration of generative AI within the realm of academic research and knowledge organization. GPT's AI-driven classifications of Islamic topics, which span from Qur'anic studies to modern-day issues, displayed a unique blend of speed, comprehensiveness, and linguistic adeptness. Compared to traditional human-led methods, ChatGPT's approach showcased both its advantages, primarily in speed and vastness and its limitations, particularly in areas that demanded deep cultural understanding and experiential knowledge. A noteworthy observation was the model's adaptability following the temporary deactivation of the Bing browsing feature. The AI model, reliant on its training data up until September 2021, continued to generate contextually relevant insights. Nonetheless, the importance of real-time data sourcing was evident, emphasizing the added value of up-to-date information in rendering contemporary insights. However, the challenges posed by "hallucination" incidents served as a cautionary tale. These episodes, where ChatGPT produced seemingly accurate but in fact erroneous data, underline the imperative for meticulous validation and human oversight in AI-driven scholarly endeavors.

Looking forward, the project is keen to embrace the capabilities of other advanced language models, notably Google Bard. A collaborative approach utilizing both ChatGPT and Bard is envisaged to yield a more intricate categorization system, possibly even branching into multilingual domains, for instance, integrating resources in Chinese alongside Arabic, Persian, and Azerbaijani. This dual-model approach holds the promise of supporting and enriching human-led efforts, helping to create a digital library that is adaptive, inclusive, and informed by the latest in AI-driven academic organizations. As we advance in our research, there's potential in harnessing newer features like "custom instructions", which during our initial experiments were unavailable. This could refine our inquiries and potentially elevate the outcomes. Moreover, detailed discussions on expansive classifications and specific topics, such as Islam in China<sup>9</sup>, could further enrich our understanding and provide depth to the library's offerings. Lastly, the collaborative efforts from AI generative entities from the Arab countries, including Jais (MBZUAI) and Ejaba, suggest a larger, collective movement towards AI-augmented academic endeavors. This merging of collective endeavors with the continuously advancing AI domain hints at a promising future for Islamic studies in the age of AI-enhanced research and academia.

---

<sup>9</sup>See the file named "ChatGPT inquiry ..." p. 92-99, within the Github repository, <https://github.com/projectmaktabah/Bridging-Islamic-Knowledge-and-AI>

## Acknowledgments

This work was supported by the PNRR project Italian Strengthening of Esfri RI Resilience (ITSERR) funded by the European Union – NextGenerationEU (CUP:B53C22001770006).

## References

- [1] S. Bergamaschi, R. Martoglia, F. Ruozi, R. A. Vigliermo, S. De Nardis, L. Sala, M. Vanzini, Preserving and conserving culture: First steps towards a knowledge extractor and cataloguer for multilingual and multi-alphabetic heritages, in: Proceedings of the Conference on Information Technology for Social Good, GoodIT '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 301–304. URL: <https://doi.org/10.1145/3462203.3475927>. doi:10.1145/3462203.3475927.
- [2] S. Bergamaschi, S. De Nardis, R. Martoglia, F. Ruozi, L. Sala, M. Vanzini, R. A. Vigliermo, Novel perspectives for the management of multilingual and multialphabetic heritages through automatic knowledge extraction: The digitalmaktaba approach, *Sensors* 22 (2022). URL: <https://www.mdpi.com/1424-8220/22/11/3995>. doi:10.3390/s22113995.
- [3] R. Martoglia, L. Sala, M. Vanzini, R. A. Vigliermo, A tool for semiautomatic cataloguing of an islamic digital library: A use case from the digital maktaba project, in: A. Paschke, G. Rehm, C. Neudecker, L. Pintscher (Eds.), Proceedings of the Third Conference on Digital Curation Technologies (Qurator 2022), Berlin, Germany, Sept. 19th-23rd, 2022, volume 3234 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022. URL: <https://ceur-ws.org/Vol-3234/paper1.pdf>.
- [4] R. Martoglia, S. Bergamaschi, F. Ruozi, M. Vanzini, L. Sala, R. A. Vigliermo, Knowledge extraction, management and long-term preservation of non-Latin cultural heritages - Digital Maktaba project presentation, in: B. Alessia, F. Alex, F. Stefano, M. Stefano, R. Domenico (Eds.), Proceedings of the 19th Conference on Information and Research Science Connecting to Digital and Library Science, volume 3365 of *CEUR Workshop Proceedings*, CEUR, Bari, Italy, 2023, pp. 153–161. URL: <https://ceur-ws.org/Vol-3365/#short11>, iSSN: 1613-0073.
- [5] M. Büchler, S. Riegert, F. Alpi, F. Cadeddu, Towards big religious data: Resilience research infrastructure for data on religion in the digital age, in: Proceedings of the 2nd International Conference on Digital Tools & Uses Congress, DTUC '20, Association for Computing Machinery, New York, NY, USA, 2020. URL: <https://doi.org/10.1145/3423603.3424007>. doi:10.1145/3423603.3424007.
- [6] A. S. 'Ishāq, muğam muşannafāt al-qur'ān al-karīm, volume 1, Dār al-rifāi, Riyād, 1983.
- [7] M. A. Anees, A. N. Athar, Guide to Sira and Hadith Literature in Western Languages, London, 1986.
- [8] H. Idrees, Library classification systems and organization of islamic knowledge: Current global scenario and optimal solution, *Library Resources & Technical Services* 56 (2012) 171–182. URL: <https://api.semanticscholar.org/CorpusID:62539423>.
- [9] M. S. Arianto, Islamic knowledge classification scheme in islamic countries' libraries, *Al-Jami'ah: Journal of Islamic Studies* 44 (2006) 295–323. doi:10.14421/ajis.2006.442.295-323.

- [10] H. Idrees, K. Mahmood, Classification of Islamic literature in Pakistani libraries: a survey, *Library Review* 59 (2010) 149–164.
- [11] Soltani, P., Translation and expansion of classification systems in the Arab countries and Iran, in: *Proceedings of the 61st IFLA General Conference*, 1995.
- [12] M. Shafi, Expansions of Dewey Decimal Classification relating to oriental, Islamic and Pakistani topics, *Pakistani Library Review* 4 (1962) 42–88.
- [13] S. Riazuddin, *Classification of Islamic Literature*, Royal Book Company, Karachi, 2002.
- [14] K. Fani, *Development of Islamic philosophy in Congress Classification*, National Library of Iran, Tehran, 2000.
- [15] Z. Sardar, *Islam, outline of a classification scheme*, Clive Bingley, London, 1979.
- [16] McGill University, J. Weryho, *Smith Islamic classification in use at McGill University Islamic Studies Library*, Montreal, McGill University Islamic Studies Library, Montreal, 1979.
- [17] G. A. Sabzwari, *Universal Islamic Classification*, *Pakistan Library Bulletin* 13 (1982) 1–20.
- [18] Z. Muttaqin, Implementation of Islamic Education Learning with Artificial Intelligence (CHATGPT), in: *6th International Conference on Islamic Studies*, proceeding.uingusdur.ac.id, Pekalongan, Indonesia, 2023, pp. 1–9.
- [19] H. Leunard, R. Rachmawati, B. Zani, K. Maharjan, GPT Chat: Opportunities and Challenges in the Learning Process of Arabic Language in Higher Education, *Journal International of Lingua and Technology* 2 (2023) 10–22. doi:10.55849/jiltech.v2i1.439.
- [20] T. Mali, R. Deshmukh, USE OF CHAT GPT IN LIBRARY SERVICES (2023). doi:10.1729/Journal.33816.
- [21] B. D. Lund, T. Wang, Chatting about ChatGPT: how may AI and GPT impact academia and libraries?, *Library Hi Tech News* 40 (2023) 26–29. URL: <https://doi.org/10.1108/LHTN-01-2023-0009>. doi:10.1108/LHTN-01-2023-0009, publisher: Emerald Publishing Limited.
- [22] C. Cox, E. Tzoc, ChatGPT: Implications for academic libraries, *College & Research Libraries News* (2023) 99–102. URL: <https://crln.acrl.org/index.php/crlnews/article/view/25821>. doi:<https://doi.org/10.5860/crln.84.3.99>.
- [23] ChatGPT in Libraries? A Discussion | IFLA Continuing Professional Development and Workplace Learning (CPDWL) Section Blog, 2023. URL: <https://blogs.ifla.org/cpdwl/2023/05/14/chatgpt-in-libraries-a-discussion/>.
- [24] E. Sexton, From the desk of the Chief Librarian: What does ChatGPT mean for libraries?, ???? URL: <https://www.lib.jjay.cuny.edu/content/desk-chief-librarian-what-does-chatgpt-mean-libraries>.
- [25] A. Papini, ChatGPT: A Library Perspective, 2023. URL: <https://library.bryant.edu/chatgpt-library-perspective>.
- [26] R. Brzustowicz, From chatgpt to catgpt: The implications of artificial intelligence on library cataloging, *Information Technology and Libraries* 42 (2023). doi:10.5860/ital.v42i3.16295.
- [27] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, P. Fung, Survey of hallucination in natural language generation, *ACM Comput. Surv.* 55 (2023). URL: <https://doi.org/10.1145/3571730>. doi:10.1145/3571730.