Journal of Integrated Sciences
Volume 5, Issue 4, September 2025
ISSN: 2806-4801

ARTIFICIAL INTELLIGENCE (AI) AND NATURAL LANGUAGE PROCESSING (NLP) FOR RELIGIOUS TEXTS – A SYSTEMATIC MAPPING STUDY

Khurram Shahzad & Adnan Ashraf

International Open University, The Gambia

ABSTRACT

The rapid advancement of technology is profoundly reshaping many aspects of human life. Artificial Intelligence (AI) and Natural Language Processing (NLP) are emerging as transformative forces, altering problem-solving approaches and solution derivation. Given the profound importance of religion and religious texts in individuals' lives, the impact of AI on this domain is a significant inquiry. This study presents a comprehensive and unbiased overview of the application of AI and NLP works for religious texts, focusing on religious texts containing Islamic religious text and biblical text. A Systematic Mapping Study was done of AI works on religious texts following established guidelines to examine the existing corpus of AI and NLP work focused on religious texts. 41 studies were analyzed in relation to the two research inquiries. The lack of research on the implementation of AI in religious texts is highlighted, and the limitations of AI usage in this area are discussed. The outcomes reveal that the current methodologies, lacking the employment of any Large Language Models (LLMs), and GPT (Generative Pre-training Transformer) could address complex issues.

Keywords: *Artificial Intelligence, AI, Natural Language Processing, NLP, Religious Text, Fiqh Text, Islamic Jurisprudence text, Systematic Mapping Study, LLM.*

Corresponding author: Khurram Shahzad can be contacted at khurram063@gmail.com

1. INTRODUCTION

The rapid advancement of technology is significantly impacting various aspects of human life. The upcoming integration of Artificial Intelligence (AI) and Natural Language Processing (NLP) has the capacity to transform problem-solving methods and solutions. AI, a specialized field within computer science, aims to give machines human-like intelligence (Merriam-Webster, 2023a). Similarly, NLP, another area of computer science, involves analyzing textual data using specific algorithms. This field uses various computational techniques which potentially can enable computers to understand, interpret, and analyze language like human linguistic comprehension (Elizabeth, 2001).

AI introduces a novel perspective for tackling issues that require cognitive abilities. AI algorithms empower researchers to cultivate AI and uncover solutions to with completely new perspective. Notably, an Algorithm constitutes a defined sequence of steps designed to address mathematical quandaries utilizing computational devices (Merriam-Webster, 2023b).

This paper presents a comprehensive Systematic Mapping Study (SMS) that investigates the application of AI within the context of religious texts and examines its broader societal implications. Despite the rapid advancements in AI research across various domains, the application of AI on religious texts has not been thoroughly explored, revealing a significant gap in

literature. The structure of the Paper is organized as follows. Section 2 provides essential background information that contextualizes the study, detailing the evolution of AI technologies and their relevance to the analysis of religious texts. This section also discusses the methodologies used in the field and the potential benefits and challenges associated with integrating AI into the study of these texts. Section 3 presents the results of the Systematic Mapping Study, summarizing the key findings from the literature reviewed. Here, the paper categorizes various approaches and applications of AI within religious studies, including NLP, and machine learning techniques, as well as highlighting key findings. Section 4 offers a critical analysis of the results obtained from the mapping study. This analysis examines the implications of using AI in interpreting religious texts, addressing ethical considerations and possible limitations in various algorithms. It also looks at the impact such technologies can have on societal understanding of religion.

Finally, the Paper concludes with reflections on the importance of further interdisciplinary research in this area and suggests potential future directions for academics and practitioners interested in the convergence of AI and religious studies.

2. REVIEW OF LITERATURE

2.1 Natural Language Processing (NLP)

Natural Language Processing, a subfield of AI, encompasses language processing through various stages. It typically operates on three distinct levels to accomplish its objectives: 1) lexical analysis, 2) syntax analysis, and 3) semantic analysis. Numerous technologies are available that can execute one or

more of these stages within the domain of natural language processing (Graham, 2019).

Lexical analysis means parsing the sentence into tokens and correcting any token if it is misspelled, while syntax analysis tries to understand the sentence's grammatical structure (Graham, 2019). For example, if we write "Can Muslim man wear shorts for prayars?", most programs such as Google and Microsoft Word would parse the sentence into tokens, understand its grammatical structure, and find out if any of the words are misspelled.

Google can find that "prayers" is not correctly spelled, and there is a grammatical mistake as well; it will immediately show that the sentence should be written as "Can Muslim men wear shorts for prayers." The same applies to Microsoft Word as well; it can detect spelling and grammatical mistakes, but Google performs an additional step by suggesting sentences with better synonyms stored in its text repository. It might suggest the new sentence as "Can Muslim men wear shorts for Salah?".

Semantic analysis in NLP is even more complicated because it tries to understand the meaning of the sentence and allows the program to recreate the sentence better. Grammarly is a tool that does this analysis and restructures the sentence in a more formal and better way. Grammarly might suggest rewriting the sentence as "Is it allowed for Muslim men to pray in Shorts?" Word processing software like Microsoft Word and Google usually do not perform semantic analysis.

Question-answering (QA) systems have been explored in the technological world for the past few decades. IBM has

developed IBM Watson, a system that uses intelligent NLP and AI to answer any question related to the data fed to it. The system is trained on the data and produces models that can help find the answers (Ferrucci et al., 2010; Ferrucci, 2012).

2.2 Supervised and Unsupervised Machine Learning

2.2.1 Supervised Learning

Supervised learning is a method used in machine learning where the computer is given examples that are already "answered." Think of it like a student learning from a teacher who gives both the questions (inputs) and the correct answers (outputs). The aim is for the computer to understand how to connect questions with their answers so that when it sees new, similar questions, it can predict the answers correctly (Delua, 2023).

Let's simplify this with an everyday example: imagine you are learning to identify different types of fruits. Someone shows you several pictures of apples and tells you they are apples. You're also shown pictures of bananas with the label "banana." Over time, you learn to distinguish apples from bananas based on features like color, shape, and size. The next time you see some fruit you've been taught about; you can identify it correctly. Supervised learning works in a similar way but with data. It's like teaching the computer to recognize patterns by giving it lots of labeled examples (Delua, 2023).

2.2.2 Unsupervised Learning

Unsupervised learning, a pivotal branch of machine learning, leverages algorithms to sift through and scrutinize data autonomously. This approach diverges from supervised learning by eschewing the need for pre-labeled output data. Instead, it

empowers the algorithm to independently discover patterns and deduce insights, making it a self-reliant and insightful method of data analysis. This form of learning is not without its challenges; it demands extensive datasets to derive meaningful and applicable outcomes effectively. As highlighted by Delua (2023), unsupervised learning's essence lies in its ability to navigate through data without predefined guidance, thereby revealing underlying structures and associations that might not be immediately apparent.

2.2.3 Semi-supervised Learning

Semi-supervised learning is a type of AI training that mixes two different styles 1) Supervised learning (where the computer learns from labeled examples) and 2) Unsupervised learning (where the computer finds patterns in data without any answers provided). Semi-supervised learning utilizes a limited set of labeled data accompanied by a substantially larger collection of unlabeled data for model training (Hady & Schwenker, 2013). Semi-supervised learning techniques often involve strategies like self-training, where the model uses labeled and unlabeled data itself and then retrains on this newly labeled set, and co-training, where multiple models are trained on separate views of the data and then collaborate to label the unlabeled data (Hady & Schwenker, 2013).

2.2.4 Large Language Models (LLMs)

A Large Language Model (LLM) is a special class of AI system that is designed in processing and generating human language. LLMs are designed to understand, interpret, and produce text, allowing them to execute diverse linguistic operations, including

answering questions, generating content, and translating languages (Bengio et al., 2016; Radford et al., 2018; Tamkin et al., 2021).

LLMs are considered "large" due to their use of vast amounts of data for training. They are built using advanced machine learning techniques and trained in large corpora of text data such as books, articles, and websites. This extensive exposure allows them to learn the structures, patterns, and relationships between words, phrases, and concepts in human language. Through training, LLMs become capable of generating human-like text and responding to user inputs in a coherent and contextually appropriate manner (Brown et al., 2020). LLMs are versatile tools that can perform a range of tasks, including:

- Answering Questions: LLMs can provide responses to factual or explanatory queries based on their training data.
- Text Generation: They can generate essays, stories, articles, and other forms of written content.
- Language Translation: LLMs are capable of translating text from one language to another by understanding and replicating linguistic patterns.
- Conversation Simulation: LLMs can engage in real-time dialogues, offering responses that simulate human conversation.
- Educational Assistance: They can explain complex concepts, provide study help, and assist with homework in various subjects.

LLMs represent a powerful tool in the ongoing intersection of technology and human language, capable of transforming many industries and daily activities (Brown et al., 2020).

2.3 Related Review Studies

This section outlines the exciting efforts that have been made to review existing studies. While the work on Islamic Religious text NLP is not new. Linguistic computation has been investigated in numerous studies since 2000 to the texts of the Quran, hadith, and jurisprudence.

Atwell and Sawalha (2011) presented their contribution to Quranic NLP. Their research examines how their framework contributes to interpretations of Islamic concepts in Western contexts (Atwell & Sawalha, 2011). They aim to utilize machine learning to achieve a completely unbiased understanding and analysis of Islamic religious texts.

Kammani and Safeena (2014) analyzed multiple approaches and technologies associated with Arabic NLP. Their study encompasses efforts in Arabic and Quranic NLP, focusing on contributions made from 1997 to 2011. Their discussion covered Quranic corpus, annotation schemes, grammar treebanks, semantic search, translation evaluation, morphological analysis, and Quranic ontology (Kammani & Safeena, 2014). A notable limitation of their review is missing the discussion regarding reasoning applications in NLP, such as question answering or text summarization, which remain unhighlighted.

Bashir et al. (2022) provide a broad overview of NLP studies related of the Hadith text. Additionally, Ahmed et al. (2022)

provided an overview of open-access Arabic language corpora, which includes the Quran and Hadith texts.

This study specifically concentrates on the domain of religious texts, rather than the broader field of Arabic NLP and application of AI. This targeted approach allows us to delve deeper into the unique linguistic and semantic features of Quranic verses, Hadith literature, and Islamic jurisprudence text. By focusing on these religious texts, we aim to gain a more profound understanding of their status within the NLP landscape. Ultimately, our work aspires to bridge the gap between religious text and advanced computational methods, paving the way for contributions that can enrich both academic discourse and practical applications in the study of Islamic texts.

3. RESEARCH METHODOLOGY

3.1 Research Design

To present a structured study of current AI and NLP applications in religious texts, specifically focusing on Islamic Jurisprudence and the "Fiqh" text, a Systematic Mapping Study (SMS) method was conducted (Petersen et al., 2008, 2015). This approach helps in identifying, categorizing, and analyzing the breadth of AI and NLP applications within this niche area. The primary objectives of the systematic mapping study included: 1) To identify and categorize the existing body of literature on AI and NLP applications in religious texts. 2) To assess the type of machine learning methodologies employed in these studies and the extent of their findings. 3) And to highlight gaps in current research and suggest potential areas for future exploration.

We employed the well-established guidelines put forth by Petersen et al. (2008) for systematic mapping studies (SMS) for this study. The methodology mentioned for this research is as follows: first, we define a search string aligned with the research questions outlined in Section 3.1 to gather pertinent leading studies from various major online research databases. The duplicates were removed after reviewing the abstracts that do not meet the relevance criteria as specified in Section 3.3. After this step, the snowballing method was applied on remaining studies which is mentioned by Jalali and Wohlin (2012). This steps helps finding additional related studies which were not captured in the initial searches (Jalali & Wohlin, 2012). Following this, we assessed the quality of the studies based on the criteria mentioned in the Section 3.4 and excluded those that fail to meet the minimum quality standards. Finally, we gather the essential data from the selected studies to answer our research questions. Figure 1 shows this process of including and excluding studies at each stage.

3.2 Research Questions

These core research questions were established to guide the Systematic Mapping Study (SMS).

RQ1: How are AI technologies utilized to develop intelligent systems that facilitate the interpretation and understanding of Islamic Jurisprudence (Fiqh) texts?

RQ2: What are the potential benefits of AI on the role of religious scholars and the broader understanding of religious teachings within contemporary culture?

RQ3: What are the potential challenges of using AI and NLP to analyze religious texts?

RQ4: What are the ethical considerations and challenges of using AI to interpret sacred texts?

3.3 Search Strategy

First, the most important keywords related to our topic were picked. Then, these were combined into a search phrase that helps us find the best results.

3.3.1 Search Strings

The development of a query string is an iterative process. Initially, we adhered to the SMS guidelines mentioned by Kitchenham and Charters (2007) to craft an initial string employing Boolean OR/AND operators. This initial string was evaluated as a pilot on various search engines, including various databases such as IEEE, ACM Digital Library, Science Direct, Springer, Scopus, and JSTOR, to assess its ability to retrieve relevant primary studies. Any keywords present in both the known primary studies and the newly discovered ones were integrated into the query string if they were not already included. Additionally, we examined the titles, abstracts, and author-assigned keywords from previously identified primary studies to identify additional search terms.

The final search string used is displayed below. It is worth noting that this string had to be customized to according to the requirements of each database we accessed.

("Artificial Intelligence" OR AI" OR "Natural Language processing" OR NLP OR "Machine Learning" OR "statistical analysis" OR GPT OR "Large Language models") AND (Bible OR "biblical text" OR Religio OR "Historical text" OR Quran)*

The terms "Artificial Intelligence," "Natural Language Processing," and "Machine Learning" have many synonyms and alternate terms used in literature and many results returned. However, when combined with search terms such as "biblical text," "Quran," or "Historical text," the desired studies were quite few in the returned result set. As an integral component of the initial search strategy, search queries were executed across various databases to retrieve primary research studies.

3.3.2 Selection of Sources

To gain a more comprehensive view, we conducted a systematic search across electronic databases rather than focusing on a limited range of journals and conference proceedings. These databases cover a wide spectrum of conferences, and journal articles that are relevant to the software engineering domain.

Journal of Integrated Sciences
Volume 5, Issue 4, September 2025
ISSN: 2806-4801

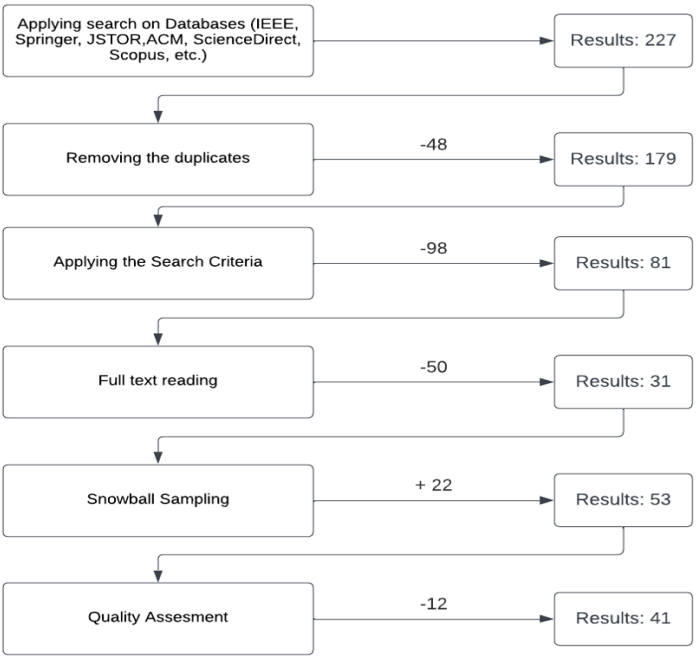

Figure 1. Studies found in various databases and total selected studies

3.4 Study Selection Criteria

The goal was to focus solely on peer-reviewed and published research papers providing ample technical details. We conducted a comprehensive study selection to identify the most pertinent primary studies for the Systematic Mapping Study (SMS). We also employed snowballing, as recommended by Petersen et al. (2015), to supplement the electronic search.

Inclusion criteria: We selected studies in which AI, NLP, or any form of machine learning were applied to analyze religious or historical texts. Additionally, we considered studies examining these technologies' broader societal impact.

We used the following inclusion criteria in our selection process:

- Publications with titles, keywords or abstracts addressing topics such as NLP, machine learning, AI, or any alternate terms outlined in Section 2.2.1.
- Papers which are written in English.
- In cases where an extended version (e.g., a book chapter or journal Paper) of a conference Paper provided more technical details, only the extended version was considered.
- If multiple versions of a study existed, only the most comprehensive version was included.

Studies were excluded if they did not meet the above requirements fell under the following conditions:

- Publications categorized as secondary studies.
- Non-peer review publications.
- Duplicated publications.

3.5 Quality Assessment of Studies

To evaluate the quality of selected publications, we employed a questionnaire based on the criteria outlined by Jalali and Wohlin (2012). The questionnaire encompassed the following key aspects:

- Clarity of goals description.
- Clarity of method/algorithm description.
- Clarity of assumptions/restrictions.
- Validation of the method through a case study.
- Discussion of any tools or techniques used.
- Qualitative benchmarking against existing approaches.
- Quantitative comparison with alternative methods.

Each primary study underwent evaluation using the above questionnaire, with scoring as per the criteria established by Usman et al. (2014):

- 4 points for complete answers.
- 2 points for partial answers.
- 0 point for irrelevant answers.

The relevance and contribution of each primary study to our overall findings corresponded directly with its assigned score. The scoring ranged from a minimum of 0 points to a maximum of 8 points. We set the cut-off point at 3 points on our quality assessment scale, following the guidelines of Usman et al.

(2014). Studies scoring 1 or less were excluded from our analysis, resulting in the inclusion of 41 primary studies.

3.6 Study Timeline

Table 1 shows a detailed schedule of our study, emphasizing that while the Systematic Mapping Study (SMS) steps seem sequential, many were iterative based on our experience. Organizing the results was particularly challenging and time intensive. Additionally, we found that presenting and organizing the results in a comprehensive and coherent manner proved to be a time-intensive and demanding task. It is important to note that we finalized our primary studies on Sep 15th, 2024.

Table 1. Schedule of the Study

No	Step	Date	Number of business days it took
1	Plan the study and formulate the protocol for the study.	July 1, 2024	9
2	Conduct database searches.	July 25, 2024	7
3	Filter studies based on title and abstract and then by full text.	July 2, 2024	23
4	Implement snowball sampling.	Aug 3, 2024	3
5	Extraction and processing of data.	August 15, 2024	11
6	Perform quality assessment.	September 13, 2024	4
7	Prepare the first draft of the Paper.	September 30, 2024	11

Source: Author(s) own work.

3.7 Data Extraction and Synthesis

A data extraction form was created, incorporating both general fields such as publication year, conference or journal, and author details, along with specific fields aligned with each research question such as the AI and NLP methods used and the type of religious text analyzed. The extracted data were then systematically synthesized to address and provide answers to the research questions.

4. RESULTS AND ANALYSIS

This section provides an overview of the outcomes for the comprehensive Systematic Mapping Study process, including findings pertaining to each individual research question.

4.1 RQ1: How are AI technologies utilized to develop intelligent systems that facilitate the interpretation and understanding of Islamic Jurisprudence (Fiqh) texts?

The Bible has been incorporated into several AI experiments to uncover diverse patterns and conduct textual analyses. For instance, Hu (2012) conducted a study in which Proverbs were analyzed, leading to the creation of clusters to explore their patterns. The objective of this research was to offer a unique perspective on Biblical text. In another instance, Tschuggnall and Specht (2016) utilized AI to conduct grammar-based text analysis to assess the authorship and authenticity of biblical text.

Similarly, another study applied AI supervised learning to scrutinize language features, with the goal of precise dating

when the text was written mentioned by Moldwin (2013). Researchers sought to categorize the text as Early Biblical Hebrew or Late Biblical Hebrew, providing insights into whether it was composed before or after the Babylonian exile.

Furthermore, certain researchers employed AI image-processing techniques to detect hidden authorship and verify the authenticity of Biblical text (Faigenbaum-Golovin et al., 2016). In addition, AI algorithms were employed in various studies that leveraged multiple translations of the biblical text to identify patterns and diverse writing styles (Büchler et al., 2014).

Figure 2 illustrates how primary studies are grouped into different AI and NLP learning methodologies. It shows that most of the research has relied on traditional statistical models to achieve results in areas such as pattern recognition, concept clustering, and theme identification. While Neural Networks and LLMs are often used for reasoning tasks in text, other NLP techniques are typically applied to classification, machine translation, and sentiment analysis. There is a noticeable gap, as very few studies have utilized LLMs. The figure emphasizes the need for more research using LLMs, particularly in the context of religious texts.

Journal of Integrated Sciences
Volume 5, Issue 4, September 2025
ISSN: 2806-4801

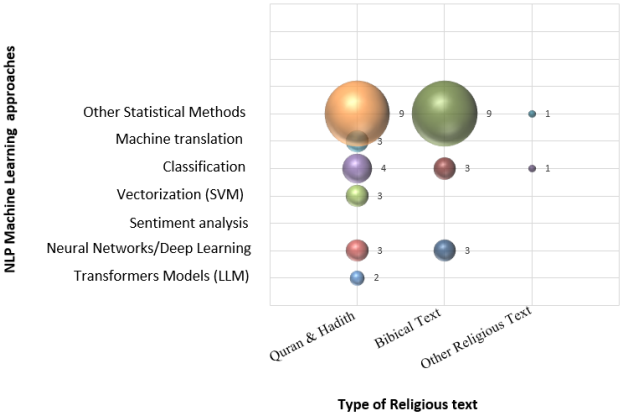

Figure 2. Number of the Primary studies with respect to various Machine Learning methods in Islamic, Biblical and other religions

Table 2 shows the list of the studies classified by the type of machine Learning techniques used in specified studies.

Table 2. List of Studies classified by the type of Machine Learning technique used

Machine Learning Techniques	List of Studies
Classification	S16, S17, S39
Neural Networks/Deep Learning	S12, S23, S29, S41, S7
Other Statistical Methods	S1, S10, S13, S18, S2, S21, S22, S25, S26, S28, S33, S34, S35, S36, S37, S38, S40, S5, S6, S8
Support Vector Machines	S19, S27, S4
Transformers Models (LLM)	S11, S24, S3, S30

Source: Author(s) own work.

Table 3 shows the type of text and related studies.

Table 3. List of Studies classified by the type of text

Type of Text	List of Studies
Islamic Text (Quran, Hadith, Jurisprudence)	S1, S12, S16, S18, S19, S2, S22, S24, S27, S29, S30, S34, S36, S37, S38, S39, S4, S5, S6, S7, S8, S9
Biblical Text	S10, S13, S14, S15, S17, S21, S25, S26, S28, S31, S32, S33, S35, S40, S41, S11, S20, S23, S3

Source: Author(s) own work.

Table 4 lists of the studies and type of problems they solve.

Table 4. List of Studies classified by the type of Problem they are solving or discussing

Type of Problem being Solved	List of Studies
Classification of Concepts for similar themes retrieval	S10, S16, S17, S18, S19, S2, S22, S27, S28, S33, S34, S35, S36, S37, S38, S39, S4, S40, S5
Multiple Machine Learning Models to check accuracy	S29
Pattern Matching	S13, S21
Pattern Matching and Classification	S25, S26
Question Answering System	S1, S11, S12, S24, S30, S41
Semantic search	S7, S8
Semantic search and Classification of Concepts	S6
Sentiment analysis	S23
Summarization	S3

Source: Author(s) own work.

Some initiatives have been undertaken to develop reasoning-based solutions through the creation of question-answering systems, particularly within the contexts of biblical and Islamic religious texts. One notable contribution to this field was made by Zhao and Liu (2018), who developed a sophisticated question-answering system employing Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), and the Bidirectional Attention Flow (BiDAF) model. Zhao utilized an expansive dataset containing thousands of question-answer pairs, which served as the foundation for training a highly effective deep learning model (Zhao & Liu, 2018).

In parallel, similar research efforts have been directed toward understanding and interpreting Quranic texts, hadith literature, and Islamic jurisprudence. Abdi et al. (2020) made significant progress in this area by employing a Graph-Based Model to create a question-answering system specifically tailored to hadith texts. Their approach encompassed comprehensive architecture that includes crucial preprocessing steps such as sentence segmentation, stop-word removal, and stemming.

Rizqullah, Purwarianti, and Aji (2023) developed an impressive question-answering system using LLMs focused on the Prophetic Biography, known as the Seerah text. They deployed models such as mBERT and ARABERT, among others, to evaluate the accuracy and quality of the results. The system they created is called QASiNa. This research provides evaluation metrics for various LLMs used in question answering. If the system can generate insightful questions and reasoning based on the Seerah text, it will offer scholars a new perspective for exploring the prophetic biography.

Hamoud and Atwell (2016) developed an equation-answering system focused on the Quranic text. They utilized data mining and machine learning techniques, employing specific research software called WEKA (Waikato Environment for Knowledge Analysis). This software originated from a research initiative at the University of Waikato's Department of Computer Science, New Zealand, designed specifically for scholarly applications.

This software was created by the University of Waikato in New Zealand's Department of Computer Science for research purposes. The researchers applied multiple data mining techniques, including stemming, data transformation, and pattern evaluation. Although this research did not use the most recent LLMs, it still achieved commendable results from the perspective of question-answering systems.

Alqahtani and Atwell (2016) employed machine learning to create a concept-based search system, often referred to as semantic search, specifically for Quranic text. This approach allows users to match concepts rather than simply conducting keyword searches. Their Quranic Search Tool, which is based on ontology and NLP, enables users to search through the concepts and themes present in the Quran.

Alshammeri, Atwell, and Alsalka (2021) conducted a semantic search using Document embeddings and the Doc2Vec machine learning algorithm. Their goal was to represent the Arabic Quranic verses as numerical vectors that capture semantic properties. By utilizing these vectors, they were able to measure the similarity between different concepts effectively (Alshammeri et al., 2021, 2021).

Atwell and Sawalha (2011) aimed to develop semantic search capabilities by employing different machine learning models to analyze Arabic text, enabling the identification of patterns for similar concepts within the Arabic language.

Shahzadi (2012) explored a semantic search approach to automate the text categorization of Quranic texts. This research focuses on the automated categorization of themes and concepts within the Quran, which can help identify and cluster useful patterns. Abu Nada et al. (2020) conducted research on various semantic-based approaches utilizing multiple machine learning algorithms to compare their results and accuracy with Arabic text. Their study employed LLMs such as Transformers and the AraBERT model, examining the outcomes through text summarization and executive summaries. The research aimed to compare these approaches to demonstrate the accuracy of each algorithm.

Munshi et al. (2022) focused on deriving reasoning from Islamic jurisprudence texts by utilizing LLMs. Their research employed transfer learning techniques on various NLP models. This study showcases a question-answering system capable of providing verdicts based on Islamic law. The model was trained using an extensive dataset gathered from various Islamic verdict websites.

The selected studies predominantly utilize NLP and machine learning techniques to address a range of complex problems. These include innovative applications such as intelligent search capabilities, pattern matching, clustering of data, semantic search, and text summarization.

4.2 RQ2: What are the potential benefits of AI on the role of religious scholars and the broader understanding of religious teachings within contemporary culture?

The selected studies and proposed systems significantly enhance the ability to engage with and interpret religious texts. Utilizing various methodologies and technologies, these studies provide innovative tools that make analyzing, understanding, and applying religious teachings more accessible. This advancement not only aids scholars and practitioners in their work but also invites a broader audience to explore and connect with religious literature in meaningful ways. These systems offer several benefits:

a. They enable users to perform searches based on semantic meaning or themes.

b. They allow users to cluster ideas or themes and create patterns from similar data.

c. They help users deduce reasoning and gain understanding from the text.

d. They facilitate the interpretation of the text by users.

Table 5. Categories of benefits or problems discussed or solved by the studies

List of the benefits discussed and/or problems solved	List of the Studies
They enable users to perform searches based on semantic meaning or themes.	S6, S7, S8, S15, S36
They allow users to cluster ideas or themes and create patterns from similar data.	S2, S4, S5, S17, S18, S19, S21, S22, S23, S25, S26, S27, S33, S35, S37, S39, S40
They help users deduce reasoning and gain understanding from the text.	S1, S3, S9, S11, S12, S16, S29, S30, S34, S38, S41
They facilitate the interpretation of the text by users.	S20, S24

Source: Author(s) own work.

These technical advancements will certainly help both scholars and the wider community in utilizing them. However, access to these tools may also have negative consequences for both groups. All machine learning algorithms come with a margin of error and relying on NLP and AI to make decisions could result in potential inaccuracies. Additionally, depending on these tools, without a thorough understanding of how they function can lead to further issues.

While there is a limited number of research studies that have utilized NLP or AI technologies for religious text analysis, there are, nonetheless, a few studies that focus on the biblical text and employ machine learning techniques to explore and examine it from a unique perspective.

One study introduced the idea of using AI to extract knowledge from Quranic texts and provide answers to the public's questions. The authors have named this proposed system the

"Impartial Online Quran Expert" (Atwell et al., 2010). This system would involve creating intelligent technologies capable of answering inquiries based on Quranic knowledge, thus assisting both Muslims and non-Muslims in understanding its teachings.

One of the prominent advantages of AI is its semantic understanding, which allows it to interpret and deduce important knowledge from text. AI, particularly LLMs focused on text, can assist scholars while also presenting challenges in understanding religious texts. In the study S24 Munshi et al. (2022) discusses the automated generation of jurisprudential verdicts. A fine-tuned LLM trained on jurisprudence texts could help create a question-answering system, greatly enhancing access to this information for everyone. Another noteworthy project is QASiNa, developed by Rizqullah et al. (2023).

The iQurNet system has been proposed Putra et al. (2022) to assist non-Arabic speakers in interacting with the Quran in their own language. Currently, its accuracy stands at 84%, but there is potential for improvement by further refining the AI techniques. Additionally, a question-answering system developed for the Quran, which utilizes the data mining software WEKA, can help generate intelligent responses.

4.3 RQ3: What are the potential challenges of using AI and NLP to analyze religious texts?

Each machine learning technique has its own advantages and disadvantages. While AI techniques have shown great promise, they also come with certain limitations. Tufail et al. (2023) has discussed the limitations in detail and highlights the

shortcomings of various machine learning methods. It is important to carefully consider which technique to use for specific use cases.

Table 6. Table outlining the key challenges for some of the machine learning approaches

Machine Learning Approach	Challenges	Related Studies
Simple Statistical Methods	• Limited scalability for large datasets. • Struggles with complex patterns and high-dimensional data. • Requires Assumptions (e.g., linearity, normality) may not hold true for all datasets. • Often underperforms compared to more advanced methods (e.g., deep learning).	S1, S2, S9, S18, S21, S25, S26
Sentiment Analysis	• Domain-specific vocabulary and context can lead to inaccurate results without proper fine-tuning. • Requires labelled data for training, which can be labor-intensive to prepare.	S23
Classification	• Imbalanced class distributions can skew results (e.g., biased predictions). • Feature extraction and selection can significantly impact model performance. • Overfitting and underfitting when not carefully tuned.	S2, S4, S5, S10, S16, S17, S18, S19, S23, S28, S29, S33, S35, S38
Support Vector Machine (SVM)	• Computationally expensive for large datasets. • Difficulty in choosing the right kernel and tuning the hyperparameters.	S4, S19, S27

	• High memory usage for large-scale problems.	
Machine Translation	• Ambiguities in language can cause inaccurate translations (e.g., idiomatic expressions). • Requires large amounts of parallel training data for high-quality translation models. • Can have difficulty with sentence structure and grammar in complex translations.	S29
Neural Networks and Deep Learning	• Requires vast amounts of labeled data for effective training. • Computationally expensive and requires significant hardware (e.g., GPUs). • Hard to interpret and understand, leading to the "black-box" issue. • Training can be time-consuming and difficult to optimize	S7, S12, S11, S18, S23, S29, S36, S41
Transformers Models (LLM)	• Requires extensive computational resources (e.g., high-end GPUs). • Expensive to train from scratch due to large model size and data requirements. • High risk of biases in pre-trained models leading to ethical and fairness concerns.	S11

Source: Author(s) own work.

Given the limitations discussed, understanding and interpreting religious texts with the help of AI can be quite challenging, especially considering the potential societal impact. There are two main types of challenges: 1) technical challenges related to

AI techniques, and 2) the social impact of AI on society. Each selected study highlights the technical limitations of various machine learning approaches. While some of these limitations are comprehensible, others remain opaque due to the nature of the AI techniques employed, such as neural networks, deep learning, and LLMs.

The studies S11, S14, S20, S21, S24, and S26 examine social challenges in detail. Rossano (2001) points out that future developments in AI will enhance the convergence between humans and machines; however, ethical challenges will also arise from advanced AI technologies and moral evaluations. It emphasizes the need to establish a moral framework for assessing AI technologies. Goltz, Dowdeswell, and Zeleznikow (2020) discuss and highlight concerns related to AI agents operating as independent entities within the community, raising questions about responsibility.

4.4 RQ4: What are the ethical considerations and challenges of using AI to interpret sacred texts?

Most of the research conducted aims to enhance our interaction with religious texts and to help us understand the issues from various perspectives, leading to a deeper comprehension. When research is carried out, several aspects may be affected, such as the general community and its reaction to the latest findings.

When it comes to interpreting religious texts through the lens of AI, there can be significant issues within society. Currently, AI does not produce results with 100% accuracy and occasionally experience hallucinations, which can lead to

negative consequences and these incorrect results can ethically be incorrect.

The table below highlights various ethical issues associated with the studies.

Table 7. List of studies related to ethical issues

Types of Ethical issues	List of Studies
Bias and discrimination	S1, S2, S3, S5, S8, S10, S11, S14, S17, S20, S24, S28, S31, S35, S38, S41
Ethical decision-making in autonomous systems	S1, S8, S10, S14, S20, S24, S31, S38, S41
Misinterpretation of religious texts	S1, S8, S14, S20, S24, S31, S38, S41
Lack of cultural and contextual understanding	S14, S24, S31
Over-reliance on AI interpretations	S1, S8, S10, S14, S20, S24, S31, S38, S41
Sensitivity to diverse theological perspectives	S14, S24, S31
Impact on traditional scholarship	S14, S24, S31
Authority of AI in religious discourse	S14, S24, S31
Challenges in multilingual religious text analysis	S29, S30

Source: Author(s) own work.

Hu (2012) mentioned that unsupervised learning is being applied to various books of the Bible to analyze Proverbs and Psalms. This approach identifies similarities between the chapters and verses of these two books. Such studies could help pinpoint differences in the biblical text. AI may be capable of detecting where the texts have been altered, modified, or supplemented by examining the different versions of the

scriptures. This can raise concerns or discussions among the scholars and bring a lot of ethical challenges.

Using unsupervised learning techniques on various books of the Bible, particularly focusing on Proverbs and Psalms, presents an intriguing opportunity for conducting in-depth textual analysis. By analyzing the similarities and patterns among the chapters and verses in both books, we can uncover nuances that might have otherwise gone unnoticed. This type of study could illuminate not only the thematic connections between Proverbs and Psalms but also highlight the distinctiveness of each book.

Moreover, employing AI in this analysis allows for a comprehensive examination of the biblical text across different versions and translations. By comparing variations, AI could potentially identify sections that have been altered, modified, or added over time. This could lead to insights into historical context, authorship, and the evolution of religious thought. Ultimately, such studies may contribute to a richer appreciation to analyze the biblical.

Büchler et al. (2014) concentrates on the intricate task of detecting historical text re-use, specifically focusing on variations in how texts are re-used over time. By examining seven English editions of the Holy Bible, it evaluates various techniques aimed at identifying paraphrased relationships within these texts. The findings suggest that when instances of text re-use are detected, it opens the possibility for rewriting these passages. However, this rewriting raises important concerns among religious scholars, as it could impact the interpretation, authenticity, and theological implications tied to

sacred texts. This could raise a serious ethical concern if AI starts to rewrite religious text.

Moldwin (2013) looks at the biblical text dating using the machine learning algorithms. It uses the earliest and latest Biblical text and explore linguistic features for authorship and dating. Studies like these could allow solving some of the unknown questions on the authenticity of the Biblical text and allows the scholars to identify the tempering in the text.

Goltz, Dowdeswell, and Zeleznikow (2020) discuss surrounding AI regulations, consider various perspectives and highlights how AI and AI agents can potentially cause chaos in society. The text emphasizes the moral and legal personhood of artificial agents, asserting that any decisions made by AI should ultimately hold a human responsible for moral accountability. The author stresses the importance of establishing clear guidelines and frameworks to ensure these technologies serve the greater good and contribute positively to human welfare.

Rossano (2001) noted that future advancements in AI will promote closer integration between humans and machines, but these developments will also bring about ethical challenges. For instance, neural implants could enhance cognitive functions and social contributions, raising questions about how these improvements might affect comparisons of human abilities. Moreover, religion plays a significant role in maintaining community stability and this highlights the ethical dilemmas posed by advancements in AI. Therefore, establishing a moral framework for evaluating AI technologies is essential.

5. CONCLUSION AND LIMITATIONS

5.1 Conclusion

In summary, this research provides the Systematic Mapping Study (SMS) of the usage of Artificial Intelligence (AI) and Natural language Processing (NLP) on religious text. While research in this area is still relatively limited, it holds the promise of unlocking new insights, patterns, and perspectives within these sacred texts.

The selected studies showcase various approaches, from clustering and grammar-based analysis to supervised learning and image processing, demonstrating the versatility of AI and NLP techniques. These endeavors not only contribute to the understanding of religious texts but also offer opportunities to solve modern-day Jurisprudence problems.

Furthermore, a reasoning-based question-answering system will undoubtedly have a considerable impact on the broader community. These innovative approaches to information retrieval and understanding not only promise to enhance the accessibility and precision of information but also aim to revolutionize the way we interact with knowledge databases

While these initiatives are in various stages of development, they underscore the significance of embracing technological advancements to gain fresh perspectives on texts that have shaped cultures and societies for centuries. As technology continues to evolve, it is likely that AI and NLP will play an increasingly important role in the study and interpretation of religious texts, enriching our understanding of these enduring sources of knowledge and wisdom.

5.2 Limitations

The limitations outlined below could have influenced the outcomes of this study.

5.2.1 Limitations in Publication Selection

Our selection of publications may not be complete. While we followed a systematic and unbiased process relying on guidelines from Petersen et al. (2008, 2015). Our inclusion criteria focused on publications with the terms 'machine learning', or 'deep learning' or 'Natural Language Processing' in titles or abstracts, possibly missing out on other valuable studies. We also did not reach out to researchers for unpublished results and excluded gray literature like working papers due to uncertain quality. This exclusion might mean some publications were wrongly left out. We extensively searched well-known electronic databases containing numerous journals, conferences, and workshops to minimize this risk.

5.2.2 Potential Data Extraction Inaccuracy

Due to the subjective nature of the data extraction process, there is a possibility of gathering inaccurate results. To mitigate the inherent subjectivity in the data extraction process, which may lead to inaccuracies, we implemented a structured approach. Specifically, in Section 3.6, we developed a detailed data extraction form. This form provides a comprehensive and standardized framework for identifying and collecting the relevant data from the primary studies, ensuring consistency and precision in the extraction process. Each research question

is aligned with the specific data points to be extracted, reducing ambiguity and promoting the reliability of the results.

5.2.3 Potential Biases in Quality Appraisal

The evaluation of the quality of selected studies posed a difficulty. To address this challenge impartially, we have meticulously defined our criteria for quality assessment in Section 3.

REFERENCES

Abdi, A., Hasan, S., Arshi, M., Shamsuddin, S. M., & Idris, N. (2020). A question answering system in Hadith using linguistic knowledge. *Computer Speech & Language*, *60*, 101023. https://doi.org/10.1016/j.csl.2019.101023

Abu Nada, A. M., Alajrami, E., Alsaqqa, A., Al-Saqqa, A. A., & Abu-Naser, S. S. (2020). Arabic text summarization using AraBERT model using extractive text summarization approach. *ResearchGate*. https://www.researchgate.net/publication/344175857

Ahmed, A., Ali, N., Alzubaidi, M., Zaghouani, W., & Househ, M. (2022). Free and accessible Arabic corpora: A scoping review. *Computer Methods and Programs in Biomedicine Update*, *2*, 100049. https://doi.org/10.1016/j.cmpbup.2022.100049

Alqahtani, M., & Atwell, E. (2016). Arabic Quranic search tool based on ontology. In *Lecture Notes in Computer Science* (Vol. 9612, pp. 478–485). Springer. https://doi.org/10.1007/978-3-319-41754-7_52

Alshammeri, M., Atwell, E., & Alsalka, M. A. (2021). Detecting semantic-based similarity between verses of the Quran with Doc2vec. *Procedia Computer Science*, *189*, 351–358. https://doi.org/10.1016/j.procs.2021.05.104

Atwell, E., Habash, N., & Abu Shawar, B. (2010). Understanding the Quran: A new grand challenge for computer science and artificial intelligence. *ResearchGate*. https://www.researchgate.net/publication/215442740

Atwell, E., & Sawalha, M. (2011). An artificial intelligence approach to Arabic and Islamic content on the Internet. https://doi.org/10.13140/2.1.2425.9528

Bashir, M., Azmi, A., Nawaz, H., Zaghouani, W., Diab, M., Al-Fuqaha, A., & Qadir, J. (2022). Arabic natural language processing for Qur'anic research: A systematic review. *Artificial Intelligence Review*, *56*, 1–54. https://doi.org/10.1007/s10462-022-10313-2

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., et al. (2020). Language models are few-shot learners. *arXiv*. https://doi.org/10.48550/arXiv.2005.14165

Büchler, M., Burns, P. R., Müller, M., Franzini, E., & Franzini, G. (2014). Towards a historical text re-use detection. In *Research and Advanced Technology for Digital Libraries* (pp. 221–238). Springer. https://doi.org/10.1007/978-3-319-12655-5_11

Delua, J. (2023). Supervised vs. unsupervised learning. *IBM*. https://www.ibm.com/topics/supervised-learning

Faigenbaum-Golovin, S., Shaus, A., Sober, B., Levin, D., Na'aman, N., Sass, B., Turkel, E., Piasetzky, E., & Finkelstein, I. (2016). Algorithmic handwriting analysis of Judah's military correspondence sheds light on composition of biblical texts. *Proceedings of the National Academy of Sciences*, *113*(17), 4664–4669. https://doi.org/10.1073/pnas.1522200113

Goltz, N., Dowdeswell, T., & Zeleznikow, J. (2020). From the tree of knowledge & the Golem of Prague: The ethics of artificial intelligence through Jewish eyes. *Emerj.* https://emerj.com/ai-sector-overviews/machine-learning-medical-diagnostics-4-current-applications/

Graham, B. (2019). Using natural language processing to search for textual references. In *Digital Biblical Studies* (Vol. 3, pp. 115–132). Brill. https://doi.org/10.1163/9789004399297_008

Hamoud, B., & Atwell, E. (2016). Quran question and answer corpus for data mining with WEKA. In *2016 Conference of Basic Sciences and Engineering Studies (SGCAC)* (pp. 211–216). IEEE. https://doi.org/10.1109/SGCAC.2016.7458032

Hu, W. (2012). Unsupervised learning of two Bible books: Proverbs and Psalms. *Sociology Mind*, *2*(3), 325–334. https://doi.org/10.4236/sm.2012.23043

Jalali, S., & Wohlin, C. (2012). Systematic literature studies: Database searches vs. backward snowballing. In *2012 International Symposium on Empirical Software Engineering and Measurement* (pp. 29–38). ACM. https://doi.org/10.1145/2372251.2372257

Kammani, A., & Safeena, R. (2014). A review of Quranic computation for e-learning. *International Journal of Web Science*, *2*(1), 127–139. https://doi.org/10.1504/IJWS.2014.066415

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (Version 2.3). *EBSE Technical Report*.

Moldwin, T. Y. (2013). Linguistic dating of biblical texts using supervised machine learning. [Unpublished manuscript].

Munshi, A. A., AlSabban, W. H., Farag, A. T., Rakha, O. E., Al Sallab, A., & Alotaibi, M. (2022). Automated Islamic jurisprudential legal opinions generation using artificial intelligence. *Pertanika Journal of Science and Technology*, *30*(2), 1135–1156. https://doi.org/10.47836/pjst.30.2.16

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. In *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering* (pp. 68–77). BCS.

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, *64*, 1–18. https://doi.org/10.1016/j.infsof.2015.03.007

Putra, O. V., Elfianita, F. D., Harmini, T., Trisnani, A., & Nugroho, A. (2022). IQurNet: A deep convolutional neural network for text classification on the Indonesian Holy Quran translation. In *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)* (pp. 86–91). IEEE.

https://doi.org/10.1109/ISMODE53584.2022.9743132

Rizqullah, M. R., Purwarianti, A., & Aji, A. F. (2023). QASiNa: Religious domain question answering using Sirah Nabawiyah. In *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*. IEEE. https://doi.org/10.1109/ICAICTA59291.2023.10390123

Rossano, M. J. (2001). Artificial intelligence, religion, and community concern. [Unpublished manuscript].

Shahzadi, N. (2012). Semantic network based classifier of Holy Quran. *International Journal of Computer Applications*, *39*(1), 39–45.

Tschuggnall, M., & Specht, G. (2016). From plagiarism detection to Bible analysis: The potential of machine learning for grammar-based text analysis. In *Proceedings of the International Conference on Theory and Practice of Digital Libraries* (Vol. 9853, pp. 245–248). Springer. https://doi.org/10.1007/978-3-319-46131-1_27

Tufail, S., Riggs, H., Tariq, M., & Sarwat, A. (2023). Advancements and challenges in machine learning: A comprehensive review of models, libraries, applications, and algorithms. *Electronics*, *12*(8), 1789. https://doi.org/10.3390/electronics12081789

Usman, M., Mendes, E., Weidt, F., & Britto, R. (2014). Effort estimation in agile software development: A systematic

literature review. In *Proceedings of the 10th International Conference on Predictive Models in Software Engineering* (pp. 82–91). ACM. https://doi.org/10.1145/2639490.2639503

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *arXiv*. https://doi.org/10.48550/arXiv.1706.03762

Zhao, H., & Liu, J. (2018). Finding answers from the Word of God: Domain adaptation for neural networks in biblical question answering. In *Proceedings of the IEEE International Conference on Big Data (Big Data)*. IEEE.