MDPI

Article

Challenges of Using Artificial Intelligence in the Process of Shi'i *Ijtihad*

Hasan Latifi

School of Religion, Queen's University, Kingston, ON K7L 3N6, Canada; hasan.latifi@queensu.ca

Abstract: This article aims to explore the potential challenges that may arise when employing generative AI models in the process of Shi'i *ijtihad*. By drawing upon academic literature and relevant primary sources, the essay surveys the most critical AI-related hurdles in this field, including issues of accessibility, privacy concerns, the problem of "AI hallucination" and the generative nature of AI models, biases in AI systems, the lack of transparency and inexplicability, the intricacies of interpreting and understanding sensitive topics, accountability, authority, trust and acceptance among lay believers. Using discourse and content analysis as method, the article concludes that, given these challenges, generative AI models are not yet suitable for utilization in this process. However, the rapid progress in AI may eventually make it an effective tool for this purpose.

Keywords: Shi'a Islam; *ijtihad*; generative AI; authority of AI; transparent AI; AI hallucination

1. Introduction

The death of an Iranian woman in September 2022 sparked a prolonged period of unrest in Iran and triggered a profound debate on the mandatory nature of hijab within Islamic law. While the country's authorities argue that the Islamic government must uphold Islamic values, including the requirement of hijab, in order to be considered truly Islamic, a faction of intellectual Muslim dissenters has sought to demonstrate that hijab should not be exclusively interpreted as a set of dress code regulations.

Throughout the history of Shi'a Islamic law, there have been instances of significant and minor alterations in the rulings, indicating that change is not an inconceivable notion ('Abediyan 1381). The primary catalyst for such change lies in the comprehensive understanding of factors that influence religious decrees, ranging from the fundamental sources like the Qur'an and hadith to detailed historical accounts and biographical reports concerning the narrators of hadiths. Finding an individual with expertise across these diverse fields of knowledge, which contribute to the deduction of religious rulings, is challenging and requires decades of dedicated study and research. Even with such knowledge, some scholars, like Ayatollah Muballighi, believe that the advent of AI-enhanced software can facilitate discussions and achievements that were previously almost impossible (Muballighi 2022).[1] Given the current advancements in technology, the prospect of developing an artificial intelligence trained on these diverse sources appears feasible.

In relation to the matter of hijab in Islam, I consulted ChatGPT (the GPT-4 version) to inquire about the process of issuing an Islamic ruling on the permissibility of hijab based on Islamic sources. While acknowledging that hijab holds significant importance in Islamic rulings, and cannot be readily disregarded due to the existence of substantial evidence within Islamic sources, GPT-4 stated that the only way to determine the permissibility of hijab is by presenting various interpretations of the relevant Qur'anic verses, a task that is far from straightforward (OpenAI 2023). This response highlights the impressive breadth of knowledge possessed by this artificial intelligence model, despite not being specifically trained on Islamic sources. Nevertheless, it is crucial to consider potential

concerns regarding interpretations of the relevant Qur'anic verses and the issuance of *fatwa*s through such technology.

This essay investigates the potential challenges associated with employing generative artificial intelligence (AI) models in the process of Shi'i *ijtihad*, examining both the positive and negative aspects associated with their use. The issues discussed in this study are explored from two perspectives: first, when the AI model is used as an independent tool (i.e., the only tool) to undertake the *ijtihad* process, and second, when the AI model functions as an assistant to the Muslim jurist. Certain issues arise only when the model is employed independently, while others manifest when used either way.

To accomplish this objective, this essay draws upon the findings of various projects and experiments conducted on the application of AI systems to religious content. While one might expect that the most relevant sources to this research would be those that directly pertain to the use of AI in an Islamic jurisprudential context, it is noteworthy that projects that have successfully implemented working AI models in religious contexts may have more significant contribution to this research. The primary aim of this research is to identify the challenges that may arise if a generative AI model is employed in the process of Shi'i *ijtihad*.

This study employs discourse analysis and content analysis to examine the possibility of using AI in the process of issuing a *fatwa* (*ijtihad*) in the Shi'i jurisprudential school. The research design involves finding projects in any religion that have utilized AI for complicated inferential tasks that can be compared to that of *Ijtihad* in its Shi'i meaning. The ideal projects for this study are those in which an AI model is trained on a specific database of religious corpora.

1.1. What Is Artificial Intelligence (AI)?

AI, or Artificial Intelligence, refers to the field of computer science that focuses on creating intelligent machines that can perform tasks that typically require human intelligence. AI involves developing algorithms and systems that can learn, reason, perceive and problem-solve, similar to human cognitive abilities (Winston 1992, p. 13).

There are two primary types of AI: Narrow AI and General AI. Narrow AI, also known as Weak AI, is designed to perform specific tasks and is limited to those specific domains. Examples of narrow AI include virtual assistants, image recognition systems and recommendation algorithms.[2] General AI, on the other hand, refers to theoretical autonomous systems that possess the ability to understand, learn and apply knowledge across multiple domains, essentially possessing human-level intelligence. However, General AI is still largely a theoretical concept and does not yet exist in practice (Goertzel and Pennachin 2007, p. 1).

In this research, the focus is solely on generative AI. Generative AI is a type of narrow AI and refers to a subset of artificial intelligence techniques that involve generating new content, such as images, music, text or even video, that is original and not directly copied from existing examples. It focuses on creating new data that resembles a particular training dataset in terms of style, structure or other characteristics. Generative AI models are designed to learn patterns and generate outputs that are similar to the data they were trained on. These models typically use machine learning, specifically deep learning, techniques to learn and mimic the underlying patterns and distribution of the training data. They learn to capture the essence of the data and then use that knowledge to generate new samples (Goodfellow et al. 2016, pp. 542–43).

Machine learning focuses on developing algorithms and models that enable computers to learn and make predictions or decisions without being explicitly programmed to develop certain algorithms and models. In machine learning, computers are trained on large amounts of data to recognize patterns, extract insights and make predictions or take actions based on that data. The fundamental idea behind machine learning is to build mathematical models that can automatically learn from data and improve their performance over time. Instead of relying on explicit instructions, machine learning algorithms learn patterns and

relationships within the data, enabling them to generalize and make predictions on new, unseen data (Goodfellow et al. 2016, pp. 98–110).

As some of the most important research in employing AI in religion are done with language models, it is a must to talk about that here. A language model is a type of generative artificial intelligence model that is designed to understand, generate and predict human language. It is trained on large amounts of text data to learn the statistical patterns, relationships and structures of language. Language models are used in a variety of natural language processing (NLP) tasks, such as machine translation, text generation, sentiment analysis, chatbots and more (Russell and Norvig 2021, p. 824). The goal of a language model is to generate coherent and contextually relevant language based on a given input or prompts. It learns to predict the probability of a word or sequence of words based on the context provided by the preceding words in a sentence. This ability to predict the next word or sequence of words allows language models to generate text that is coherent and syntactically correct (Russell and Norvig 2021, p. 824). Language models are typically trained on large corpora of text data, such as books, articles, websites or even entire internet archives. During training, the model learns to assign higher probabilities to more frequent word sequences and lower probabilities to less common ones. This enables the model to generate text that is both fluent and contextually relevant (Russell and Norvig 2021, p. 824). Large-scale language models (LLMs), such as OpenAI's GPT-3, have demonstrated impressive capabilities in generating human-like text and performing a wide range of language-related tasks. They have been used for tasks like text completion, question answering, summarization and even creative writing.

1.2. What Is Shi'i Ijtihad?

There are two major denominations within Islam: Sunni and Shi'a. The most mentioned point that distinguishes these two from each other is the successorship of the Prophet Muhammad. While Sunnis believe that he did not mention anyone as his successor and left the issue with the Muslims, Shi'a maintain that he has appointed Ali as his successor. The title for the successors of the Prophet in the Sunni school of thought is Khalifa (or Calif, Caliph) while the successors of the prophet among the Shi'a are called Imam. It is also noteworthy that what is referred to as Shi'a in this paper are the Twelver or Imami Shi'a. The rationale behind selecting this particular sect amidst other Shi'a sects is due to its prevailing position as the dominant sect within the current global Shi'a landscape. Furthermore, my academic pursuits have involved an in-depth study of Twelver Shi'a tradition, rendering it a significant and relevant aspect of my scholarly endeavors.

Ijtihad is an Arabic word which literally means to "try hard and do whatever you can to accomplish a task or gain something" (Ibn Manzur 1414, p. 133). However, the technical meaning of *ijtihad* has not always denoted the same meaning. Throughout various historical periods and geographical locations, as well as across diverse schools of thoughts, a multitude of interpretations and understandings of *ijtihad* have emerged. Nonetheless, in a general look at these definitions, there are two major understandings of this term. First, or what is known as "*Ijtihad* in the general sense", is the utilization of all efforts and endeavors to obtain a ruling from Islamic sources such as the Qur'an and hadith. This type of *Ijtihad* is claimed to be accepted unanimously by all Muslim scholars from various Islamic denominations (Raja'i 2014, p. 6). The second type, or what is known as "*Ijtihad* in the specific sense", involves employing and accepting "valid conjectures" as evidence for religious rulings in cases where there is no explicit textual evidence (Ibn Qudama al-Maqdisi 1415, p. 141). It is important to note that in the Shi'i scholarly atmosphere the term for issuing a *fatwa* (*ifta'*) and *ijtihad*, are often used interchangeably, as the reader may notice in this essay as well.

What makes Shi'i *Ijtihad* a very good case for employing AI models is that Shi'a scholars do not accept personal opinion (al-*Ijtihad bi al-ra'y*) as a source for acquiring the religious ruling, thus they have to analyze the vast corpora on various fields of study, including detailed discussions and opinions to find the most accurate religious ruling

(Al-Sadr 1419, pp. 155–59). For example, to deduce a religious ruling, a faqih (i.e., a Muslim jurist) must first refer to the Qur'an for possibly related verses, then the jurist should check the numerous works of exegesis to find the most accurate meaning. A faqih should also refer to various Arabic dictionaries and observe the rules of Arabic syntax, morphology and rhetoric to be able to confidently say whether the verse is establishing a legal ruling or is just a piece of moral advice. Next, the jurist should go to the hadiths—sayings, acts and tacit approval of the Prophet Muhammad and the Imams that have been transmitted through the narrators in different generations. In this stage, the chain of narrators should be studied thoroughly through various comprehensive biographical works to make sure that they are reliable narrators. Then the text of the hadith should be studied by referring to commentaries, dictionaries, syntax, morphology and rhetoric.

In both cases (i.e., referring to the Qur'an and the hadiths), the historical context of the issuance of the text must be studied carefully and it should be determined whether that context can narrow down the meaning of the text or whether the text is laying out a general principle that can go beyond that historical incident. With regard to issues that have not been addressed either in the Qur'an or in the hadith, the Shi'i jurist relies on what are called "valid conjectures", which are the general rules that were obtained through sayings, acts and/or mainly tacit approval of the Prophet Muhammad and Imams. There are of course extensive debates on which kind of these conjectures are valid and in what condition and with what characteristics. Therefore, the opinions of almost all the jurists, throughout approximately one thousand two hundred years after the time of the Imams should be checked. These opinions have also an important role in understanding the verses of the Qur'an and the hadith.

All in all, mastery of more than 12 scholarly disciplines is needed during the process of *ijtihad*. Needless to say, it is a very difficult task and that is the reason there are only a few Shi'a jurists capable of performing *ijtihad* at a particular time.

1.3. What Artificial Intelligence Can Do in Ijtihad?

1.3.1. The Current Ijtihad

The current way of practicing *ijtihad* among Shi'i jurists has not very much changed from the traditional way. Although the introduction of some narrow AI projects, such as finding similar hadiths with different wordings (Computer Research Center of Islamic Sciences 2014b), exploring and applying various syntaxes to the text (Computer Research Center of Islamic Sciences 2014c) and determining the identity of narrators with the same names (Computer Research Center of Islamic Sciences 2014a), have helped the jurists meaningfully during the last decade, the main process is still practiced by the jurist.

Based on the complexity of the issue and availability of the sources, issuing a *fatwa*, (i.e., a religious ruling about a detailed issue) can take hours to days. Time is not the only thing that matters here, in some cases due to the vast amount of sources and fields of study related to the process, some important points may be overlooked. Moreover, a jurist's opinion on one issue can change during the course of his life, which is very common.

1.3.2. The AI-Enhanced Ijtihad

Employment of artificial intelligence in the process of *ijtihad* can bring about several changes the least of which is efficiency. An AI model that is finely trained on the above-mentioned sources can synthesize and analyze them in seconds.

The most important change that might be hoped to happen by utilizing AI in this field is reformation of some religious rulings. Even if some do not agree that Islam is a law-based religion, no one can deny the significance of laws and rulings in Islam. Therefore, reformation in Islamic laws can cause important changes in this religion. The possible reformations that can be initiated by AI in *ijtihad* are the following: more comprehensive data collection and analysis, deeper study of the historical background of not only the Qur'anic verses and hadith but also the *fatwa*s issued by other jurists, tracing the chains of narrators to find more details of the narrators and possibly narrators who had authored the early books on hadiths, and

finding and potentially resolving the inconsistencies within the Islamic jurisprudence (fiqh) and the principle of jurisprudence (usul al-fiqh) (Muballighi 2022).

Nonetheless, the development of an artificial intelligence model for application in the process of *ijtihad* raises notable concerns. Lack of creativity among students, disconnection of researchers from the sources, lack of deep and profound contemplation on the sources, weakening of deduction (*istinbat*), analysis and theorizing ability among scholars and deviation from the ultimate purpose of jurisprudential discussion (which is proximity to God) are some the important concerns expressed by famous Shi'i scholars about using AI in the process of *ijtihad* (ShabZendeDar 2023; Rajabi 2020). it is imperative to acknowledge that these concerns were not voiced in outright opposition to employing AI in this domain. On the contrary, during the same sessions, these scholars underscored the necessity and potential benefits of utilizing AI in *ijtihad*. However, they also exercised caution, warning against possible drawbacks that must be proactively addressed and mitigated. Such awareness underscores the importance of approaching the integration of AI in *ijtihad* with a balanced perspective, carefully weighing its advantages against the potential challenges it may introduce.

This research pursues two overarching objectives. Firstly, it seeks to draw Shi'a Muslim researchers and scholars' attention to the challenges involved in employing AI in Shi'i *ijtihad*. It strives to dispel the oversimplifying notion that the utilization of AI in this context is a straightforward matter, as some researchers have assumed (Ostadi 2023), and to demonstrate that the application of AI in this field is fraught with crucial and formidable issues that demand consideration. Secondly, the essay endeavors to make a contribution to one of the most fundamental discussions surrounding the implementation of AI in the process of *ijtihad*; the potential challenges of such implementation. The topic of AI and *ijtihad* has been progressively garnering heightened attention from scholars and practitioners alike. By delving into this discourse, the study aims to enrich the ongoing conversation and, thus, the scholarly landscape on this critical subject matter.

2. Challenges of AI in *Ijtihad*

This is the main part of this research, focusing on the challenges that a generative AI system may encounter when applied to the process of Shi'i *ijtihad*. The following text aims to enumerate and delve into these challenges in detail. Each challenge will be scrutinized in terms of whether it arises solely when using AI independently,[3] as an assistant, or in both scenarios. Additionally, potential solutions to overcome these challenges will be proposed after the study of how each aspect could impede the successful integration of AI in the *ijtihad* process. An essential aspect to bear in mind when engaging with this part is the inherent interconnectedness of the discussed challenges, leading to their mutual influence, interdependence and, in some cases, overlapping implications.

2.1. Accessibility

The Internet is a phenomenon that has changed almost all aspects of our life, and it could be the biggest turning point in the whole history of human life. The advent of the Internet stands as one of the most crucial technological advancements of the last century. It has also paved the way for the development of numerous other inventions. One of its most notable benefits is the easy access to content available on the web, anytime and from anywhere on the globe, as long as you have an internet connection. This accessibility applies to most AI projects as well, entailing similar advantages and challenges that the Internet offers. Furthermore, even digital projects not available online still provide much more accessibility compared to traditional methods of finding and analyzing data.

The accessibility of AI projects pertaining to religion can be examined from various angles. Firstly, these projects are available at any time, offering convenience and availability round the clock. For instance, AI providing pastoral care (Young 2022, pp. 6–22) can be accessed even during late hours, when reaching out to a physical pastor or other religious leader might be challenging. Similarly, AI projects like *Virtual Ifta'* in Dubai (Tsourlaki 2022,

p. 12), offering answers to religious inquiries, are accessible 24/7, providing continuous support. On the other hand, engaging with AI for spiritual guidance or seeking answers is also time-efficient. There is no need for individuals to physically go anywhere or wait for a service, reducing the time consumption. AI allows for prompt responses and assistance, making it convenient for those seeking religious guidance or answers to their queries. Many of these AI supports are free of financial cost to those with a digital device and internet access.

The second aspect of accessibility is linked to location. AI utilized for conducting religious rituals, ceremonies or providing spiritual care and comfort can, in many cases, be accessed from anywhere provided that internet access is available. This includes remote villages nestled behind mountains, religious communities in the diaspora and even challenging locations like battlefields and intensive care units in hospitals. The presence of AI enables access to religious services and support regardless of physical distance, ensuring that individuals in various locations can benefit from such assistance.

Furthermore, AI has the potential to enhance the accessibility of content. While simple literal searches may not require artificial intelligence, scholars often encounter information expressed in different words or phrases. In such cases, AI can play a significant role in making content more accessible to users by assisting in finding relevant information even when phrased differently. Additionally, AI-enhanced software can analyze large datasets faster, making big data more accessible to researchers and expediting their work. Cost is indeed an essential aspect of accessibility. With the availability of devices with internet connectivity and decent internet connections, accessing religious content or services generally requires little to no additional cost.

However, it's crucial to recognize that there is a negative side to this accessibility. While internet access may be taken for granted in urban centers of developed countries, it remains a significant challenge in some nations. Many underserved communities, particularly in rural or developing areas, lack the necessary infrastructure and resources to access AI-powered applications and services. This disparity creates a digital divide, hindering the potential benefits of AI in these regions.

To address this issue, collaborative efforts from governments, non-profit organizations and private companies are necessary. They should work together to expand broadband coverage and provide affordable access to technology, thereby bridging the gap and ensuring that these communities are not left behind. The lack of access not only results in the underrepresentation of these regions in the AI landscape but also contributes to AI models' biases. Biases can emerge in AI systems due to skewed data, and if certain demographics or regions are excluded from the training data, it can lead to biased AI models. Biased AI is another critical challenge that needs to be addressed to ensure that AI is fair, inclusive and beneficial to all.

Another challenge posed by high accessibility to AI services is the potential fading of the role and significance of religious communities. What is the main purpose of a religious community? One of the most prevalent motivations for joining a religious community is to connect with fellow believers, receive support and empathy, deepen the knowledge of the religion, participate in rituals and more—almost all of which can be found in some form through AI. Increased accessibility means an increased threat to the position of traditional in-person religious communities, which are a vital aspect of religion even in its modern form.

Another challenge related to accessibility is language and cultural barriers. AI applications often rely on natural language processing (NLP) to interact with users. However, language and cultural diversity pose challenges in developing inclusive AI interfaces. Many languages, especially indigenous and lesser-known ones, lack sufficient NLP support, limiting access to AI-driven services for speakers of these languages. To overcome this barrier, AI developers must prioritize multilingual support and invest in research to include underrepresented languages and dialects.

The final accessibility challenge we explore here pertains to the complexities of regulation and law. The dynamic landscape of AI regulations poses a significant hurdle to accessibility. Varying rules and restrictions across countries can impede the smooth development and deployment of AI. An impactful example of this challenge emerged when I relocated from Canada to my home country, Iran, and attempted to use ChatGPT. While accessing the website and using it posed no issues in Canada, in Iran, a disheartening message appeared at the center of the page, stating, "unable to load the site". Though some claim the Iranian government has banned this service (Ishaq 2023), the truth lies in the restriction of Iran's IP addresses due to sanctions (Naragh 2023; Borhani 2023). There were more than 300 websites that could not be accessed by Iran's IP due to sanctions in July 2020 and the list has been growing since (Borhani 2023). Adding to the frustration, I discovered that even registering on the OpenAI website (which is the provider of the ChatGPT service) proved impossible in Iran due to the non-acceptance of Iranian phone numbers for authentication (Naragh 2023). As I discussed the remarkable capabilities of this Natural Language Processing (NLP) model with my friends, I couldn't help but feel the privilege of my access. Regrettably, such discriminatory barriers to accessing AI services have fostered misconceptions about AI, fueling various conspiracy theories surrounding its use and implications.

As is evident, the challenges related to accessibility can jeopardize both the independent and assistant applications of AI software in the process of *ijtihad*. The solutions to these challenges vary accordingly. In some aspects, individuals themselves must take the initiative to overcome obstacles, particularly those related to language barriers. While employing translation AI services could potentially mitigate the problem, it is essential to acknowledge that these services also present their own set of challenges. In certain cases, these challenges might even exacerbate the issues related to language barriers. On the other hand, certain accessibility challenges require the intervention of governments and/or other authorities, who possess the ability to mitigate issues through various measures, such as developing infrastructure or implementing policy changes.

2.2. Bias

Despite the various potentials of AI to enhance efficiency and accuracy, AI systems are not immune to bias. The primary consequence of biased AI in religion is the distortion of the interpretation of sacred texts and religious sources. This outcome raises concerns about the accuracy and integrity of the insights provided by AI systems within religious contexts. Such bias can lead to unfair and discriminatory outcomes, perpetuating existing societal inequalities and even giving rise to new ones, thereby potentially deepening divisions among various groups. This poses a significant challenge to the authenticity of *ijtihad* conducted by an AI model. There are at least four primary causes of bias in AI: first, the utilization of biased or unrepresentative datasets for training the AI model; second, intentional or unintentional algorithm designs; third, the lack of diversity in AI developing teams, which may lead to overlooking potential sources of bias; and fourth, the human-centric data collection, which implies that AI systems are often trained on data reflecting human behavior, thereby requiring them to learn and replicate this behavior, some of which may be inherently biased (Kantayya 2020). All of these causes of bias pose significant threats to the impartiality of the outcome of the AI model used in the process of *ijtihad*.

In the context of AI and religion, one should be aware of at least two instances of biased artificial intelligence. The first pertains to facial recognition technology, as also brought up in Kantayya's movie, *Coded Bias* (Kantayya 2020). Because algorithms used in facial recognition technology are predominantly trained on data featuring individuals who do not wear religious head coverings, such as hijabs or turbans, this technology is less accurate in identifying those who wear such head coverings, resulting in biased outcomes against individuals who do. On numerous occasions, I have observed that the camera on my mobile phone has encountered difficulty in identifying the facial features of my wife

while she is adorned in a hijab; yet, upon her removal of the hijab, her facial features are immediately detected, even at non-frontal angles.

Another pertinent example, which also underscores the deleterious impact of AI bias, is my interaction with ChatGPT. It is well-known that there are two predominant Islamic sects, namely Sunni and Shi'a. Given that the majority of Muslims identify as Sunni (approximately 90%) (Cavendish 2010, p. 130), and that many Shi'a texts have not been translated into English, the corpus of information that is readily available on Islam is primarily based on the Sunni school of thought. Regrettably, the vast majority of my inquiries to ChatGPT, across various topics, were met with Sunni-centric perspectives. For instance, the term "*ijtihad*" has divergent connotations in the Sunni and Shia traditions; however, ChatGPT appears to lack recognition and knowledge of this distinction, as its response to my inquiry, "What does *ijtihad* mean in Shia?" yielded the following answer: "In Shia Islam, *ijtihad* has a similar meaning as in Sunni Islam..". Other instances of this nature, pertaining to Islamic history and doctrinal intricacies, are also discernible.

The employment of biased AI systems in the process of *ijtihad* can lead to negative implications, encompassing the following aspects:

1. Discriminatory outcomes that do not truly reflect what many understand as the intention of the religion. These outcomes may fail to align with the spirit of the faith and its principles.
2. Reinforcement and perpetuation of existing stereotypes, (such as the unfriendly attitude toward the followers of other sects or those who have failed to observe a certain religious rule) which jeopardizes one of the fundamental goals behind employing AI in this field, which is bringing about the reformation from within the Islamic jurisprudence.
3. Exclusion of marginalized opinions and scholars, contrary to the motivation of inclusivity and studying all available perspectives that come with using AI in the *ijtihad* process. Biased AI can undermine the essence of open exploration and consideration of diverse viewpoints.
4. Perhaps the most evident implication of biased AI is the loss of trust. The discovery of bias in AI can erode public trust in AI technologies and their developers. Users may become hesitant to interact with AI systems, hindering their widespread adoption and potential benefits. In the following section, the issue of trust will be discussed in detail.

It is, therefore, essential to address these concerns and work towards creating an AI system for *ijtihad*, that is as unbiased as possible to foster trust and embrace the true potential that AI offers in this field. The task of eliminating all biases from AI systems is a challenge that is on the verge of impossibility. Nevertheless, there are several steps that can be taken to diminish and alleviate such biases. These measures include, employing diverse databases, ensuring that the datasets used to train AI systems are representative and inclusive of various demographics and perspectives, identifying and modifying algorithms or datasets; actively addressing and rectifying any identified biases in algorithms or datasets aiming to minimize their impact on AI outcomes; engaging a diverse pool of developers, promoting diversity within the development teams or in ethical terms, co-design or participatory design (Mercer and Trothen 2021, p. 58), can lead to greater awareness of potential biases and foster more inclusive AI system designs; implementing ongoing monitoring of AI systems, regularly monitoring AI systems helps to prevent the gradual development of biases over time and ensures that they continue to perform fairly and accurately. By proactively implementing these steps, we can work towards building AI systems that are more equitable and unbiased, contributing to a more just and inclusive future.

The issue of the influence of prompts on the outcomes of generative AI models, especially NLP models, is of paramount importance and falls under the broader challenge of bias. The prompt is the initial input or instruction provided to the AI model, and it plays a significant role in shaping the generated response or output. The prompt serves as a guide for the AI model, helping it understand the context and purpose of the task it

needs to perform. AI models, especially language models like GPT-3 and similar models, are highly sensitive to the wording and structure of the prompt. Even small changes in the prompt can result in vastly different responses. The same AI model can generate opposing answers to a question based on slightly different phrasing in the prompt. The sensitivity of AI models to the prompt can indeed contribute to bias in their outputs. When a prompt contains biased language or reflects biased assumptions, the AI model may generate responses that perpetuate or amplify the underlying bias in the data. The internet is teeming with webpages containing "prompt tricks" or "prompt cheats" designed to elicit various responses—even those restricted by developers to reduce bias—from AI models like ChatGPT.

Despite the presence of prompt-related challenges in both AI-driven and Shi'a scholars' interactions, there are notable distinctions between the two. Firstly, AI models exhibit a heightened sensitivity to prompts, surpassing that of human scholars. Scholars, being immersed in society and exposed to diverse contexts, possess a deeper understanding of, or can infer, the underlying intent behind a question. Secondly, prominent Shia scholars, vested with the authority to issue *fatwa*, are supported by a cohort of researchers and occasionally scientists, who aid in minimizing the impact of prompts on the *fatwa* issuance process.

Sensitive Topics

Another challenge of AI models in a religious context, related to bias, is how to handle religiously sensitive issues. Insufficient data on sensitive issues can result in biased evaluation and judgment, potentially causing emotional distress among lay believers within a religious context. Controversial matters have existed in every religion, sparking debates and sometimes even conflicts. These issues range from historical details to modern matters, including LGBTQ related issues, abortion and the hijab. Developing a publicly accessible AI that can address these issues without offending the sentiments of the followers and avoiding conflicts or divisions is a highly complex task. This is the primary reason why certain AI projects, such as the *Digital Jesus* project, are not yet available to the public. This task becomes even more challenging in the context of finely-tuned AI projects, where artificial intelligence systems are trained on specialized databases. For instance, HadithGPT is an AI model that was specially trained on a database consisting of 40,000 hadiths derived from the six most authoritative Sunni hadith collections although its latest version was relatively accurate, was forcefully rejected by some Muslims due to what was perceived as "clearly incorrect" responses on religiously sensitive matters (Chowdhury 2023).

Another employment of AI in religious practices that can raise a sensitive issue is the possibility of AI occupying the position of highly revered figures in a particular religion. Throughout the early stages of prominent world religions, pivotal figures who underwent a specialized process assumed responsibility for religious acts of worship, rituals, management of religious communities and most importantly *ijtihad* as the pinnacle of Shi'a Islam authority. Traditionally, going to the religious scholar's house or meeting with him in person in a mosque was a sign of reverence and respect. Even the Prophet Muhammad has been quoted as saying that "looking at the face of an 'alim (scholar)... is an act of worship". Although this could be well interpreted as an encouragement of participation in scholarly circles and seeking knowledge, some still follow the literal understanding of this hadith. Hence, it is entirely comprehensible that certain followers may feel uneasy or refuse to accept the placement of AI in the positions traditionally held by these religious figures.

2.3. Privacy

Another aspect concerning the use of AI in issuing *fatwa* is the aspect of privacy enjoyed by users when accessing religious content. Through AI projects, users can pose private questions, share personal aspects of their lives that they may not feel comfortable discussing with others or inquire about sensitive topics they might be ashamed of. A

noteworthy instance of this is the algorithm of an AI model capable of answering Islamic rulings related to the menstrual cycle ('Alam-Huda'i and Shahbazi 2020, pp. 549–66)—a matter that some women may find uncomfortable discussing especially when a female scholar is not available. Additionally, the use of chatbots providing comfort and empathy to individuals facing challenging times in their lives offers solace for those hesitant to share their struggles with others due to social implications or other concerns (Loewen-Colón and Mosurinjohn 2022, Young 2022). In such situations, AI can prove to be a valuable although limited resource.

On the other hand, it goes without saying that privacy has always been a significant concern for anything conducted online or for apps that collect users' data. A recent example of privacy violation involved the use of data from period-tracking and pregnancy apps to persecute those suspected of having an abortion (Masunaga 2022). An AI system for issuing *fatwa* is no exception in this regard. Collecting data on frequently asked topics in each region and the phrasing of questions are some of the basic data that can be collected, potentially violating the user's privacy.

2.4. Generative AI

Generative AI models are designed to produce new data resembling a given training dataset. This creativity is an attractive force that draws people towards generative AI. For instance, in the context of NLP models trained on a vast corpus related to Jesus, scholar Randall Reed has been developing an AI that can generate responses that, while not being the exact words of Jesus, "sound like the Jesus in the Gospels" (Reed Forthcoming). The ability of generative AI to establish constant and multiple connections between different parts of the dataset is a feature that holds promise for revolutionizing *ijtihad* (Fazil Lankarani 2023). However, it is crucial to acknowledge that there are also potential consequences of generative artificial intelligence that may have negative impacts on the *ijtihad* process.

There are two important challenges related to the generative nature of AI models, which hold the potential to revolutionize Shi'i *ijtihad*. The first issue lies in the randomized responses of Generative AI models, even in finely tuned versions. In other words, the same question can yield more than one answer, not only differing in wording but, more importantly, in content. For instance, in Reed's *Digital Jesus* project, at least three responses were generated for each question. In some cases, these responses bore no resemblance to each other. For example, when asked about the greatest commandment, in one instance, Digital Jesus responded with the same response as Jesus, "The one about loving God with all your heart, soul, and mind", while in another, it stated, "The best is 'Listen, and you will be given wisdom'" (Proverbs 9:4) (Reed Forthcoming). This challenge is also evident in other NLP models like ChatGPT and HadithGPT. I have had multiple experiences with HadithGPT where the same question yielded entirely different responses. For instance, when I asked, "Among the wives of the Prophet, whom did he love the most?", I received different names each time AI generated a new response (Hadith GPT 2023).

While it is common for jurists to undergo changes and alterations in their legal opinions, it is important not to equate or confuse this process with the generation of new responses by generative AI. The primary reason for this distinction is that the evolution of a jurist's legal opinion arises from shifts in understanding or access to additional data, often requiring a significant amount of time. On the contrary, when it comes to generative AI, users can be certain that, within a minute, nothing has changed in terms of the sources or analysis of the AI model. The emergence of new responses in generative AI is simply a result of the generative nature of such AI models. Moreover, the variation in responses from an AI model is perceived as inconsistency, since different users can receive different answers to the same question simultaneously. On the other hand, when a jurist issues a modified *fatwa*, it does not imply inconsistency, as it aligns with coherent and consistent data serving as the basis not only for that specific *fatwa* but also for all other *fatwa*s issued by the same jurist.

2.5. AI "Hallucination"

The second challenge that is related to the generative nature of these AI models is AI hallucination. It refers to a phenomenon in which artificial intelligence systems, particularly language models like GPT-3, generate outputs that appear entirely believable and well-grounded in reality, but in fact, have no basis in reality. These hallucinations can be in the form of text, images or even audio generated by AI models. The inherent characteristic of language models is trying to create plausible-sounding responses without actual understanding or knowledge of the context (Athaluri et al. 2023, p. 1). Due to their immense size and training on diverse datasets, these models might produce outputs that appear to be creative or hallucinatory, often by combining unrelated concepts or generating fictional narratives.

There are numerous examples of AI hallucinations to the point that anyone who has asked questions to an AI model like ChatGPT has likely encountered a few instances. Personally, I have witnessed ChatGPT generating responses that were entirely fabricated. For example, when I inquired about the book *Strange Rites: New Religions for a Godless World*, it provided a summary of the book. Seeking more accuracy, I specified that I meant the one written by Tara Isabella Burton. In response, it apologized and generated another abstract of the book. I then asked if it could provide a summary of each chapter, and it confirmed its ability to do so. However, the titles of the chapters and their content were completely different and also incorrect. I provided additional information, mentioning the book's publisher. Once again, it apologized and provided summaries of each chapter, this time with new titles, none of which matched the book I had in front of me. This process repeated for the third time, and once more, it generated an entirely new book with no connection to the published one. Such instances highlight the challenges posed by AI hallucination and underscore the need for further refinement in AI models to ensure more accurate and reliable responses.

An intriguing example closely related to our topic is the one that occurred in the *Digital Jesus* project. When asked about the greatest commandment, in the first attempt, Digital Jesus responded with the same answer as Jesus, "The one about loving God with all your heart, soul, and mind", but in the second attempt, it provided a response, "The best is 'Listen, and you will be given wisdom' (Proverbs 9:4)". However, Proverbs 9:4 does not contain such a commandment in the Hebrew Bible. Still, the response was articulated in a way that someone unfamiliar with Christian tradition (or even familiar with Christian tradition but not have scripture memorized) might accept as valid (Reed Forthcoming). It is for such cases that for differentiating between hallucinations and reality in the process of *ijtihad*, one must be an expert in all the necessary fields of study required for *ijtihad*, and even someone with such expertise must refer to the sources to verify the generated content.

The section highlighted various challenges that significantly impact the accuracy of AI models utilized in the process of *ijtihad*. These challenges pose substantial obstacles to achieving reliable and precise results in AI applications. By acknowledging and addressing these issues, researchers and developers can strive to enhance the performance and credibility of AI systems. They are continually refining AI models to minimize these hallucinatory responses and to enhance the control and precision of the generated content. As AI technology evolves, it's likely that the capabilities of language models will improve, leading to more accurate and contextually appropriate responses while reducing hallucinatory outputs.

2.6. Authority

The concept of authority in Islam, including among Shi'a, differs significantly from that in some streams of Christianity. Unlike Roman Catholic Christianity, which has a hierarchical structure with authority flowing from the top, Islam does not follow such a system. The question of authority holds immense importance, as the outcome of the *ijtihad* process is believed to be a "ruling in accordance with divine revelation"—a crucial criterion observed in every Shi'i *fatwa* (Sheikh Anṣārī 1404, p. 303). It is also worth noting that the

challenge of authority becomes more pronounced when an AI system is independently used to derive *fatwa* from its sources. However, when AI is employed in a more modified and accountable role as an assistant for the jurist, the authority can be preserved through the presence of the jurist in the process and their supervision over it. This way, the jurist can maintain their role in ensuring the legitimacy and accuracy of the derived rulings. It is noteworthy that there is a growing body of scholarly works that argue for AI's role solely as an assistant in religious matters (Trothen 2022b, 2022a).

There are two principal ideas about the main source of religious authority. Traditionally, and as believed by many religious people from different Abrahamic religions, authority is considered to come from God. For religious statements to hold value and significance, evidence of divine appointment is typically required, often through a complex hierarchical structure, such as what is seen in Catholic Christianity. However, in the modern world, particularly after the Protestant Reformation, some believe that authority can also originate from the adherents of a religion. For example, a Muslim imam, whose community has accepted his authority, may not need an institution or a higher-ranked scholar to validate his position. It is important to note that the discussion surrounding the authority of AI primarily arises in the former situation rather than the latter. Based on the second interpretation, it is entirely plausible that a group from any religion or even without a specific religious affiliation may accept the authority of an AI system, potentially forming a new denomination or even a religious movement, like the first church of artificial intelligence, known as The Way of the Future (WOTF), which was established in late 2017 and closed in early 2021 (Harris 2017; Korosec 2021).

The process of gaining religious authority in Shi'i jurisprudence is deeply rooted in tradition, even in countries like Iran where a Shi'a Islamic government holds power. In the Shi'a tradition, achieving such authority involves embarking on a rigorous path of studies in various fields related to *ijtihad*, followed by obtaining written permission from one or more top living jurists. This chain of permissions traces back to the Imams of Shi'a Islam who lived during the 8th and 9th centuries. The "Permission of *Ijtihad*" (*Ijazat al-ijtihad*) signifies that the holder possesses the capability to deduce *fatwa*s from their sources by skillfully applying the necessary fields of study.

However, a crucial question arises: Can an AI model be given such permission? To answer this, one must understand the requirements and the process by which this permission is granted. Interestingly, there is no official or definite procedure for obtaining this certificate; rather, it mainly relies on the trust and confidence of higher-ranked scholars (*mujtahids*, qualified jurists who practice *ijtihad*) in the individual seeking this permission. The most common path to gaining this trust involves a student actively participating in the lectures of a top scholar for several years, demonstrating exceptional performance, judgment, reasoning and a profound understanding of the sources necessary for *ijtihad*. Other methods, such as extensive discussions, may also serve as a detailed test of the student's capabilities, ultimately earning that sought-after trust. Considering this, it may not be entirely impossible for an AI model to receive this certificate if it can garner the trust of a *mujtahid*. However, the possibility of AI obtaining such a certificate takes a backseat to the larger discussion of whether being human is a necessary criterion for engaging in *ijtihad*. This last point reminds me of a theological debate in Christianity, arguing that believing human beings were created in the image of God does not necessarily imply the absence of this feature in other creatures, according to some interpretations (Mercer and Trothen 2021, pp. 222–23).

2.7. Trust

With the emergence of "deepfake", the issue of trusting any content on the web has entered a new and concerning phase. Deepfakes, a combination of "deep learning" and "fake", refer to hyper-realistic videos that are digitally manipulated to depict people saying and doing things that never actually happened. These deceptive videos are challenging to detect, as they use real footage, can have authentic-sounding audio and are optimized to

rapidly spread on social media platforms (Westerlund 2019, p. 40). While deepfake may not directly impact the trust issue within the realm of using AI in the process of *ijtihad*, since *fatwa*s are almost always expressed in written form rather than orally, it serves as a pertinent example of how certain AI models can be employed for intentional deception. The prevalence of deepfake technology has contributed to the erosion of trust in AI, as individuals become increasingly cautious about the authenticity of digital content.

The issue of trust is also a crucial consideration when an AI model is employed for practicing *ijtihad*. This raises the question of whether scholars can place their trust in an AI system, which, in turn, leads us to a broader discussion about whether lay Muslims would entrust their faith to AI. Two cases, Virtual Ifta' in Dubai and "Al-Azhar Fatwa Global Centre" in Egypt, along with a survey about the same question, shed light on the issue of trust when using AI in the process of issuing *fatwa*s or *ijtihad*. The first project, Dubai's Virtual Ifta', made its debut in October 2019 during a three-day exhibition dedicated to launching "the world's first AI *fatwa* service". However, the AI model utilized in this service was non-generative. Upon entering a question, the user would receive multiple similar questions to choose from, and the corresponding answer to the chosen question would then appear (AP Archive 2019). Less than three months later, in January 2019, the Al-Azhar Fatwa Global Centre in Cairo announced its own AI *Fatwa* System. However, the system is yet to be operational as the "team of special intelligence [...] are still collecting data to support the system" (Tsourlaki 2022, p. 13). These two cases and the survey data provide valuable insights into the issue of trust concerning AI's role in issuing *fatwa*s. As AI technology continues to evolve, it remains vital to explore and address the concerns and perspectives of the Muslim community regarding the integration of AI into religious practices.

The informative survey conducted by Tsourlaki examined the attitudes of lay Muslims toward AI systems related to *fatwa* issuance. According to the author, "The participants' common characteristics were that they identified as Sunni Muslims, employed the English language in their daily communication and used Facebook. Therefore, they were familiar and comfortable with technology" (Tsourlaki 2022, p. 8). Two notable factors stand out among these common characteristics. First, the participants' familiarity and comfort with technology, as highlighted by the author, are significant. Second, their use of English as their daily language suggests that they may not have a strictly traditional background, a notion supported by 16 percent of participants who stated that they obtain their *fatwa*s from their local imams. Given the unsuitability of imams to serve as *mufti*s, this fact indicates that they do not have a deep knowledge of the exact requirements one should have to be able to produce *fatwa*s (Tsourlaki 2022, p. 12).

It is essential to consider these insights when examining the acceptance and impact of AI systems in the realm of *fatwa* issuance. As AI technology continues to be integrated into religious practices, understanding the perspectives and preferences of lay Muslims becomes crucial for developing effective and trustworthy AI-driven solutions in this context. As a part of the survey, when the participants were asked, "Is it important to know whether the *fatwa* has been created by a human or a computer?" 92.7 percent responded positively, while 7.3 percent stated that they are not concerned about it. On the aspect of *fatwa* issuance by a computer, 96.3 percent stated that they would not trust a *fatwa* that a computer had issued (Tsourlaki 2022, p. 13). The other question, "Why would you trust or reject a *fatwa* issued by a robot?" asked respondents to provide an essay-style answer. The majority of participants expressed a clear rejection of such a *fatwa*. Their rationale centered on human cognitive abilities, including reasoning, compassion and critical thinking, as well as the skill to interpret sources, grasp complex contexts and conduct comparative analyses. The significance of cultural and societal context was also emphasized in their responses.

These findings align with the conclusions of a 2015 Egyptian research study, indicating that lay believers generally tolerate and forgive mistakes made by a *mufti* (jurist who issues a religious ruling (*fatwa*)) unless they significantly disturb the public. However, the same leniency would not be extended to even minor errors made by AI (Elhalwany

et al. 2015, p. 504). Answers to this question reveal a subconscious fear of the unknown of AI's interference with traditional practices within Islam. This fear is evident in one of the responses, which stated, "I will simply reject a *fatwa* because I won't believe a computer when it comes to my faith" (Tsourlaki 2022, p. 19).

The lack of trust in computer-generated *fatwa*s explains why Virtual Ifta' received no response or serious attention and consequently had a short period of activity. The way that the project was launched and introduced its function might also have caused a misunderstanding. Virtual Ifta' utilized a repository of pre-registered *fatwa*s, and the AI aspect was limited to finding the nearest model question to the user's inquiry. However, during the launch ceremony, it was promoted as "The world's first AI *fatwa* service". Furthermore, it is surprising that before users typed their questions, an automated message informed them that "the answers to your questions are generated automatically using AI technology" (Tsourlaki 2022, pp. 12–13). According to the survey conducted by Tsourlaki, had the users known that the project relied on previously issued *fatwa*s, it would have received more attention and acceptance by the target audience (Tsourlaki 2022, p. 18). As of March 2022, no academic publication had engaged with the project, and the media coverage was limited to a few announcements during the launch week (Masudi 2019; Dajani 2019; The New Arab 2019; AP Archive 2019). It seems that Muslims either did not notice the service or rejected using it, leading to the decision by IACAD (Islamic Affairs and Charitable Activities Department in Dubai) to discontinue the project (Tsourlaki 2022, p. 19).

2.8. Acceptance

Another issue that is deeply related to the discussion of authority and trust is the issue of acceptance of artificial intelligence in religious matters. Even after gaining authority, AI systems can suffer from not having acceptance among the lay followers of a given religion. For instance, even if Mindar, the AI Zen Buddhist robot, was endorsed by the authorities at Kodiji temple, there are some Buddhists that still do not welcome this project (DW Shift 2020). Although using AI has not only been approved but also encouraged by most Sunni and Shi'a scholars (Islamweb 2023; Awais 2022; Khamenei 2021, 2023), some lay Muslims are still reluctant to use AI for different reasons including, the notion that only a human can undertake the process of *ijtihad* and issuing *fatwa*, basic and superficial understanding of AI, the problem of having no personal contact and relation with the one who is answering the question (Tsourlaki 2022, p. 19) and the idea that artificial intelligence is an imitation of God's action as it involves creating an intelligent being (Quora 2023).

Some of these rationales are not specific to Islam; for instance, considering AI as an imitation of God's act of creation is also may be viewed as an impediment to AI development in Christianity (DW Shift 2020). Another significant reason for the non-acceptance of AI systems in religious activities is the concern over its usage being sacrilegious. When reporting on Mindar, the first thought that crossed the reporter's mind was, "Isn't this sacrilegious?" (DW Shift 2020) This reaction reveals the subconscious feelings of at least a group of people towards such AI projects. This sentiment can be linked to the concept of "*Wahn*" in Shi'a Islam doctrine. "*Wahn*" refers to anything that makes Islam appear irrational, weak, inferior or insignificant to the public, regardless of their religious affiliation. It is strictly prohibited in Islam, and all Muslims have a responsibility to avoid it (Honarmand 2020, p. 13). In certain cases, the use of AI in religious matters can be perceived as demeaning to the community based on the specific act or service provided by the AI. This notion not only leads to the rejection of AI by laypeople but also has the potential to undermine the authority of AI models. All in all, these concerns illustrate the intricate interplay between technology and religious beliefs, necessitating careful consideration and understanding when integrating AI in religious contexts.

In this section, the lack of personal relationships with scholars was mentioned as one of the reasons why Muslims are reluctant to accept the outcomes of AI models. However, it is essential to recognize that this aspect can also be regarded as an independent issue worthy

of examination. There exist specific instances where a *mujtahid* has issued a tailor-made *fatwa* for an individual, drawing upon the personal acquaintance and understanding of the unique circumstances involved. For instance, in cases where an individual is grappling with an obsession (waswasa in Arabic and vasvās in Persian) related to a religiously mandated action, the *mujtahid* may issue a *fatwa* that such action, while obligatory for the general public, is deemed forbidden for that specific individual (Heidari Naraqi 1388, pp. 133–34). Such personalized *fatwa*s are not rooted in textual sources but rather arise from the jurist's intention to assist the individual in overcoming their obsessive state. Moreover, addressing the queries of an obsessed individual with conventional, established *fatwa*s applicable to the broader community, can exacerbate the obsessive condition. All in all, the generation of customized *fatwa*s through an AI model for such cases is not a straightforward task, as it necessitates considerations beyond mere textual sources, delving into the contextual nuances that can be best comprehended through in-person interactions.

2.9. Unexplainability

The unexplainability of AI models refers to the difficulty in understanding and interpreting the decision-making processes and underlying mechanisms of these models. Generative AI models often lack transparency and interpretability. The lack of interpretability and explainability in generative AI models raises concerns in critical applications, where understanding the decision-making process is crucial (Molnar 2022, pp. 13–14). The process of *ijtihad* may be counted as one of the situations in which transparency plays an important role. This issue applies to both cases of employment of artificial intelligence models in the field of *ijtihad*, i.e., using them as assistants for a jurist and using them as independent tools for deducing *fatwa*s from the sources. Particularly, the interpretability of AI systems gains more importance in the former case, as the jurist needs to know why and how the AI has come to this result to be able to assess them. In the absence of this appraisal, employing the AI model as an assistant becomes devoid of rationale. It is noteworthy that the issue of unexplainability is the most important challenge for using AI models as an assistant, because most of the aforementioned challenges, such as authority, bias and AI hallucination could be overcome by the presence of a jurist next to the AI model, but this issue cannot be solved by this presence.

Unaccountability

The lack of interpretability in AI models gives rise to another significant challenge: unaccountability. *Fatwa*s hold a crucial status in the lives of Shia Muslims, with many instances where individuals have sacrificed their lives in adherence to a *fatwa*. A recent case is Grand Ayatollah Sistani's call for war against ISIS (Reuters 2015), leading thousands of Iraqi Shia Muslims to fight, with many losing their lives in fulfillment of this *fatwa*. *Fatwa*s wield immense power, but when an AI model issues an erroneous *fatwa* resulting in adverse outcomes, who bears responsibility? If these AI models lack interpretability and are solely trained on datasets through machine learning, it implies that developers are not directly involved in the issuance of such *fatwa*s, hence absolving them of accountability. This, in turn, may lead to *fatwa*s for which no one can be held responsible, even in the event of severe consequences.

All in all, researchers are actively working on developing techniques to enhance the transparency and interpretability of these models, but it remains an ongoing challenge in the field of artificial intelligence (Molnar 2022, pp. 13–14).

3. Conclusions

Using trained generative artificial intelligence models in the process of Shi'i *ijtihad* is a fascinating topic and holds great promise for the future. However, getting to the point where AI is truly ready to be utilized in this field is no simple task, given the various types of challenges these AI models must overcome. The primary objective of this research was to identify and examine some of the most significant challenges, highlighting that while

the idea of employing AI in this domain may be alluring and intriguing to both religious scholars and some laity, it must first surmount significant obstacles.

The present research deliberated upon a range of critical issues pertaining to the utilization of trained generative artificial intelligence models in the context of Shi'i *ijtihad*. The discussed concerns encompass limitations of accessibility to this technology, concerns about the privacy of user information, the problem of "AI hallucination" and generative nature of AI models studied in this essay, biases in the training data and algorithms, lack of transparency and inexplicability in judgments and decision making, concerns about the interpretation of sensitive or controversial topics, as well as persistent questions about trust and the authority of non-human generated religious interpretations and legal determinations. For instance, the challenges associated with trust and acceptance present formidable obstacles, deeply rooted in the psyche of lay individuals who are supposed to engage with such AI systems. Effecting transformative shifts in their perspectives may prove quite arduous judging by the survey research cited in the previous part. Another salient matter of concern is the presence of bias in AI, a phenomenon that transcends the sphere of *ijtihad* and extends to numerous other domains, such as facial recognition AI models. Similarly, the aspect of the unexplainability of AI models poses a momentous predicament when employing them in the *ijtihad* process. Simply put, generative AI models are functioning akin to enigmatic black boxes. This issue will eventually render AI models unsuitable for employment in the process of *ijtihad*, where a comprehensive understanding of the reasoning behind issued *fatwa*s is of paramount significance, particularly in scholarly circles. Moreover, the absence of interpretability in AI models undermines their utility as indispensable assistants in the *ijtihad* process, as jurists necessitate a lucid grasp of the interconnections between various topics and the manner in which a particular *fatwa* finds its basis in specific sources.

Some of these challenges are common between using AI in *ijtihad* and other fields where AI is utilized. For example, privacy concerns and limitations in accessibility are pervasive issues that have preoccupied users across various AI services and even non-AI internet-based services. For instance, there are more than 300 websites, offering various AI and non-AI services, that are blocked for Iranian IPs due to sanctions (Borhani 2023). Nevertheless, some of these challenges are more specific to the use of AI models in the process of *ijtihad*, such as the challenge of generating new responses while no changes had happened in the dataset. Moreover, the challenges expounded upon in this research were examined through a dual prism: the utilization of AI models as an independent tool for issuing *fatwa*s and using them as assistants to jurists in the *fatwa* issuance process or *ijtihad*. Certain issues present heightened challenges in the former scenario, whereas others manifest greater complexities in the latter. For instance, as elucidated, matters of authority, trust and acceptance pose more formidable obstacles when an AI model is used as the only tool for practicing *ijtihad*. Conversely, unexplainable AI poses a greater challenge in the deployment of AI as an assistant to jurists. The AI system's inability to provide explanations for its decisions hinders its usefulness in aiding the jurist to arrive at a specific religious decision or *fatwa*. Without a clear understanding of how and why the AI arrived at a particular conclusion and the sources and reasoning behind it, the jurist's confidence in relying on the AI's assistance may be compromised. Furthermore, certain issues, such as privacy and accessibility, present comparable levels of challenge in both deployment cases.

Although the foregoing challenges were enumerated analytically as distinct issues, it is important to recognize that there are important interrelationships and overlaps between these challenges and concerns among scholars and practitioners. In fact, these issues are closely interconnected with each other and their interwoven nature merits due consideration. To illustrate the interrelation of these aspects, it is pertinent to draw attention to the fact that restricted access to AI systems can result in under-representation and eventual bias against certain groups. For instance, as mentioned in the previous part the responses of ChatGPT are more based on a Sunni understanding of Islam as it is the dominant Islamic denomination. Conversely, increased accessibility may compromise privacy, as users are more

inclined to divulge personal information to avail themselves of these services. Moreover, a biased AI holds the potential to render unthoughtful judgments on sensitive matters, as addressing and effectively dealing with such issues requires access to more comprehensive information from specific communities. The aftermath of addressing sensitive issues with such bias may lead to the question of accountability. Moreover, the endeavor to enhance the interpretability of AI models can potentially encroach upon privacy, as the process of interpretation may entail accessing information submitted by users. The interrelation and interdependence among these challenges will further evolve with additional exploration and deeper contemplation and reflection. Accordingly, viewing these challenges as discrete, isolated issues would be an oversimplification, as they are merely presented in list format to facilitate the cohesive flow of discussion in this essay.

What the Future Holds

An important aspect of these challenges is that all of them are issues that current generative AI is facing. These challenges do not necessarily warrant complete abandonment of using AI in this field. Believed by many specialists, AI is still in its nascent stage and has a long way to go (Bostrom 2015). This paper, at best, can only provide a discussion of the current state of AI. Predicting the rapid development of AI is almost an impossible task. Even with the introduction of GPT-4 and its plugins, there are more interesting advancements that demonstrate how AI is overcoming some of the challenges outlined in this study. Ultimately, it is conceivable that AI will be capable of surmounting some of these challenges or at least mitigating them in the future. However, there are questions that remain unanswered: how long it will take and whether new challenges will emerge during this time; or whether certain fundamental dimensions of religious authority will always resist full acceptance of non-human, computer-based religious determinations. Even if these significant challenges are overcome, there are still questions regarding how religious culture will be shaped in new directions through AI's influence on religion or in response to it.

With regard to future studies and research related to the title of this essay, I think there are two topics that need to be given more attention than others. The first topic is a theoretical discussion that is more related to Islamic jurisprudence and that is finding the special criteria of a jurist that cannot be dehumanized. In other words, are there any characteristics essential for the one who practices *ijtihad* that cannot be obtained by an AI model? What are these characteristics and is there any way to make a substitute for them in an AI model that is designed to be used in the process of *ijtihad*? The second proposal for future studies in this field is more practical and that is building or training an AI model for the purpose of using in the process of Shi'i *ijtihad* and getting detailed feedback on the advantages and shortcomings. A project similar to *Digital Jesus*, trained on the vast data that is used in the process of *ijtihad*, starting from the Qur'an and its commentaries, other scholarly fields needed for this process, to numerous books of jurisprudence and its principle authored by the jurists during approximately 12 centuries. The responses of such a project will demonstrate the challenges of using AI in this field more specifically and will give researchers in both computer engineering and Islamic studies a better vision of the pros and cons of employment of AI in the process of *ijtihad*. Although the Najaf project was introduced in 2018 for the purpose of "application of artificial intelligence in Islamic Sciences", such a project on using AI in the process of *ijtihad* has not yet been done or at least been reported or discussed in scientific journals. Needless to say, other than these two topics, there are many other issues related to the title of this research, such as the parts of Islamic jurisprudence that can be revolutionized by the use of AI, the possibility of giving an AI model the permission for *ijtihad* and employment of AI in each scholarly fields that will affect the process of *ijtihad*.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

Notes

1   More details about Ayatollah Mubalighi's views can be found in this paper, specifically at lines #187–196.

2   Virtual assistants: AI-based software that performs tasks, answers questions, and interacts with users through voice commands or text input.Image recognition systems: AI-powered technology that analyzes visual data, identifying objects, people, and activities in images or videos.Recommendation algorithms: AI algorithms that personalize content or product suggestions based on user data and behavior to enhance user experience.

3   Using AI as the only tool to undertake the process of *ijtihad*. In this case, the AI model is provided with the data, including Islamic sacred texts and other sources, for machine learning. Subsequently, when a user poses a question, the AI model generates a full-fledged *fatwa* in response.

References

'Abediyan, Mir Hossein. 1381. پژوهشنامه متن (Influencing fators in changeing the [religious] ruling). عوامل مؤثر در تغیر حکم 16–15: 38–105.

'Alam-Huda'i, Seyyed Muhammad Hasan, and Alireza Shahbazi. 2020. افزاری پاسخگویخودکار احکام بانوان بر اساس الگوریتم مسائل شرعی طرح نرم (The Design of the Automatic Answering Software for Women's Rulings Based on the Algorithm of Islamic Rulings). *Conference of Artificial Intelligence and Islamic Sciences*. Available online: https://www.noormags.ir/view/fa/articlepage/1795250 (accessed on 30 March 2023).

Al-Sadr, Sayyid Muhammad Baqir. 1419. دروس فی علم الاصول (*Lessons in the Principles of Jurisprudence*), 2nd ed. Qom: Majma' al-Fikr al-Islami, vol. 1.

AP Archive. 2019. *AI-Powered Chatbot Gives Muslims Religious Guidance*. Available online: https://www.youtube.com/watch?v=-V4yRuEgaAA (accessed on 15 March 2023).

Athaluri, Sai Anirudh, Sandeep Varma Manthena, V. S. R. Krishna Manoj Kesapragada, Vineel Yarlagadda, Tirth Dave, and Rama Tulasi Siri Duddumpudi. 2023. Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing through ChatGPT References. *Cureus* 15: e37432. [CrossRef] [PubMed]

Awais, Ammar. 2022. Islam on Artificial Intelligence. *Islam Explained*. November 5. Available online: https://islamexplained.info/2022/11/05/islam-on-artificial-intelligence/ (accessed on 27 July 2023).

Borhani, Ali. 2023. List of Sites Which Block IPs Come from Iran [UPDATING... (July 16, 2020)] #SANCTIONS. GitHub Gist. Available online: https://gist.github.com/alibo/dfd7c258bcc44a0e8c9f7c5bfd3bd2c3 (accessed on 5 August 2023).

Bostrom, Nick. 2015. *What Happens When Our Computers Get Smarter than We Are?* Available online: https://www.youtube.com/watch?v=MnT1xgZgkpk. (accessed on 31 March 2023).

Cavendish, Marshall. 2010. *Islamic Beliefs, Practices, and Cultures*. Tarrytown: Marshall Cavendish Reference.

Chowdhury, Muajul I. 2023. Hadith GPT|DARUL IFTAA NEW YORK. March 3. Available online: https://askthemufti.us/hadith-gpt/ (accessed on 28 July 2023).

Computer Research Center of Islamic Sciences. 2014a. Dirayat al-Noor 1.2 Software. December 16. Available online: https://www.noorshop.ir/en/product/6185/software (accessed on 2 June 2023).

Computer Research Center of Islamic Sciences. 2014b. Jami' al-Ahadith 3.5 Software. December 16. Available online: https://www.noorshop.ir/en/product/6320/software (accessed on 2 June 2023).

Computer Research Center of Islamic Sciences. 2014c. Jami' al-Tafasir Software, Comprehensive Commentary Collection 3. December 16. Available online: https://www.noorshop.ir/en/product/12636/software (accessed on 2 June 2023).

Dajani, Haneen. 2019. Virtual Fatwas Delivered in Dubai to Better Guide the Faithful. *The National*. October 30. Available online: https://www.thenationalnews.com/uae/government/virtual-fatwas-delivered-in-dubai-to-better-guide-the-faithful-1.930365 (accessed on 27 July 2023).

DW Shift. 2020. *Mindar: Can a Robot Be Religious?|Buddhist Robot Priest Mindar|Japanese Robots*. Available online: https://www.youtube.com/watch?v=Y3VuHpYPU6Y (accessed on 14 March 2023).

Elhalwany, Islam, Ammar Mohammed, Khaled Tawfik Wassif, and Hesham A. Hefny. 2015. Using Textual Case-Based Reasoning in Intelligent Fatawa QA System. *The International Arab Journal of Information Technology* 12: 503–9.

Fazil Lankarani, Muhammad Javad. 2023. تواند تحول در اجتهاد إیجاد کندهوش مصنوعی می (AI Can Revolutionize the Ijtihad). Available online: https://fazellankarani.com/persian/lecture/24116/ (accessed on 12 July 2023).

Goertzel, Ben, and Cassio Pennachin. 2007. *Artificial General Intelligence*. Berlin and Heidelberg: Springer Science & Business Media.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. Cambridge: MIT Press.

Hadith GPT. 2023. Available online: https://www.hadithgpt.com/ (accessed on 25 March 2023).

Harris, Mark. 2017. Inside the First Church of Artificial Intelligence|Backchannel. *Wired*. November 15. Available online: https://www.wired.com/story/anthony-levandowski-artificial-intelligence-religion/ (accessed on 5 August 2023).

Heidari Naraqi, Ali Mohammad. 1388. های درمان‌وسواس، شناخت و راه (*Obsession, Recognition, and Treatment Methods*). Qom: Meytham Tammar Publication. Available online: https://hawzah.net/fa/Book/View/45317/ (accessed on 8 July 2023).

Honarmand, Ali-Asghar. 2020. واکاوی ادله عقلی و نقلی وهن دن (Analyzing the intellectual and narrative evidences of weakening the religion). فقه و اجتهاد 13: 73–98.

Ibn Manzur, Muhammad ibn Mukrim. 1414. *Lisān al-ʿArab*. Beirut: Dar al-Fikr, vol. 3.

Ibn Qudama al-Maqdisi, ʾAbdullah ibn Ahmad. 1415. الامام احمد بن حنبل روضة الناظر و جُنة المُناظر فی اصول الفقه علی مذهب (*The Garden of the Observer and the Shield of the Debater in the Principles of Jurisprudence according to the School of Imam Ahmad ibn Hanbal*). Riyadh: Maktabat al-Rushd, vol. 2.

Ishaq, Rana. 2023. What Countries Is ChatGPT Not Available In? *PC Guide*. April 30. Available online: https://www.pcguide.com/apps/countries-chatgpt-not-available/ (accessed on 23 July 2023).

Islamweb. 2023. *Using Artificial Intelligence*. Available online: https://www.islamweb.net/en/fatwa/211585/using-artificial-intelligence (accessed on 27 July 2023).

Kantayya, Shalini, dir. 2020. *Coded Bias*. Brooklyn: 7th Empire Media.

Khamenei, Seyyed Ali. 2021. We Should Move on the Path to Making Iran a Source of Science within 50 Years. *Khamenei.ir*. November 17. Available online: http://english.khamenei.ir/news/8767/We-should-move-on-the-path-to-making-Iran-a-source-of-science (accessed on 27 July 2023).

Khamenei, Seyyed Ali. 2023. Significance of Propagation in Post Internet, AI Era. *Khamenei.ir*. July 12. Available online: http://english.khamenei.ir/news/9927/Significance-of-propagation-in-post-internet-AI-era (accessed on 27 July 2023).

Korosec, Kirsten. 2021. Anthony Levandowski Closes His Church of AI. *TechCrunch*. February 19. Available online: https://techcrunch.com/2021/02/18/anthony-levandowski-closes-his-church-of-ai/ (accessed on 5 August 2023).

Loewen-Colón, Jordan, and Sharday Mosurinjohn. 2022. Fabulation, Machine Agents, and Spiritually Authorizing Encounters. *Religions* 13: 333. [CrossRef]

Masudi, Faisal. 2019. Dubai Launches 'World's First' Artificial Intelligence Fatwa Service|Uae—Gulf News. *Gulf News*. October 29. Available online: https://gulfnews.com/uae/dubai-launches-worlds-first-artificial-intelligence-fatwa-service-1.67466584 (accessed on 27 July 2023).

Masunaga, Samantha. 2022. How Data from Period-Tracking and Pregnancy Apps Could Be Used to Prosecute Pregnant People. *Los Angeles Times*. August 17. Available online: https://www.latimes.com/business/story/2022-08-17/privacy-reproductive-health-apps (accessed on 30 March 2023).

Mercer, Calvin, and Tracy J. Trothen. 2021. *Religion and the Technological Future: An Introduction to Biohacking, Artificial Intelligence, and Transhumanism*. Cham: Springer International Publishing. Available online: https://books.google.com.sg/books?id=Db4fEAAAQBAJ (accessed on 30 March 2023).

Molnar, Christoph. 2022. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed. Munich: Leanpub. Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 27 July 2023).

Muballighi, Ahmad. 2022. شودهوش مصنوعی جایگزین فقه و مجتهد نمی (Artificial Intelligence Will Not Replace the Jurist (Faqih) and the Mujtahid). December 11. Available online: https://iqna.ir/fa/news/4106133/ (accessed on 2 June 2023).

Naragh, Mohsen. 2023. ChatGPT Login Access Denied (Iran Ip Address). *OpenAI Developer Forum*. February 17. Available online: https://community.openai.com/t/chatgpt-login-access-denied-iran-ip-address/65139 (accessed on 23 July 2023).

OpenAI. 2023. *ChatGPT: AI Language Model*. San Francisco: OpenAI. Available online: https://chat.openai.com (accessed on 5 April 2023).

Ostadi, Kazem. 2023. آه الله العظمی اآی (Grand Ayatollah AI). *WhatsApp*. June 6. Available online: https://chat.whatsapp.com/C9PEgZyIiDO66yHthasuft (accessed on 4 August 2023).

Quora. 2023. Is Artificial Intelligence Haram in Islam? Artificial Intelligence Is to Create an Intelligent Being and It Means Imitating Allah. Available online: https://www.quora.com/Is-artificial-intelligence-haram-in-Islam-Artificial-intelligence-is-to-create-an-intelligent-being-and-it-means-imitating-Allah (accessed on 27 July 2023).

Raja'i, Mahdi. 2014. بررسی مفهوم اجتهاد (Examining the Concept of Ijtihad). Research Institute of the Guardian Council, January. Available online: https://ccri.ac.ir/files/fa/news_files/12387.pdf (accessed on 2 June 2023).

Rajabi, Mahmood. 2020. بررسی مزایای استفاده از هوش مصنوعی در فرآیند اجتهاد (Examining the Advantages of Using Artificial Intelligence in the Process Of Ijtihad). October 20. Available online: http://rasanews.ir/fa/news/666502 (accessed on 4 August 2023).

Reed, Randall. Forthcoming. *Digital Jesus: An Experiment in Artificial Intelligence*.

Reuters. 2015. Grand Ayatollah Ali Al-Sistani Urges Global War Against ISIS. *NBCNews*. October 2. Available online: https://www.nbcnews.com/storyline/isis-terror/grand-ayatollah-ali-al-sistani-urges-global-war-against-isis-n437421 (accessed on 6 August 2023).

Russell, Stuart Jonathan, and Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach*. London: Pearson.

ShabZendeDar, Mahdi. 2023. حرکت به سوی استفاده از هوش مصنوعی همراه با آسیب شناسی این عرصه ضروری است (Moving towards the Use of Artificial Intelligence, Accompanied by a Critique of This Field, Is Essential). *Shora-gc.ir*. April 18 World. Available online: http://www.shora-gc.ir/fa/news/9229 (accessed on 4 August 2023).

Sheikh Anṣārī, Murtaḍā b. Muḥammad Amīn. 1404. مطارح الانظار (*Topics of Discussion*) *(Old Print)*. Edited by Muhammad Ali Kalāntarī. Muʾassisat Āl al-Bayt. Available online: https://noorlib.ir/book/view/2758 (accessed on 25 July 2023).

The New Arab. 2019. Dubai Launches 'first Ever' Artificial Intelligence-Powered Fatwa Service. *The New Arab*. October 31. Available online: https://www.newarab.com/news/dubai-launches-first-ever-artificial-intelligence-powered-fatwa-service (accessed on 27 July 2023).

Trothen, Tracy J. 2022a. Intelligent Assistive Technology Ethics for Aging Adults: Spiritual Impacts as a Necessary Consideration. *Religions* 13: 452. [CrossRef]

Trothen, Tracy J. 2022b. Replika: Spiritual Enhancement Technology? *Religions* 13: 275. [CrossRef]

Tsourlaki, Sofia. 2022. Artificial Intelligence and Fatwa Issuance: A Case Study of Dubai and Egypt. *Islamic Inquiries*, November 23. [CrossRef]

Westerlund, Mika. 2019. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review* 9: 40–53. [CrossRef]

Winston, Patrick Henry. 1992. *Artificial Intelligence*. Boston: Addison-Wesley Publishing Company.

Young, William. 2022. Virtual Pastor: Virtualization, AI, and Pastoral Care. *Theology and Science* 20: 6–22. [CrossRef]