Spektrum Iran | Jg. 38 | Nr. 2 | Sommer und Herbst-2025 | pp. 213-242 |

Original-Forschungsarbeit

# Künstliche Intelligenz und digitale Hermeneutik: Datenbias, algorithmische Ethik und gesellschaftliche Implikationen

**Fatemeh Abdollahpour Sangachi** [1], **Hossein Rahnamaei** [2*], **Ali Asgari Yazdi** [3], **Mehran Rezaei** [4]

[1] *Doktorandin im Fach Didaktik der theoretischen Grundlagen des Islam, Universität Teheran, Teheran, Iran*
[2] *Assistenzprofessor an der Fakultät für Islamisches Denken und Religionswissenschaften, Universität Teheran, Teheran, Iran*
[3] *Ordentlicher Professor an der Fakultät für Islamisches Denken und Religionswissenschaften, Universität Teheran, Teheran, Iran*
[4] *Assistenzprofessor an der Fakultät für Theologie und Islamische Studien, Universität Mazandaran, Babolsar, Iran*

**Empfangen:** 15. März 2025 **Akzeptiert:** 10. Juni 2025

**Zusammenfassung:**

Der vorliegende Beitrag untersucht das komplexe Wechselverhältnis zwischen Datenbias, algorithmischer Ethik und den gesellschaftlichen Konsequenzen digitaler Hermeneutik. Mit der zunehmenden Präsenz künstlicher Intelligenz in interpretativen Praxisfeldern insbesondere in der Auslegung religiöser und philosophischer Texte, geraten die Annahmen von Datenneutralität und algorithmischer Objektivität zunehmend in die Kritik. In einem analytisch-erklärenden Zugriff zeigt die Studie, dass Trainingsdaten keineswegs neutrale Träger von Information darstellen, sondern vielmehr kulturelle Vorverständnisse, Wertsetzungen und implizite Annahmen in sich tragen, die in algorithmischen Prozessen fortgeschrieben und reproduziert werden. Diese Reproduktion kann zu semantischer Verengung, zur Marginalisierung interpretativer Pluralität und in bestimmten Fällen sogar zur Verzeichnung heiliger Texte führen. Aus hermeneutischer Perspektive betont der Beitrag die Notwendigkeit, strikt zwischen menschlichem Vorverständnis und maschineller Datenverarbeitung zu unterscheiden. Es wird argumentiert, dass das Fehlen von Bewusstsein, kritischer Selbstreflexion und gelebter Erfahrung in algorithmischen Systemen die Möglichkeit eines eigentlichen Verstehens grundsätzlich ausschließt. Die gesellschaftlichen Implikationen dieser Begrenzung reichen weit über den Bereich der Textinterpretation hinaus und umfassen Gefährdungen der Privatsphäre, die Verstärkung sozialer Ungleichheiten sowie eine schleichende Erosion kultureller Vielfalt. Abschließend wird die These vertreten, dass digitale Hermeneutik nur dann ein konstruktives Potenzial entfalten kann, wenn die technischen Leistungsfähigkeiten künstlicher Intelligenz in einen Rahmen ethischer Governance, religiöser Reflexionsinstanzen und einer kontinuierlichen Rückbindung an etablierte hermeneutische Traditionen eingebettet werden.

**Schlüsselwörter:** künstliche Intelligenz, Datenbias, algorithmische Ethik, digitale Hermeneutik, rechnerische Gerechtigkeit, gesellschaftliche Implikationen von Technologie

* Korrespondierender Autor ✉ rahnamaei@ut.ac.ir https://orcid.org/0000-0002-8098-1675

***Wie dieser Artikel zu zitieren ist:***
Abdollahpour Sangachi, F., Rahnamaei, H., Asgari Yazdi, A., & Rezaei, M. (2025). Artificial intelligence and digital hermeneutics: Data bias, algorithmic ethics, and social implications. *Spektrum Iran*, 38(2), 213-242.
https://doi.org/10.22034/spektrum.2026.562967.1050

مقاله پژوهشی

# هوش مصنوعی و هرمنوتیک دیجیتال: سوگیری داده‌ها، اخلاق الگوریتمی و پیامدهای اجتماعی

**فاطمه عبداله‌پور سنگچی [۱]، حسین رهنمائی [۲*]، علی عسگری یزدی [۳]، مهران رضائی [۴]**

*۱ دانشجوی دکتری مدرسی مبانی نظری اسلام، دانشگاه تهران، تهران، ایران*

*۲ استادیار دانشکده اندیشه و معارف اسلامی دانشگاه تهران، تهران، ایران*

*۳ استاد تمام دانشکده اندیشه و معارف اسلامی دانشگاه تهران، تهران، ایران*

*۴ استادیار دانشکده الهیات و معارف اسلامی دانشگاه مازندران بابلسر، بابلسر، ایران*

**دریافت:** ۱۴۰۳/۱۲/۲۵؛ **پذیرش:** ۱۴۰۴/۳/۲۰

**چکیده:**

این مقاله به بررسی نسبت میان سوگیری داده‌ها، اخلاق الگوریتمی و پیامدهای اجتماعی هرمنوتیک دیجیتال می‌پردازد. با گسترش روزافزون ورود هوش مصنوعی به عرصه‌های تفسیری به‌ویژه در حوزه متون دینی و فلسفی، مسئله بی‌طرفی داده‌ها و عینیت الگوریتم‌ها به یکی از دغدغه‌های اساسی بدل شده است. پژوهش حاضر با رویکردی تحلیلی ـــ تبیینی نشان می‌دهد که داده‌های آموزشی، عناصری خنثی و بی‌پیش‌فرض نیستند، بلکه حامل مفروضات فرهنگی، ارزش‌ها و پیش‌انگاره‌هایی‌اند که در فرآیندهای الگوریتمی بازتولید می‌شوند. این بازتولید می‌تواند به تقلیل معنایی، حذف تنوع تفسیری و در مواردی، به تحریف معنای متون مقدس بینجامد. مقاله حاضر با بهره‌گیری از رویکرد هرمنوتیکی، بر ضرورت تمایز میان «پیش‌فهم انسانی» و «داده‌های ماشینی» تأکید می‌کند و استدلال می‌نماید که فقدان آگاهی، تأمل انتقادی و تجربه زیسته در سامانه‌های الگوریتمی، امکان تحقق فهم اصیل را منتفی می‌سازد. پیامدهای اجتماعی این محدودیت صرفاً به حوزه تفسیر متن محدود نمی‌شود، بلکه تهدیدهایی چون نقض حریم خصوصی، بازتولید نابرابری‌های اجتماعی و تضعیف تنوع فرهنگی را نیز دربر می‌گیرد. در نهایت، مقاله بر این نکته تأکید دارد که هرمنوتیک دیجیتال تنها در صورتی می‌تواند نقشی سازنده ایفا کند که توانمندی‌های فنی هوش مصنوعی در چارچوب حکمرانی اخلاقی، نظارت دینی و پایبندی مستمر به سنت‌های معتبر تفسیری به‌کار گرفته شود.

**واژگان کلیدی:** هوش مصنوعی، سوگیری داده‌ها، اخلاق الگوریتمی، هرمنوتیک دیجیتال، عدالت محاسباتی، پیامدهای اجتماعی فناوری

* نویسنده مسئول

https://orcid.org/0000-0002-8098-1675

rahnamaei@ut.ac.ir

https://doi.org/10.22034/spektrum.2026.562967.1050

Spektrum Iran | Vol. 38 | Issue 2 | Summer & Autumn 2025 | pp. 213-242 |

Original Research Paper

# Artificial intelligence and digital hermeneutics: Data bias, algorithmic ethics, and social implications

**Fatemeh Abdollahpour Sangachi** [1], **Hossein Rahnamaei** [2*], **Ali Asgari Yazdi** [3], **Mehran Rezaei** [4]

[1] *PhD Candidate in Teaching the Theoretical Foundations of Islam, University of Tehran, Tehran, Iran*

[2] *Assistant Professor, Faculty of Islamic Thought and Knowledge, University of Tehran, Tehran, Iran*

[3] *Full Professor, Faculty of Islamic Thought and Knowledge, University of Tehran, Tehran, Iran*

[4] *Assistant Professor, Faculty of Theology and Islamic Studies, University of Mazandaran, Babolsar, Iran*

**Received:** Mar. 15, 2025 **Accepted:** Jun. 10, 2025

## Abstract

This study examines the relationship between data bias, algorithmic ethics, and the social consequences of digital hermeneutics. As artificial intelligence increasingly influences interpretive domains—particularly religious and philosophical texts—the question of data neutrality and algorithmic objectivity has become a fundamental concern. Using an analytical-explanatory approach, the study demonstrate that training data, contrary to common assumptions, are not neutral. Instead, they embody cultural values and presuppositions that are reproduced within algorithmic processes. This reproduction can result in semantic simplification, the reduction of interpretive diversity, and even the distortion of sacred texts. Drawing on a hermeneutical perspective, the article emphasizes the need to distinguish between "human pre-understanding" and "machine data," showing that the absence of awareness, critical reflexivity, and lived experience in algorithms prevents the attainment of authentic understanding. Moreover, the study indicates that the social implications of this condition extend beyond textual interpretation, posing risks to privacy, intensifying social inequalities, and undermining cultural diversity. Ultimately, the article argues that digital hermeneutics can be constructive only when the technical capacities of artificial intelligence are accompanied by ethical principles, religious oversight, and the preservation of interpretive traditions.

**Keywords:** artificial intelligence, data bias, algorithmic ethics, digital hermeneutics, computational justice, social implications of technology

* Corresponding Author ✉ rahnamaei@ut.ac.ir https://orcid.org/0000-0002-8098-1675

***How to Cite this Article:***
Abdollahpour Sangachi, F., Rahnamaei, H., Asgari Yazdi, A., & Rezaei, M. (2025). Artificial intelligence and digital hermeneutics: Data bias, algorithmic ethics, and social implications. *Spektrum Iran*, 38(2), 213-242.
https://doi.org/10.22034/spektrum.2026.562967.1050

## 1. Introduction

The rapid expansion of artificial intelligence—particularly large language models—has profoundly transformed textual interpretation in the humanities and social sciences. In recent years, AI-driven systems have increasingly been employed in domains traditionally governed by human interpretive judgment, including philosophy, law, and religious studies. These developments have generated both enthusiasm and concern. While algorithmic tools promise greater efficiency, accessibility, and large-scale textual analysis, they also raise fundamental questions about understanding, meaning, and epistemic responsibility.

Within this context, the emerging field of digital hermeneutics seeks to examine how digital technologies reshape the conditions of interpretation. Far from being a purely technical issue, digital hermeneutics addresses the philosophical question of whether algorithmic systems can meaningfully participate in acts of interpretation or whether they merely simulate understanding through statistical pattern recognition. This question becomes particularly urgent when artificial intelligence is applied to religious and philosophical texts, in which interpretation is historically embedded, normatively charged, and inseparable from ethical accountability.

One of the most contested assumptions underlying contemporary AI applications is the notion of algorithmic objectivity. AI systems are often presented as neutral and unbiased tools capable of processing data without the limitations of human subjectivity. However, a growing body of research challenges this view by demonstrating that training datasets are shaped by cultural, linguistic, and ideological presuppositions. Rather than eliminating bias, algorithms often reproduce and even amplify existing power structures embedded within their data sources. In interpretive contexts, this tendency risks reducing meaning to probabilistic correlations, thereby marginalizing historical depth, contextual nuance, and interpretive plurality.

From the perspective of philosophical hermeneutics, this problem can be reframed through the concept of pre-understanding (*Vorverständnis*). Human interpretation is always situated within historical horizons, linguistic traditions, and lived experience. These pre-understandings, while influential, remain open to critique, revision, and ethical responsibility. Algorithmic systems, by contrast, rely on fixed datasets that function as a non-reflexive

form of pre-structuring. Unlike human interpreters, algorithms lack historical consciousness, intentionality, and the capacity for self-critical revision. As a result, their outputs may resemble interpretation at a linguistic level while remaining epistemically deficient from a hermeneutic standpoint.

This distinction becomes especially consequential in the interpretation of religious texts. Within Islamic exegetical traditions, understanding the Qur'an has never been conceived as a purely linguistic operation. Rather, interpretation unfolds within a living continuum of transmitted knowledge, historical context, ethical accountability, and communal responsibility. The replacement—or even the uncritical supplementation—of human interpretation with algorithmic outputs therefore raises serious epistemic and theological concerns. In the absence of sensitivity to interpretive tradition, contextual depth, and normative responsibility, AI-assisted interpretation risks semantic reduction, doctrinal distortion, and the erosion of interpretive diversity.

Against this background, the present article argues that the algorithmic interpretation of religious texts is epistemically limited and ethically problematic because data bias functions as a non-reflexive form of pre-understanding. Unlike human pre-understanding, algorithmic data lack historical awareness, reflexivity, and responsibility. For this reason, digital hermeneutics can be legitimate only if human interpretive agency, ethical oversight, and established interpretive traditions remain epistemically primary. Artificial intelligence, on this account, may serve as an auxiliary analytical tool, but it cannot function as an autonomous interpreter of meaning.

Methodologically, this study adopts an analytical-explanatory approach grounded in qualitative documentary analysis. Drawing on philosophical hermeneutics, contemporary AI ethics, and Islamic exegetical sources, the article seeks to integrate these discourses within a conceptually disciplined framework rather than juxtaposing them. Philosophical hermeneutics provides the theoretical lens for analyzing understanding and pre-understanding; Islamic exegetical traditions offer a historically grounded model of interpretive practice; and AI ethics establishes the normative horizon for evaluating social and epistemic consequences.

The article proceeds in four conceptual steps. First, it clarifies data bias and critically examines claims of algorithmic objectivity. Second, it distinguishes human pre-understanding from machine data through the framework of philosophical hermeneutics. Third, it explains why this distinction is especially critical in the interpretation of religious texts, with particular attention to Islamic exegetical traditions. Finally, it addresses the ethical and social implications of digital hermeneutics and proposes normative conditions for responsible AI-assisted interpretation.

## 2. Research Method

This study is theoretical in scope and employs an analytical-explanatory methodology grounded in qualitative documentary analysis. This methodological choice is justified by the conceptual and normative nature of the research questions, which address meaning, interpretation, and epistemic responsibility rather than empirical measurement. Quantitative or experimental approaches would be less suitable for engaging these issues, as they presuppose measurable variables rather than interpretive frameworks.

Data were collected through library-based research and include authoritative sources in artificial intelligence, algorithmic ethics, the philosophy of understanding, hermeneutics, and Qur'anic studies, as well as recent scholarship in data science and digital technologies. The research corpus consists of books, academic articles from ISI- and Scopus-indexed journals, as well as other peer-reviewed journals, relevant dissertations, and classical works of Islamic exegesis. Sources were selected on the basis of their scholarly authority, relevance to the research questions, and their recognized standing within their respective disciplinary traditions.

In line with the objectives of the study, the following research questions were formulated:

1. To what extent are the training datasets of artificial intelligence systems objective, and how much cultural and linguistic bias do they reflect?

2. What epistemic, ethical, and social consequences arise from machine-based interpretation of religious texts?

3. How can the relationship between human understanding and algorithmic processing be conceptualized within the framework of digital hermeneutics?

4. What principles are necessary to prevent distortion or semantic reduction in the digital interpretation of sacred texts?

Because the study does not rely on numerical data or quantitative testing, statistical tools were not employed. Instead, analysis was conducted using qualitative methods, including thematic analysis and comparative examination of relevant theories. During the analytical phase, key concepts were coded, categorized, and subsequently interpreted within the frameworks of philosophical hermeneutics and the ethics of technology. Interpretive claims were assessed on the basis of conceptual coherence, consistency with established hermeneutical theories, and sustained engagement with authoritative philosophical and exegetical sources. This procedure is intended to reduce the risk of arbitrary or purely subjective interpretation.

Alternative perspectives—particularly instrumentalist views that treat AI as a neutral interpretive tool—are addressed through critical comparison rather than exclusion, allowing their underlying assumptions to be examined through the proposed hermeneutical framework. Methodologically, the study distinguishes among three analytical registers: descriptive analysis, interpretive-hermeneutical analysis, and normative evaluation. While the first two levels aim to clarify conceptual and textual structures, the third explicitly advances ethical and theological judgments. This distinction is maintained throughout the article in order to avoid conflating analytical description with normative prescription.

**2.1. Data Analysis**

To ensure methodological transparency and analytical rigor, data analysis in this study was conducted through a structured three-level qualitative framework. These levels are analytically distinct yet conceptually interconnected, and each serves a specific argumentative function. The first level is descriptive and aims to clarify key concepts and dominant positions within the existing literature. The second level is explanatory, focusing on the causal and structural relationships among data bias, algorithmic processes,

and interpretive outcomes. The third level is interpretive and normative, in which the findings are critically evaluated within the frameworks of philosophical hermeneutics and Islamic exegetical traditions.

This stepwise analytical structure enables the study to proceed systematically from description to explanation and, finally, to normative evaluation, while clearly demarcating shifts in analytical register and avoiding the conflation of empirical description with ethical or theological judgment.

**2.2. Descriptive Analysis**

At the first level, the analysis remains strictly descriptive. The aim of this stage is not to evaluate or judge, but to clarify the conceptual landscape within which discussions of artificial intelligence and digital hermeneutics take place. Key concepts such as "data bias," "algorithmic objectivity," "digital hermeneutics," "machine interpretation," and "interpretive tradition" were extracted from the selected sources and systematically categorized.

This level of analysis demonstrates how contemporary artificial intelligence systems, particularly large language models, are trained on datasets derived from news media, academic publications, digital platforms, and human-curated resources. The descriptive findings indicate that these datasets are shaped by linguistic dominance, cultural hierarchies, and historically situated assumptions. At this stage, no normative claims are made; rather, the analysis provides a conceptual and empirical foundation for understanding how bias can emerge within data-driven systems.

**2.3. Explanatory (Causal-Structural) Analysis**

The second level of analysis moves beyond description and adopts an explanatory orientation. At this stage, the study examines the causal and structural relationships among training data, algorithmic design, and interpretive outcomes. The central explanatory claim is that algorithmic outputs are not neutral products of computation but are structurally conditioned by the nature of their input data and the assumptions embedded in their design.

The analysis demonstrates that when training datasets contain cultural, linguistic, or ideological presuppositions, these presuppositions are

systematically reproduced in algorithmic outputs. As a result, outputs that appear objective or scientifically grounded may in fact reflect dominant cultural perspectives or narrowed interpretive horizons. This causal dynamic is particularly significant in the interpretation of religious texts, where meaning is deeply dependent on historical context, semantic depth, and interpretive plurality.

Moreover, the absence of lived experience, historical consciousness, and critical reflexivity in artificial intelligence systems explains why such systems can generate coherent text without achieving genuine understanding. This explanatory level therefore accounts for how and why processes of semantic reduction, interpretive simplification, and the marginalization of alternative traditions arise in machine-based interpretation.

**2.4. Interpretive and Normative Analysis**

The third level of analysis explicitly moves into the interpretive and normative domain. Building on the descriptive foundations and explanatory relationships established in the previous stages, this level evaluates the implications of machine-based interpretation within the frameworks of philosophical hermeneutics and Islamic exegetical traditions.

From a hermeneutical perspective, the findings indicate that interpretation is inseparable from historical situatedness, linguistic depth, and participation in a living tradition. Philosophical hermeneutics emphasizes that understanding arises through the interaction between text, interpreter, and tradition, a process that cannot be replicated by algorithmic systems lacking awareness and reflexivity.

Within the Islamic exegetical tradition—both Shiʿi and Sunni—interpretation has always been grounded in methodological principles such as attention to context, occasions of revelation, linguistic sciences, and transmitted knowledge. Evaluated against these criteria, the replacement of the human interpreter with an algorithm constitutes not merely a technical limitation but an epistemic and ethical distortion. Accordingly, this level of analysis advances the normative claim that artificial intelligence may function as an auxiliary analytical tool, but cannot serve as an autonomous interpreter of sacred texts without undermining interpretive plurality and religious responsibility.

## 3. Findings

The findings reported in this section summarize the results of the descriptive and explanatory stages of the analysis and do not yet advance normative or prescriptive claims, which are developed in subsequent interpretive discussions. Based on an analysis of the sources and a comparative examination of philosophical, religious, and technological perspectives, the study yields the following findings:

**Training data are not neutral.** Data drawn primarily from dominant cultures, major world languages, or Western digital resources introduce implicit biases into artificial intelligence models. As a result, although algorithmic outputs may appear neutral, they often reflect underlying structures of cultural power.

**Algorithms generate text, not genuineunderstanding.** While language models are capable of organizing data and producing coherent textual output, they lack lived experience, intentionality, historical awareness, and critical reflexivity. Their outputs therefore constitute linguistic production rather than epistemic interpretation.

**Risk of semantic reduction in religious texts.** Because algorithms rely on statistical pattern recognition, they tend to overlook deeper semantic layers, historical context, occasions of revelation, and established interpretive traditions. This can result in superficial or, at times, inaccurate interpretations of Qur'anic verses and narrations.

**Risk of monolithic interpretation.** When models are trained on limited or homogenized sources, the diversity of Shiʿi and Sunni exegetical traditions may be obscured or erased, reducing religious understanding to a uniform and mechanistic interpretation—an outcome that is epistemically and socially problematic.

**Social and cultural consequences.** Data bias in machine-based interpretation of sacred texts may lead to:

A) the reinforcement of dominant discourses;

B) the marginalization of local languages and indigenous cultures;

C) the reduction of religious meaning to statistical representation; and

D) the potential for ideological or propagandistic misuse.

**Key analytical finding.** Digital hermeneutics can be legitimate and effective only under specific conditions:

A) the extraction of data and the use of models are supervised by experts in religious studies and ethics;

B) established interpretive traditions function as evaluative criteria; and

C) the role of the human interpreter as critical agent and overseer is preserved.

The conceptual sections that follow are not independent thematic essays but analytically grounded extensions of the multi-level framework employed in this study. Each section elaborates a particular dimension of the findings developed through the descriptive, explanatory, and interpretive stages of analysis. In particular, discussions of ethical oversight, cultural bias, multilinguality, and interpretive traditions emerge from the interpretive–normative level, where the epistemic limits of algorithmic interpretation are evaluated in dialogue with philosophical hermeneutics and Islamic theology. This methodological continuity ensures that conceptual reflection remains anchored in the study's analytical structure rather than proceeding through mere thematic accumulation.

### 3.1. Ethical and Social Dimensions in Digital Hermeneutics and Moral–Cultural Challenges

This section develops the ethical and social implications of the interpretive limitations identified in the preceding analysis, particularly those related to data bias, claims of algorithmic objectivity, and the absence of genuine pre-understanding in artificial intelligence systems. From a methodological perspective, the discussion operates at the interpretive–normative level of the study and should be read as an evaluative extension of the analytical framework rather than as an independent ethical discourse.

Digital hermeneutics lies at the intersection of two fundamental domains: information ethics, on the one hand, and the philosophy of human understanding, on the other. In this study, philosophical hermeneutics provides the analytical framework for understanding interpretation as a historically situated and horizon-dependent process; Islamic interpretive theology supplies the normative and semantic criteria governing the interpretation of sacred texts; and contemporary AI ethics functions as the

evaluative domain through which algorithmic practices are assessed. These three discourses are not merely juxtaposed but methodologically integrated. Philosophical hermeneutics clarifies how understanding occurs, Islamic theology determines what counts as legitimate understanding in religious contexts, and AI ethics specifies the conditions under which digital mediation becomes ethically and epistemically acceptable.

This positioning prevents digital hermeneutics from being reduced to technical questions of algorithmic efficiency. Instead, it prompts a deeper reconsideration of the relationship between human beings and the digital world, as well as the ways in which self-understanding is shaped. As Capurro observes, "the digitalization of human life has affected every dimension of human existence—from the body and individual autonomy to the experience of time and space, emotional states, social structures, imagination, scientific understanding, and even religious beliefs" (Capurro, 2010, p. 39). Consequently, the ethics of digital hermeneutics emerges as a genuine necessity in the contemporary era—one that extends beyond behavioral norms in cyberspace to encompass ontological and meaning-constituting questions.

With the widespread integration of code and algorithms into social life, ethical questions have shifted to an unprecedented level. Whereas earlier discussions of information ethics focused primarily on regulating conduct in digital environments, current concerns penetrate the core of philosophical inquiry into the nature of the human being and the world. Digital hermeneutics reminds us that technology is not merely a neutral instrument for understanding but forms part of the very horizon through which reality is interpreted. Contemporary empirical research further illustrates this dynamic by showing that global discourse on artificial intelligence is structured by geopolitical competition and asymmetries of power. Dominant technological actors disproportionately shape how AI is framed and understood, while less powerful regions frequently articulate concerns of dependency and exclusion. Such findings reinforce the claim that artificial intelligence does not operate within a neutral communicative space but is embedded within contested structures of global power (Salehi et al., 2025). The pervasive presence of artificial intelligence in everyday life has blurred the boundary between human agents and non-human actors, reshaping even the lived experience of the body and the world (Capurro, 2010, p. 39). From

a hermeneutical perspective, such transformations do not merely produce social consequences; they actively redefine the horizon of self-understanding.

Within this framework, data bias becomes ethically significant precisely because algorithms function as quasi-interpretive agents that mediate meaning rather than simply process information. One of the most critical ethical dimensions of digital hermeneutics concerns bias in training data. Contrary to common assumptions, the datasets used to train algorithms are neither neutral nor objective; they contain embedded values, presuppositions, and cultural prejudices. These biases do not remain confined to the statistical domain but are reproduced within the machine's interpretive processes (Youvan, 2024, p. 9). The decisive issue here is the distinction between human pre-understanding and machine data. Human pre-understanding emerges within contexts of awareness and critical reflexivity and remains open to revision. Machine data, by contrast, lack self-awareness and reflexive capacity. Consequently, when training data are biased, artificial intelligence systems systematically reproduce those same distortions (O'Neil, 2016, p. 181).

When algorithms trained on biased datasets are applied to Qur'anic interpretation, these distortions acquire particular gravity. A human interpreter approaches the text with a pre-understanding shaped by faith, rational inquiry, and critical engagement—an orientation that can be refined through interpretive dialogue. Machine presuppositions, however, are inflexible and non-reflexive. If training sources privilege a single interpretive orientation or a narrow cultural framework, the algorithm mechanically reproduces that orientation. This dynamic risks producing a "monologization" of Qur'anic understanding in digital spaces and threatens the diversity of Islamic interpretive traditions.

For this reason, a fundamental condition for preventing distortion or semantic reduction in digital hermeneutics is scholarly and religious oversight of training datasets, together with the inclusion of diverse interpretive sources in algorithmic development. Methodologically, this claim arises from qualitative analysis of philosophical hermeneutics and Islamic exegetical traditions rather than empirical testing. At this stage, ethical and theological references are introduced not as independent

normative assertions but as interpretive frameworks that illuminate the implications of biased data practices.

The Qur'anic and hadith references cited here function as hermeneutic analogies that demonstrate how distortion and interpretive responsibility have historically been conceptualized within Islamic thought. The Qur'an states: "And indeed, among them is a group who distort their tongues with the Book so that you may think it is from the Book, while it is not from the Book…" (Āl ʿImrān 3:78). This verse describes deliberate manipulation of revealed meaning. In the digital context, biased training data can produce an analogous form of distortion, albeit through mechanical rather than intentional means.

Similarly, the Prophet Muhammad (peace be upon him) warned: "Whoever interprets the Qur'an according to his own opinion, let him prepare for himself a seat in the Fire" (Fayḍ Kāshānī, 1364/1985, vol. 1, p. 35). This hadith condemns interpretation grounded in unfounded presuppositions. When algorithms rely on narrow datasets and lack reflexive capacity, the resulting outputs resemble a form of "machine-based interpretation by opinion."

Ethical responsibility therefore rests decisively with human agents, particularly developers and scholars. They must address bias not only at a technical level but also through culturally and socially informed decisions about data selection, model design, and output governance. Neglecting these dimensions carries serious ethical risks, as algorithmic outputs increasingly inform social, legal, and religious decision-making and may reinforce structural discrimination (O'Neil, 2016, p. 183).

The Qur'an warns against textual fabrication: "So woe to those who write the Book with their own hands and then say, 'This is from God'…" (al-Baqarah 2:79). In digital hermeneutics, a parallel danger arises when algorithms generate distorted meanings that may be mistaken for authoritative interpretation. Imam ʿAlī (peace be upon him) similarly observed: "This Qur'an is nothing but writing inscribed between two covers; it does not speak with a tongue, and it must have an interpreter" (Nahj al-Balāgha, Sermon 125). Without a knowledgeable and just interpreter, mechanical processing risks superficiality.

Artificial intelligence also produces wide-ranging social and cultural effects. Among the most serious is the erosion of privacy: large-scale data extraction enables pervasive surveillance and behavioral profiling (Bostrom, 2014, p. 60). In addition, AI-driven automation may contribute to job displacement, the deepening social inequality, and the concentration of informational power within a limited set of institutions (Brynjolfsson & McAfee, 2014, p. 143). These developments demonstrate that AI is not merely a technical tool but a transformative force reshaping social structures.

From a cultural perspective, algorithmic systems risk identity homogenization and the marginalization of local cultures. Because dominant datasets disproportionately reflect hegemonic languages and perspectives, minority narratives and indigenous traditions often receive limited representation. Over time, this process can erode cultural diversity and weaken lived communal experience (Floridi, 2014, p. 263). Digital hermeneutics therefore entails not only technical evaluation but also cultural critique.

In the religious domain, the central concern remains how algorithms interact with sacred texts. Although AI can assist in organizing sources and facilitating access, it cannot grasp semantic depth, historical context, or the subtlety of religious language. As a result, it may generate superficial or erroneous interpretations that affect religious belief and practice (Capurro, 2010, p. 39). The Qur'an states: "We sent no messenger except [speaking] in the language of his people, so that he might clarify [the message] for them" (Qur'an 14:4). Neglecting linguistic and cultural specificity in digital mediation leads to distortion and loss of meaning.

Ibn ʿArabī emphasizes that authentic knowledge arises from lived experience and spiritual unveiling, not abstract pattern recognition. Algorithms, by contrast, operate primarily on statistical correlations. This contrast underscores why artificial intelligence can never replace the human interpreter in the domain of religious understanding.

A final issue concerns interpretive agency itself. In machine-assisted interpretation, who bears responsibility for meaning? From a hermeneutical perspective, meaning cannot emerge without a historically situated, self-aware subject (Capurro, 2010, p. 38). Algorithmic outputs must therefore be treated not as interpretation in the full hermeneutic sense, but as preliminary

materials whose evaluation and responsibility remain fundamentally human (Floyd et al., 1992, p. 289). The Qur'an cautions: "Do not pursue that of which you have no knowledge" (Qur'an 17:36), a principle that applies directly to uncritical reliance on machine-generated outputs.

**3.2. Cultural Bias and Multilingual Challenges in Machine Interpretation of Texts**

This section develops the central argument of the paper by showing how cultural bias and multilingual limitations concretely expose the non-hermeneutic character of machine interpretation.

One of the fundamental challenges within digital hermeneutics is the issue of cultural bias and the limitations posed by multilinguality in artificial intelligence systems. Machine interpretation, developed primarily to meet the growing need for rapid cross-lingual communication, confronts not merely technical issues but profound contextual constraints as well. Every text derives its meaning within a specific cultural, historical, and social framework; therefore, without an understanding of this context, the likelihood of misinterpretation becomes significant.

This challenge is especially pronounced in the Arabic language. Issues such as orthographic ambiguity, lexical ambiguity, and the role of iʿrāb (grammatical vowel markings), combined with deep cultural and semantic divergences, indicate that translation is not merely a linguistic conversion but an act of interpretation shaped by traditions, belief systems, and cultural meaning structures. For example, an expression that functions as a polite taʿārof (ritual politeness) in Persian may be interpreted as a direct invitation in English; similarly, proverbs and metaphors familiar within one culture may be understood entirely differently in another.

As Nadhari (2014, p. 29) notes: "Machine translation, when approached in absolute terms, faces unique challenges in Arabic that many other languages do not possess such as orthographic ambiguity, lexical ambiguity, and the problem of iʿrāb. Thus, machine translation from a language like Arabic into Persian may be among the most difficult stages in the development of a truly global translation system."

Accordingly, any absolutist reliance on machine translation without attention to the cultural and historical layers embedded in a text can lead to incomplete or even misleading interpretations. Identifying these contextual

layers and understanding their role in textual meaning is one of the essential requirements of digital hermeneutics and a prerequisite for accurate translation analysis.

These examples are not presented as isolated technical shortcomings, but as manifestations of a deeper hermeneutic limitation rooted in the absence of historically and culturally situated understanding.

Recent studies have shown that many existing language models are trained predominantly on data derived from specific languages and cultures—particularly English and Western sources. This leads to an epistemic marginalization of low-resource languages and non-dominant cultural traditions in processes of computational interpretation and natural language processing (Eder, 2016, p. 460). As a result, algorithms often resort to oversimplification or semantic distortion when encountering texts rooted in other linguistic and cultural contexts.

A clear example can be seen in research on Indonesian religious texts. Findings indicate that general-purpose models such as Sentence Transformer perform poorly in representing the Indonesian language, prompting the development of localized models like IndoBERT for more accurate processing (Lima et al., 2025, p. 14). This experience demonstrates that without culturally grounded and linguistically localized models, the transfer of linguistic and cultural biases into digital interpretation processes becomes unavoidable.

Moreover, the use of sacred texts such as the Bible and the Qur'an as bilingual corpora for training machine-translation systems has become a common strategy in cross-linguistic studies. Yet this approach faces significant challenges: machine translation of sacred texts frequently results in excessive simplification, inappropriate lexical choices, and the omission of deeper semantic layers (Lima et al., 2025, pp. 9–10). Such tendencies risk marginalizing the cultural and historical depth of these texts and may yield superficial or distorted representations of their meaning.

The intercultural dimensions of this problem are equally significant. In multilingual contexts, cultural differences and linguistic presuppositions can lead to divergent interpretations of the same text. For instance, a question regarding a specific Qur'anic verse posed in English may yield a response

different from the same question asked in Arabic or Persian, because the models are influenced not only by linguistic data but also by the hidden cultural assumptions embedded within those data. This demonstrates that machine interpretation is consistently accompanied by a layer of cross-cultural bias and cannot be detached from its cultural–linguistic contexts.

From a philosophical and hermeneutical perspective, the limitations of artificial intelligence in reproducing human and cultural identity have been repeatedly emphasized. As Wang, a philosopher of logic and information, and Floridi argue, although digital technologies can serve as powerful tools for expanding understanding, they are incapable—at the existential and cultural level—of replacing the historically situated and lived horizons of human beings. This means that machine translation and interpretation of texts, especially sacred texts, will always remain incomplete and limited in the absence of human interpreters and an appropriate cultural contexts (Wang, 2021, p. 43; Floridi et al., 2018, p. 129).

At this point, the discussion moves from descriptive analysis of multilingual bias to a hermeneutic reflection on meaning, responsibility, and linguistic justice, drawing on Islamic thought as an interpretive framework rather than as an external normative authority.

From an Islamic perspective, this situation can be reconsidered in light of the Qur'anic verse: "We sent no messenger except in the language of his people" (Ibrāhīm 14:4). Divine revelation has always been delivered within a specific linguistic and cultural framework so that effective communication with its audience may be achieved. If artificial intelligence systems attempt to translate and interpret sacred texts without taking these frameworks into account, essential dimensions of the original meaning of revelation as well as its cultural and historical context will inevitably be lost. Accordingly, the issue of cultural bias and multilinguality is not merely a technical challenge; rather, it is a hermeneutical and existential problem intrinsically linked to linguistic justice, the preservation of interpretive traditions, and epistemic responsibility toward sacred texts.

Alongside international research on machine translation, a number of Iranian studies have also addressed the challenges of this domain. For instance, the book An Introduction to Machine Translation by Tayebeh Mousavi Miyāngāh (2007) discusses the structural, lexical, and cultural

difficulties inherent in this type of translation. Such works demonstrate that the problem is not simply technical or linguistic, but is deeply intertwined with layers of cultural bias and the complexities of multilingual processing.

### 3.3. Analytical Note: Ignoring Shiʿi and Sunni Interpretive Traditions

This section advances the central argument of the paper by demonstrating that the limitations of artificial intelligence in religious interpretation become fully visible only when viewed against the background of established Shiʿi and Sunni exegetical traditions.

Here we can see one of the fundamental risks of digital hermeneutics in the interpretation of religious texts: the neglect of the Shiʿi and Sunni exegetical traditions. Throughout Islamic history, the understanding of the Qur'an has never occurred outside the framework of tradition; rather, it has always unfolded within a living, dynamic, and historical continuum that, over centuries, has encompassed diverse layers of exegetical methodologies—from transmitted and linguistic commentaries to rational, philosophical, and mystical interpretations.

In fact, tradition (al-turāth) may be viewed as a living fabric within which human experience breathes and acquires meaning. We are never detached from tradition; rather, we grow within it, and the horizons of our understanding—whether consciously or unconsciously—are woven from its very texture (Bashir et al., 2022, p. 28). Ignoring this background is effectively equivalent to severing religious understanding from its historical, epistemic, and cultural roots.

AI algorithms operate on the basis of the statistical co-occurrence of words and lack the capacity to apprehend the historical, social, and traditional contexts of texts. Exclusive reliance on artificial intelligence for interpreting scriptures without attending to these limitations can therefore lead to superficial, incomplete, or even erroneous readings (Rostam & Hassain Malim, 2021, p. 658). This warning indicates that replacing the exegetical tradition with machine-generated interpretation risks reducing sacred meanings to the most surface-level linguistic associations.

At this stage of the analysis, Qur'anic and hadith references are employed not as independent theological proofs, but as hermeneutic illustrations that

clarify how depth, context, and tradition have been historically understood as conditions of meaning within Islam.

God Almighty declares in the Qur'an: "Do they not reflect (tadabbur) upon the Qur'an…?" (al-Nisāʾ 4:82). Tadabbur entails moving beyond the surface of words to reach the depth of meaning within its historical, ethical, and divine horizons. If Qur'anic understanding is reduced merely to lexical surfaces and statistical patterns—as algorithmic models tend to do—the verses are diminished to mere adjacent sentences, whereas the Qur'an's intention is to lead the human being toward insight (basīra) and guidance.

The Prophet Muhammad (peace be upon him) stated: "Indeed, the Qur'an has an outward aspect and an inward aspect" (al-'Ayyāshī, 1395, vol. 1, p. 11). This narration indicates that understanding the Qur'an cannot be achieved merely through the outward wording; rather, it requires grasping its inner dimensions, the context of revelation, exegetical traditions, and divine guidance. If machine-based interpretation remains confined to the surface level of words, the spiritual and sacred inner meanings of the Qur'an are inevitably overlooked, resulting to misinterpretation. Thus, just as Qur'anic verses and hadiths emphasize contemplation (tadabbur), depth, and the multilayered nature of Qur'anic meanings, reliance solely on artificial intelligence without grounding in the exegetical tradition reduces sacred meaning to mere linguistic associations.

From a Shiʿi perspective, the primary foundation for understanding the Qur'an is "interpreting the Qur'an by the Qur'an" and referring to the teachings of the Infallible Imams (peace be upon them). 'Allāmah Ṭabāṭabā'ī explicitly states in the introduction to al-Mīzān that "relying on familiarity and habit when attempting to understand the meanings of the verses distorts their purposes and disrupts the process of comprehension" (Ṭabāṭabā'ī, 1374, vol. 1, p. 18). Any interpretation that departs from this methodological framework is vulnerable to deviation.

Similarly, Ayatollah Jawādī Āmulī, in Tafsīr al-Tasnīm, emphasizes that "understanding the Noble Qur'an at the level of tafsīr (interpretation) rather than ta'wīl (esoteric exegesis) and comprehending the apparent meanings of its words is accessible to all, and not exclusively reserved for the Infallibles; indeed, they encouraged and urged people to engage in such understanding" (Jawādī Āmulī, 1395, vol. 1, p. 91). These foundational principles demonstrate

that understanding the divine text is not merely the product of linguistic processing; rather, it requires participation in the prophetic tradition and the lived religious world (zist-jahān) in which revelation becomes meaningful.

In the Sunni exegetical tradition as well, particular emphasis is placed on asbāb al-nuzūl (occasions of revelation), the reports of the Companions, and the linguistic sciences. Exegetes such as al-Ṭabarī, in his Tafsīr al-Ṭabarī, and Fakhr al-Dīn al-Rāzī[1], in his Tafsīr al-Kabīr (also known as Mafātīḥ al-Ghayb), regard the understanding of the Qur'an as inseparable from its linguistic and historical contexts.

In the article "Asbāb al-Nuzūl in Tafsīr al-Ṭabarī," it is noted that al-Ṭabarī, in explaining Sūrat al-Ikhlāṣ, cites the following account regarding the occasion of revelation:

"Aḥmad ibn Manīʿ al-Marwazī and Maḥmūd ibn Khudāsh al-Ṭalaqānī narrated from Abū Saʿīd al-Ṣanʿānī, from Abū Jaʿfar al-Rāzī, from Rabīʿ ibn Abī al-ʿĀliyah, from Ubayy ibn Kaʿb, who said:

The polytheists said to the Prophet: 'Describe to us the lineage of your Lord!' Then Sūrat al-Ikhlāṣ was revealed" (Mūsavī[2], 1995, p. 9).

This report clearly specifies the occasion of revelation (the polytheists' question), cites the transmission of the Companions through Ubayy ibn Kaʿb, and demonstrates how linguistic expression ("describe the lineage of your Lord") shapes the interpretation of the verse.

Likewise, Fakhr al-Dīn al-Rāzī, in Mafātīḥ al-Ghayb, presents two major views regarding the verse of walāya (Q 5:55–56):

**1. First view:** The verse refers to all believers, supported by several transmitted reports from the Companions and Successors.

**2. Second view:** The verse refers to a particular individual, namely ʿAlī ibn Abī Ṭālib, and al-Rāzī cites narrations from the Companions to establish that the verse was revealed concerning him.

[1] Fakhr Razi
[2] Mousavi

Furthermore, al-Rāzī pays careful attention to the linguistic dimension of the phrase "wa-hum rākiʿūn", examining its semantic usage (Fakhr al-Rāzī, 2000, vol. 12, p. 385).

Thus, ignoring these historical, linguistic, and transmitted contexts and relying solely on algorithmic or machine-generated outputs amounts to severing the Qur'anic text from its living, historically grounded exegetical horizon.

Rather than treating philosophical hermeneutics and Islamic theology as parallel discourses, this comparison is intended to show how both articulate a structurally similar account of understanding grounded in tradition, pre-understanding, and historical situatedness. In Gadamer's hermeneutical theory, understanding always occurs within the framework of the "hermeneutic circle"—that is, the continuous interaction between "the whole and the part," and between "the text and the interpreter's pre-judgments" (Gadamer, 2004, pp. 283–284). Understanding is impossible without tradition, because tradition constitutes the very horizon that enables the interpreter to pose questions and receive meaning. Gadamer states that "no interpretation is ever final; every interpretation is provisional and subject to further revision with the emergence of new contexts, questions, and insights" (Bashir et al., 2022, p. 28).

If this principle is applied to the interpretation of the Qur'an, it may be said that the Islamic interpretive tradition—both in its Shi'i and Sunni dimensions—is itself the embodiment of this hermeneutic circle. Over centuries, Muslim exegetes, through a dynamic interplay between the sacred text and their religious and cultural pre-understandings, have extracted a plurality of interconnected meanings. Thus, eliminating tradition effectively means breaking the hermeneutic circle and severing the historical chain of understanding.

The crucial point here is that what Gadamer formulates philosophically has, within the Islamic tradition, been realized through principles such as "interpreting the Qur'an through the Qur'an itself" and the necessity of returning to the teachings of the Infallible Imams (a). In other words, just as Gadamer emphasizes that no understanding is possible outside tradition, ʻAllāmah Ṭabāṭabāʼī likewise insists that "the Prophet and the Imams of the Ahl al-Bayt according to the interpretive reports transmitted from them

followed the method of interpreting the Qur'an by the Qur'an" (Ṭabāṭabā'ī, 1374, vol. 1, p. 19). Hence, understanding the Qur'an without maintaining a living connection to the revelatory tradition and the Ahl al-Bayt (a) is impossible.

Imam ʿAlī (a) likewise states in Nahj al-Balāgha: "The Qur'an is such that one part of it interprets another part, and some of it bears witness to the rest" (Imam ʿAlī (a), Sermon 133, p. 226).

Imam Khomeini also affirms: "The knowledge of intellect and ignorance and their armies belongs to the divine, esoteric sciences, and true inner cognitions. Complete and comprehensive knowledge of their aspects, degrees, realities, and secrets is not attainable except for those endowed with wilāyah and certitude, the elect among the people of gnosis and faith—those who, by the light of knowledge and the path of spiritual wayfaring, have passed beyond the veil of human limitation, pierced through the barriers of the worlds of mulk and malakūt, reached the origins of existence and the sources of the unseen and witnessed realms, and have apprehended the unseen through presential vision. This occurs only for the perfected ones" (Khomeini, 2003, p. 58).

This comparison between philosophical hermeneutics and the Islamic tradition reveals that both emphasize the foundational role of tradition in the process of understanding. Yet their point of departure differs: while Gadamer conceives tradition as a historical and linguistic horizon, the Islamic tradition includes—beyond history and language—the sacred and revelatory dimension as well. Neglecting this tradition in machine-generated interpretation is, in fact, a neglect of the very hermeneutical circle that makes all understanding possible.

Furthermore, ignoring the classical exegetical traditions is not merely a theoretical danger; it carries practical and religious consequences. Machine-generated outputs may produce interpretations that conflict with theological principles or legal doctrines, potentially leading to deviation in matters of belief and practice. As the Qur'an warns: "And do not pursue that of which you have no knowledge" (al-Isrā', 17:36). This verse may be regarded as an ethical principle for engaging with machine-generated interpretations: any result lacking traditional and epistemic grounding should not be accepted uncritically.

Within this framework, the ethical responsibility of the human researcher becomes even more pronounced. The researcher is obligated not only to recognize the capacities of technology but also to remain fully aware of its limitations. Machine-based interpretation cannot replace human responsibility in the process of understanding, for meaning is realized only within the context of tradition and historical consciousness (Bashir et al., 2022, p. 28).

From the preceding hermeneutic and methodological analysis, it becomes clear that digital hermeneutics can be applied to religious texts only when the interpretive traditions of both Shi'a and Sunni Islam remain preserved. Otherwise, the dangers of reductionism, superficiality, and even religious deviation become severe. The Islamic interpretive traditions are not obstacles to technological innovation; rather, they constitute the very conditions for the possibility of any authentic understanding. Thus, any application of artificial intelligence to Qur'anic interpretation must be situated in relation to these traditions and regarded as a supplementary tool, not a substitute for human exegesis.

### 3.4. The Problem of Objectivity and Algorithmic Bias in AI Hermeneutics

This section develops the paper's central argument by examining how claims of algorithmic objectivity collapse once bias and pre-understanding are analyzed from a hermeneutic perspective.

In recent years, artificial intelligence has increasingly entered the humanities and social sciences, offering new capabilities for big data analysis, natural language processing, and the prediction of collective behaviors. Yet alongside these achievements, the issue of objectivity and algorithmic bias has emerged as one of the most complex and pressing challenges in artificial intelligence and data science, drawing growing attention from scholars and society alike.

Algorithms, which are expected to automate decision-making processes and data analysis without direct human intervention, can in practice become biased due to the nature of their input data, the design choices of programmers, or the underlying assumptions embedded within them. Recent critical analyses of AI-driven decision systems likewise highlight persistent concerns regarding opacity, embedded bias, and inadequate oversight, even in domains where predictive performance appears

technically successful. Such findings underscore the necessity of explainability and sustained human supervision, thereby further complicating claims of algorithmic objectivity (Salehi et al., 2026). These biases—especially when applied to religious and philosophical texts—may generate outputs and decisions that are not neutral, but rather partial and inequitable. Such outcomes carry serious social, ethical, and legal consequences.

"Natural Language Processing (NLP) has enabled the analysis, translation, and extraction of concepts from historical, philosophical, and literary texts, offering a more precise understanding of scientific and cultural sources. In the fields of psychology and sociology, artificial intelligence contributes to a deeper examination of human behavior by conducting sentiment analysis and predicting social changes. Nevertheless, challenges such as algorithmic bias, the inability to grasp deeply embedded human concepts, and the inherent complexity of data interpretation continue to pose significant obstacles to its development" (Saghiri, 2024).

This situation indicates that employing artificial intelligence without attention to its cultural contexts and interpretive limitations may lead to the reproduction or even intensification of existing biases.

This discussion of cognitive simulation clarifies the limits of modeling human understanding and to distinguish technical replication from hermeneutic comprehension. Within this landscape, there has emerged an interdisciplinary research domain known as cognitive simulation, whose aim is to model and replicate human or natural intelligence in order to achieve a better understanding of human cognitive capabilities. Although this field benefits from software engineering techniques and artificial intelligence, it goes beyond purely computational processes and includes efforts to reproduce intelligence through biological methods as well—an approach commonly known as wet artificial intelligence.

Despite the advancements in cognitive simulation, it remains unclear whether genuine artificial intelligence endowed with consciousness and mental states can ever be created. Simulation does not recreate intelligence itself, just as flying in a simulator is not equivalent to actual flight (Rosengrün, 2023, p. 5).

Floridi rightly emphasizes that contemporary artificial intelligence, although successful in data processing and performing narrow tasks, is still incapable of reproducing human cognitive intelligence—that is, contextual understanding, conceptual grasp, mental flexibility, and intentionality, all of which are inherent to human cognition (Floridi, 2014, p. 282). This means that AI systems, even when they generate precise or creative outputs, still lack lived experience, intentionality, and self-awareness. In other words, current AI tools may outperform humans in certain tasks, but such performance must not be mistaken for genuine understanding.

In this context, Hooke (2023, pp. 147–148) underscores that the fundamental difference between human intelligence and artificial intelligence lies in the human capacity to understand context and to adapt to complex and non-deterministic situations. Humans can disregard noise and irrelevant data, reconstruct missing elements using background knowledge, and interpret information within meaningful frameworks. By contrast, artificial intelligence is constrained by predefined algorithms and lacks the ability to respond flexibly on the basis of contextual awareness. This limitation becomes particularly pronounced when analyzing religious and philosophical texts that contain intricate semantic layers and contextual depth.

From a hermeneutical perspective, as Douglas (1916–2020) explains, the input data of algorithms can function as a form of pre-understanding (Vorverständnis). Just as a human reader approaches a text with prior beliefs, experiences, and interpretive horizons, AI models also make decisions based on their training data. If this data contains social or cultural biases, the model's output will inevitably reproduce them. Yet, the fundamental difference lies in the fact that AI data lacks lived experience and intentionality, and the so-called "internal dialogue" occurring within AI models is nothing more than a mathematical computation—not a genuine human act of interpretation (Youvan, 2024, p. 11). In essence, algorithms operate solely on formal rules and surface-level propositions; they are incapable of apprehending the deeper spiritual concepts, contextual meanings of religious rulings, or the author's intention and purpose. That is, AI may process superficial linguistic patterns, but it cannot grasp the underlying semantic layers embedded within them.

The following reference to Islamic exegetical thought is not introduced as a parallel discourse, but as an internal test case through which the hermeneutic limits of algorithmic interpretation become concretely visible.

This insight resonates with the reflections of ʿAllāmeh Ṭabāṭabāʾī, who emphasizes that true understanding is a synthesis of contextual knowledge, lived experience, and philosophical insight—elements that algorithms are fundamentally incapable of reproducing. Ṭabāṭabāʾī notes, regarding the verse "So fear God as much as you are able" (Q 26:126), that "every individual, according to his distinct level of comprehension, understands the verse differently; each person applies it in accordance with the degree of piety that is possible for him" (Ṭabāṭabāʾī, 1374, vol. 3, p. 570). This demonstrates that the understanding of religious texts requires an interaction between meaning and the human lifeworld, as well as familiarity with historical and spiritual conditions.

Consequently, this limitation explains why algorithmic outputs must be treated as interpretive inputs rather than conclusions, thereby grounding the ethical responsibility of the human interpreter. They must be treated as supplementary, conditional tools whose outputs require continuous critique and assessment by human interpreters.

In fact, artificial intelligence algorithms operate on the basis of predefined rules and mathematical logic, whereas human emotions, implicit meanings, and cultural domains are often ambiguous, contradictory, and fluid. Understanding and interpreting such concepts requires flexibility, interpretive capacity, and contextual knowledge—capabilities that current AI systems fundamentally lack (Picca, 2024, p. 15).

AI models frequently inherit biases from their training data, which can lead to problematic outcomes, unjust decisions, or inaccurate representations of information. From a hermeneutical perspective, these biases represent underlying assumptions that must be continually assessed and critically examined. As Douglas explicitly states, training data function as foundational knowledge, and any deficiency or bias embedded within them shapes the model's interpretive processes (Youvan, 2024, p. 9).

However, the fundamental differences between machine data and human pre-understanding (Vorverständnis) must be taken into account. Unlike

human preconceptions, training data lack awareness, critique, or contextual grounding; they merely serve as inputs to algorithmic procedures. The responsibility for critically evaluating these data rests entirely on human developers, while the models themselves operate solely on the basis of algorithms and the data fed into them.

Hooke demonstrates that certain AI systems, in dynamic environments, have been able to acquire a limited form of "practical knowledge," which resembles Herbert Dreyfus's (1929–2017) phenomenological prerequisite for skilled coping. This practical knowledge involves situational awareness, distinguishing between possibilities, and responding flexibly. Nevertheless, AI still lacks personal identity, emotional intelligence, and embodied experience, qualities that are essential in complex human activities (Hooke, 2023, p. 151). Hooke's analysis shows that although some AI systems exhibit a surface-level form of "practical knowledge" analogous to Dreyfus's conditions for skilled coping, this capacity remains a technical simulation rather than an authentic realization. The absence of embodiment, personal identity, and affect creates a qualitative gap between human skill and algorithmic performance—a gap that highlights the boundary between technological advancement and the existential horizons of human agency.

From a hermeneutical perspective, training data can function as a form of "pre-understanding." Just as human interpretation is shaped by prior experiences, beliefs, and cultural background, AI models make decisions based on learned statistical patterns. If the data contain social, cultural, or economic biases, these assumptions will be reproduced in the model's outputs (Youvan, 2024, p. 21). However, equating training data with human pre-understanding is ultimately an exaggeration, since models lack lived experience, intentionality, and rich cultural presuppositions. Their learning processes are grounded solely in statistical relations between inputs, without the depth that characterizes human interpretive engagement.

Artificial intelligence is fundamentally limited in its engagement with the complexities and indeterminacies of the real world. Some approaches in computer science assume that human intelligence can be reduced to formal rules and finite datasets, whereas human experience encompasses a lifeworld, contextuality, and practical knowledge that cannot be precisely modeled (Hooke, 2023, p. 148).

Floridi and colleagues emphasize that the metaphorical use of everyday concepts to describe AI—such as the term "intelligent"—must be employed with caution. In scientific discourse, metaphors often provide meaning for emerging technical concepts, yet scholars must avoid the misunderstandings that such metaphors may generate (Floridi et al., 2018, p. 284).

The question of the validity and objectivity of machine-generated interpretations—especially regarding religious and philosophical texts—draws us into the core of the interaction between human beings, language, and technology. Although AI systems are capable of processing vast quantities of data, their interpretations and analyses merely reflect the choices, design parameters, and training imposed by human developers, rather than being the product of lived experience or intentional cognition.

In summary, on the basis of the preceding hermeneutic analysis of objectivity, bias, and pre-understanding, textual interpretation is inherently historical, contextual, and dependent on horizons of meaning—horizons that algorithms are not currently capable of fully reproducing (Noble, 2018, p. 317).

## 4. Conclusion

The following conclusions synthesize the results of three-level qualitative framework employed in this study, integrating descriptive, explanatory, and normative insights developed through philosophical hermeneutics and Islamic exegetical analysis.

**1.** The findings of this study demonstrate that digital hermeneutics, in an era dominated by algorithms and large-scale data, is not merely a technological approach but a profoundly philosophical, ethical, and social concern. The interaction between data, algorithms, and human interpretation reveals that the understanding of texts, especially religious and philosophical texts, cannot be adequately reduced to statistical processing.

**2.** From an epistemological perspective, it became evident that although algorithms are capable of generating text and reproducing linguistic patterns, they lack historical horizons, lifeworld experience, and human pre-understandings. Consequently, even when machine-generated outputs appear linguistically coherent, they remain insufficient with respect to genuine understanding.

**3.** One of the most significant contributions of this research is its illumination of the decisive role of data bias in shaping algorithmic results. Training datasets carry embedded values, assumptions, and cultural structures; accordingly, algorithmic outputs are not neutral but often mirror dominant cultures and power structures.

**4.** This situation underscores the necessity of attending to algorithmic ethics. Without ethical oversight and regulatory mechanisms, algorithms may become instruments for reproducing discrimination, inequality, and semantic distortion. Thus, the design of transparent, responsible, and accountable systems is not merely optional but constitutes a socio-epistemic necessity.

**5.** From a social perspective, the reproduction of bias at the algorithmic level can lead to serious consequences, including the restriction of cultural diversity, threats to privacy, and the erosion of social justice. Particularly in the domain of religious interpretation, such bias may silence diverse voices and impose an unintended homogenization of understanding.

**6.** The present analysis underscores the need to preserve and strengthen human agency within the interpretive process. The human interpreter is not merely a passive user of AI systems but must act as an active supervisor and critic—one who shapes the horizons of meaning and evaluates machine-generated outputs in light of interpretive traditions and lived human experience.

**7.** Accordingly, three key strategies can be proposed for addressing these challenges: first, ensuring algorithmic transparency and explainability; second, establishing ethical and religious oversight to prevent semantic distortion; and third, preserving and reinforcing interpretive traditions as the ultimate reference against machine-generated data. Together, these strategies can foster a balanced synthesis between technological innovation and human responsibility.

**8.** Ultimately, the study concludes that digital hermeneutics can become a constructive approach only when a dynamic equilibrium is established between the capabilities of artificial intelligence and the philosophical–ethical requirements of interpretation. Such an approach not only mitigates the risks of bias and distortion but can also contribute to epistemic enrichment, social justice, and cultural sustainability.

## Reference

*The Holy Quran*.

Ayyashi, M. bin M. (2016). *Tafsir al-Ayyashi* (A. Salehi, Trans.). Dhu al-Qarbi.

Bashir, M. H., Azmi, A. M., Nawaz, H., Zaghouani, W., Diab, M., Al-Fuqaha', A., & Qadir, J. (2022). *Arabic natural language processing for Qur'anic research: A systematic review*. Springer.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Capurro, R. (2010). Digital hermeneutics: An outline. *AI & Society, 35*(1), 35–42.

Demichelis, R. (n.d.). The hermeneutic turn of AI: Are machines capable of interpreting? *arXiv*.

Eder, M. (2016). Rolling stylometry. *Digital Scholarship in the Humanities, 31*(3), 457–469. https://doi.org/10.1093/LLC/FQV010

Faiz-Kashani, M. ibn Sh.-M. (1985). *Tafsir al-Safi* (M. B. S. Saedi Khorasani, Trans.). Safa.

Fakhr Razi, M. ibn O. (2000). *Al-Tafsir al-Kabir (Keys to the Unseen)* (A. A. Halabi, Trans.). Asatir.

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People: An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Floyd, C., Budde, R., Züllighoven, H., & Keil-Slawik, R. (Eds.). (1992). *Software development and reality construction*. Springer-Verlag.

Gadamer, H.-G. (2004). *Truth and method* (2nd rev. ed.). Continuum.

Hooke, J. D. F. (2023). Martin Heidegger's concept of understanding (Verstehen): An inquiry into artificial intelligence. *Analecta Hermeneutica, 15*(1), 1–23.

Ibn Arabi, M. ibn A. (2007). *Fusous al-Hikam* (M. A. Movahed & S. Movahed, Trans.). Karnameh.

Imam Ali (AS). (2000). *Nahj al-Balagha* (M. Dashti, Trans.). Mashhoor.

Javadi Amoli, A. (2016). *Tasnim commentary*. Israa Publishing Center.

Khomeini, R. (2003). *Explanation of the Hadith of the Soldiers of Wisdom and Ignorance*. Institute for the Compilation and Publication of Imam Khomeini's Works.

Lima, B. C., Omar, N., Avansi, I., & Nunes de Castro, L. (2025). Artificial intelligence applied to the analysis of biblical scriptures: A systematic review. *Analytics, 4*(2), 13. https://doi.org/10.3390/analytics4020013

Moor, J. (2006). The Dartmouth College Artificial Intelligence Conference: The next fifty years. *AI Magazine, 27*(4), 87.

Mousavi Miangah, T. (2007). *An introduction to machine translation*. Samt.

Mousavi, M. (1995). Reasons for the revelation in Tabari's Tafsir. *Islamic Sciences Quarterly, 6*(2), 9–25.

Nazari, A. (2014). Typology of challenges and the position of machine translation from Arabic to Persian. *Translation Research in Arabic Language and Literature, 4*(11), 29–56.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2017). *What's at stake* [TED Talk]. https://www.ted.com/talks/cathy_o_neil_the_era_of_personalized_learning

Picca, D. (2024). Emotional hermeneutics: Exploring the limits of artificial intelligence from a Diltheyan perspective. Paper presented at the 35th ACM Conference on Hypertext and Social Media, Poznan, Poland. https://doi.org/10.1145/3648188.3680255

Qazavi, S. M. (2024). *Investigating the critical foundations and requirements of cognitive anthropology from the perspective of transcendental wisdom* [Doctoral dissertation, Baqir-ul-Uloom University]. Faculty of Philosophy and Ethics.

Rosengrün, S. (2023). What is AI, and if so, how many? Four puzzles about artificial intelligence. *Analecta Hermeneutica, 15*(1).

Rostam, N. A. P., & Hassain Malim, N. H. A. (2021). Text categorisation in Quran and Hadith: Overcoming the interrelation challenges using machine learning and term weighting. *Journal of King Saud University - Computer and Information Sciences, 33*, 638–657. https://doi.org/10.1016/j.jksuci.2019.03.007

Saghiri, A. (2024). Applications of artificial intelligence in humanities and social sciences. In *Proceedings of the First International Conference on Artificial Intelligence in the Age of Digital Transformation*. Georgia.

Salehi, K., Habib Zadeh Khiyaban, S. and Sabbar, S. (2026). Artificial intelligence and crime detection: A critical review. *Journal of Cyberspace Studies, 10*(1), 181-197. https://doi.org/10.22059/jcss.2025.402206.1179

Salehi, K., Habib Zadeh Khiyaban, S., & Sabbar, S. (2025). Artificial Intelligence and the Future of International Law and Power. *Journal of World Sociopolitical Studies*, *9*(4), 923-958.

Tabatabaei, M. H. (1995). *Translation of Tafsir al-Mizan*. Islamic Publications Office.

Wang, J. (2021). Is artificial intelligence capable of understanding? An analysis based on philosophical hermeneutics. *Cultures of Science, 4*(3), 135–146.

Youvan, D. C. (2024). *Applying hermeneutic principles to AI: Enhancing interpretability, interaction, and ethical reflection in artificial intelligence systems*. https://doi.org/10.13140/RG.2.2.32289.83047