

Evaluating the Arabic–Indonesian Translation Quality of ChatGPT on Quranic and Hadith Texts

M. Syamsul Ma'arif^{1*}, Ruston Nawawi¹, Umar Faruq¹, Moh. Badrus Sholichin¹, Hilmy Baihaqy Yussof², Prabowo Adi Widayat³

¹Universitas Islam Negeri Syekh Wasil Kediri, Indonesia

²Universiti Islam Sultan Sharif Ali, Brunei Darussalam

³Universitas Muhammadiyah Metro, Indonesia

*Correspondence: ✉ msymaarif@iainkediri.ac.id

Abstract

This study examines the reliability of ChatGPT in translating Arabic–Indonesian religious texts, addressing the concern that AI often struggles with culturally embedded, legal, and theological expressions as highlighted in recent translation studies. The issue is significant because Qur'anic and hadith texts carry normative meanings whose alterations may lead to doctrinal misinterpretation. Data were collected from ten Qur'anic legal verses and ten hadiths from al-Arba'in al-Nawawiyyah, translated by ChatGPT using a standardized prompt. The translations were evaluated through Nababan et al.'s accuracy–acceptability–readability model by expert reviewers and general readers. Findings show that ChatGPT performs well in accuracy and readability but displays limitations in rendering legal terminology, cultural nuance, and doctrinally sensitive expressions. This research contributes empirical evidence on the strengths and risks of AI-assisted translation for Islamic texts and underscores the necessity of expert post-editing to ensure theological precision and culturally appropriate interpretation.

Article History

Received: 11-09-2025

Revised: 04-12-2025

Accepted: 11-12-2025

Keywords:

ChatGPT;
Religious Texts;
Translation Evaluation;
Translation Quality.



© 2025 M. Syamsul Ma'arif, Ruston Nawawi, Umar Faruq, Moh. Badrus Sholichin, Hilmy Baihaqy Yussof, Prabowo Adi Widayat

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

INTRODUCTION

The development of artificial intelligence (AI) technology has brought about significant advancements in natural language processing (NLP). One of the most notable breakthroughs is the emergence of ChatGPT, developed by OpenAI. ChatGPT is designed to respond to textual input with natural language output that closely resembles human communication (Kalla et al., 2023). It has also been increasingly adopted in religious and academic domains, including the translation and interpretation of sacred texts. In practice, ChatGPT has been widely applied across various domains such as education, translation, and digital services, with claims of enhancing efficiency and productivity.

The selection of religious texts—particularly the Qur'an and Hadith—as the focus of this study is grounded in their position as authoritative sources and the central foundation of Islamic teachings. Translating these texts requires an exceptionally high level of precision, as even the slightest shift in meaning may profoundly affect theological, legal, and ethical understanding within the Muslim community. In this regard, translation accuracy is not merely a linguistic concern, but also a matter of spiritual and cultural legitimacy (Gunawan et al., 2023). Qur'anic translation, in particular, is deeply influenced by the cultural context of the translator; therefore, cultural sensitivity is essential to ensure accuracy and avoid bias (AlJahsh, 2023). Recent evaluations of AI-based Qur'an translation further underline that domain-specific religious nuances are easily lost if models are treated as general-purpose translators (AlSukhni et al., 2016; Gaanoun & Alsuhaibani, 2025).

Furthermore, the linguistic and narrative complexity of the Qur'an (Dani, 2023) and Hadith—characterized by metaphorical expressions, distinctive morphological structures, and strong contextual significance—makes them highly challenging test cases for translation technology

(Zitouni et al., 2022). The ability of AI to produce adequate translations of texts with such complexity serves as a strong indication of its broader potential in Islamic studies as well as in textual translation more generally (Hassan Saleh Mahdi, 2022). At the same time, recent work on AI-assisted tafsīr and digital fatwas shows that seemingly minor lexical or structural shifts can have significant doctrinal implications when applied to normative Islamic sources (M. Faidh Fasyani et al., 2024; Mu'tashim Billah et al., 2023).

In the context of translation, ChatGPT demonstrates the ability to produce cross-linguistic translations with remarkable speed and fluency (Latifah & Djamilah, 2024). Several studies even suggest that its performance surpasses other translation engines. For instance, ChatGPT has been shown to be more accurate than Google Translate (Framesthia et al., 2024; Nurfaiza, 2024; Setyawan et al., 2025) and Gemini AI (Majid et al., 2024), and more effective than Bing Translator. It is capable of translating Arabic–Indonesian texts with a high degree of accuracy, free from grammatical or semantic errors, although its lexical choices tend to be more concise compared to Google Translate (Arju et al., 2025). Similar findings report that, supported by GPT-4, ChatGPT is able to compete with commercial translation products while generating the fewest errors (Jiao et al., 2023). Comparative studies on AI and human translations of religious texts further confirm that ChatGPT often matches or even exceeds conventional machine translation systems at the surface level, while still raising questions about deeper semantic and doctrinal adequacy (Abdaloussein, 2025; Sabrin Azzam Abdulrahman Fadhel., 2025; Shormani & Alfahad, 2025).

Nevertheless, a number of studies highlight significant limitations. Ruhmadi observes that ChatGPT remains weak in handling morphological aspects (Ruhmadi & Farisi, 2023; Saimin et al., 2024), which may lead to shifts in meaning, including in the translation of the Qur'an and Hadith. Maram further emphasizes that ChatGPT's translations of proverbs often encounter cultural, lexical, and literalness issues (El-Saadany, 2024). Zhaokun points out that although ChatGPT can rival Neural Machine Translation (NMT) systems, automatic evaluation metrics often diverge from human assessment, underscoring the continued necessity of human involvement (Jiang et al., 2024). Meanwhile, other studies reveal that online translation systems, including Google Translate and Bing, remain ineffective for religious texts such as the Qur'an, as reflected in their low ATEC evaluation scores (AISukhni et al., 2016). More recent evaluations of English Qur'anic translations produced by ChatGPT confirm this discrepancy between automatic scores and expert-based judgment, particularly for verses with dense legal or theological content (Dahia & Belbacha, 2024).

In addition, research on Arabic–Indonesian translation indicates that classical challenges such as interference, cultural vocabulary, metaphor, transliteration, morphology, syntax, and idioms continue to hinder translation accuracy (Huda, 2018; Maulana, 2022). Syihabudin shows that variations in the accuracy of Qur'anic term translations are strongly influenced by translation techniques, cultural differences, and lexical understanding (Syihabuddin, 2016). Salma identifies as many as 18 Arabic–Indonesian translation techniques, each requiring translator flexibility to maintain semantic precision (Istiqomah et al., 2023). On the other hand, Tubagus proposes a local wisdom-based approach, such as the *Ngalodern* method, to improve post-editing of machine translations (Nugraha et al., 2022). In the specific case of AI-mediated Arabic–Indonesian translation, recent studies also show that ChatGPT tends to produce fluent output while still struggling with culturally embedded religious terminology and genre conventions (Abidah et al., 2024).

These findings collectively affirm that the quality of translation, particularly Arabic–Indonesian texts, is determined not solely by the sophistication of the machine but also by methodological factors and human involvement. In this regard, holistic models of translation quality assessment, such as that developed by Nababan et al. (Nababan & Nuraeni, n.d.) are essential as they provide a comprehensive evaluation of accuracy, acceptability, and readability. Previous applications of this model, however, have largely focused on general, literary, or technical texts rather than AI-generated translations of sacred or normative religious sources (Ruth Suci Putri et al., 2022; Trisakti Agriani et al., 2018). However, previous studies have rarely applied this model

specifically to AI-generated translations of authoritative Islamic texts. The present study offers a distinct contribution by using Nababan's framework to evaluate ChatGPT's translations of ten Qur'anic legal verses (*āyāt al-ahkām*) and ten hadiths from al-Arba'īn al-Nawawī, assessed by multidisciplinary expert reviewers and lay readers, thereby providing a more fine-grained picture of how an AI system handles normative legal–theological content. By focusing on highly normative verses and prophetic traditions that are central in Islamic education, this study extends prior work on AI and Qur'an/Hadith translation, which has typically examined either individual verses, general religious prose, or non-legal passages (AlSukhni et al., 2016; M. Faidh Fasyani et al., 2024).

Despite the growing interest in AI translations, there remains a significant gap in empirical studies that systematically evaluate ChatGPT's Indonesian translations using a human-centered assessment model. The majority of existing research tends to focus on general texts, rely on automated evaluation scores, or explore isolated lexical items, often overlooking the involvement of multidisciplinary expert reviewers. This study seeks to rectify this gap by analyzing ChatGPT's translations of ten Qur'anic legal verses (*āyāt al-ahkām*) and ten prophetic traditions from al-Arba'īn al-Nawawī. These texts have been intentionally selected due to their profound legal and doctrinal significance, their widespread application in Islamic education, and their compact yet conceptually rich nature, making them well-suited for evaluating AI translation capabilities. In undertaking this research, we not only assess the surface quality of the AI's output but also explore the extent to which a readily available large language model can be trusted to accurately render texts that have binding legal and ethical implications for Muslim audiences.

This study seeks to assess the quality of ChatGPT's translations of selected Qur'anic legal verses and hadiths from Arabic to Indonesian, utilizing Nababan's translation quality assessment model. The analysis will concentrate on three primary dimensions: accuracy, acceptability, and readability, in order to determine how well ChatGPT's outputs convey the original meanings while remaining accessible to Indonesian readers. The anticipated findings aim to provide theoretical insights into the role of AI-assisted translation in Islamic studies, as well as practical guidance for the application of AI tools in digital Islamic education and Arabic-Indonesian translation policy. Furthermore, this study will contribute to ongoing discussions about the epistemic and ethical boundaries associated with employing general-purpose AI systems for the translation of sacred and norm-setting texts in Muslim contexts.

METHODS

This study employs a qualitative descriptive approach with the aim of evaluating the quality of Arabic–Indonesian translations produced by ChatGPT for legal verses (*āyāt al-ahkām*) in the Qur'an and selected Hadiths from Arba'īn al-Nawawī. The choice of this approach is based on the need to gain an in-depth understanding of the linguistic, semantic, and cultural dimensions of machine-generated translations. Rather than relying solely on quantitative measurement, the study seeks to assess meaning, acceptability, and readability within the context of the Indonesian language, following established practices in translation quality assessment research that prioritize rich qualitative judgment over purely automatic metrics (Nababan et al., 2012; Ruth Suci Putri et al., 2022).

The data in this study consist of twenty normative Islamic texts. Ten Qur'anic verses on legal themes (*āyāt al-ahkām*) were purposively selected from various chapters to represent major domains of Islamic law, including family law (marriage and inheritance), economic transactions, criminal law, constitutional matters, international relations, judicial procedures, and ritual worship (ablution, prayer, and fasting). In parallel, ten hadiths were chosen from the first ten narrations of Arba'īn al-Nawawī, covering key doctrines and ethical principles such as intention (*niyyah*), Islam–*īmān–ihsān*, lawful and unlawful matters, sincere counsel (*naṣīḥah*), and the sanctity of Muslim life. These texts were selected because they combine high semantic and conceptual density, are widely used in Islamic education, and carry significant legal and ethical implications, making them particularly suitable for testing AI-based translation systems. This kind of purposive sampling of

legally and theologically salient Qur'anic verses and hadiths is consistent with recent AI-and-religion studies that focus on norm-bearing texts rather than purely narrative material (Abdaloussein, 2025; Dahia & Belbacha, 2024; Sabrin Azzam Abdulrahman Fadhel., 2025).

The Arabic texts were not retrieved from ChatGPT but rather from verified primary sources. The Qur'anic verses were taken directly from the muṣḥaf, while the Hadiths were obtained from a fully vocalized *tahqīq* edition of Arba'īn al-Nawawī. This measure was taken to avoid potential distortions of wording, since ChatGPT does not have direct access to original textual databases and operates solely on linguistic predictions derived from training data, which may not always be valid in terms of sanad and matn. This step also aligns with methodological cautions raised in recent work on AI-assisted tafsīr and fatwa generation, which stresses the need to anchor AI output in authenticated scriptural sources (M. Faidh Fasyani et al., 2024).

The translation process was carried out using the latest version of ChatGPT available at the time of research (GPT-4o, free version). To ensure consistency, a fixed prompt was employed as follows: */Terjemahkan teks Arab berikut ke dalam bahasa Indonesia dengan memperhatikan tiga aspek kualitas terjemahan: (1) keakuratan (kesetiaan pada makna teks sumber), (2) keberterimaan (kesesuaian dengan norma kebahasaan dan budaya Indonesia), dan (3) keterbacaan (kejelasan dan kelancaran dibaca oleh penutur Indonesia)/*. Prompting the model with explicit quality dimensions mirrors recent experimental designs in AI translation research, which often guide large language models toward specific evaluation criteria (Abidah et al., 2024; Shormani & Alfahad, 2025).

The prompt was provided once at the beginning of the session, followed by block inputs (each consisting of five verses or Hadiths) without repetition of the prompt, in order to simulate typical user interaction and to test the stability of ChatGPT's responses in a natural context. Varying input blocks in this way also allows observation of potential consistency issues that have been reported when large language models process religious texts across different sessions or batching conditions (Gaanoun & Alsuhaibani, 2025).

Additional variables examined in this study included access time (across different days), input size (1, 5, or 10 texts), and the consistency of prompt structure. ChatGPT's translations were then compared with human translations from authoritative sources. For Qur'anic verses, the official version from the Indonesian Ministry of Religious Affairs (quran.kemenag.go.id) was used, while for the Hadiths of Arba'īn al-Nawawī, the translation was taken from the widely used website www.alkhoirot.org. This comparative design is in line with earlier evaluations of machine translation for Qur'anic and religious texts, which typically benchmark system output against recognized human translations (AlSukhni et al., 2016; Shormani & Alfahad, 2025).

Translation quality was evaluated through triangulation of both methods and sources. The assessment involved focus group discussions and in-depth interviews with experts from multiple disciplines: lecturers of Qur'anic exegesis and Hadith (accuracy), lecturers of Arabic language and translation (lexical and syntactic equivalence), lecturers of Indonesian language (acceptability), as well as students and general readers (readability). For the readability assessment, informants were only provided with ChatGPT's translations without the Arabic texts or their human-rendered equivalents, ensuring that the evaluation remained unbiased. The use of expert panels alongside non-expert readers follows best practices in translation quality studies that seek to balance specialist judgment with end-user reception (Nababan et al., 2012; Trisakti Agriani et al., 2018).

The analysis was conducted descriptively using Nababan et al.'s translation quality assessment model, which comprises three main aspects: accuracy, acceptability, and readability. The evaluation employed numerical scoring, supplemented by qualitative descriptions based on the evaluators' perceptions and interpretations. All assessment results were synthesized independently to produce a comprehensive overview of the strengths and weaknesses of ChatGPT's translations of religious texts. By applying this model—previously used mainly for human translations of general or literary texts—to AI-generated translations of Qur'anic legal verses and hadiths, the study extends its scope into the domain of sacred and norm-producing discourse (Mustofa et al., 2025; Ruth Suci Putri et al., 2022).

RESULTS AND DISCUSSION

Accuracy, Acceptability, and Readability Scores for Chatgpt's Translation

Once the corpus had been established, the quality of ChatGPT's translations was evaluated using Nababan's model, which comprises three main aspects: accuracy, acceptability, and readability. The assessment was carried out by a panel of experts and lay readers, as described in the Methods section. A summary of the scores for the Qur'anic verses is presented in table 1, while table 2 summarises the salient differences between the official translation of the Indonesian Ministry of Religious Affairs (Kementerian Agama or Kemenag) and ChatGPT's translation for each verse. Accordingly, the conclusions in this subsection are grounded in both quantitative data (Nababan scores) and qualitative notes from the focus group discussions (FGDs).

Table 1. Accuracy, Acceptability, and Readability Scores for Chatgpt's Translation of the *Āyāt Al-Aḥkām*

| No. | Verse | Theme | Accuracy | Acceptability | Readability |
|-----|------------------|---------------------------|----------|---------------|-------------|
| 1 | Al-Baqarah 2:221 | Marriage law | 2 | 3 | 3 |
| 2 | Al-Baqarah 2:240 | Inheritance law | 2 | 2 | 2 |
| 3 | Āl 'Imrān 3:130 | Economic/financial law | 3 | 3 | 2 |
| 4 | Ash-Shūrā 42:40 | Islamic criminal law | 3 | 3 | 3 |
| 5 | An-Nisā' 4:59 | Constitutional law | 3 | 3 | 3 |
| 6 | Al-Anfāl 8:39 | International law | 3 | 3 | 3 |
| 7 | Al-Mā'idah 5:8 | Judicial procedures | 3 | 3 | 3 |
| 8 | Al-Mā'idah 5:6 | Ablution (<i>wuḍū'</i>) | 3 | 2 | 2 |
| 9 | Al-Baqarah 2:238 | Prayer (<i>ṣalāh</i>) | 3 | 2 | 2 |
| 10 | Al-Baqarah 2:183 | Fasting (<i>ṣawm</i>) | 3 | 3 | 3 |

Table 2. Summary of Differences between Kementerian Agama's Translation and ChatGPT's Translation for the *Āyāt Al-Aḥkām*

| No. | Sūrah-Verse | Excerpt from Kemenag Translation | Excerpt from ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|-----|-------------|---|--|--|
| 1 | 2:221 | <i>Janganlah kamu menikahi perempuan musyrik hingga mereka beriman... hamba sahaya perempuan yang beriman lebih baik</i> | <i>Janganlah kalian menikahi perempuan musyrik sebelum mereka beriman... seorang hamba perempuan yang beriman lebih baik</i> | ChatGPT adds a tafsīr-based gloss / <i>anak perempuan kalian</i> / to the phrase <i>wa lā tunkihū al-mushrikīn</i> and uses / <i>hamba</i> / without the qualifier / <i>sahaya</i> /. The core meaning remains correct, but there is an interpretive expansion and a slight blurring of the socio-historical context of slavery; this underlies the accuracy score of 2, while acceptability and readability are both rated 3. |
| 2 | 2:240 | <i>Orang-orang yang akan mati di antaramu dan meninggalkan istri-istri hendaklah membuat wasiat... nafkah sampai setahun tanpa mengeluarkannya dari rumah</i> | <i>Dan orang-orang di antara kalian yang wafat dan meninggalkan istri-istri, hendaklah memberikan wasiat... nafkah selama satu tahun tanpa mengusir mereka</i> | Kemenag clearly presents <i>waṣīyyah</i> as a legal act undertaken before death / <i>Orang-orang yang akan mati</i> / 'who are going to die', / <i>hendaklah membuat wasiat</i> / 'make a bequest'. ChatGPT's wording suggests that the act occurs after |

| No. | Sūrah-Verse | Excerpt from Kemenag Translation | Excerpt from ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|-----|-------------|--|---|--|
| | | | | death <i>orang-orang di antara kalian yang wafat</i> 'who have died', /hendaklah memberikan wasiat/ 'give a bequest', making the timing and subject of the action less accurate. This reduces accuracy and also affects acceptability and readability (scores 2-2-2). |
| 3 | 3:130 | <i>Wahai orang-orang yang beriman! Janganlah kamu memakan riba dengan berlipat ganda, dan bertakwalah kepada Allah agar kamu beruntung</i> | <i>Wahai orang-orang yang beriman! Janganlah kalian memakan riba yang berlipat ganda, dan bertakwalah kepada Allah agar kalian beruntung</i> | Lexically and structurally the two versions are almost identical; differences are limited to the pronoun /kamu/ versus /kalian/ and minor phrasing /dengan berlipat ganda/ versus /yang berlipat ganda/. Accuracy and acceptability are therefore rated very high (3-3); readability is slightly below maximum (2) because the sentence remains relatively long. |
| 4 | 42:40 | <i>Balasan suatu keburukan adalah keburukan yang setimpal. Akan tetapi, siapa yang memaafkan dan berbuat baik (kepada orang yang berbuat jahat), maka pahalanya dari Allah</i> | <i>Balasan bagi suatu kejahatan adalah kejahatan yang serupa. Tetapi barang siapa memaafkan dan berdamai, maka pahalanya ada di sisi Allah</i> | The main difference lies in /memperbaiki/ versus /berdamai/. Both derive from the root <i>islāḥ</i> and are theologically acceptable, although /berdamai/ is somewhat narrower in scope than 'improve/rectify the situation'. Overall, accuracy, acceptability, and readability are all rated very good (3-3-3). |
| 5 | 4:59 | <i>Wahai orang-orang yang beriman, taatilah Allah dan taatilah Rasul (Nabi Muhammad) serta ululamri (pemegang kekuasaan) di antara kamu. Jika kamu berbeda pendapat</i> | <i>Wahai orang-orang yang beriman! Taatilah Allah, taatilah Rasul, dan pemimpin di antara kalian. Jika kalian berselisih...kembalikanlah kepada Allah dan Rasul</i> | ChatGPT replaces the technical term <i>ulī l-amr</i> with the more communicative phrase /pemimpin di antara kalian/ 'leaders among you'. This increases acceptability and readability for general readers but slightly weakens the specificity of the fiqh/theological term. The panel nonetheless rates accuracy as 3 because the notion of legitimate |

| No. | Sūrah-Verse | Excerpt from Kemenag Translation | Excerpt from ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|-----|-------------|--|---|---|
| | | | | authority is preserved; acceptability and readability are also rated 3. |
| 6 | 8:39 | <i>Perangilah mereka sampai tidak ada lagi fitnah (penganiayaan atau syirik) dan agama seutuhnya hanya bagi Allah. Jika mereka berhenti (dari kekufuran), sesungguhnya Allah Maha Melihat apa yang mereka kerjakan</i> | <i>Dan perangilah mereka sampai tidak ada lagi fitnah dan agama hanya milik Allah semata. Jika mereka berhenti, maka sungguh Allah Maha Melihat apa yang mereka kerjakan</i> | There is no substantive change in meaning. Consequently, the scores for accuracy, acceptability, and readability are all 3. |
| 7 | 5:8 | <i>Wahai orang-orang yang beriman! Jadilah kamu penegak keadilan karena Allah, sebagai saksi dengan adil. Dan janganlah kebencian suatu kaum mendorongmu</i> | <i>Wahai orang-orang yang beriman! Jadilah kalian penegak keadilan karena Allah, menjadi saksi dengan adil. Janganlah kebencian terhadap suatu kaum mendorong kalian untuk berlaku tidak adil</i> | The translations are almost parallel; differences concern the pronoun singular / <i>kamu</i> / versus plural / <i>kalian</i> / and slight variation in phrasing / <i>mendorongmu untuk tidak berlaku adil</i> / versus / <i>mendorong kalian untuk berlaku tidak adil</i> /. The panel assigns the maximum score (3–3–3) for all three aspects. |
| 8 | 5:6 | <i>Apabila kamu berdiri hendak melaksanakan salat, maka basuhlah wajahmu dan tanganmu sampai ke siku serta usaplah kepalamu dan (basuh) kedua kakimu sampai kedua mata kaki</i> | <i>Apabila kalian hendak melaksanakan salat, maka basuhlah wajah kalian dan tangan kalian sampai ke siku, sapulah kepala kalian, dan basuhlah kaki kalian sampai kedua mata kaki</i> | In terms of meaning, the sequence of ablution requirements and conditions is correct, so accuracy is rated 3. However, several technical phrases are somewhat long and rigid, and the repeated use of / <i>kalian</i> / makes the text feel dense; as a result, some evaluators assigned scores of 2 for acceptability and readability. |
| 9 | 2:238 | <i>Peliharalah semua salat (fardu) dan salat Wustā... (catatan kaki: menurut pendapat masyhur, salat Wustā adalah salat Asar)</i> | <i>Peliharalah salat-salat kalian, terutama salat tengah (salat pertengahan), dan berdirilah untuk Allah dengan khusyuk</i> | Kemenag preserves the technical term <i>ṣalāt al-wuṣṭā</i> in transliteration and clarifies its meaning (‘ <i>Aṣr</i> prayer’) in a footnote. ChatGPT renders it literally as / <i>salat tengah</i> / “the middle prayer” or “the prayer in the middle”, which is not a |

| No. | Sūrah–Verse | Excerpt from Kemenag Translation | Excerpt from ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|-----|-------------|---|---|---|
| 10 | 2:183 | Wahai orang-orang yang beriman, diwajibkan atas kamu berpuasa sebagaimana diwajibkan atas orang-orang sebelum kamu agar kamu bertakwa | Wahai orang-orang yang beriman! Diwajibkan atas kalian berpuasa sebagaimana diwajibkan atas orang-orang sebelum kalian agar kalian bertakwa | conventional expression in Indonesian fiqh discourse. At the level of general meaning, the idea of a “central” prayer is correct, but theologically and linguistically the expression is considered less precise. This explains why acceptability and readability drop to 2, even though accuracy is rated 3 at the level of overall content. The only differences are the pronoun /kamu/ versus /kalian/ and punctuation style. In terms of content and contemporary Indonesian style, ChatGPT’s translation is rated very highly on all three aspects (3–3–3). |

Assessment Results for the Translation of Qur’anic Legal Verses

The ten *āyāt al-aḥkām* analysed in this study cover various domains of Islamic law: family law (Qur’an 2:221, 2:240), economic matters and usury (Qur’an 3:130), criminal law and *qiṣāṣ* (Qur’an 42:40), constitutional obedience (Qur’an 4:59), international relations and conflict (Qur’an 8:39), judicial procedures and justice (Qur’an 5:8), as well as ritual acts of worship such as ablution, prayer, and fasting (Qur’an 5:6; 2:238; 2:183).

Based on the panel’s assessment, the average accuracy score for these verses was approximately 2.8, indicating that, in general, ChatGPT was able to convey the core meaning of the verses adequately. Verses related to public law—such as criminal law (Qur’an 42:40), obedience to *ulī l-amr* (Qur’an 4:59), conflict and aggression (Qur’an 8:39), and judicial justice (Qur’an 5:8)—tended to receive high scores for accuracy, acceptability, and readability. Their relatively straightforward sentence structures, the use of already well-established terminology (e.g. *‘adil*, *zalim*, “aggression”), and a more “general” legal context made it comparatively easy for ChatGPT to preserve meaning while still producing natural Indonesian.

By contrast, lower scores were observed for verses dealing with family law (Qur’an 2:221; 2:240) and ritual worship (Qur’an 5:6; 2:238). In this group, accuracy remained in the “good” category, but acceptability and readability declined. This was primarily due to the presence of technical fiqh terms (such as *‘iddah*, *ṭahārah*, *al-ṣalāt al-wuṣṭā*) and socio-historical contexts (slavery, guardian–female relationships) that were not always fully captured in ChatGPT’s lexical choices.

To illustrate these patterns in greater detail, the following subsection presents an analysis of three verses that were judged particularly salient in terms of accuracy, acceptability, and readability, namely Qur’an 2:221, 2:240, and 2:238 (Al-Baqarah). Summary comparisons between the Kemenag translation and ChatGPT’s translation for each of these verses are presented in tables 3, 4, and 5.

Analysis of Quran Surah Al-Baqarah Verse 221

One of the verses that received an accuracy score of 2 and an acceptability score of 3 was Surah Al-Baqarah verse 221, which contains the prohibition of marrying polytheists. This verse was selected because, substantively, it addresses fundamental principles of marital law, while linguistically it employs diction bearing significant socio-cultural and legal weight.

وَلَا تَنْكِحُوا الْمُشْرِكَةَ حَتَّىٰ تُؤْمِنَ ۗ وَلَأَمَةٌ مُّؤْمِنَةٌ خَيْرٌ مِّنْ مُّشْرِكَةٍ وَلَا تُنْكَحُوا الْمُشْرِكِينَ حَتَّىٰ يُؤْمِنُوا ۗ
وَلَعِبْدٌ مُّؤْمِنٌ خَيْرٌ مِّنْ مُّشْرِكٍ وَلَا أَعْجَبَكُمْ ۗ أُولَٰئِكَ يَدْعُونَ إِلَى النَّارِ ۗ وَاللّٰهُ يَدْعُوا إِلَى الْجَنَّةِ وَالْمَغْفِرَةِ بِإِذْنِهِ ۗ وَيُبَيِّنُ آيَاتِهِ
لِلنَّاسِ لَعَلَّهُمْ يَتَذَكَّرُونَ

Kementerian Agama Translation:

Janganlah kamu menikahi perempuan musyrik hingga mereka beriman! Sungguh, hamba sahaya perempuan yang beriman lebih baik daripada perempuan musyrik, meskipun dia menarik hatimu. Jangan pula kamu menikahkan laki-laki musyrik (dengan perempuan yang beriman) hingga mereka beriman. Sungguh, hamba sahaya laki-laki yang beriman lebih baik daripada laki-laki musyrik meskipun dia menarik hatimu. Mereka mengajak ke neraka, sedangkan Allah mengajak ke surga dan ampunan dengan izin-Nya. (Allah) menerangkan ayat-ayat-Nya kepada manusia agar mereka mengambil pelajaran.

ChatGPT Translation:

Janganlah kalian menikahi perempuan musyrik sebelum mereka beriman. Sesungguhnya seorang hamba perempuan yang beriman lebih baik daripada perempuan musyrik, walaupun ia menarik hati kalian. Dan janganlah kalian menikahkan (anak perempuan kalian) dengan laki-laki musyrik sebelum mereka beriman. Sesungguhnya seorang hamba laki-laki yang beriman lebih baik daripada laki-laki musyrik, walaupun ia menarik hati kalian. Mereka mengajak kepada neraka, sementara Allah mengajak ke surga dan ampunan dengan izin-Nya. Dan Allah menjelaskan ayat-ayat-Nya kepada manusia agar mereka

Table 3. Comparison of the Translation of Qur'an Surah Al-Baqarah verse 221 between Kementerian Agama and ChatGPT

| Aspect Compared | Kemenag Translation | ChatGPT Translation |
|------------------------------------|---|---|
| Subject of <i>tunkihū</i> | Does not explicitly state the human subject / <i>Jangan pula kamu menikahkan</i> / 'Do not marry' | Makes the subject explicit / <i>janganlah kalian menikahkan (anak perempuan kalian)</i> / 'Do not marry off (your daughters)' |
| Terms <i>amah</i> and ' <i>abd</i> | /hamba sahaya perempuan/laki-laki/ 'female slave... male slave' | /seorang hamba perempuan/laki-laki/ 'a female servant a male servant' |
| Style and pronouns | Uses / <i>kamu</i> /; more formal style, close to standard <i>muṣḥaf</i> translation | Uses / <i>kalian</i> /; more communicative style, closer to contemporary Indonesian usage |
| Conjunctions and cohesion | Uses / <i>sedangkan</i> /, / <i>meskipun</i> / | Uses / <i>sementara</i> /, / <i>walaupun</i> / |

Qur'an Surah Al-Baqarah verse 221 sets out the prohibition of marrying believing women to polytheist men and vice versa, and contrasts the virtue of a believing slave with that of a free but polytheist person. This verse was selected because it contains sensitive family-law terminology and vocabulary linked to the socio-historical context of slavery.

As summarised in table 3, ChatGPT generally succeeds in preserving the prohibitive structure and the evaluative emphasis on faith: the prohibition of marrying or marrying off to polytheists, the assertion that a believing slave is better than a polytheist, and the contrast between the call to the Fire and Allah's call to Paradise and forgiveness. However, the panel identified two critical points.

First, in the phrase *wa lā tunkihū al-mushrikīn*, the Kemenag translation does not explicitly identify the human subject, whereas ChatGPT renders it as */janganlah kalian menikahkan (anak perempuan kalian) dengan laki-laki musyrik/* ‘do not marry off (your daughters) to polytheist men...’. The insertion */anak perempuan kalian/* ‘your daughters’ is consistent with many tafsīr works that interpret the subject as the guardians (*awliyā*), yet in the context of Qur’anic translation this insertion of exegetical material into the main text risks blurring the boundary between translation and interpretation. For this reason, the panel judged the accuracy of this verse to be only moderately accurate with notes (score 2), even though, in terms of clarity for lay readers, the move is helpful.

Second, with regard to the terms *amah* and *‘abd*, Kemenag translates them as */hamba sahaya/* ‘female slave/male slave’, whereas ChatGPT chooses */hamba/* ‘female servant/male servant’. Omitting the word */sahaya/* weakens the marker of the slavery context that these terms are intended to convey, so that modern readers may lose sight of the fact that what is meant is not merely ‘servant’ in a metaphorical sense, but a specific socio-legal status at the time of revelation.

From the standpoint of acceptability and readability, ChatGPT’s translation is in fact rated very highly. The use of the plural pronoun */kalian/*, conjunctions such as */sementara/* and */walaupun/*, and the overall rhythm of the sentences make the text sound natural to contemporary Indonesian readers. Thus, Qur’an 2:221 illustrates a pattern in which ChatGPT performs strongly in terms of acceptability and readability, but requires close monitoring with respect to the accuracy of legal concepts and the social background of the verse.

Analysis of Quran Surah Al-Baqarah Verse 240

Arabic Text:

وَالَّذِينَ يَتَّقُونَ مِنكُمُ وَيَذَرُونَ أَزْوَاجًا وَصِيَّةً لِأَزْوَاجِهِمْ مَتَاعًا إِلَى الْحَوْلِ غَيْرَ إِخْرَاجٍ فَإِنْ خَرَجْنَ فَلَا جُنَاحَ عَلَيْكُمْ فِي مَا فَعَلْنَ فِي أَنفُسِهِنَّ مِنْ مَّعْرُوفٍ وَاللَّهُ عَزِيزٌ حَكِيمٌ

Kementerian Agama Translation:

Orang-orang yang akan mati di antara kamu dan meninggalkan istri-istri hendaklah membuat wasiat untuk istri-istrinya, (yaitu) nafkah sampai setahun tanpa mengeluarkannya (dari rumah). Akan tetapi, jika mereka keluar (sendiri), tidak ada dosa bagimu mengenai hal-hal yang patut yang mereka lakukan terhadap diri mereka sendiri. Allah Maha Perkasa lagi Maha Bijaksana.

ChatGPT Translation:

Orang-orang di antara kalian yang wafat dan meninggalkan istri-istri, hendaklah memberikan wasiat kepada istri-istri mereka berupa nafkah selama setahun tanpa mengusir mereka. Tetapi jika mereka keluar (dari rumah), maka tidak ada dosa bagi kalian atas apa yang mereka lakukan terhadap diri mereka secara patut. Dan Allah Maha Perkasa lagi Maha Bijaksana.

Table 4. Comparison of the Translation of Qur’an Al-Baqarah verse 240 between Kementerian Agama and ChatGPT

| Aspect Compared | Kemenag Translation | ChatGPT Translation |
|---|---|--|
| Subject phrase <i>alladzīna yutawaffawna</i> | <i>/Orang-orang yang akan mati di antara kamu/</i> ‘Those among you who will die’ | <i>/Orang-orang di antara kalian yang wafat/</i> ‘Those among you who have passed away’ |
| Verb related to <i>wasīyyah</i> | <i>/hendaklah membuat wasiat untuk istri-istrinya/</i> ‘should make a will for his wives’ | <i>/hendaklah memberikan wasiat kepada istri-istri mereka/</i> ‘should give a will to their wives’ |
| Temporal nuance of the legal act | Indicates a legal act performed before death <i>/pembuatan wasiat saat hidup/</i> | Reads as if the act occurs after death <i>/yang wafat/</i> → then <i>/memberikan wasiat/</i> |

| | | |
|------------------------|--|---|
| Diction and politeness | <i>/mengeluarkannya (dari rumah)/</i> ‘take it out (from the house)’ | <i>/mengusir mereka (dari rumah)/</i> ‘evict them (from the house)’ |
| Sentence style | Closer to a literal style and classical fiqh formulation | More communicative and polite <i>/wafat/</i> , <i>/mengusir/</i> , but less precise juridically |

Qur'an Al-Baqarah verse 240 regulates the *waṣiyyah* of maintenance and housing for a wife whose husband has died for the duration of one year, as well as the possibility for the wife to leave the house of her own accord. This verse was selected because it contains detailed rulings on inheritance and *'iddah* that require precision regarding timing and the subject of the legal act.

Table 6 shows that the global meaning concerning the obligation to provide a *waṣiyyah* of maintenance for one year and the prohibition of */mengusir istri dari rumah/* ‘evicting one's wife from the house’ is preserved in both translations. However, there is a significant difference in how the timing of the *waṣiyyah* is represented. Kemenag translates the opening part with */orang-orang yang akan mati di antara kamu... hendaklah membuat wasiat.../* ‘Those of you who are about to die... should make a will...’, which implicitly marks the legal act as taking place during the person’s lifetime. ChatGPT, by contrast, uses the formulation */orang-orang di antara kalian yang wafat dan meninggalkan istri-istri, hendaklah memberikan wasiat.../* ‘Those of you who die and leave wives behind should make a will...’. Pragmatically, this wording suggests that the act */memberikan wasiat/* occurs after the subject has died, thereby creating ambiguity regarding who performs the legal act and when it is carried out.

From a fiqh perspective, this shift in nuance is judged to be problematic, as it may affect readers’ understanding of the correct procedure for *waṣiyyah*. For this reason, the panel assigned an accuracy score of 2 to this verse, even though the global message regarding financial support and the prohibition of */mengusir istri/* ‘evicting one's wife’ remains understandable.

On the other hand, ChatGPT’s lexical choices, such as */wafat/* ‘passed away’ instead of */mati/* ‘die’, and its shorter sentence structure were considered to enhance acceptability and readability. Lay readers find it easier to follow the flow of the sentences and to grasp the general message. The case of Qur'an 2:240 thus shows that ease of reading does not always coincide with precision in legal meaning, and that the involvement of fiqh experts remains necessary in the post-editing of AI-generated outputs.

Analysis of Quran Surah Al-Baqarah Verse 238

Arabic Text:

حَافِظُوا عَلَى الصَّلَوَاتِ وَالصَّلَاةِ الْوُسْطَىٰ وَقُومُوا لِلَّهِ قَانِتِينَ

Kementerian Agama Translation:

Peliharalah semua salat (fardu) dan salat Wusṭā. Berdirilah karena Allah (dalam salat) dengan khushyuk. Catatan kaki: Menurut pendapat yang masyhur, salat Wusṭā adalah salat Asar.

ChatGPT Translation:

Peliharalah salat-salat kalian, terutama salat tengah (salat pertengahan), dan berdirilah untuk Allah dengan khushyuk.

Table 5. Comparison of the translation of Qur'an 2:238 between Kementerian Agama and ChatGPT

| Aspect Compared | Kemenag Translation | ChatGPT Translation |
|--|---|--|
| Key term <i>al-ṣalāt al-wusṭā</i> | Retained as <i>/salat Wusṭā/</i> and explained in a footnote as <i>/salat Asar/</i> | Translated literally as <i>/salat tengah (salat pertengahan)/</i> |
| Strategy for translating technical terms | Combination of transliteration + exegetical explanation in a footnote | Direct translation into a general equivalent without reference to the tafsir tradition |
| Space for | Opens room for readers to consult tafsir | Simplifies the term so that it appears as |

| | | |
|--------------------|---|--|
| interpretation | and fiqh literature through the preserved technical term | a generic concept without specific markers |
| Style and pronouns | <i>/Peliharalah semua salat (fardu)... Berdirilah karena Allah.../</i> 'Perform all prayers (fardu)... Stand up for Allah...' | <i>/Peliharalah salat-salat kalian... berdirilah untuk Allah.../</i> 'Maintain your prayers... stand up for Allah' |

Qur'an Surah Al-Baqarah verse 238 emphasises the command to maintain the obligatory prayers and, in particular, mentions */al-ṣalāt al-wuṣṭā/*, which in the tradition of the majority of scholars is generally identified with the */salat Asar/*. This verse provides an important example of how ChatGPT handles technical terms that have a long-standing history of interpretation.

In the Kemenag translation, the term is retained in transliterated form as */salat Wuṣṭā/*, and then explained in a footnote stating that, according to the predominant opinion, what is meant is *salat Asar*. This strategy signals that the expression is not merely a descriptive phrase such as */salat yang berada di tengah/*, but a technical term with specific theological and fiqh implications.

By contrast, as shown in Table 5, ChatGPT translates *al-ṣalāt al-wuṣṭā* as */salat tengah (salat pertengahan)/*. Lexically, this choice is defensible, since *wuṣṭā* does mean 'middle/central'. However, from the standpoint of terminological accuracy, the translation is considered less precise because it ignores the tafsir consensus that associates the term with *salat Asar* and shifts a technical term into a general descriptive expression.

The panel also noted that the formulation */salat tengah/* is not customary in Indonesian fiqh discourse or in religious pedagogy, and therefore may cause confusion for readers who are familiar with the expressions */salat Asar/* or */salat Wuṣṭā/*. In terms of readability, ChatGPT's sentence is simple and easy to understand, but this readability remains superficial: lay readers may quickly grasp the string of words, yet fail to obtain the crucial theological information that the verse refers to a specific prayer with particular merits.

Thus, Qur'an Surah Al-Baqarah verse 238 illustrates how a literal translation strategy for technical terms can yield a text that appears easy to read while weakening conceptual precision, leading to lower accuracy and acceptability scores compared with verses that do not contain technical terms as strongly marked as *al-ṣalāt al-wuṣṭā*.

The three examples above reinforce the general pattern observed in tables 3 and 4. For verses with relatively general public-law themes, ChatGPT is able to produce translations with high accuracy, acceptability, and readability. Conversely, in verses that contain technical fiqh terminology or sensitive socio-historical contexts, ChatGPT's tendency to simplify terms and soften social context results in a reduction in conceptual accuracy, even though the text remains pleasant to read.

Theoretically, these findings underscore the importance of domain adaptation for language models when dealing with normative religious texts such as *āyāt al-aḥkām*. Practically, the results suggest that AI-based Qur'anic translations should be positioned as initial drafts that require subsequent editing by experts in tafsir, fiqh, and language, so that the legal and theological messages of the verses are not diminished by lexical choices that are overly literal or overly communicative.

Assessment Results for the Translation of Ḥadīths from al-Arba'īn al-Nawawī

The ten initial ḥadīths of al-Arba'īn al-Nawawī analysed in this study cover the core themes of Islamic teaching: intention and sincerity in action, the pillars of Islam and *īmān*, the concept of *iḥsān*, divine decree (*qadar*), the prohibition of *bid'ah*, the boundaries of *ḥalāl-ḥarām* and doubtful matters (*shubuhāt*), the centrality of sincere counsel (*al-dīn al-naṣīḥah*), the inviolability of a Muslim's blood and property, the command to obey within one's capacity, and the obligation to avoid doubtful or suspect matters. Together, these themes represent the dimensions of creed

(*‘aqīdah*), worship (*‘ibādah*), transactions (*mu‘āmalah*), and ethics (*akhlāq*) that form the ethical foundation of a Muslim’s daily life.

A summary of the translation quality scores for the ten ḥadīths is presented in table 6, while table 7 summarises the main differences between the reference translation (alkhoirot) and ChatGPT’s translation for each ḥadīth, based on the FGDs and the panel’s notes.

Table 6. Accuracy, Acceptability, and Readability Scores for ChatGPT’s Translation of the Ḥadīths

| No. | Ḥadīth (Reference) | Theme | Accuracy | Acceptability | Readability |
|-----|-------------------------|--------------------------------|----------|---------------|-------------|
| 1 | Ṣaḥīḥ al-Bukhārī-Muslim | Deeds Depend on Intentions | 3 | 3 | 3 |
| 2 | Ṣaḥīḥ Muslim no. 8 | Islam, Faith, and <i>Iḥsān</i> | 2 | 3 | 3 |
| 3 | Muttafaquun ‘Alayh | Pillars of Islam | 3 | 2 | 3 |
| 4 | Muttafaquun ‘Alayh | Human Destiny (<i>Qadar</i>) | 3 | 3 | 2 |
| 5 | Muttafaquun ‘Alayh | Rejection of Innovations | 3 | 3 | 2 |
| 6 | Ṣaḥīḥ al-Bukhārī-Muslim | Halal and Haram | 3 | 3 | 3 |
| 7 | Muslim and others | Religion as Sincere Counsel | 3 | 3 | 2 |
| 8 | Muttafaquun ‘Alayh | Sanctity of Muslim Blood | 3 | 3 | 2 |
| 9 | Muttafaquun ‘Alayh | Observing Commands | 3 | 3 | 3 |
| 10 | Ṣaḥīḥ Muslim no. 1015 | Abstaining from the Prohibited | 3 | 3 | 3 |

Table 7. Summary of Differences between the Reference Translation and ChatGPT’s Translation of Ḥadīths

| No | Ḥadīth Theme | Reference Translation (alkhoirot) | ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|----|---|---|---|--|
| 1 | Ḥadīth of intention (<i>innamal a‘mālu bin-niyyāt...</i>) | <i>Sesungguhnya segala amal tergantung niatnya, dan setiap orang hanya mendapatkan apa yang ia niatkan. Barangsiapa hijrahnya karena Allah dan Rasul-Nya... barangsiapa hijrahnya karena dunia atau wanita... ”</i> | <i>Sesungguhnya setiap amal perbuatan bergantung pada niat. Setiap orang akan memperoleh sesuai apa yang ia niatkan. Jika seseorang berhijrah karena dunia atau karena seorang wanita....</i> | The core meaning is preserved (high accuracy), but the illustrative part on <i>hijrah</i> tends to be condensed and the style is more communicative. Acceptability and readability are high; some narrative detail is reduced. |
| 2 | Ḥadīth of Jibrīl (<i>Islam–īmān–iḥsān</i>) | <i>sangat hitam rambutnya, tidak nampak kalau sedang bepergian, dan tidak ada seorang pun dari kami yang mengenalnya. Kemudian dia duduk menghadap Nabi shallallahu ‘alaihi wa sallam lalu menyandarkan lututnya kepada lutut beliau, dan meletakkan kedua telapak tangannya di atas paha beliau.</i> | <i>sangat hitam rambutnya, tidak tampak padanya tanda-tanda perjalanan, dan tidak ada seorang pun dari kami yang mengenalnya. Ia duduk di hadapan Nabi ﷺ, merapatkan lututnya ke lutut Nabi dan meletakkan kedua tangannya di atas pahanya,</i> | The key messages (definitions of <i>Islam–īmān–iḥsān</i>) remain accurate, but parts of the narrative detail and dramatic nuance are attenuated. Acceptability and readability are high because the sentences are shorter and easier to follow. |
| 3 | Ḥadīth on the pillars of Islam | <i>Islam dibangun atas lima perkara: bersaksi bahwa tiada Tuhan selain Allah dan</i> | <i>Islam tegak di atas lima: bersaksi bahwa tidak ada Tuhan selain Allah dan bahwa Muhammad adalah</i> | The order of the pillars and the core meaning are identical (accuracy 3); differences lie only in lexical variants |

| No | Ḥadīth Theme | Reference Translation (alkhoirot) | ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) |
|----|---|---|--|--|
| | | Muhammad utusan Allah, mendirikan salat, menunaikan zakat, puasa Ramadan, dan haji bagi yang mampu. | utusan-Nya, melaksanakan salat, menunaikan zakat, berpuasa di bulan Ramadan, dan berhaji bagi yang mampu. | /dibangun/ → /tegak/, /mendirikan/ → /melaksanakan/. Acceptability and readability are very good; ChatGPT's style is slightly more contemporary. |
| 4 | Ḥadīth on <i>qadar</i> and the creation of the foetus | Sesungguhnya seorang dari kalian dikumpulkan penciptaannya di perut ibunya selama empat puluh hari berupa sperma, kemudian menjadi segumpal darah selama itu pula, kemudian menjadi segumpal daging selama itu pula, kemudian diutus seorang malaikat kepadanya untuk meniupkan ruh padanya, dan diperintahkan empat kalimat: menulis rezekinya, ajalnya, amalnya, dan celaka atau bahagia. | Sesungguhnya salah seorang di antara kalian dikumpulkan penciptaannya dalam perut ibunya selama empat puluh hari dalam bentuk nutfah (air mani), kemudian menjadi segumpal darah selama itu juga, lalu menjadi segumpal daging selama itu juga. Kemudian diutuslah malaikat yang meniupkan ruh ke dalamnya, dan diperintahkan untuk menulis empat hal: rezekinya, ajalnya, amalnya, dan apakah ia termasuk orang yang celaka atau bahagia. | The global meaning of <i>qadar</i> and the relationship between human effort and destiny is preserved, but theological depth and rhetorical emphasis are reduced. Accuracy is adequate; acceptability and readability are high for lay readers, though for <i>'aḡdah</i> study the narrative detail would need to be supplemented. |
| 5 | Ḥadīth on <i>bid'ah</i> (<i>man aḥdatsa...</i>) | Barangsiapa yang mengada-mengada dalam urusan kami ini yang bukan bagian darinya, maka ia tertolak. | Barangsiapa mengada-adakan suatu perkara baru dalam urusan (agama) kami ini yang tidak termasuk darinya, maka ia tertolak. | The legal message (<i>bid'ah</i> is rejected) is fully preserved; ChatGPT tends to replace /mengada-adakan perkara baru/ with other, more generic formulations. Acceptability is good; readability is high. The technical nuance of the term / <i>bid'ah</i> / still needs to be clarified outside the translation itself. |
| 6 | Ḥadīth on ḥalāl-ḥarām-shubuhāt | Sesungguhnya yang halal itu jelas dan yang haram itu jelas, dan di antara keduanya ada perkara-perkara yang samar (syubhat)... Barangsiapa menjauhi syubhat, ia telah menyelamatkan agama dan kehormatannya... Di dalam jasad ada | Yang halal itu jelas dan yang haram itu jelas. Di antara keduanya ada hal-hal yang samar yang banyak orang tidak mengetahuinya. Siapa yang menjauhi yang samar, ia menjaga agama dan kehormatannya... Di tubuh ada segumpal daging... | The argument structure and the heart metaphor remain the same (high accuracy). ChatGPT slightly simplifies the section on / <i>syubhat</i> / and the metaphor of the / <i>segumpal daging</i> / with shorter sentences. Acceptability and readability are very good; the metaphorical depth can still be enriched through further |

| No | Ḥadīth Theme | Reference Translation (alkhoirot) <i>segumpal daging...</i> | ChatGPT Translation | Main Differences (Accuracy / Acceptability / Readability) explanation. |
|----|---|--|---|--|
| 7 | Ḥadīth / <i>al-dīn al-naṣīḥah/</i> | Agama adalah nasihat. Kami bertanya, Untuk siapa? Beliau menjawab, Untuk Allah, Kitab-Nya, Rasul-Nya, dan pemimpin kaum muslimin dan orang awamnya. | Agama itu adalah nasihat. Kami bertanya: Untuk siapa, wahai Rasulullah? Beliau menjawab: Untuk Allah, untuk kitab-Nya, untuk Rasul-Nya, untuk para pemimpin kaum Muslimin, dan untuk seluruh kaum Muslimin. | The content (the five objects of <i>naṣīḥah</i>) is fully preserved. ChatGPT condenses the dialogic format, so the ḥadīth reads more as expository discourse than as dialog. Acceptability and readability are very good; the pedagogical flavour of question-answer is somewhat reduced. |
| 8 | Ḥadīth on the sanctity of blood, wealth, and honour | Aku diperintahkan untuk memerangi manusia hingga mereka bersyahadat <i>lâ ilâha illâllâh</i> dan <i>muhammadur rasûlûllâh</i> , menegakkan <i>shalat</i> , dan membayar zakat. Jika mereka melaksanakan hal tersebut, ... | Aku diperintahkan untuk memerangi manusia sampai mereka bersaksi bahwa tiada tuhan selain Allah dan bahwa Muhammad adalah utusan Allah, menegakkan <i>shalat</i> , dan menunaikan zakat. Jika mereka telah melakukannya, | The message of protection for blood, property, and honour is maintained (accurate). ChatGPT tends to replace the term / <i>haram</i> / with more explicit prohibitive formulations that are more communicative. Acceptability and readability are high; the rhetorical tone and the legal term / <i>haram</i> / are softened and may require reinforcement in the <i>sharḥ</i> . |
| 9 | Ḥadīth on commands and capability | Apa yang aku larang bagi kalian maka <i>jauhilah</i> , dan apa yang aku perintahkan kepada kalian maka kerjakan semampu kalian. Sesungguhnya yang membinasakan orang-orang sebelum kalian adalah banyak bertanya dan menyelisih para nabi mereka.” | Apa yang aku larang bagi kalian, maka <i>jauhilah</i> . Dan apa yang aku perintahkan kepada kalian, maka laksanakanlah semampu kalian. Sesungguhnya yang membinasakan orang-orang sebelum kalian adalah karena banyaknya pertanyaan mereka dan perselisihan mereka terhadap nabi-nabi mereka. | Accuracy is very high; the differences are restricted to stylistic variation and omission of supporting context. Acceptability and readability are very good; the practical message is easily understood by lay readers. |
| 10 | Ḥadīth on avoiding the prohibited / doubtful | Sesungguhnya Allah itu baik, tidak mau menerima kecuali yang baik. Dan sesungguhnya Allah telah memerintahkan kepada orang-orang mukmin seperti yang telah diperintahkan kepada para rasul, | Sesungguhnya Allah Ta'ala itu Maha Baik, dan tidak menerima kecuali yang baik. Allah memerintahkan kepada orang-orang beriman sebagaimana Dia memerintahkan kepada para rasul-Nya. | The general meaning of caution in avoiding doubtful matters is preserved, but concrete examples that help readers understand everyday application are not always fully present. Readability is high; accuracy of the core meaning is good, though practical detail is reduced. |

Overall, the panel judged the accuracy of ChatGPT's translations of the core meanings of the ḥadīths to be good: the main messages on intention, the conditions for valid deeds, the clarity of

ḥalāl-ḥarām and *syubhat*, and the imperative to avoid what is prohibited are all successfully conveyed. Most ḥadīths received high accuracy scores, particularly when the *matan* is relatively straightforward and not heavily laden with technical fiqh terminology.

In terms of acceptability and readability, ChatGPT tends to perform very strongly. The Indonesian sentences it generates are relatively short, coherent, and employ diction that is familiar to contemporary readers. Compared with the reference translation, which sometimes preserves classical structures and more formal style, ChatGPT's renderings are easier to process for lay readers, especially students and general congregants.

However, table 7 also reveals several recurring weaknesses. First, in ḥadīths with extended narratives (for example, the Ḥadīth of Jibrīl on Islam-*īmān-iḥsān*), ChatGPT tends to compress descriptive passages and omit some of the dramatic detail that is characteristic of the original mode of transmission. Second, in ḥadīths that contain key technical terms (such as *bid'ah*, *syubhat*, or the metaphor of the heart as a */segumpal daging/*), the chosen equivalents are often too general, so that the technical nuance recognised in ḥadīth and fiqh scholarship is not fully captured.

To illustrate these patterns further, the next subsection presents a concise analysis of three representative ḥadīths, namely the Hadith on Islam, *Īmān*, and *Iḥsān*, the ḥadīth on *bid'ah*, and the hadith on *al-dīn al-naṣīḥah*.

Analysis of the Second Hadīth on Islam, Īmān, and Iḥsān

The second ḥadīth in al-Arba'īn al-Nawawīyyah presents the core teachings of Islam, encompassing the dimensions of practice (*Islam*), belief (*Īmān*), and spirituality (*Iḥsān*). ChatGPT's translation of this ḥadīth was assessed as assessed as generally accurate (score 2) with some noted issues, though with certain limitations in acceptability and readability

عَنْ عُمَرَ رَضِيَ اللَّهُ عَنْهُ أَيْضاً قَالَ: بَيْنَمَا نَحْنُ جُلُوسٌ عِنْدَ رَسُولِ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ ذَاتَ يَوْمٍ، إِذْ طَلَعَ عَلَيْنَا رَجُلٌ شَدِيدُ بَيَاضِ الثِّيَابِ، شَدِيدُ سَوَادِ الشَّعْرِ، لَا يُرَى عَلَيْهِ أَثَرُ السَّفَرِ، وَلَا يَعْرِفُهُ مِنَّا أَحَدٌ حَتَّى جَلَسَ إِلَى النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَأَسْنَدَ رُكْبَتَيْهِ إِلَى رُكْبَتَيْهِ وَوَضَعَ كَفَيْهِ عَلَى فَخْذَيْهِ وَقَالَ: يَا مُحَمَّدُ أَخْبِرْنِي عَنِ الْإِسْلَامِ! فَقَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: «الْإِسْلَامُ أَنْ تَشْهَدَ أَنْ لَا إِلَهَ إِلَّا اللَّهُ وَأَنَّ مُحَمَّدًا رَسُولُ اللَّهِ، وَتُقِيمَ الصَّلَاةَ، وَتُؤْتِيَ الزَّكَاةَ، وَتَصُومَ رَمَضَانَ، وَتُحُجَّ الْبَيْتَ إِنْ اسْتَطَعْتَ إِلَيْهِ سَبِيلاً». قَالَ: صَدَقْتَ. فَعَجَبْنَا لَهُ يَسْأَلُهُ وَيُصَدِّقُهُ، قَالَ: فَأَخْبِرْنِي عَنِ الْإِيمَانِ! قَالَ: «أَنْ تُؤْمِنَ بِاللَّهِ، وَمَلَائِكَتِهِ، وَكُتُبِهِ، وَرُسُلِهِ، وَالْيَوْمِ الْآخِرِ، وَتُؤْمِنَ بِالْقَدَرِ خَيْرِهِ وَشَرِّهِ». قَالَ: صَدَقْتَ، قَالَ فَأَخْبِرْنِي عَنِ الْإِحْسَانِ! قَالَ: «أَنْ تَعْبُدَ اللَّهَ كَأَنَّكَ تَرَاهُ، فَإِنْ لَمْ تَكُنْ تَرَاهُ فَإِنَّهُ يَرَاكَ» قَالَ: فَأَخْبِرْنِي عَنِ السَّاعَةِ! قَالَ: «مَا الْمَسْئُولُ عَنْهَا بِأَعْلَمَ مِنَ السَّائِلِ» قَالَ: فَأَخْبِرْنِي عَنْ أَمَارَاتِهَا! قَالَ: «أَنْ تَلِدَ الْأُمَةُ رَبَّتَهَا، وَأَنْ تَرَى الْخُفَاةَ الْعُرَاةَ الْعَالَةَ رِعَاءَ الشَّيْءِ يَتَطَاوَلُونَ فِي الْبُنْيَانِ» ثُمَّ انْطَلَقَ فَلَبِثْتُ مَلِيًّا ثُمَّ قَالَ: «يَا عُمَرُ! أَتَدْرِي مَنْ السَّائِلُ؟» قُلْتُ: اللَّهُ وَرَسُولُهُ أَعْلَمُ قَالَ: «فَإِنَّهُ جِبْرِيلُ أَتَاكُمْ يُعَلِّمُكُمْ دِينَكُمْ» رواه مسلم.

ChatGPT Translation

Ketika kami sedang duduk bersama Rasulullah ﷺ pada suatu hari, tiba-tiba datang seorang lelaki yang sangat putih pakaiannya, sangat hitam rambutnya, tidak tampak padanya tanda-tanda perjalanan, dan tidak ada seorang pun dari kami yang mengenalnya. Ia duduk di hadapan Nabi ﷺ, merapatkan lututnya ke lutut Nabi dan meletakkan kedua tangannya di atas pahanya, lalu berkata: "Wahai Muhammad, beritahukan kepadaku tentang Islam." Nabi ﷺ menjawab: "Islam adalah engkau bersaksi bahwa tidak ada tuhan selain Allah dan bahwa Muhammad adalah utusan Allah, menegakkan salat, menunaikan zakat, berpuasa di bulan Ramadan, dan berhaji ke Baitullah jika engkau mampu menempuh perjalanannya." Lelaki itu berkata, "Engkau benar." Kami pun heran, dia yang bertanya tapi membenarkannya. Kemudian ia bertanya lagi, "Beritahukan kepadaku tentang iman." Nabi ﷺ menjawab: "Engkau beriman kepada Allah, kepada para malaikat-Nya, kepada kitab-kitab-Nya, kepada para rasul-Nya, kepada hari akhir, dan engkau beriman kepada takdir, baik dan buruknya." Ia berkata, "Engkau benar." Kemudian ia bertanya, "Beritahukan kepadaku tentang ihsan." Nabi ﷺ menjawab: "Ihsan adalah engkau beribadah kepada Allah seakan-akan engkau melihat-Nya. Jika engkau tidak melihat-Nya, sesungguhnya Dia melihatmu." Ia bertanya lagi, "Kapan hari kiamat itu?" Nabi ﷺ menjawab: "Orang yang ditanya tidak lebih

tahu daripada yang bertanya.” Ia pun berkata, “Beritahukan kepadaku tentang tanda-tandanya.” Nabi ﷺ menjawab: “Jika seorang budak perempuan melahirkan tuannya, dan jika engkau melihat orang-orang miskin, telanjang, tidak beralas kaki, penggembala kambing, berlomba-lomba dalam meninggikan bangunan.” Kemudian lelaki itu pergi, dan aku (Umar) pun terdiam beberapa saat. Lalu Rasulullah ﷺ berkata: “Wahai Umar, tahukah engkau siapa yang bertanya itu?” Aku menjawab, “Allah dan Rasul-Nya lebih mengetahui.” Beliau bersabda: “Itu adalah Jibril yang datang kepada kalian untuk mengajarkan agama kalian.”

ChatGPT successfully conveyed the core meaning of the ḥadīth with completeness and precision. All essential elements—including the dialogical structure, definitions, and narrative flow—were preserved without omissions. However, a point of ambiguity arises in the phrase */meletakkan kedua tangannya di atas pahanya/* ‘putting both hands on his thighs’. In some narrations, Jibrīl is described as placing his hands on the Prophet’s thighs rather than on his own. The absence of clear referent for the pronoun */nya/* may lead to misunderstanding, particularly among lay readers unfamiliar with variant narrations.

The translation is reasonably acceptable, with sentences arranged in a coherent sequence and dialog that flows naturally. The wording avoids stiffness and is generally consistent with standard Indonesian-English religious translation norms. Nonetheless, minor adjustments in diction could further enhance alignment with traditional Islamic discourse.

The readability is relatively good. The sentence structures follow the original dialogical style, making the narration accessible and easy to follow. However, given the length and richness of the text, the translation would benefit from the inclusion of footnotes or explanatory annotations, particularly for key terms such as *Ihsān* and *qadar*, to ensure comprehension among readers less familiar with Islamic theological vocabulary.

Analysis of the Fifth Hadith: Rejection of Innovations (Bid‘ah)

Arabic Text:

مَنْ أَحَدَثَ فِي أَمْرِنَا هَذَا مَا لَيْسَ مِنْهُ فَهُوَ رَدٌّ

ChatGPT Translation:

Barangsiapa mengada-adakan suatu perkara baru dalam urusan (agama) kami ini yang tidak termasuk darinya, maka ia tertolak.

The ChatGPT translation demonstrates a high degree of accuracy. The phrase */mengada-adakan perkara baru/* ‘fabricate new cases’ aptly conveys the meaning of the verb *aḥdatha*. The parenthetical addition */(agama)/* ‘religious’ in rendering *amrinā hādhā* reflects an interpretive effort to clarify the phrase, which is commonly understood in the exegetical tradition as */urusan agama/* ‘religious affairs’. This addition effectively reduces the risk of misinterpretation.

From the standpoint of acceptability, the sentence structure is broadly acceptable to Muslim Indonesian readers. However, the closing phrase */maka ia tertolak/* ‘then it was rejected’ may create ambiguity, as it remains unclear whether what is rejected is the actor or the act itself.

The translation is generally readable and accessible. Nevertheless, slight refinement could enhance clarity. An alternative phrasing such as */maka amalnya tertolak/* ‘then his deeds are rejected’ would more explicitly communicate the intended meaning, especially for lay readers, without altering the original sense of the ḥadīth.

Overall, ChatGPT’s translation of this ḥadīth is both accurate and sufficiently acceptable. However, refinement of the final diction is necessary to ensure that the ḥadīth’s legal and theological message is conveyed with greater clarity, leaving no room for multiple interpretations.

Analysis of the Seventh Hadith: Religion is Sincere Counsel (al-Dīn al-Naṣīḥah)

Arabic Text:

«الدِّينُ النَّصِيحَةُ» قُلْنَا: لِمَنْ؟ قَالَ: «لِللَّهِ، وَلِكِتَابِهِ، وَلِرَسُولِهِ، وَلِأَيِّمَةِ الْمُسْلِمِينَ، وَعَامَّتِهِمْ»

ChatGPT Translation:

“Agama itu adalah nasihat.” Kami bertanya: “Untuk siapa, wahai Rasulullah?” Beliau menjawab: “Untuk Allah, untuk kitab-Nya, untuk Rasul-Nya, untuk para pemimpin kaum Muslimin, dan untuk seluruh kaum Muslimin.”

The translation is highly accurate, as the essential meaning of the ḥadīth is conveyed completely, with all key components of the Arabic text faithfully rendered. However, the phrase */nasihat untuk Allah/* ‘advice for Allah’ may pose a risk of theological misunderstanding for lay readers, since Allah does not require counsel from His creation. In the exegetical tradition, this phrase is understood as a call to believe in Allah, to obey His commands, and to abstain from His prohibitions.

In terms of acceptability, the lexical choices and sentence structures are appropriate and flow naturally in Indonesian. Nevertheless, the theologically nuanced phrase */nasihat untuk Allah/* ‘advice for Allah’ ideally requires explanatory annotation to ensure that the intended meaning aligns with the established understanding of the Muslim community.

The readability of the translation is good, with a clear sentence structure and a logical dialogical sequence that is easy for readers to follow. However, since this ḥadīth encapsulates dense conceptual terminology, the translation would be more effective if accompanied by exegetical notes from ḥadīth scholars, particularly to elaborate on the concept of *naṣīḥah* in the context of religion.

Overall, ChatGPT’s translation of this ḥadīth is accurate and communicative. Nonetheless, explanatory notes remain essential to guide readers toward a precise understanding of the terminology in accordance with Islamic teachings.

The results of the analysis indicate that ChatGPT’s translations of both Qur’anic verses and ḥadīth generally demonstrate a high degree of accuracy, particularly in preserving the literal meaning and technical Islamic terminology. For the Qur’anic texts, most accuracy scores were rated in the highest category, with only minor lexical differences that did not alter the core meaning. Similarly, in the ḥadīth, nine out of ten texts achieved the highest score, reflecting the system’s capacity to capture the essential message of the source texts.

Nevertheless, the aspects of acceptability and readability showed greater variation. In both types of texts, several translations appeared rigid or overly literal, making them less aligned with the linguistic norms of Indonesian religious discourse. This was evident, for instance, in the ḥadīth on the pillars of Islam and the ḥadīth “Religion is sincere counsel,” where lexical choices and sentence structures required further editing to clarify the intended meaning and to avoid potential misinterpretation.

In terms of emerging patterns, Qur’anic texts tended to present challenges in selecting contextually appropriate equivalents that are consistent with classical tafsīr, whereas ḥadīth texts more frequently struggled with producing a fluent and communicative narrative—particularly in passages containing dialog or complex theological concepts. These findings resonate with Nababan et al. (2012), who emphasized the importance of integrating both semantic and pragmatic dimensions in the translation of religious texts.

Both sets of data also underscore the continuing necessity of human intervention, not only for refining diction and adjusting stylistic choices but also for adding explanatory annotations to clarify key concepts. Accordingly, the role of the translator extends beyond simply post-editing machine outputs; it encompasses acting as a mediator of meaning who ensures that the translation remains faithful to theological intent while also culturally accessible to Indonesian Muslim readers.

The findings of this study are consistent with a range of recent works showing that ChatGPT excels in the fluency and readability of its translations, often surpassing or at least matching

conventional machine translation systems such as Google Translate and other NMT engines in terms of target-language fluency and contextual coherence (Sadiq, 2025; Siregar et al., 2025). However, as demonstrated by studies that specifically evaluate translations of the Qur'an and other religious texts, ChatGPT tends to exhibit a decline in accuracy when dealing with religious terminology that is heavily loaded with theological and cultural meaning (Abdalhussein, 2025; Dahia & Belbacha, 2024; Shormani & Alfahad, 2025). The present findings reinforce this pattern: in the case of *āyāt al-aḥkām* and the *ḥadīths* of *al-Arba'īn al-Nawawī*, the model appears strong in terms of acceptability and readability, yet more prone to simplifying technical fiqh terms, socio-historical contexts, and theological metaphors, so that certain legal and doctrinal nuances risk being reduced.

The main novelty of this study lies in its specific focus on ten *āyāt al-aḥkām* and ten normative *ḥadīths* from *al-Arba'īn al-Nawawī*, evaluated simultaneously using Nababan's assessment model (accuracy, acceptability, readability). This approach complements earlier studies that have predominantly evaluated Qur'anic translation into English (Dahia & Belbacha, 2024) or compared ChatGPT with human and machine translators on non-normative general texts (Abdalhussein, 2025). The use of a multidisciplinary panel (experts in *tafsīr-ḥadīth*, Arabic–Indonesian language, and lay readers) and the presentation of “real data” in the form of comparative tables for each verse and *ḥadīth* make the assessment process more transparent than many previous studies that relied solely on automatic scores or broad impressionistic judgments.

At the theoretical and methodological level, this study demonstrates that Nababan's framework can be adapted to assess the quality of AI translations of normative religious texts, while enriching ongoing discussions about how large language models need to be “domain-adapted” when applied to sensitive Islamic materials. The finding that ChatGPT is fluent but not always normatively accurate aligns with other research evaluating ChatGPT's responses in *tafsīr*, *ḥadīth*, and *fatwā*, all of which emphasise the importance of religious scholars' involvement in monitoring AI outputs (El Ganadi et al., 2025; M. Faidh Fasyani et al., 2024; Mu'tashim Billah et al., 2023). Accordingly, the contribution of this study is to provide a more fine-grained empirical picture of ChatGPT's strengths and weaknesses on the Arabic–Indonesian pair for Qur'anic and *ḥadīth* texts, and to offer an example of an assessment procedure that can be replicated in future research.

Practically, these results imply that ChatGPT-based translations should be positioned as preliminary drafts that facilitate rapid access to the global meaning of verses and *ḥadīths*, rather than as stand-alone references for legal rulings, *fatwā* formulation, or formal instruction. This recommendation is in line with several studies that advocate hybrid models in which AI assists the initial analysis, while religious authorities retain final decision-making power in matters of Islamic law (El Ganadi et al., 2025; Mu'tashim Billah et al., 2023). For subsequent researchers, the design and findings of this study can serve as a basis for developing broader corpora, domain-specific AI models for Islamic studies, and additional indicators—such as “normative conformity”—in the evaluation of translation quality for sacred texts.

CONCLUSION

This study set out to evaluate the quality of ChatGPT's translations from Arabic into Indonesian on ten Qur'anic legal verses (*āyāt al-aḥkām*) and several *ḥadīth* from *al-Arba'īn al-Nawawīyyah*, using three key parameters: accuracy, acceptability, and readability. The findings indicate that ChatGPT is generally able to produce translations that are accurate, communicative, and accessible. In terms of accuracy, ChatGPT successfully conveyed the core meanings of the source texts, though minor errors were observed in the rendering of legal terminology, emphatic particles, and contextual nuances. These inaccuracies did not obscure the central meaning but highlight the importance of involving Arabic linguists and religious scholars to ensure that translations remain theologically sound. With respect to acceptability, the use of neutral diction and syntactically appropriate Indonesian structures made the translations broadly acceptable to general readers. Nevertheless, adaptation to established translational conventions within Indonesian Muslim

communities remains necessary in order to align with linguistic norms and cultural sensitivities. Regarding readability, ChatGPT produced sentences that were concise, coherent, and communicative. Its more dynamic style, compared with many conventional translations, proved effective as an entry point for initial understanding. However, given the sacred and normative character of religious texts, translation must rest not only on linguistic proficiency but also on an informed grasp of theological context and spiritual values. This study is not without limitations. First, the corpus is relatively small and highly specific, being restricted to several *āyāt al-aḥkām* and several ḥadīth from a single collection (al-Arba'īn al-Nawawīyyah), so the findings cannot be generalized to all genres of Qur'anic and ḥadīth literature. Second, the evaluation focuses on one configuration of a single model (ChatGPT, GPT-4o), at one point in time, whereas AI systems are continually updated and may behave differently across versions and settings. Third, the assessment relies on expert and lay judgments within the Indonesian context, which, although methodologically robust, still contains an element of subjectivity and may not fully reflect perceptions in other linguistic or cultural environments. Future research should therefore broaden the range of texts, models, and evaluators to test the robustness of these findings. Overall, this study contributes to a deeper understanding of both the potential and the limitations of AI technology in the translation of religious texts. It also underscores the promise of AI as a supportive partner in Qur'anic exegesis, Arabic language pedagogy, and the development of digital religious literacy. Future research may expand this scope by comparing the performance of different AI models, analyzing additional genres of religious texts, or exploring hybrid translation approaches that integrate AI outputs with the interpretive authority of scholars and exegetes.

REFERENCES

- Abdalthussein, H. F. (2025). Comparative Analysis Of AI And Human Translations Of Qur'anic Legal Verses In Surah Al-Ma'idah. *Ijaz Arabi Journal of Arabic Learning*, 8(1). <https://doi.org/10.18860/ijazarabi.v8i1.28804>
- Abidah, N. K. K., Hikmawati, R., & Erawanto, V. (2024). ChatGPT in Indonesia-Arabic Translation: A Quality Analysis. *Al-Arabi: Jurnal Bahasa Arab Dan Pengajarannya = Al-Arabi: Journal of Teaching Arabic as a Foreign Language*, 8(1), 33–52. <https://doi.org/10.17977/um056v8i1p33-52>
- AlJahsh, M. A. I. (2023). Influence Of Cultural Context On Qur'anic Translation: Analyzing Social Justice Interpretations In Sura An-Nisā' Verse 58. *Ma'ālim Al-Qur'ān Wa al-Sunnah*, 19(2), 366–387. <https://doi.org/10.33102/jmq.v19i2.446>
- AlSukhni, E., Al-Kabi, M. N., & Alsmadi, I. M. (2016). An Automatic Evaluation for Online Machine Translation: Holy Quran Case Study. *International Journal of Advanced Computer Science and Applications (Ijacs)*, 7(6), Article 6. <https://doi.org/10.14569/IJACSA.2016.070614>
- Arju, S., Husein, M. F., Sa'adah, B., & Rejeki, A. M. (2025). Analisis Keakuratan Hasil Terjemahan Chatgpt Menggunakan Teknik Back Translation. *El-Tsaqafah : Jurnal Jurusan PBA*, 24(1), Article 1. <https://doi.org/10.20414/tsaqafah.v24i1.10604>
- Dahia, I., & Belbacha, M. (2024). Machine-Learning-based English Quranic Translation: An Evaluation of ChatGPT. *International Journal of Linguistics, Literature and Translation*, 7, 128–136. <https://doi.org/10.32996/ijllt.2024.7.8.17>
- Dani, A. R. (2023). Gaya Bahasa Dialog Al-Qur'an "Analisis Stilistika Atas Dialog Al-Qur'an Mengenai Hari Kebangkitan. *Jurnal Sosial Dan Sains*, 3(3), 258–270. <https://doi.org/10.59188/journalsosains.v3i3.703>
- El Ganadi, A., Aftar, S., Gagliardelli, L., & Ruozzi, F. (2025). Generative AI for Islamic Texts: The EMAN Framework for Mitigating GPT Hallucinations: *Proceedings of the 17th International Conference on Agents and Artificial Intelligence*, 1221–1228. <https://doi.org/10.5220/0013312800003890>

- El-Saadany, M. R. (2024). A Comparative Study between Chat GPT and Human Translation in Translating English Proverbs into Arabic. *مجلة البحث العلمي في الآداب*. <https://doi.org/10.21608/jssa.2024.257874.1592>
- Framesthia, L. M., Khoerunnisa, S., Izzah, S. A., Farisi, M. Z. A., & Supriadi, R. (2024). Analisis Perbandingan Teknik Penerjemahan Arab-Indonesia pada Google Translate dan ChatGPT. *Ihtimam : Jurnal Pendidikan Bahasa Arab*, 7(02), 114–127. <https://doi.org/10.36668/jih.v7i02.716>
- Gaanoun, K., & Alsuhaibani, M. (2025). Sentiment preservation in Quran translation with artificial intelligence approach: Study in reputable English translation of the Quran. *Humanities and Social Sciences Communications*, 12. <https://doi.org/10.1057/s41599-024-04181-00>
- Gunawan, F., Batmang, B., & Boulahnane, S. (2023). Translating the Qur'an in Indonesia: Marrying the Concept of Transcreation and Critical Discourse Analysis. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 8(2). <https://doi.org/10.21093/ijeltal.v8i2.1523>
- Hassan Saleh Mahdi. (2022). The use of web-based translation software in translating religious terms. *Saudi Journal Of Language Studies*. <https://doi.org/10.1108/SJLS-11-2021-0023>
- Huda, K. (2018). Problematika Kebudayaan dalam Penerjemahan Bahasa Arab ke Bahasa Indonesia. *Al-Fathin: Jurnal Bahasa Dan Sastra Arab*, 1(02), 136–149. <https://doi.org/10.32332/al-fathin.v1i2.1270>
- Istiqomah, S. N., Nurhaliza, T. N., Nafis, Z., & Supriadi, R. (2023). Teknik Penerjemahan Bahasa Arab ke Bahasa Indonesia: Taqniyyāt al-Tarjamah min al-‘Arabiyyah ila al-Indūnīsiyyah. *Aphorisme: Journal of Arabic Language, Literature, and Education*, 4(2), Article 2. <https://doi.org/10.37680/aphorisme.v4i2.4500>
- Jiang, Z., Lv, Q., Zhang, Z., & Lei, L. (2024). *Convergences and Divergences between Automatic Assessment and Human Evaluation: Insights from Comparing ChatGPT-Generated Translation and Neural Machine Translation* (No. arXiv:2401.05176). arXiv. <https://doi.org/10.48550/arXiv.2401.05176>
- Jiao, W., Wang, W., Huang, J., Wang, X., Shi, S., & Tu, Z. (2023). *Is ChatGPT A Good Translator? Yes With GPT-4 As The Engine* (No. arXiv:2301.08745). arXiv. <https://doi.org/10.48550/arXiv.2301.08745>
- Kalla, D., Smith, N., Samaah, F., & Kuraku, S. (2023). *Study and Analysis of Chat GPT and its Impact on Different Fields of Study* (SSRN Scholarly Paper No. 4402499). Social Science Research Network. [Google](https://www.google.com)
- Latifah, S. N., & Djamilah, W. I. F. (2024). Penggunaan Chat Gpt Dalam Penerjemahan Bahasa Arab. *Al-Dirosah : Jurnal Pendidikan Agama Islam*, 1(02). [Google](https://www.google.com)
- M. Faidh Fasyani, Abdul Muqsid, & Abdal Wahab Mujtaba. (2024). Artificial Intelligence And Digital Tafsīr: Assessing the Interpretive Accuracy of ChatGPT's Engagement with Tafsīr al-Qurṭūbī. *Journal of Ushuluddin and Islamic Thought*, 2(1). <https://doi.org/10.15642/juit.2024.2.1.86-118>
- Majid, M. N., Akmaliyah, A., & Hermawan, M. B. (2024). Perbandingan Hasil Terjemah Arab – Indonesia Antara ChatGPT dan Gemini AI. *Ihtimam : Jurnal Pendidikan Bahasa Arab*, 7(02), 1–21. <https://doi.org/10.36668/jih.v7i02.1004>
- Maulana, M. A. (2022). Problematika Penerjemahan Arab-Indonesia. *Al Ibrah: Journal of Arabic Language Education*, 5(2), Article 2. <https://doi.org/10.24256/jale.v5i2.3417>
- Mustofa, R., Mustofa, A., & Surur, M. (2025). Variations of Arabic-Indonesian Translation Strategies in Automatic Machine Translation: A Comparative Study. *Journal of Language and Literature Studies*, 5(3), 522–540. <https://doi.org/10.36312/jolls.v5i3.3075>
- Mu'tashim Billah, Susiknan Azhari, Nurdhin Baroroh, & Vivi Rahma. (2023). Evaluating AI-Generated Fatwas: A Quality Assessment of ChatGPT and Google Bard Against Authoritative Islamic Rulings. *Asy-Syir'ah*, 57(1). <https://doi.org/10.14421/ajish.v57i1.1032>

- Nababan, M., & Nuraeni, A. (n.d.). Pengembangan Model Penilaian Kualitas Terjemahan. *Kajian Linguistik dan Sastra*, 24(1). [Google](#)
- Nababan, M., Nuraeni, A., & Sumardiono. (2012). Pengembangan Model Penilaian Kualitas Terjemahan. *Kajian Linguistik Dan Sastra*, 24(1), Article 1. <https://doi.org/10.23917/cls.v24i1.101>
- Nugraha, T. C., Lukman, F., & Amalia, R. M. (2022). Learning Model for Local Wisdom-based Prophet's Hadith Translation. *Islamic Research*, 5(2), Article 2. <https://doi.org/10.47076/jkps.v5i2.170>
- Nurfaiza. (2024). Pengaruh Penggunaan Chat Gpt Dalam Pembelajaran Terjemah Indonesia-Arab. *Ibtidaiyah Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, 70–89. <https://doi.org/10.52491/ibtidaiyah.v2i2.228>
- Ruhmadi, A., & Farisi, M. Z. A. (2023). Analisis Kesalahan Morfologi Penerjemahan Arab-Indonesia pada ChatGPT: Tahlīl al-Akhtā' al-Ṣarfīyyah 'ibra Tarjamah al-Arabiyyah wa al-Indūnīsiyyah fī ChatGPT. *Aphorisme: Journal of Arabic Language, Literature, and Education*, 4(1), Article 1. <https://doi.org/10.37680/aphorisme.v4i1.3148>
- Ruth Suci Putri, Rudy Sofyan, & Ely Hayati Nasution. (2022). Translation Quality Assessment on Medan City Tourism Official Web Pages. *LingPoet: Journal of Linguistics and Literary Research*, 2(3). <https://doi.org/10.32734/lingpoet.v3i1.6500>
- Sabrin Azzam Abdulrahman Fadhel. (2025). The Use of the AI Model Chat GPT in Translating Religious texts and Compare it with Expert Translation (Comparative Study). *Albaydha University Journal*, 7(1). <https://doi.org/10.56807/buj.v7i1.758>
- Sadiq, S. (2025). Evaluating English-Arabic translation: Human translators vs. Google Translate and ChatGPT. *Journal of Languages and Translation*, 12(1), 67–95. <https://doi.org/10.21608/jltmin.2025.423147>
- Saimin, A. A., Supriadi, R., & Farisi, M. Z. A. (2024). Analisis Kesalahan Penerjemahan Teks Bahasa Indonesia ke dalam Bahasa Arab pada ChatGPT (Studi Analisis Morfologi dan Sintaksis). *Jurnal Naskhi Jurnal Kajian Pendidikan Dan Bahasa Arab*, 6(2), 1–12. <https://doi.org/10.47435/naskhi.v6i1.2668>
- Setyawan, C. E., Zulaeha, Rauhillah, S., & Mu'at. (2025). Analisis Penerjemahan Indonesia-Arab Google Translate dan CHAT-GPT di Tinjau dari Linguistik Struktural dan Semantik. *Bilingua*, 2(2), 1–11. <https://doi.org/10.33752/bilingua.v2i2.8946>
- Shormani, M. Q., & Alfahad, A. (2025). Artificial Intelligence or Human: The Use of ChatGPT in the Academic Translation for Religious Texts. *Sage Open*, 15(3), 21582440251343954. <https://doi.org/10.1177/21582440251343954>
- Siregar, R., Subagiharti, H., Handayani, D. S., Hutagaol, D., Hasibuan, A. L., Nuraida, N., & Sipahutar, E. (2025). A Comparative Analysis of Google Translate and ChatGPT Results Based on Translation Quality Assessment (TQA). *International Journal of Educational Research Excellence (IJERE)*, 4(1), 328–337. <https://doi.org/10.55299/ijere.v4i1.1175>
- Syihabuddin, S. (2016). Penerjemahan Teks Suci: Analisis Ketepatan Terjemahan Istilah Kecendekiaan Dalam Alquran dan Terjemahnya. *Adabiyāt: Jurnal Bahasa Dan Sastra*, 15(1), Article 1. <https://doi.org/10.14421/ajbs.2016.15105>
- Trisakti Agriani, M. R. Nababan, & Djatmika Djatmika. (2018). Translation Quality on Words Containing Spelling, Punctuation and Grammatical Error. *International Journal of Multicultural and Multireligious Understanding*, 5(5). <https://doi.org/10.18415/ijmmu.v5i5.314>
- Zitouni, M., Alzahrani, A., Kous, N. A., Almutlaq, S., Abdul-Ghafour, A.-Q., & Zemni, B. (2022). The Translation of Selected Cultural Items in Al-Nawawi's Forty Hadiths: A Descriptive and Analytical Study. *Journal of Intercultural Communication*, 22(3), 43–53. <https://doi.org/10.36923/jicc.v22i3.74>