Journal of Practice Learning and Educational Development
https://digitalpress.gaes-edu.com/index.php/jpled
E-ISSN: 2809-1205
Vol 6, No. 3 (2026) 2726-2735
DOI: 10.58737/jpled.v6i3.1345

Submitted: 15-06-2026
Revised: 16-08-2026
Accepted: 26-08-2026

Comparative Study of Arabic-Indonesian Translation Quality Using Dola AI and Grok AI

Syahla[1], Angga Frananda[2]
Institut Agama Islam Imam Syafii Indonesia[1,2]
*E-mail: syahlarem@gmail.com

Abstract

The rapid proliferation of Artificial Intelligence (AI) in Arabic-Indonesian translation has created an urgent need to evaluate the capabilities of emerging platforms. This study analyzes and compares the translation quality of two latest-generation AI platforms, Dola AI and Grok AI, using the accuracy, acceptability, and readability parameters of the Nababan et al. (2012) evaluation model within Nida's dynamic equivalence framework. A qualitative descriptive-comparative method was employed, with 20 Arabic texts selected through purposive sampling across four genres—basic structure, religious studies, journalistic, and literary—validated by domain experts using an identical prompt. Results reveal that Grok AI consistently outperforms Dola AI with an overall mean score of 2.57 versus 2.33, with the largest gap in religious texts (0.54 points) due to Dola AI's critical terminological confusion in fiqh vocabulary. Both platforms failed equally on figurative language in literary texts. This study concludes that Grok AI is more reliable for academic Arabic translation, though human verification remains essential for Islamic legal and literary texts.

Keywords: Translation Quality, Dola AI, Grok AI, Nababan Model, Arabic-Indonesian

Licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they give the author or licensor the credits (attribution) in the manner specified by these. Licensees may copy, distribute, display, and perform the work and make derivative works and remixes based on it only for non-commercial purposes.

Pendahuluan

Revolusi digital telah mengubah cara manusia berinteraksi dengan bahasa secara fundamental. Salah satu perubahan paling nyata dalam kehidupan akademis saat ini adalah meningkatnya ketergantungan pada *platform Artificial Intelligence* (AI) sebagai alat bantu penerjemahan, termasuk dalam penerjemahan teks bahasa Arab ke bahasa Indonesia. Fenomena ini mencerminkan kebutuhan nyata yang dirasakan oleh mahasiswa, akademisi, dan praktisi Pendidikan Bahasa Arab dalam menghadapi kompleksitas linguistik teks-teks berbahasa Arab. Kebutuhan ini menyentuh dimensi yang jauh lebih luas dari sekadar aktivitas akademis biasa, karena bahasa Arab adalah bahasa sumber utama bagi Al-Qur'an, hadis, dan khazanah literatur Islam klasik yang menjadi rujukan jutaan umat di Indonesia. Realitas ini menjadikan akurasi terjemahan berbasis AI persoalan yang berimplikasi langsung pada kebenaran pemahaman keagamaan dan keilmuan masyarakat (Azhar et al., 2025; Nurullawasepa et al., 2023).

Di tengah kondisi tersebut, dua *platform* AI generasi terbaru hadir dan menarik perhatian luas: Dola AI dan Grok AI. Dola AI mendapatkan sorotan signifikan setelah melalui proses *rebranding* dari aplikasi yang sebelumnya dikenal sebagai Cici, di mana transformasi identitas tersebut terbukti memengaruhi persepsi dan kepuasan pengguna melalui pembaruan fitur yang lebih cerdas (Kholik et al., 2025). Sementara itu, Grok AI hadir dengan keunggulan kompetitif pada aspek pengecekan fakta (*fact-checking*) dan kemampuan pemrosesan data secara *real-time* yang berkontribusi terhadap peningkatan kesadaran verifikasi informasi penggunanya (Samet, 2025). Kemunculan keduanya

menciptakan urgensi akademis untuk mengkaji efektivitasnya dalam mendukung transmisi makna antarbahasa secara akurat.

Secara teoretis, penerjemahan bertujuan mewujudkan ekuivalensi makna, yakni kondisi di mana respons pembaca dalam bahasa sasaran sepadan dengan pembaca bahasa sumber, sebagaimana dirumuskan dalam prinsip ekuivalensi dinamis Nida (Nuur & Hamzah, 2025). Setiadi (2017) menegaskan bahwa penerjemahan menuntut pemahaman mendalam terhadap unsur gramatika, sehingga pencapaian ekuivalensi tidak dapat dipisahkan dari penguasaan morfologi (*sharaf*) dan sintaksis (*nahwu*). Berbagai studi mengungkapkan bahwa meskipun *Large Language Models* (LLM) seperti ChatGPT memiliki potensi besar, keterbatasan pada teks kompleks tetap memerlukan intervensi manusia (Arju et al., 2025; Latifah & Djamilah, 2024). Kesalahan morfologis—seperti salah mengidentifikasi kelas kata atau kekeliruan menentukan kasus gramatikal (*i'rab*)—masih menjadi permasalahan mendasar pada AI saat ini (Saimin et al., 2024). Untuk mengukur kualitas terjemahan secara terstandarisasi, model penilaian Nababan, Nuraeni, dan Sumardiono (2012) dengan parameter keakuratan, keberterimaan, dan keterbacaan tetap menjadi instrumen paling objektif dan banyak diadopsi (Ma'rifah et al., 2025; Aminudin et al., 2025).

Penelusuran literatur mengungkap kesenjangan penelitian (*research gap*) yang signifikan: mayoritas studi terdahulu masih terfokus pada *platform* mapan seperti Google Translate, ChatGPT, dan Gemini (Syam et al., 2023; Framesthia et al., 2024; Majid et al., 2024; Abas et al., 2025), sementara kajian komparatif yang khusus mengeksplorasi performa Dola AI dan Grok AI belum ditemukan. Di samping itu, penelitian terdahulu umumnya hanya menguji satu genre teks, padahal setiap genre memiliki karakteristik linguistik yang berbeda—teks sastra misalnya menuntut kepekaan estetis yang sering gagal ditangkap AI secara harfiah (Kintami et al., 2025; Putri & Tatang, 2024). Evaluasi terhadap teknologi terbaru sangat penting untuk memahami kemajuan AI di masa mendatang (Mohammed et al., 2025). Oleh karena itu, penelitian ini bertujuan menganalisis dan membandingkan kualitas terjemahan bahasa Arab ke bahasa Indonesia yang dihasilkan Dola AI dan Grok AI menggunakan parameter model Nababan et al. (2012) dalam kerangka ekuivalensi dinamis Nida, lintas empat genre teks, guna memberikan panduan praktis bagi akademisi dan praktisi Pendidikan Bahasa Arab (Yusuf, 2023; Lutfiyatun et al., 2023).

Metode

Penelitian ini menggunakan pendekatan deskriptif kualitatif dengan rancangan komparatif, dipilih karena data yang dikaji berupa satuan-satuan linguistik yang memerlukan interpretasi mendalam terhadap fenomena kebahasaan (Creswell, 2014). Meskipun proses penilaian menggunakan skala numerik 1 hingga 3, angka tersebut tidak difungsikan sebagai data statistik inferensial, melainkan semata-mata sebagai penanda komparatif kualitas berdasarkan parameter Nababan et al. (2012), sehingga orientasi analisis tetap bersifat kualitatif-interpretatif. Penelitian ini bersifat pustaka digital dengan sumber data diakses secara daring melalui *platform* Dola AI dan Grok AI selama empat minggu.

Sampel penelitian ditetapkan sebanyak 20 teks Bahasa Arab yang dipilih menggunakan teknik *purposive sampling*, didistribusikan ke dalam empat genre: struktur dasar (teks 1–5, dari *Durus al-Lughah al-'Arabiyyah* karya Dr. F. Abdurrahim), keilmuan keagamaan (teks 6–10, dari *Minhaj al-Muslim* karya Abu Bakar Jabir al-Jazairi), jurnalistik (teks 11–15, dari artikel Al-Jazeera.com), dan sastra/balaghah (teks 16–20, dari *Al-Binayah fi Syarh al-Bidayah fi 'Ulum al-Balaghah* karya Khalid bin Mahmud al-Juhani). Penetapan 20 teks ini dilandasi pertimbangan metodologis untuk memastikan tercapainya titik jenuh data (data saturation) yang mencakup keragaman karakteristik linguistik secara memadai (Saimin et al., 2024; Framesthia et al., 2024).

Instrumen utama adalah peneliti sebagai *human instrument*, didukung lembar penilaian model Nababan et al. (2012) dan perangkat validasi ahli di bidang linguistik Arab dan penerjemahan, yakni Angga Frananda, M.Pd. Pengumpulan data dilaksanakan melalui dokumentasi digital dengan memasukkan seluruh teks sumber ke dalam masing-masing platform menggunakan *prompt* identik: "Terjemahkan teks bahasa Arab berikut ke dalam bahasa Indonesia yang akurat, berterima, dan mudah

dipahami." Seluruh hasil terjemahan kemudian dianalisis melalui analisis isi (*content analysis*) secara bertahap: reduksi data berdasarkan genre, penyajian dalam tabel komparatif, analisis morfosintaksis pada level *sharaf* dan *nahwu*, penskoran berdasarkan model Nababan et al. (2012) yang telah divalidasi ahli, hingga penarikan kesimpulan komparatif (Ma'rifah et al., 2025; Mas'udah, 2024).

Hasil dan Pembahasan

1. Hasil

Penelitian ini mengujikan 20 teks Bahasa Arab kepada platform Dola AI dan Grok AI, dinilai oleh validator ahli menggunakan tiga parameter model Nababan et al. (2012) seperti pada Tabel 1

Tabel 1.
Parameter Penilai Terjemahan Model Nababan

INSTRUMEN PENILAI KEAKURATAN TERJEMAHAN		
Kategori Terjemahan	Skor	Parameter Kualitatif
Akurat	3	Makna kata, istilah teknis, frasa, klausa, kalimat atau teks bahasa sumber dialihkan secara akurat ke dalam bahasa sasaran; sama sekali tidak terjadi distorsi makna
Kurang Akurat	2	Sebagian besar makna kata, istilah teknis, frasa, klausa, kalimat atau teks bahasa sumber sudah dialihkan secara akurat ke dalam bahasa sasaran. Namun, masih terdapat distorsi makna atau terjemahan makna ganda (taksa) atau ada makna yang dihilangkan, yang mengganggu keutuhan pesan.
Tidak Akurat	1	Makna kata, istilah teknis, frasa, klausa, kalimat atau teks bahasa sumber dialihkan secara tidak akurat ke dalam bahasa sasaran atau dihilangkan (deleted).
INSTRUMEN PENILAI TINGKAT KEBERTERIMAAN TERJEMAHAN		
Kategori Terjemahan	Skor	Parameter Kualitatif
Berterima	3	Terjemahan terasa alamiah; istilah teknis yang digunakan lazim digunakan dan akrab bagi pembaca; frasa, klausa dan kalimat yang digunakan sudah sesuai dengan kaidah-kaidah bahasa Indonesia
Kurang Berterima	2	Pada umumnya terjemahan sudah terasa alamiah; namun ada sedikit masalah pada penggunaan istilah teknis atau terjadi sedikit kesalahan gramatikal.
Tidak Berterima	1	Terjemahan tidak alamiah atau terasa seperti karya terjemahan; istilah teknis yang digunakan tidak lazim digunakan dan tidak akrab bagi pembaca; frasa, klausa dan kalimat yang digunakan tidak sesuai dengan kaidah-kaidah bahasa Indonesia
INSTRUMEN PENILAI TINGKAT KETERBACAAN TERJEMAHAN		
Kategori Terjemahan	Skor	Parameter Kualitatif
Tingkat Keterbacaan Tinggi	3	Kata, istilah teknis, frasa, klausa, kalimat atau teks terjemahan dapat dipahami dengan mudah oleh pembaca.
Tingkat Keterbacaan Sedang	2	Pada umumnya terjemahan dapat dipahami oleh pembaca; namun ada bagian tertentu yang harus dibaca lebih dari satu kali untuk memahami terjemahan.
Tingkat Keterbacaan Rendah	1	Terjemahan sulit dipahami oleh pembaca

Rekapitulasi skor keseluruhan disajikan pada Tabel 2.

Tabel 2.
Rekapitulasi Skor Total dan Rata-Rata Kualitas Terjemahan

Parameter	Total Dola AI	Rata-rata Dola AI	Total Grok AI	Rata-rata Grok AI	Skor Maksimal
Keakuratan	46	2,30	52	2,60	60
Keberterimaan	45	2,25	49	2,45	60
Keterbacaan	49	2,45	53	2,65	60
Keseluruhan	140	2,33	254	2,57	180

Data pada Tabel 1 menunjukkan bahwa Grok AI secara konsisten menghasilkan skor lebih tinggi dibandingkan Dola AI pada ketiga parameter. Selisih terbesar pada keakuratan sebesar 0,30 poin, disusul keberterimaan dan keterbacaan masing-masing 0,20 poin. Secara keseluruhan, rerata skor Grok AI (2,57) melampaui Dola AI (2,33). Untuk gambaran yang lebih rinci, Tabel 3 menyajikan perbandingan per genre.

Tabel 3.
Perbandingan Skor Per-Genre

Genre	Dola AI (Ak/Kb/Tb)	Rata-rata Dola AI	Grok AI (Ak/Kb/Tb)	Rata-rata Grok AI	Selisih
Struktur Dasar	13/14/15	2,80	14/14/15	2,87	0,07
Keilmuan Keagamaan	12/10/10	2,13	14/13/13	2,67	0,54
Jurnalistik	13/14/15	2,80	14/14/14	2,80	0,00
Sastra/Balaghah	8/7/9	1,60	10/8/11	1,93	0,33

Keterangan: Ak = Keakuratan, Kb = Keberterimaan, Tb = Keterbacaan; skor maksimal 15 per parameter per genre

Tabel 2 mengungkapkan tiga pola utama: genre sastra/balaghah adalah genre terlemah bagi kedua *platform* secara konsisten; kesenjangan terlebar terjadi pada genre keilmuan keagamaan (selisih 0,54 poin); dan genre jurnalistik menjadi satu-satunya genre di mana keduanya mencapai rata-rata identik (2,80).

2. Pembahasan

a. Kualitas Terjemahan Genre Struktur Dasar (Teks 1-5)

Tabel 4.
Hasil Terjemahan Teks 1-5 dari Dola AI dan Grok AI

No	Teks Sumber	Terjemahan Dola AI	Terjemahan Grok AI
1	وصلت الطّائرة مطار جدة	Pesawat mendarat di Bandara Jeddah.	Pesawat telah tiba di Bandara Jeddah.
2	الطّلّاب الكبار في الملعب والطّلّاب الصّغار في الفصل	Siswa-siswa yang lebih besar di lapangan, dan siswa-siswa yang lebih kecil di kelas.	Siswa-siswa yang lebih besar berada di lapangan, sedangkan siswa-siswa yang lebih kecil berada di kelas.
3	هذه الكتب لا تباع	Buku-buku ini tidak dijual.	Buku-buku ini tidak dijual.
4	يحبّ الله المتّقين	Allah mencintai orang-orang yang bertakwa.	Allah menyukai orang-orang yang bertakwa.
5	لمن الدّفتر ذو غلاف جميل؟	Untuk siapa buku catatan bersampul indah ini?	Untuk siapa buku catatan yang bersampul indah itu?

Genre struktur dasar menjadi genre dengan performa terbaik bagi kedua *platform* (Dola AI: 2,80; Grok AI: 2,87). Teks-teks dalam genre ini memuat struktur gramatikal dasar seperti *jumlah fi'liyah, khabar syibhul jumlah, isim isyarah,* dan *maf'ul bih* yang relatif lebih mudah diproses AI. Perbedaan kapabilitas morfologis keduanya sudah terlihat pada teks pertama (وصلت الطّائرة مطار جدة): Dola AI menghasilkan "Pesawat mendarat di Bandara Jeddah" (keakuratan: 2) karena hilangnya aspek temporal *fi'il madhi*, sedangkan Grok AI menghasilkan "Pesawat telah tiba di Bandara Jeddah" (keakuratan: 3) yang berhasil mempertahankan aspek waktu lampau sesuai prinsip Setiadi (2017) bahwa penerjemahan menuntut pemahaman mendalam terhadap gramatika.

Pada teks kedua (الطّلّاب الكبار في الملعب والطّلّاب الصّغار في الفصل), Grok AI unggul dengan menggunakan konjungsi "sedangkan" yang secara semantis lebih tepat untuk menyampaikan hubungan perbandingan, serta menambahkan kata kerja "berada" yang meningkatkan keberterimaan—sejalan dengan prinsip ekuivalensi dinamis Nida bahwa yang diprioritaskan adalah kesepadanan efek komunikatif (Nuur & Hamzah, 2025). Sementara itu, kegagalan identik keduanya pada teks kelima (لمن الدّفتر ذو غلاف جميل؟) dalam mengidentifikasi fungsi *lam al-milki*—keduanya salah menggunakan "untuk siapa" alih-alih "milik siapa"—mencerminkan keterbatasan sistemik AI dalam membedakan fungsi huruf *jar* yang bergantung konteks, sebagaimana dicatat Saimin et al. (2024).

b. Kualitas Terjemahan Genre Keilmuan Keagamaan

Tabel 5.
Hasil Terjemahan Teks 6-10 dari Dola AI dan Grok AI

No	Teks Sumber	Terjemahan Dola AI	Terjemahan Grok AI
6	فإنّ معنى يطيقونه: يطيقونه بمشقّة شديدة، فإن هم أفطروا قضوا أو أطعموا مسكينا	Adapun makna "mereka mampu menanggungnya" adalah: mereka mampu menanggungnya dengan kesulitan yang sangat berat. Jika mereka berbuka (puasa), mereka wajib menggantinya atau memberi makan seorang miskin.	Makna "mereka menanggungnya" adalah mereka menanggungnya dengan kesulitan yang sangat berat. Jika mereka membatalkan puasa, maka mereka harus mengganti (qadha) atau memberi makan orang miskin.
7	الخيار في النّكاح: الإعسار بدفع الصّداق الحالّ	Pilihan dalam pernikahan: ketidakmampuan membayar mahar yang harus dibayar segera.	Khiyar dalam pernikahan: kesulitan (ketidakmampuan) membayar mahar yang harus dibayarkan saat itu (mahar mu'ajjal).
8	وأمّا المسح على الجبائر فإنّه يبلّ يده ومسح فوق الجبيرة كلّها مرّة واحدة	Adapun mengusap khuf (sepatu kulit) dengan perban, maka ia membasahi tangannya dan mengusap seluruh perban sekaligus.	Adapun mengusap atas pembalut (jibairah), maka ia membasahi tangannya lalu mengusap seluruh pembalut tersebut satu kali.
9	الافتراش: هو أن يجلس على باطن رجله اليسرى وينصب اليمنى	Iftirasy (duduk): adalah ia duduk di atas bagian dalam kaki kirinya dan menegakkan kaki kanannya.	Al-Iftirash: yaitu duduk dengan menempelkan bagian dalam kaki kiri (paha dan betis menempel ke lantai) dan menegakkan kaki kanan.
10	الوليمة: لقوله ﷺ لعبد الرّحمن بن عوف لمّا تزوّج: ((أولم ولو بشاة))	Walimah (jamuan makan): berdasarkan sabda Nabi ﷺ kepada Abdurrahman bin Auf ketika ia menikah: "Adakanlah walimah,	Walimah: berdasarkan sabda Rasulullah ﷺ kepada Abdurrahman bin Auf ketika ia menikah: "Berwalimahlah,

		meskipun hanya dengan seekor kambing."	meskipun hanya dengan seekor kambing.”

Genre keilmuan keagamaan menjadi genre dengan kesenjangan performa terbesar (selisih 0,54 poin). Genre ini memiliki kekhasan berupa terminologi fikih yang sangat teknis dan nuansa hukum yang terkandung dalam setiap pilihan diksi. Titik pembeda paling dramatis terjadi pada teks kedelapan (وأمّا المسح على الجبائر...): Dola AI melakukan kesalahan fatal dengan menerjemahkan *jabair* (perban/gips) sebagai "*khuf* (sepatu kulit)"—mencampuradukkan dua entitas hukum fikih yang berbeda, yakni *masah 'alal khuffain* dan *masah 'alal jabair*—sehingga meraih skor terendah mutlak (1-1-1). Grok AI sebaliknya mengidentifikasi *jabair* dengan presisi sebagai "pembalut" dan meraih skor sempurna (3-3-3). Kasus ini sepenuhnya mengkonfirmasi temuan Arju et al. (2025) bahwa intervensi manusia masih sangat diperlukan untuk teks keagamaan teknis, serta memperkuat argumentasi Latifah & Djamilah (2024) tentang keniscayaan verifikasi penutur kompeten.

Pada teks kesembilan (الافتراش...), Grok AI kembali unggul dengan menambahkan keterangan kontekstual dalam kurung "(paha dan betis menempel ke lantai)" yang mentransformasi instruksi abstrak menjadi panduan visual—sebuah strategi *explicitation* yang merupakan wujud konkret ekuivalensi dinamis Nida (Nuur & Hamzah, 2025). Sebaliknya, Dola AI menghasilkan struktur "adalah ia duduk..." yang terasa kaku dan hanya meraih skor keberterimaan dan keterbacaan masing-masing 2. Ketidakmampuan kedua platform pada teks ketujuh (الخيار في النّكاح...) dalam menangkap konsep hukum *fasakh* yang terkandung dalam kata *al-khiyar*—menghasilkan skor (2-1-1) bagi keduanya—membuktikan bahwa tanpa pemahaman sharaf dan nahwu yang diperkuat konteks hukum Islam, terjemahan AI pada teks fikih kehilangan ruh maknanya (Setiadi, 2017; Baihaqi & Syarifah, 2023).

c. Kualitas Terjemahan Genre Jurnalistik

Tabel 6.
Hasil Terjemahan Teks 11-15 dari Dola AI dan Grok AI

No	Teks Sumber	Terjemahan Dola AI	Terjemahan Grok AI
11	أغلب مدارسنا لا تزال تشتغل وفق ما يمكن تسميته بـ"التعليم الأفقي"، حيث يظل التلميذ متلقيًا سلبيًا	Sebagian besar sekolah kita masih bekerja sesuai dengan apa yang dapat disebut sebagai "pendidikan horizontal", di mana siswa tetap menjadi penerima pasif.	Sebagian besar sekolah kita masih beroperasi sesuai dengan apa yang dapat disebut “pendidikan horizontal”, di mana siswa tetap menjadi penerima yang pasif.
12	التعليم المبني على أيديولوجيات متطرفة قد يؤدي إلى إنتاج أفراد يحملون أفكارا تهدد الأمن الإقليمي والدولي	Pendidikan yang didasarkan pada ideologi ekstrem dapat menyebabkan lahirnya individu-individu yang membawa gagasan yang mengancam keamanan regional dan internasional.	Pendidikan yang dibangun di atas ideologi-ideologi ekstrem dapat menghasilkan individu-individu yang membawa pemikiran-pemikiran yang mengancam keamanan regional maupun internasional.
13	تُعدّ العربية من بين الثقافات القليلة التي تعاني من ازدواجية لغوية حادة بين الفصحى والعامية	Bahasa Arab termasuk di antara sedikit budaya yang menderita dualisme linguistik yang tajam antara Bahasa Arab Baku (Fusha)	Bahasa Arab termasuk salah satu budaya yang sedikit mengalami dualitas linguistik yang sangat tajam antara bahasa fusha (standar) dan bahasa amiyah (sehari-hari).

		dan Bahasa Arab Ammiyah (dialek sehari-hari).	
14	نحن أمام خيارين: إما أن نظل أسرى تعليم أفقي ينتج موظفين مطيعين، أو نجرؤ على الصعود لنصنع جيلا من المبدعين	Kita dihadapkan pada dua pilihan: apakah kita tetap menjadi tawanan pendidikan horizontal yang menghasilkan pegawai yang patuh, atau kita berani bangkit untuk menciptakan generasi para inovator.	Kita dihadapkan pada dua pilihan: tetap menjadi tawanan pendidikan horizontal yang menghasilkan pegawai-pegawai yang patuh, atau berani naik ke tingkat yang lebih tinggi untuk menciptakan generasi para pembuat (kreatif/inovator).
15	يستخدم الذكاء الاصطناعي لتكييف المحتوى مع مستوى كل طالب، مما يسمح بتعلم مخصص يتناسب مع احتياجاتهم الفردية	Kecerdasan buatan digunakan untuk menyesuaikan konten dengan tingkat setiap siswa, yang memungkinkan pembelajaran yang dipersonalisasi sesuai dengan kebutuhan individu mereka.	Kecerdasan buatan digunakan untuk menyesuaikan konten sesuai dengan tingkat masing-masing siswa, sehingga memungkinkan terjadinya pembelajaran yang dipersonalisasi sesuai kebutuhan individu mereka.

Genre jurnalistik menjadi satu-satunya genre di mana kedua platform mencapai rata-rata identik (2,80), mengindikasikan bahwa teks berbahasa Arab kontemporer dengan register jurnalistik merupakan kekuatan bersama—selaras dengan Framesthia et al. (2024) yang mencatat kemampuan AI generasi terkini lebih baik dalam menangani bahasa Arab modern. Namun perbedaan halus tetap terdeteksi: pada teks kesebelas, Dola AI menggunakan frasa "bekerja sesuai" untuk mendeskripsikan sistem sekolah (keberterimaan: 2), sementara Grok AI memilih "beroperasi" yang lebih profesional (keberterimaan: 3). Sebaliknya, Dola AI justru mengungguli Grok AI pada teks ketigabelas: Dola AI mengonstruksi kalimat dengan akurasi skor 3, sedangkan Grok AI menghasilkan frasa "sedikit mengalami" yang menggeser penekanan teks asli dan meraih akurasi skor 2—membuktikan keunggulan Grok AI tidak bersifat absolut pada semua konteks (Ma'rifah et al., 2025).

Inkonsistensi penerjemahan bentuk jamak pada Dola AI juga tampak pada genre ini: bentuk jamak *aydiyulujiyyat* diterjemahkan secara tunggal pada teks keduabelas, dan *muwazzhafin* diterjemahkan sebagai "pegawai" (singular) pada teks keempatbelas—pola pergeseran morfologis yang oleh Saimin et al. (2024) diidentifikasi sebagai salah satu problematik inti LLM dalam konteks Bahasa Arab. Teks keempatbelas juga menunjukkan dinamika menarik: Dola AI unggul pada keberterimaan dengan diksi "bangkit" yang puitis (skor 3), sementara Grok AI lebih akurat secara semantis dengan skor keakuratan 3 namun kalah pada keberterimaan (skor 2) karena frasa "naik ke tingkat yang lebih tinggi" yang terasa kaku secara stilistis.

d. Kualitas Terjemahan Genre Sastra/Balaghah

Tabel 7.
Hasil Terjemahan Teks 16-20 dari Dola AI dan Grok AI

No	Teks Sumber	Terjemahan Dola AI	Terjemahan Grok AI
16	زيد بحر السماح والجود	Zaid adalah lautan kemurahan dan kedermawanan.	Zaid adalah lautan kemurahan hati dan kedermawanan.

17	قول الشاعر: تسيل على حدّ السيوف نفوسنا	Perkataan penyair: "Jiwa kami mengalir di atas ujung pedang."	Seperti kata penyair: “Jiwa-jiwa kita mengalir di tepi pedang-pedang.”
18	رأيت أسدا في المعركة	Aku melihat singa di medan perang.	Aku melihat seekor singa di medan pertempuran. (biasanya bermakna kiasan: melihat seorang pemberani/srikandi di medan perang)
19	محمد كثير الرمّاد	Muhammad banyak membuang abu. (Ungkapan yang berarti ia sangat pemurah).	Muhammad banyak abunya. (kiasan: sangat dermawan, sering memasak untuk tamu sehingga banyak abu tungku)
20	في حديث الهجرة قيل لرسول الله ﷺ: ممّن أنتما؟ فقال رسول الله ﷺ: ((نحن من ماء))	Dalam hadis tentang hijrah, Rasulullah ﷺ ditanya: "Kalian berasal dari mana?" Maka Rasulullah ﷺ menjawab: "Kami berasal dari air."	Dalam kisah hijrah, ada yang bertanya kepada Rasulullah ﷺ: “Kalian berdua dari mana?” Rasulullah ﷺ menjawab: “Kami dari air.”

Genre sastra/balaghah secara konsisten menjadi genre terlemah bagi kedua *platform* (Dola AI: 1,60; Grok AI: 1,93), memperlihatkan kegagalan ekuivalensi yang paling fatal. Temuan ini menegaskan argumentasi Kintami et al. (2025) dan Putri & Tatang (2024) bahwa AI belum mampu menangani bahasa figuratif secara memadai. Kegagalan paling absolut terjadi pada teks ketujuhbelas (تسيل على حدّ السيوف نفوسنا) yang mengandung *majaz mursal*: kedua *platform* menerjemahkan *nufus* sebagai "jiwa" secara denotatif, padahal dalam konteks syair perang tersebut *nufus* adalah metonimi untuk darah yang mengucur—menghasilkan skor terendah mutlak (1-1-1) bagi keduanya. Ketidakmampuan menangkap lapisan makna figuratif ini merupakan wujud paling nyata kegagalan ekuivalensi dinamis Nida: pembaca teks sasaran tidak mendapat efek komunikatif yang setara dengan pembaca teks sumber (Nuur & Hamzah, 2025; Baihaqi & Syarifah, 2023).

Pada teks kedelapanbelas (رأيت أسدا في المعركة) yang mengandung *isti'arah*, Grok AI menunjukkan kesadaran metalinguistik yang lebih baik dengan menyertakan keterangan dalam kurung "(biasanya bermakna kiasan: melihat seorang pemberani)" sehingga meraih keakuratan skor 3, sementara Dola AI hanya menerjemahkan secara literal (keakuratan: 2). Pada teks kesembilanbelas tentang *kinayah* (محمد كثير الرمّاد), Grok AI kembali sedikit lebih baik: penjelasan kontekstual dalam kurung "(kiasan: sangat dermawan, sering memasak untuk tamu sehingga banyak abu tungku)" menjadi penyelamat keterbacaan (skor 3), sementara ungkapan "membuang abu" versi Dola AI tidak memiliki korelasi makna dengan kedermawanan dalam budaya Indonesia (keberterimaan: 1). Kegagalan absolut pada kasus *tauriyah* (teks 20) menunjukkan bahwa ambiguitas yang disengaja dalam teks sumber—yang merupakan inti nilai tekstualnya—tidak dapat dijangkau oleh sistem AI manapun saat ini (Latifah & Djamilah, 2024; Mas'udah, 2024).

e. Implikasi Teoritis dan Keterbatasan Penelitian

Secara keseluruhan, penelitian ini membuktikan bahwa Grok AI lebih konsisten dalam menghasilkan ekuivalensi dinamis Nida melalui ketepatan morfologis dan strategi *explicitation*. Pergeseran morfosintaksis dominan pada Dola AI meliputi: penghilangan aspek temporal *fi'il madhi*, inkonsistensi bentuk jamak, dan konfusi terminologi fikih—ketiga pola ini selaras dengan temuan Saimin et al. (2024). Meski demikian, terdapat beberapa keterbatasan yang perlu diakui: sampel 20 teks tidak dapat merepresentasikan seluruh variasi linguistik Bahasa Arab yang kaya; penilaian keberterimaan dan keterbacaan tetap mengandung unsur subjektivitas meskipun telah

melalui validasi ahli (Aminudin et al., 2025); dan kedua *platform* bersifat dinamis sehingga hasil ini merepresentasikan *snapshot* kapabilitas pada satu titik waktu—yang menurut Mohammed et al. (2025) menjadikan evaluasi berkala sebagai keniscayaan ilmiah. Implikasi praktisnya adalah rekomendasi kepada akademisi dan praktisi Pendidikan Bahasa Arab untuk memanfaatkan Grok AI sebagai alat bantu terjemahan yang lebih andal—khususnya untuk teks keilmuan dan jurnalistik—sambil tetap memberlakukan verifikasi manusia untuk teks sastra dan hukum keagamaan (Yusuf, 2023; Lutfiyatun et al., 2023).

Simpulan

Penelitian ini menyimpulkan bahwa Grok AI menghasilkan kualitas terjemahan Arab-Indonesia yang lebih tinggi dibandingkan Dola AI berdasarkan seluruh parameter model Nababan et al. (2012), dengan rerata skor keseluruhan 2,57 berbanding 2,33. Keunggulan Grok AI bersifat konsisten pada ketiga parameter—keakuratan (2,60 vs. 2,30), keberterimaan (2,45 vs. 2,25), dan keterbacaan (2,65 vs. 2,45)—di mana Grok AI unggul pada 11 dari 20 teks dan Dola AI hanya pada 2 teks. Kesenjangan terlebar terjadi pada genre keilmuan keagamaan (selisih 0,54 poin) akibat konfusi terminologi fikih pada Dola AI, sementara genre jurnalistik menghasilkan rata-rata identik (2,80) bagi keduanya.

Dalam kerangka ekuivalensi dinamis Nida, Grok AI terbukti lebih mampu menjembatani jarak budaya melalui strategi *explicitation*—penambahan keterangan kontekstual dalam kurung—yang meningkatkan efek komunikatif teks sasaran. Pergeseran morfosintaksis dominan pada Dola AI mencakup penghilangan aspek temporal *fi'il madhi*, inkonsistensi bentuk jamak, dan konfusi terminologi fikih. Kedua *platform* sama-sama gagal pada genre sastra/balaghah—khususnya *majaz mursal* dan *tauriyah*—sebagai keterbatasan sistemik LLM yang belum terpecahkan. Secara formal, penelitian ini berkontribusi dengan mengisi kesenjangan literatur melalui kajian komparatif Dola AI dan Grok AI serta membuktikan bahwa pengujian lintas genre lebih representatif dibandingkan pengujian genre tunggal (Framesthia et al., 2024; Majid et al., 2024). Penelitian lanjutan disarankan mencakup korpus yang lebih luas, instrumen khusus untuk bahasa figuratif, dan evaluasi longitudinal seiring pembaruan *platform* (Mohammed et al., 2025).

Daftar Rujukan

Abas, A. D. S., Saniyah, B., Filza, L. A. Y., & Salsabilah, Y. A. (2025). Perbandingan hasil terjemahan bahasa Arab ke bahasa Indonesia dalam penggunaan ChatGPT dan DeepL Translator. IDENTIK: Jurnal Ilmu Ekonomi, Pendidikan dan Teknik, 2(6), 38–52.

Aminudin, M. R., Syihabbudin, & Nurmala, M. (2025). Studi komparatif terjemahan buku Mulakhkhas Qawa'id al-Lughah al-'Arabiyyah karya Fuad Ni'mah oleh Abu Ahmad dan Quillbot.ai. Ihya Al-Arabiyah: Jurnal Pendidikan Bahasa dan Sastra Arab, 11(1).

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), 1–15. https://doi.org/10.20414/tsaqafah.v24i1.10604.

Azhar, M., Yolanda, D., Frananda, A., Saputra, M. R., Mulyani, R., & Nurdinah, S. (2025). Revolusi pembelajaran bahasa Arab di era digital: Analisis peran Artificial Intelligence dalam pengembangan kompetensi berbahasa. Al-Ma'lumat: Jurnal Ilmu-Ilmu Keislaman, 3(2), 70–82. https://doi.org/10.56184/jam.v3i2.506.

Baihaqi, M., & Syarifah, A. (2023). Menjadi praktisi penerjemah Arab-Indonesia. Penerbit JDS.

Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). Sage Publications.

Framesthia, L. M., Khoerunnisa, S., Al Izzah, S., Al Farisi, M. Z., & Supriadi, R. (2024). Analisis perbandingan teknik penerjemahan Arab-Indonesia pada Google Translate dan ChatGPT (Kajian sintaksis: Kalimat verba dan nomina). Jurnal Pendidikan Bahasa Arab, 7(2).

Kholik, A., Fauzi, M. D., & Kesuma, H. D. (2025). Analisis sentimen rebranding aplikasi Cici menjadi Dola pada Google Play Store menggunakan metode IndoBERT. Jurnal Ilmiah Informatika Global, 16(3), 370–378.
Kintami, R., et al. (2025). The Eden Poem by Kate Llewellyn: The comparison case on Sapardi Djoko Damono's and Grok's translation. Poetics and Linguistics Journal, 11(1), 34–49.
Latifah, S. N., & Djamilah, W. I. F. (2024). Penggunaan Chat GPT dalam penerjemahan bahasa Arab. Al-Dirosah: Jurnal Pendidikan Agama Islam, 1(2).
Lutfiyatun, E., Kurniati, D., & Fajriah, N. (2023). Pemanfaatan Artificial Intelligence (AI) dalam meningkatkan kompetensi pembelajaran gramatikal, tarjamah dan muhadatsah di perguruan tinggi. SEULANGA: Jurnal Pendidikan dan Pelatihan, 2(2).
Ma'rifah, S. N., Syihabuddin, & Supriadi, R. (2025). Analisis kualitas terjemahan dalam Qasidad Al-Burdah: Studi komparatif antara terjemahan Abdullah Azzam bin Azlan dan DeepL Translator. Jurnal Onoma: Pendidikan, Bahasa dan Sastra, 11(3), 3632–3646.
Majid, M. N., Akmaliyah, & Hermawan, M. B. (2024). Perbandingan hasil terjemah Arab-Indonesia antara ChatGPT dan Gemini AI. Jurnal Pendidikan Bahasa Arab, 7(2).
Mas'udah. (2024). Efektivitas penggunaan alat instan dalam penerjemahan teks bahasa Arab. Ilmuna: Jurnal Studi Pendidikan Agama Islam, 7(1), 363–375. https://doi.org/10.54437/ilmuna.
Mohammed, A. A. Q., Mudhsh, B. A., Bin-Hady, W. R. A., & Al-Tamimi, A. S. (2025). DeepSeek and Grok in the spotlight after ChatGPT in English education: A review study. Journal of English Studies in Arabia Felix, 4(1), 13–22. https://doi.org/10.56540/jesaf.v4i1.114.
Nababan, M., Nuraeni, A., & Sumardiono. (2012). Pengembangan model penilaian kualitas terjemahan. Kajian Linguistik dan Sastra, 24(1), 39–57.
Nurullawasepa, M., Mandani, N. Z., Adawiyah, R., Al Ayyubi, S., & Abdillah, A. A. (2023). AI (Artificial Intelligence) dalam penerjemahan teks bahasa Arab. 3rd E-Proceeding SENRIABDI 2023, 3(1), 141–157.
Nuur, K. N., & Hamzah, A. A. (2025). Analisis ekuivalensi dalam penerjemahan Arab-Indonesia: Sintesis teori Nida, Koller, dan Newmark. Loghat Arabi: Jurnal Bahasa Arab dan Pendidikan Bahasa Arab, 6(2), 630–645. https://doi.org/10.36915/la.v6i2.522.
Putri, A., & Tatang. (2024). Variasi genre teks dalam evaluasi mesin penerjemah: Analisis perbandingan teks akademik dan sastra. Jurnal Linguistik Komputasional, 7(2), 85–99.
Saimin, A. A., Supriadi, R., & Al Farisi, M. Z. (2024). Analisis kesalahan penerjemahan teks bahasa Indonesia ke dalam bahasa Arab pada ChatGPT (Studi analisis morfologi dan sintaksis). NASKHI: Jurnal Kajian Pendidikan dan Bahasa Arab, 6(2), 110–125. https://doi.org/10.47435/naskhi.v6i1.2668.
Samet, U. (2025). The positive influence of large language models on fact-checking practices: A case study of Grok. World Journal of Advanced Engineering Technology and Sciences, 15(3), 1727–1738. https://doi.org/10.30574/wjaets.2025.15.3.1123.
Setiadi, S. (2017). Penerjemahan Arab-Indonesia. Maninjau Press.
Syam, M. N., Isnaini, R. L., Rohmah, L., & Sa'adah, S. N. (2023). The analysis of Google Translate translation error from Indonesian to Arabic and tips for using it. IJAZ ARABI: Journal of Arabic Learning, 6(1), 247–257. https://doi.org/10.18860/ijazarabi.v6i1.16299.
Yusuf, K. (2023). Bringing ChatGPT into Arabic language learning. ICONITIES: International Conference on Islamic Civilization and Humanities, 1(1), 465–478.

Ucapan Terima Kasih

Peneliti mengucapkan terima kasih kepada Institut Agama Islam Imam Syafii Indonesia atas dukungan akademis dalam pelaksanaan penelitian ini. Apresiasi khusus disampaikan kepada Validator Ahli yang telah meluangkan waktu untuk menilai seluruh data terjemahan secara objektif, serta kepada seluruh pihak yang telah memberikan masukan konstruktif dalam proses penyempurnaan naskah ini.