KSU-JRLT

Journal of Research in Language & Translation
Volume 6, Special Issue (December 2026)

جامعة الملك سعود
King Saud University

مراجعة نطاقية للأدلة المنشورة بالإنجليزية حول موارد بيانات الترجمة وتقنيات اللغة للعربية في المملكة العربية السعودية

كايد السلطان
جامعة الإمام محمد بن سعود الإسلامية
الرياض، المملكة العربية السعودية
ORCID: https://orcid.org/0009-0009-7224-5588
kayedalsultan@gmail.com
Received: 6 October 2026; Revised: 31 October 2026; Accepted: 10 November 2026
DOI: 10.33948/JRLT-KSU-6-S-7

الملخص

تستعرض هذه المراجعة البحوثَ المنشورة باللغة الإنجليزية بين يناير 2016 ويونيو 2026 في موارد بيانات الترجمة وتقنيات اللغة الخاصة بالعربية ذات الصلة بالمملكة العربية السعودية، وقد اتُّبعت في عرض تقريرها إرشادات PRISMA-ScR. وجُمعت المصادر من عمليتي بحث في قاعدة Web of Science، ومن قاعدة Scopus، ومن عشرين بحثًا على الويب، ومن مسح يدوي لأعداد مجلة سعودية متخصصة في الترجمة، وتولّى باحث واحد فرزها واستخراج بياناتها مستعينًا بمساعد ببليوغرافي مدعوم بالذكاء الاصطناعي. واستوفى 139 مصدرًا معايير الإدراج، منها 14 في موارد البيانات والمدونات، و32 في تطوير تقنيات الترجمة، و72 في تقييمها واستخدامها، و21 في المصطلحات، ونُشر 105 منها في عام 2022 أو بعده. واستُخرجت نتائج 71 مصدرًا، أما المصادر الـ68 التي أضافتها عمليات البحث في قواعد البيانات في أكتوبر 2026 فاقتُصر فيها على الحقول الوصفية. وتتألف الموارد التي استُخرجت نتائجها في معظمها من مدونات ومجموعات بيانات بُني كل منها لمهمة واحدة، واختُبرت الأنظمة على معايير قياس منفصلة، وأغلب دراسات التقييم صغيرة تتناول Google Translate وروبوتات المحادثة في نصوص قانونية أو دينية أو أدبية أو مرتبطة بثقافة بعينها. وترصد دراسات المصطلحات تباينًا وأخطاءً في المصطلحات القانونية والحاسوبية والطبية، وقلّما تذكر أوراق الموارد طريقة الوصول إلى بياناتها أو الجهة التي تتولى صيانتها. وتشير النتائج إلى أربع أولويات للبحث والتشاور، هي توثيق الموارد وإعادة استخدامها، والبيانات اللازمة لمهام محددة، وممارسات التقييم المشتركة، ودعم المصطلحات. وتناقش المراجعة منصةً مشتركة للمصطلحات بوصفها خيارًا واحدًا ينبغي اختباره. ويحدّ من نتائجها اقتصارُها على المصادر الإنجليزية، وتولّي باحث واحد الفرز، وعدم استخراج النتائج من جميع المصادر.

الكلمات المفتاحية: الترجمة الآلية؛ المملكة العربية السعودية؛ المراجعة النطاقية؛ المصطلحية؛ مدونات الترجمة؛ تدريب المترجمين

KSU-JRLT

Journal of Research in Language & Translation
Volume 6, Special Issue (December 2026)

جامعة الملك سعود
King Saud University

A Scoping Review of English-Language Evidence on Translation Data Resources and Language Technologies for Arabic in Saudi Arabia

Kayed Alsultan
Imam Mohammad Ibn Saud Islamic University
Riyadh, Saudi Arabia
ORCID: https://orcid.org/0009-0009-7224-5588
kayedalsultan@gmail.com
Received: 6 October 2026; Revised: 31 October 2026; Accepted: 10 November 2026
DOI: 10.33948/JRLT-KSU-6-S-7

Abstract

The present review maps English-language research published between January 2016 and June 2026 on translation data resources and language technologies for Arabic with a Saudi connection. Reporting is guided by PRISMA-ScR. Sources came from two Web of Science searches, Scopus, 20 web searches, and a hand-search of one Saudi translation journal, and one reviewer screened and charted them with an AI-assisted bibliography assistant. The review included 139 sources of evidence. Fourteen concern data resources and corpora, 32 the development of translation technology, 72 its evaluation and use, and 21 terminology, and 105 date from 2022 or later. Findings were charted for 71 sources, and the 68 added by the database searches of October 2026 were charted on descriptive fields only. The charted resources are mostly corpora and datasets built for single tasks, systems were tested on separate benchmarks, and most evaluations are small studies of Google Translate and chatbots on legal, religious, literary, or culture-bound texts. Terminology studies report variation and errors in legal, computing, and medical terms. Few resource papers state how their data can be accessed or who maintains them. The findings point to four priorities for research and consultation, namely resource documentation and reuse, data for particular tasks, shared evaluation practice, and terminology support. A shared terminology platform is discussed as one option to be tested. Coverage is limited by the English-only scope, single-reviewer screening, and partial charting.

Keywords: machine translation; Saudi Arabia; scoping review; terminology; translation corpora; translator training

Introduction

Saudi Arabia created several public bodies with responsibilities for Arabic, translation, or language technology during the period covered by this review. Saudi Vision 2030 was launched in April 2016 (Saudipedia, n.d.-e). The Saudi Data and Artificial Intelligence Authority was established in 2019 (Saudipedia, n.d.-d), and its ALLaM series of Arabic and English language models was published in conference proceedings in 2025 (Bari et al., 2025). The Literature, Publishing and Translation Commission was established in 2020 (Saudipedia, n.d.-c) and launched the Tarjim translation initiative in the same year (Saudipedia, n.d.-f). The King Salman Global Academy for Arabic Language was also established in 2020 (King Salman Global Academy for Arabic Language, n.d.), and its stated functions include approving and publishing linguistic terms (Saudipedia, n.d.-b). Saudi universities produced corpora, machine translation systems, evaluations of large language models, and analyses of specialized terms in the same period.

The resulting studies appear in computing, translation studies, applied linguistics, and health journals, and earlier reviews cover parts of them. Bibliometric studies counted translation research from Saudi Arabia and the Gulf states (Alangari, 2024; Alyami & Qassem, 2024), and reviews examined Arabic and English machine translation, post-editing, and the quality of neural translation (AlGhamedi, 2024; Almaaytah & Alzobidy, 2023; L. I. Omar & Salih, 2024). Boulaares (2026) discussed Saudi investment in language technology and terminology as language planning. None of them maps resources, technologies, and terminology studies with a Saudi connection together. Translators of specialized texts, translator education programs, and planners need such a map, and a scoping review suits the task because it charts the extent and nature of mixed evidence (Pollock et al., 2026; Tricco et al., 2018). Three research questions guided the review.

RQ1. Which translation data resources and language technologies for Arabic were developed or studied with a Saudi connection between January 2016 and June 2026, and which institutions or teams are associated with them?

RQ2. What do the studies report about the coverage, quality, use, and limitations of these resources and technologies?

RQ3. Which gaps do the studies identify, and how do these gaps relate to the documented functions of relevant Saudi institutions?

The window begins in January 2016 so that the review covers about ten years of research, a span long enough to show change over time and small enough for one reviewer to screen. It ends in June 2026, the last full month before the first database search on 24 July 2026.

Literature Review

Large parallel collections that include Arabic exist. The United Nations Parallel Corpus contains aligned documents of the United Nations in its six official languages (Ziemski et al., 2016), and OPUS gathers parallel corpora built for different purposes (Tiedemann, 2012), so the domain coverage of the two differs. Arabic terminology resources are older. The BASM term bank of King Abdulaziz City for Science and Technology began in 1983 within a machine translation project and held more than 200,000 terms by 1989 (Sieny, 1989). Arabterm, a technical dictionary in four languages verified by the Arab League Educational, Cultural and Scientific Organization, completed volumes on water technology and renewable energy with a grant approved in 2015 (OPEC Fund for International Development, n.d.), and the King Abdullah Institute for Translation and Arabization (n.d.) lists the unification and dissemination of scientific terminology among its objectives. These sources describe aims and contents. They report nothing on how widely the resources are used or how they reach translators' working tools.

Methods

Design and Reporting

The study is a scoping review. Its reporting is guided by PRISMA-ScR (Tricco et al., 2018) and the search items of PRISMA-S (Rethlefsen et al., 2021), and the JBI guidance (Pollock et al., 2026) informed screening and charting. The unit of inclusion is a source of evidence, and versions of one work were linked and counted once. The synthesis is narrative and thematic. No protocol was registered. Supplementary File 4 gives the extended methods and results and lists the decisions changed during revision, among them the operational definitions, the re-screening of all database records, the audit of the keyword filter, the journal hand-search, and the two added database searches.

Eligibility Criteria

No restriction was placed on participants. The concept had three elements, namely translation data resources involving Arabic and another language, technologies used to produce, assist, or assess translation involving Arabic, and terminology involving Arabic and another language where particular terms, texts, or participants were analyzed. The context was a Saudi connection, defined as an author with a Saudi institutional affiliation, a setting or participants in Saudi Arabia, or Saudi texts or Saudi Arabic data. Funding alone did not qualify, and no form of connection was taken as evidence of who owns a resource. Journal articles, conference papers, dissertations, technical reports, and preprints published between January 2016 and June 2026 with an English full text were eligible, and each source was dated by its earliest publication, online or in print. Each source was assigned one primary cluster by its main contribution, namely data resources (A), technology development (B), evaluation and use of technology (C), or terminology (D). The 13 exclusion codes are defined in Supplementary File 4.

Information Sources and Search

Web of Science Core Collection, accessed through the Saudi Digital Library, was searched on 24 July 2026 with ALL=(arabic AND ("machine translation" OR "post-editing" OR "translation memory" OR corpus OR terminology OR arabization OR localization) AND (saudi OR "saudi arabia")), which returned 594 records from six indexes. During revision, on 20 October 2026, a second Web of Science search widened the terms to arabization and its spelling variants, translat*, interpreting, post-edit* and its variants, translation memor*, corpora, terminolog*, termbase*, term bank*, CAT tool*, computer-assisted and computer-aided translation, localization and localisation, large language model*, and chatgpt, combined with saudi or KSA, and returned 2,108 records. Scopus was searched with the same terms in titles, abstracts, and keywords, combined with a Saudi affiliation or the terms saudi or ksa, and returned 2,182. Twenty web queries were run through an AI-assisted bibliography assistant in June and July 2026, and the 13 issues of the *Journal of Research in Language & Translation* published between 2021 and October 2026 were hand-searched. The databases were searched without language or document-type filters. Supplementary File 4 (Tables S4 and S5) and Supplementary File 1 give the exact queries, dates, and yields. Arabic-language databases were not searched, because the review was restricted to English full texts from the start.

Selection and Charting

The records of the first Web of Science search were filtered by a keyword script (Supplementary File 3), which kept a record when its title, abstract, or author addresses contained a Saudi-link term and a translation-relevance term and its year lay between 2016 and 2026. The script retained 261 of the 594 records. During revision, all 261 were screened again against written definitions, and the titles of the 333 records rejected by the filter were read, with abstracts read where needed for the 74 that a broader term list flagged. The audit found no eligible record among them. Before the October records were screened, 589 records of the

second Web of Science search that the first search had already retrieved were set aside, and 1,533 Scopus records that matched Web of Science records by DOI (1,405) or by normalized title (128) were removed. The 2,167 October records left after the removal of duplicates were screened on title, 124 on abstract, and 80 on the full record. Every decision was logged with a reason code (Supplementary File 1), and the date, language, and Saudi connection of each candidate were checked at the publisher or in Crossref. The Crossref records of all included sources with a Crossref DOI were checked for retractions, corrections, and expressions of concern on 20 and 25 October 2026, and none was found.

One reviewer carried out screening, verification, and charting with an AI-assisted bibliography assistant. The review had a single author and no funding, so no second reviewer was available. The charting form recorded bibliographic details, primary cluster, form of Saudi connection, Saudi institutions, design, publication status, sample or dataset, language direction, technology or resource, domain, main finding, and, for resources, the stated access conditions. All fields were charted for the 71 sources from the first Web of Science search, the web searches, and the journal hand-search. The 68 sources added by the October searches were charted on all fields except the main finding and the access conditions, because charting those two fields was not possible within the revision period. Charting drew mostly on database records and published abstracts, and the charting workbook is Supplementary File 2. Findings are reported as stated by the sources, missing values are recorded as not reported, and the current availability, version, license, and host of resources were not checked.

Appraisal and Ethics

No formal critical appraisal was carried out, which is consistent with a scoping review. Publication status, design, and sample were recorded as descriptive dimensions. The review analyzed published documents only, so the Saudi Law of Ethics of Research on Living

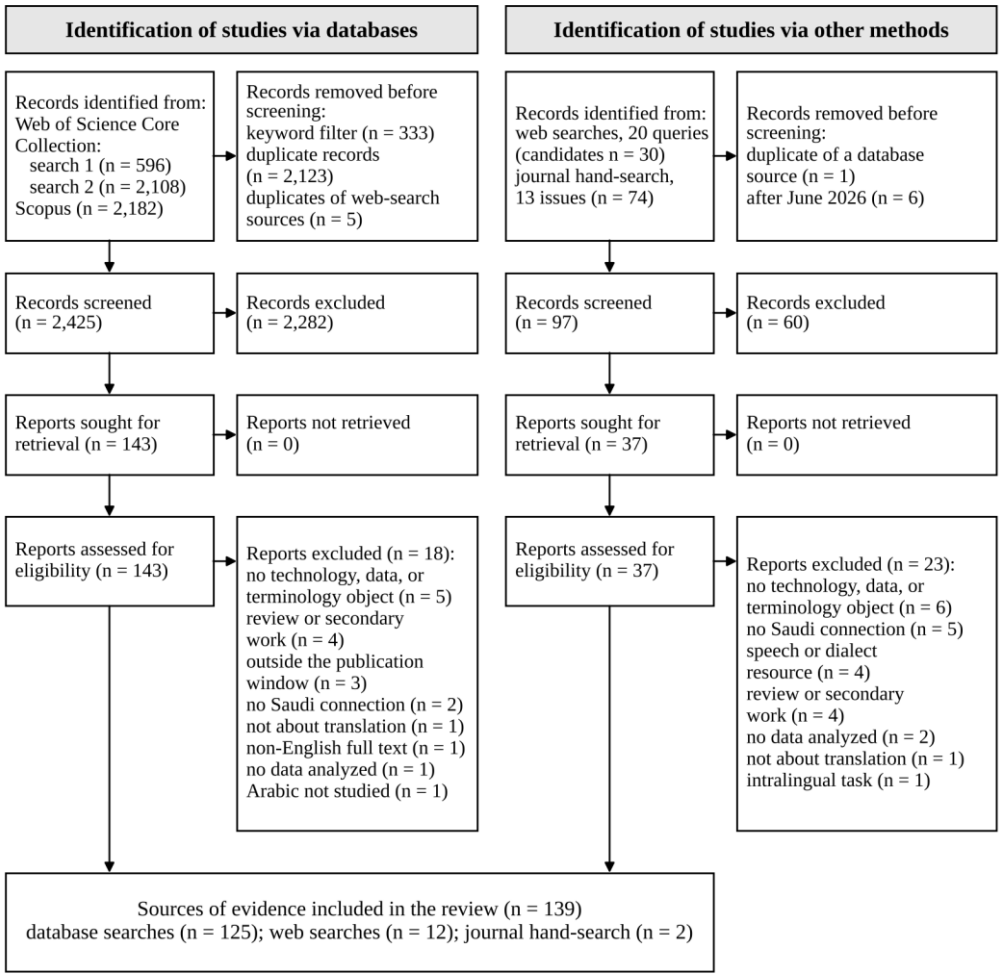
Creatures did not apply (Law of Ethics of Research on Living Creatures, 2010), and no ethics review was sought.

Results

Selection of Sources

Figure 1 shows the selection process. The database searches identified 4,886 records, 2,425 of which were screened after the keyword filter and the removal of duplicates, and 125 sources were included. Titles excluded 2,043 of the October records and abstracts 44 more, and 12 of the 80 assessed in full were excluded, among them three published online after June 2026 and two without a Saudi connection, one of which corrected a Scopus indexing error. The journal hand-search covered 74 articles, 6 of them published after June 2026, and 8 were assessed in full. The web searches added 12 sources and the journal hand-search 2, for a total of 139. Supplementary File 4 (Table S6) lists the sources excluded after full-record assessment with their reasons.

Figure 1
Selection of Sources of Evidence



Note. Adapted from "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews," by M. J. Page et al., 2021, *BMJ, 372*, Article n71 (https://doi.org/10.1136/bmj.n71). CC BY 4.0. Web of Science search 1 comprises the 594 records exported in July 2026 and 2 records added when the search was run again in October 2026. The keyword filter was applied to the July export only. The duplicate records are 589 search 2 records also retrieved by search 1, 1,533 Scopus records that matched Web of Science records, and 1 repeated Scopus record. The web-search candidate that duplicated a database source is counted once, in the database column. The records screened under other methods are the 29 remaining web-search candidates and the journal articles published by June 2026.

Overview of the Included Evidence

Table 1 summarizes the 139 sources by cluster. Of these, 105 were published in 2022 or later, and 43, all from 2023 onward, developed, used, or evaluated large language models or chatbots. At least one author has a Saudi affiliation in 136 sources, and for 100 of them affiliation is the only Saudi connection. Twenty-seven sources involve a setting or participants in Saudi Arabia, and 13 analyze Saudi texts or Saudi Arabic. Authors with Saudi affiliations

came from 46 institutions, most often King Saud University (23 sources), Prince Sattam bin Abdulaziz University (16), King Abdulaziz University (14), Princess Nourah bint Abdulrahman University (14), and Najran University (11). Fifty-six sources concern translation into Arabic, 42 translation from Arabic, and 26 both directions, while three involve other language combinations and the records of 12 do not state a direction. Supplementary File 2 and Table S1 in Supplementary File 4 give the characteristics of every source.

Table 1
Included Sources of Evidence by Cluster

Cluster	Sources	Findings charted	Published 2022 or later	Main kinds of source
A. Data resources and corpora	14	11	7	Parallel and comparable corpora, learner translation corpora, and datasets made with machine translation
B. Translation technology development	32	15	20	Machine translation systems and models, alignment and benchmarks, language and speech models, quality estimation, and teaching tools
C. Evaluation and use of translation technology	72	31	61	Evaluations of machine translation and large language models, post-editing and workflow studies, and studies of learners and users
D. Terminology	21	14	17	Analyses of legal, computing, medical and health, and Qur'anic embryology terms
All clusters	139	71	105	109 journal articles, 25 proceedings papers, 3 preprints, and 2 doctoral dissertations

Note. Findings were charted for the sources from the first Web of Science search, the web searches, and the journal hand-search. The sources added by the October searches were charted on descriptive fields only. Supplementary File 2 lists every source, and Table S1 in Supplementary File 4 gives the characteristics of each.

Data Resources and Corpora

The charted resource papers describe corpora and datasets built for particular tasks. They include comparable and parallel corpora for training or teaching (Abouammoh et al., 2016; Alotaibi, 2017; Althobaiti, 2022), a bilingual Hadith corpus (Altammami et al., 2020), and two learner translation corpora connected with one university (Alfuraih, 2020; Al-Harthi et al., 2024). Althobaiti (2022) extracted 105,010 sentence pairs from Wikipedia that trained a system comparable to systems trained on Wiki-OPUS and TED, Bounhas and Ben Guirat (2019) built a classical Arabic retrieval collection with English translations of its documents, and Ehab et al. (2026) released Arabic and English thesis titles for scientific translation. Two datasets were

made with machine translation or AI, Arabic ImageNet labels of which 65.6% were judged accurate (Alsudais, 2019) and SauDial, game-localization data in four Saudi dialects reviewed by linguists (Alanazi et al., 2025). The two learner corpora differ in language coverage, design, and annotation, and their shared university is not evidence of duplicated collection. Access is reported unevenly. The papers name a web concordancer (Alotaibi, 2017), a repository (Altammami et al., 2020), data on request (Abouammoh et al., 2016), and an application and website (Al-Harthi et al., 2024), while most other records give no access route, license, version, or maintainer. The ten-million-word size of the King Saud University corpus is a project target (Alotaibi, 2017), and current availability was not checked for any resource.

Translation Technology Development

Development studies built rule-based, example-based, statistical, and neural systems, alignment methods, quality estimation, and teaching tools. Alrajeh (2018) found a neural system 13 BLEU points above a phrase-based baseline, Alghamdi et al. (2024) improved the translation of financial news by fine-tuning on a small in-domain set and released their data and models, and Huang and Asiri (2026) introduced an alignment benchmark whose hard subset exposed weaknesses that one-to-one data had hidden. Mahmoud and Mengash (2021) reported that their example-based system outperformed OmegaT and Apertium once the corpus exceeded 1.6 million phrases, Alamri (2026) reported a BLEU score of 47.8 for an Arabic-specific transformer, and Alrashidi and Mathkour (2026) found that BLEU alone gave misleading conclusions about tokenization under low-resource conditions. Almarzoqi and Alsuhaibani (2025) used large language models to judge the emotion of English translations of the Quran and found agreement with Arabic readers between kappa 0.29 and 0.35. Bari et al. (2025) reported that parallel and translated data helped to align Arabic and English knowledge in ALLaM during training. Aleedy et al. (2025) reused the Saudi Learner Translation Corpus

to generate sentences for teaching tools, the one documented case of a Saudi corpus being reused. The scores come from different test sets, metrics, and directions and cannot be ranked across studies. The 17 sources added in October include seven further systems or models, two language or speech models trained with translated or synthetic data, and three studies that localized material with large language models or semi-automated pipelines.

Evaluation and Use of Translation Technology

Cluster C is the largest. Of its 31 charted sources, 10 evaluate machine translation output, 7 evaluate large language models, 7 concern post-editing and workflow, and 7 concern learners and users. Machine translation evaluations reported errors with literary texts, dialect, and dialogue (A. Omar & Gomaa, 2020; Qassem & Aldaheri, 2023; Sabtan et al., 2021), and the ranking of systems depended on the metric (Al-Rukban & Saudagar, 2017). Fakih et al. (2024) judged Instagram's machine translation deficient for 90% of 30 literary captions, and Alwazna (2024) compared five systems on articles of the Saudi Commercial Law and cautioned that HTER alone can miss term errors that change legal meaning. Evaluations of large language models began in 2023. They report BLEU scores of 35.05 for GPT-3.5 and 38.83 for GPT-4 on a United Nations test set (Alyafeai et al., 2023), better ChatGPT than Google Translate output for academic abstracts (Mohsen, 2024), literal renderings of idioms most often from Google Translate (Obeidat et al., 2024), and better legal translations when prompts included the translation commission (Aldosari & Altuwairesh, 2025). Bing AI Chat outperformed Google Translate and Bing Translator on colloquial Arabic sentences (Aldawsari, 2024), and GPT-4 had the highest BLEU score (46.56) among the systems tested on Saudi Arabic (Alahmari et al., 2025). Alhumsi (2025) reported BLEU scores of 0.77 for ChatGPT and 0.69 for Google Translate on four research abstracts. Among the studies of practice and training, 17 of 18 translators in certified agencies post-edited machine translation (Alshehri & Alowedi,

2023), and training raised students' acceptance of post-editing (Samman, 2022). Fifteen students post-editing Saudi legal texts in Trados Studio struggled with precision, extra-linguistic meaning, and specialized terms (Zitouni et al., 2024). Human post-editors outperformed ChatGPT-4o on most quality measures, although ChatGPT-4o was faster (Algaraady & Mahyoob, 2025), and two translators using ChatGPT with post-editing translated 15 psychiatric scales in 24 hours, against 120 hours for four translators using translation and back-translation (Aboraya et al., 2026). Students surveyed by Alkhatnai (2025) valued the speed of AI tools but corrected idioms and cultural content by hand. The 41 October sources in the cluster add evaluations of chatbots on religious texts, proverbs, idioms, and legal articles, evaluations of Google Translate on drug information, signs, and health texts, and surveys and experiments with students and translators. The designs limit comparison. The smallest test sets are five sentences (Ali, 2016) and one text (Aldawsari, 2022), participant samples run between 2 and 229 people, metrics vary, and model versions and test dates are not always reported.

Terminology

The charted terminology studies concern legal, computing, and medical terms. Alwazna (2019) located the difficulty of legal terms in conceptual mismatch between legal systems, El-Zeiny (2017) compared how three bilingual dictionaries handle lexical gaps for Islamic terms, Alhomoud (2022) documented semantic, syntactic, pragmatic, and cultural problems with Sharia terms in Saudi contracts, and Jehjah (2025) reported high equivalence with seven detrimental errors in legal terms concerning children. Bloggers glossed computing terms more often where standardized forms existed without common use (Arrabai, 2022), and loan translations produced frequent synonymous variants (Al-Athwary & Ali, 2024). Elgobshawi (2022) found that 159 computer science students preferred translated forms of 25 computer terms to Arabized ones, and translators of information technology texts used amplification,

explication, and transliteration where standardized equivalents were missing (Muftah, 2026). El-Kassas et al. (2025) surveyed physicians, linguists, and patients from 16 Arab countries to agree on a term for steatotic liver disease, Sarsak (2024) built a free English and Arabic occupational therapy dictionary app with more than 2,400 terms, Assiri (2026) found ChatGPT-4 renderings of ten medical terms less accurate than human translations, and Qassem (2026), a preprint, found course titles named and translated differently at two universities. These findings describe conceptual non-equivalence, variation, errors, and consensus procedures, which need to be kept apart, and they show no general failure of standardization. The seven October sources add corpus studies of legal terms in Saudi laws and analyses of other legal, medical, and Qur'anic embryology terms.

Search Boundaries

Related work was screened and excluded. Among the database records, these were 15 speech and dialect resources not directed at translation, 60 sign language translation studies, 44 studies of intralingual tasks, and audiovisual and other translation studies without a technology, data-resource, or terminology object. No study of interpreting technology in professional practice was found, although such research may exist outside the sources searched. Among the database records screened, 16 from the first search and 29 from the October searches had full texts in Arabic or other languages and were excluded.

Institutional Context

Table 2 summarizes the documented functions of public bodies with a published mandate concerning Arabic terminology, translation, or language technology. The table is contextual. Its sources were found during charting and by searches of official and reference websites on 19 and 20 October 2026, and they lie outside the included set.

Table 2

Documented Functions of Public Bodies and Resources Relevant to Arabic Terminology and Translation

Body or initiative	Dates stated	Documented function or activity	Source
Literature, Publishing and Translation Commission (Ministry of Culture)	Established 2020	Regulates and develops the literature, publishing, and translation sectors	Saudipedia (n.d.-c)
Tarjim initiative of the Commission	Launched September 2020	Translation grants open to Saudi publishing houses and translation of academic journals in cultural, social, and humanities fields	Saudipedia (n.d.-f)
Arab Observatory for Translation (listed by Saudipedia as the Arabic Observatory of Translation)	Launched 2022 by the Commission; grant program announced 5 September 2022; memorandum with ALECSO signed 26 October 2022	Bibliographies of translated books and of translation research, a directory of publishing houses and translators, and grants for translation research	Saudipedia (n.d.-a)
King Salman Global Academy for Arabic Language	Council of Ministers Decree No. 34, 1 September 2020	Stated functions include approving, updating, and publishing linguistic terms	King Salman Global Academy for Arabic Language (n.d.); Saudipedia (n.d.-b)
Saudi Data and Artificial Intelligence Authority	Royal Decree of 30 August 2019; ALLaM reported 2024 and published 2025	National Center for Artificial Intelligence among its executive entities; ALLaM Arabic and English model series	Saudipedia (n.d.-d); Bari et al. (2025)
BASM term bank, King Abdulaziz City for Science and Technology	Experimental version from June 1983	Term bank in Arabic, English, French, and German built for a machine translation project, with more than 200,000 terms by 1989; used as the reference for standardized technology terms by Arrabai (2022)	Sieny (1989); Arrabai (2022)
King Abdullah Institute for Translation and Arabization, Imam Mohammad Ibn Saud Islamic University	Establishment approved by Council of Higher Education decision 39/66/1432; operating from 30 Muharram 1433 AH (December 2011)	Objectives include translating, unifying, and publicizing scientific terminology and creating dictionaries	King Abdullah Institute for Translation and Arabization (n.d.)
Arabterm (regional)	OPEC Fund grant approved 2015 for the completion of two volumes	Technical dictionary in four languages developed with GIZ and verified by ALECSO, with volumes on water technology and renewable energy	OPEC Fund for International Development (n.d.)

Note. Information comes from official and reference websites accessed on 19 and 20 October 2026, from Sieny (1989), and from two included sources (Arrabai, 2022; Bari et al., 2025). Apart from those two sources, the table is outside the included set and was not appraised. Dates are those stated by the sources. ALECSO = Arab League Educational, Cultural and Scientific Organization; GIZ = German development cooperation agency.

The included sources show several links between public bodies and translation resources. ALLaM is a model series of the Saudi Data and Artificial Intelligence Authority (Bari et al., 2025), Arrabai (2022) used BASM as the reference for standardized terms, and the Commission supports translation research through grants of the Arab Observatory for Translation (Saudipedia, n.d.-a). The Academy and the King Abdullah Institute list the approval or unification of terms among their functions, but none of the sources consulted describes how

approved terms reach translators' tools or how often translators consult them. A connection that the sources do not report may still exist.

Synthesis of Gaps

The charted evidence points to four priorities for investigation. Resource documentation comes first, since few resource papers state an access route, license, version, or steward, and only one case of reuse is documented. Data needs are task-specific, covering domain adaptation, scientific translation, dialect data, and alignment benchmarks, and the evidence shows no uniform shortage. Evaluation practice needs shared reporting of samples, metrics, model versions, and test dates. Terminology support is the fourth, since legal, computing, and medical studies report mismatch, variation, and errors, while evidence on how translators use existing term resources is thin. These are implications drawn from the literature. The sources do not measure duplicated effort, expenditure, or the effect of any remedy (Table S3 in Supplementary File 4).

Discussion

Interpreting the Gaps

The sources describe what was built and tested at 46 Saudi institutions during a decade in which public bodies acquired mandates for Arabic terminology, translation, and language technology. They say little about how resources can be found, reused, or connected to translators' work, so the question they raise concerns continuity between resource development and professional use. The review cannot judge the adequacy of existing coordination, and a shared university or a missing citation between two corpus papers is not evidence of duplicated cost. The learner corpora show different design choices and a possible opportunity to compare or connect resources, and the professional and economic consequences of the terminology problems would need to be measured directly.

Responses Considered

Several responses could be compared before any new platform is built. An existing term bank such as BASM could be updated, collaboration with Arabterm could bring in verified multilingual terms, and a federated arrangement could leave resources with their holders while adding shared metadata and search. Federation here means planned coordination, and resources held by different institutions without shared metadata do not amount to it. The included sources do not compare these options, so a needs assessment and an open comparison of coverage, licensing, governance, cost, and uptake would have to come first.

A Proposed Platform

One option is a national platform for terminology and translation resources. The proposal is offered for development and evaluation and goes beyond what the review itself can show. The terminology findings suggest concept-based entries with definitions, sources, domain and jurisdiction labels, and documented acceptable variants, since a single equivalent per concept could erase legitimate legal or register distinctions. Scientific validation could rest with a qualified reference body, and the King Salman Global Academy for Arabic Language, whose stated functions include approving terms, could be consulted. Approved entries could reach translators through computer-assisted translation tools and an application programming interface, within the copyright, license, and consent conditions attached to learner and institutional data. AI-assisted functions could propose term candidates and flag inconsistent usage, with expert review, versioned entries, and recorded provenance. A limited pilot that compares a prototype with existing resources and with federated search, under criteria set in advance, would have to precede any national step. Possible measures include term accuracy, retrieval success, task completion time, consistency, the rate of expert disagreement, and uptake by users. Supplementary File 4 gives the extended proposal.

Limitations

One reviewer screened and charted the sources with an AI-assisted bibliography assistant, and no second reviewer checked the decisions, so no agreement statistic can be reported. The two rounds of database searching differed in procedure, the later searches added 68 sources (55 of them from Scopus), and venues outside both databases may still be missing. ScienceDirect was not searched separately, institution names were not used as search terms, only one journal was hand-searched, the result lists of eight web queries were not retained, and the keyword filter removed 333 records before human screening, although the audit found no eligible record among them. The English-language restriction excludes Arabic-language research, which bears directly on Arabization and terminology. The Saudi connection rests mostly on author affiliation, and the concept boundaries required judgment at the margins, for example for studies that used machine translation only to produce data for another task. Charting relied mainly on records and abstracts, the findings of the 68 October sources were not charted, and no formal appraisal was carried out.

Conclusion

The 139 sources, 71 of them with charted findings, describe task-specific corpora and datasets, systems tested on separate benchmarks, a growing number of small evaluations of machine translation and large language models, and terminology problems in legal, computing, and medical domains. They rarely report access, maintenance, or how translators use term resources, so they cannot show duplication, the uptake of national term banks, or the adequacy of coordination. An inventory of resources and surveys of translators and terminologists would supply that evidence, and a design study that compares a prototype platform with improved term banks and federated search would test the proposal. Updates of this review should chart the remaining findings, add Arabic-language sources, and use dual screening.

Acknowledgements

An earlier version of this work was presented at the First Conference of the Saudi Translation Association (SATA) on 1 October 2026.

Funding Statement

This research received no external funding.

Conflict of Interest Statement

The author declares no conflict of interest.

Data Availability Statement

The data supporting the findings of this study, including the extraction workbook, the Web of Science export, and the search log, are available from the corresponding author upon reasonable request.

AI Disclosure

During the preparation of this paper, the author used Grammarly (Grammarly Inc.) for proofreading and formatting, and Claude (Anthropic) to locate and format bibliographic references. Each reference was checked against its original source, and all tool output was reviewed and edited by the author. All data were collected by the author from publicly available sources. The analytical claims and theoretical arguments are the author's own work, and the author takes full responsibility for the accuracy and originality of the content.

Bio

Kayed Alsultan is affiliated with Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, and is a PhD student at King Saud University. His research interests are forensic phonetics, forensic pragmatics, corpus linguistics, and syntactic theory.

References

Aboraya, A., Elnour, O. A., Ghoneim, A., Sakr, M., Nayel, I., Ghossoub, E., Khoury, R., Mansour, H., Ibrahim, I., Naga, H. A., & Rizvi, A. (2026). A comparative analysis of translation and back-translation vs. machine translation with post-editing. *Annals of Clinical Psychiatry, 37*(1), 35–40. https://doi.org/10.1177/10401237261417729

Abouammoh, M., Shah, K., & Aker, A. (2016). Creation of comparable corpora for English-Urdu, Arabic, Persian. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 4193–4196). European Language Resources Association. https://aclanthology.org/L16-1663/

Alahmari, S., Atwell, E., Alsalka, M., & Saadany, H. (2025). Evaluating the performance of LLMs when translating Saudi Arabic as low resource language. In M. Bramer & F. Stahl (Eds.), *Artificial intelligence XLI: SGAI 2024* (Lecture Notes in Artificial Intelligence, Vol. 15447, pp. 264–269). Springer. https://doi.org/10.1007/978-3-031-77918-3_22

Alamri, F. (2026). Improve Arabic-English translation using Arabic Language Attention Transformer. *Data Technologies and Applications, 60*(1), 192–217. https://doi.org/10.1108/DTA-02-2025-0110

Alanazi, N., Al-Batineh, M., & Abu-Rayyash, H. (2025). SauDial: The Saudi Arabic dialects game localization dataset. *Data in Brief, 62*, Article 111906. https://doi.org/10.1016/j.dib.2025.111906

Alangari, E. (2024). Translation and interpreting research in Saudi Arabia: A bibliometric analysis (1990–2019). *The Translator, 30*(2), 210–226. https://doi.org/10.1080/13556509.2023.2200054

Al-Athwary, A. A. H., & Ali, H. K. (2024). Arabicization via loan translation: A corpus-based analysis of neologisms translated from English into Arabic in the field of information technology. *Open Cultural Studies, 8*(1), Article 20240005. https://doi.org/10.1515/culture-2024-0005

Aldawsari, H. A. H. (2022). The applicability of Munday's systemic model for descriptive translation studies on human and machine Arabic translations of an English text. *Journal of Research in Language & Translation, 2*(1), 25–43. https://doi.org/10.33948/JRLT-KSU-2-1-2

Aldawsari, H. A. H. (2024). Evaluating translation tools: Google Translate, Bing Translator, and Bing AI on Arabic colloquialisms. *Arab World English Journal* (Special Issue on ChatGPT), 237–251. https://doi.org/10.24093/awej/ChatGPT.16

Aldosari, L. A., & Altuwairesh, N. (2025). Assessing the effects of translation prompts on the translation quality of GPT-4 Turbo using automated and human evaluation metrics: A case study. *Perspectives: Studies in Translation Theory and Practice*. Advance online publication. https://doi.org/10.1080/0907676X.2025.2464120

Aleedy, M., Alshihri, F., Meshoul, S., Al-Harthi, M., Alramlawi, S., Aldaihani, B., Shaiba, H., & Atwell, E. (2025). Designing AI-powered translation education tools: A

framework for parallel sentence generation using SauLTC and LLMs. *PeerJ Computer Science, 11*, Article e2788. https://doi.org/10.7717/peerj-cs.2788

Alfuraih, R. F. (2020). The undergraduate learner translator corpus: A new resource for translation studies and computational linguistics. *Language Resources and Evaluation, 54*(3), 801–830. https://doi.org/10.1007/s10579-019-09472-6

Algaraady, J., & Mahyoob, M. (2025). Exploring ChatGPT's potential for augmenting post-editing in machine translation across multiple domains: Challenges and opportunities. *Frontiers in Artificial Intelligence, 8*, Article 1526293. https://doi.org/10.3389/frai.2025.1526293

Alghamdi, E. A., Zakraoui, J., & Abanmy, F. A. (2024). Domain adaptation for Arabic machine translation: Financial texts as a case study. *Applied Sciences, 14*(16), Article 7088. https://doi.org/10.3390/app14167088

AlGhamedi, N. (2024). Constraints to neural machine translation quality, human and automated evaluation, and quality improvement across language pairs: A systematic literature review. *Journal of Research in Language & Translation*, Special Issue on AI and Translation Quality, 38–79. https://doi.org/10.33948/JRLT-KSU-S-1-4

Al-Harthi, M., Alsaif, A., Al-Nafjan, E., Alshihri, F., & Saleh, M. (2024). Saudi Learner Translation Corpus: The design and compilation of an English-Arabic learner translation corpus. *PLOS ONE, 19*(10), Article e0303729. https://doi.org/10.1371/journal.pone.0303729

Alhomoud, G. N. J. (2022). On translating and drafting Sharia legal terms and expressions in Saudi contracts. *Education Research International, 2022*, Article 7167970. https://doi.org/10.1155/2022/7167970

Alhumsi, M. H. (2025). A comparative transitivity analysis of human vs. artificial intelligence-based translations: An evaluation of abstracts from humanities and social science studies. *Journal of Research in Language & Translation, 5*(2), 164–192. https://doi.org/10.33948/JRLT-KSU-5-2-7

Ali, M. A. (2016). Artificial intelligence and natural language processing: The Arabic corpora in online translation software. *International Journal of Advanced and Applied Sciences, 3*(9), 59–66. https://doi.org/10.21833/ijaas.2016.09.010

Alkhatnai, M. (2025). The role of artificial intelligence tools in mediating Sino-Arab cultural exchanges through intercultural translation. *Babel, 71*(6), 740–769. https://doi.org/10.1075/babel.25101.alk

Almaaytah, S. A., & Alzobidy, S. A. (2023). Challenges in rendering Arabic text to English using machine translation: A systematic literature review. *IEEE Access, 11*, 94772–94779. https://doi.org/10.1109/ACCESS.2023.3309642

Almarzoqi, A., & Alsuhaibani, M. (2025). Leveraging large language models for detecting and preserving emotions in Quran translations. *Journal of King Saud University Computer and Information Sciences, 37*(9), Article 271. https://doi.org/10.1007/s44443-025-00269-y

Alotaibi, H. M. (2017). Arabic-English parallel corpus: A new resource for translation training and language teaching. *Arab World English Journal, 8*(3), 319–337. https://doi.org/10.24093/awej/vol8no3.21

Alrajeh, A. (2018). *A recipe for Arabic-English neural machine translation* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1808.06116

Alrashidi, F., & Mathkour, H. I. (2026). An empirical study of transformer-based neural machine translation for English to Arabic. *Information, 17*(2), Article 198. https://doi.org/10.3390/info17020198

Al-Rukban, A., & Saudagar, A. K. J. (2017). Evaluation of English to Arabic machine translation systems using BLEU and GTM. In *Proceedings of the 9th International Conference on Education Technology and Computers (ICETC 2017)* (pp. 228–232). Association for Computing Machinery. https://doi.org/10.1145/3175536.3175570

Alshehri, B. A., & Alowedi, N. A. (2023). Saudi translation agencies and translation centers: A study of post-editing practices. *World Journal of English Language, 13*(3), 118. https://doi.org/10.5430/wjel.v13n3p118

Alsudais, A. (2019). Image classification in Arabic: Exploring direct English to Arabic translations. *IEEE Access, 7*, 122730–122739. https://doi.org/10.1109/ACCESS.2019.2926924

Altammami, S., Atwell, E., & Alsalka, A. (2020). Constructing a bilingual Hadith corpus using a segmentation tool. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 3390–3398). European Language Resources Association. https://aclanthology.org/2020.lrec-1.415/

Althobaiti, M. J. (2022). A simple yet robust algorithm for automatic extraction of parallel sentences: A case study on Arabic-English Wikipedia articles. *IEEE Access, 10*, 401–420. https://doi.org/10.1109/ACCESS.2021.3137830

Alwazna, R. Y. (2019). Translation and legal terminology: Techniques for coping with the untranslatability of legal terms between Arabic and English. *International Journal for the Semiotics of Law, 32*(1), 75–94. https://doi.org/10.1007/s11196-018-9580-y

Alwazna, R. Y. (2024). The use of automation in the rendition of certain articles of the Saudi Commercial Law into English: A post-editing-based comparison of five machine translation systems. *Frontiers in Artificial Intelligence, 6*, Article 1282020. https://doi.org/10.3389/frai.2023.1282020

Alyafeai, Z., Alshaibani, M. S., AlKhamissi, B., Luqman, H., Alareqi, E., & Fadel, A. (2023). *Taqyim: Evaluating Arabic NLP tasks using ChatGPT models* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2306.16322

Alyami, N., & Qassem, M. (2024). Examining translation and interpretation studies in GCC states: Data from Web of Science (1991–2023). *Cogent Arts & Humanities, 11*(1), Article 2421083. https://doi.org/10.1080/23311983.2024.2421083

Arrabai, A. M. (2022). A corpus-based study on the use of glossing in Saudi technology blogs. *Arab World English Journal, 13*(4), 562–573. https://doi.org/10.24093/awej/vol13no4.36

Assiri, A. A. (2026). A comparative linguistic analysis of human and AI medical term translations into Arabic. *Ijaz Arabi: Journal of Arabic Learning, 9*(1), 11–28. https://doi.org/10.18860/ijazarabi.V9i1.32531

Bari, M. S., Alnumay, Y., Alzahrani, N., Alotaibi, N., Alyahya, H., AlRashed, S., Mirza, F., Alsubaie, S., Alahmed, H., Alabduljabbar, G., Alkhathran, R., Almushayqih, Y., Alnajim, R., Alsubaihi, S. I., Al Mansour, M., Hassan, S., Alrubaian, M., Alammari, A., Alawami, Z., . . . Khan, H. (2025). ALLaM: Large language models for Arabic and English. In *Proceedings of the International Conference on Learning Representations (ICLR 2025)* (pp. 34179–34214). https://proceedings.iclr.cc/paper_files/paper/2025/hash/54c15a3033686e7999aecd2740c5a7c4-Abstract-Conference.html

Boulaares, D. M. (2026). Enhancing Arabic in scientific research: Saudi Arabia's case for linguistic infrastructure and cognitive sovereignty. *Language Problems and Language Planning*. Advance online publication. https://doi.org/10.1075/lplp.25028.bou

Bounhas, I., & Ben Guirat, S. (2019). KUNUZ: A multi-purpose reusable test collection for classical Arabic document engineering. In *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)* (pp. 1–8). IEEE. https://doi.org/10.1109/AICCSA47632.2019.9035212

Ehab, R., El-Sawy, A., Aldawsari, M., & Nayel, H. (2026). DEAST: A dataset for English-Arabic scientific translation and vice versa. *Data in Brief, 64*, Article 112381. https://doi.org/10.1016/j.dib.2025.112381

Elgobshawi, A. E. (2022). Translation and Arabization of computer terminology: A study of learners' preferences and attitudes. *Eurasian Journal of Applied Linguistics, 8*(1), 288–297. https://ejal.info/menuscript/index.php/ejal/article/view/196

El-Kassas, M., Al-Naamani, K., Elbadry, M., Awad, A., Tharwat, M., Debzi, N., Zemmouchi, S., Abdulla, M., Zakaria, D., Esmat, G., El-Karaksy, H., Waked, I., Shaltout, I., Medhat, M. A., El-Shabrawi, M., Abdeen, N., Al-khairalla, M., Akroush, M. W., Alali, A. A., . . . Alswat, K. (2025). Establishing consensus on Arabic medical terminology for steatotic liver disease: A mixed-methods approach. *Arab Journal of Gastroenterology, 26*(2), 143–148. https://doi.org/10.1016/j.ajg.2025.01.009

El-Zeiny, I. T. (2017). Lexicographic approaches to bridging the lexical gap in translating Islamic juristic terms/concepts into English. *International Journal of Lexicography, 30*(2), 187–224. https://doi.org/10.1093/ijl/ecv035

Fakih, A., Ghassemiazghandi, M., Fakih, A.-H., & Singh, M. K. M. (2024). Evaluation of Instagram's neural machine translation for literary texts: An MQM-based analysis. *GEMA Online Journal of Language Studies, 24*(1), 213–233. https://doi.org/10.17576/gema-2024-2401-13

Huang, B., & Asiri, A. (2026). AlignAR: Generative sentence alignment for Arabic-English parallel corpora of legal and literary texts. In H. Al-Khalifa, M. El-Haj, & S. Ezzini (Eds.), *The 7th Workshop on Open-Source Arabic Corpora and Processing Tools*

(OSACT7) with 5 Shared Tasks (pp. 59–65). Association for Computational Linguistics. https://doi.org/10.63317/2c9daup5j4k7

Jehjah, A.-J. B. A. (2025). The quality of translated legal terms regarding children in the MENA region: A comparative study. *International Journal for the Semiotics of Law, 38*(6), 2111–2135. https://doi.org/10.1007/s11196-024-10243-y

King Abdullah Institute for Translation and Arabization. (n.d.). *About*. Imam Mohammad Ibn Saud Islamic University. Retrieved October 20, 2026, from https://units.imamu.edu.sa/colleges/en/kaiftaa/profile/Pages/default.aspx

King Salman Global Academy for Arabic Language. (n.d.). *About*. Retrieved October 20, 2026, from https://ksaa.gov.sa/en/about-us

Law of Ethics of Research on Living Creatures, Royal Decree No. M/59 (2010) (Saudi Arabia).

Mahmoud, H. A. H., & Mengash, H. A. (2021). Machine translation utilizing the frequent-item set concept. *Sensors, 21*(4), Article 1493. https://doi.org/10.3390/s21041493

Mohsen, M. A. (2024). Artificial intelligence in academic translation: A comparative study of large language models and Google Translate. *Psycholinguistics, 35*(2), 134–156. https://doi.org/10.31470/2309-1797-2024-35-2-134-156

Muftah, M. (2026). Translating information technology expressions from English to Arabic: Strategies for bridging linguistic gaps. *Cadernos de Tradução, 46*, Article e104242. https://doi.org/10.5007/2175-7968.2026.e104242

Obeidat, M. M., Haider, A. S., Abu Tair, S., & Sahari, Y. (2024). Analyzing the performance of Gemini, ChatGPT, and Google Translate in rendering English idioms into Arabic. *FWU Journal of Social Sciences, 18*(4), 1–18. https://doi.org/10.51709/19951272/Winter2024/1

Omar, A., & Gomaa, Y. A. (2020). The machine translation of literature: Implications for translation pedagogy. *International Journal of Emerging Technologies in Learning, 15*(11), 228–235. https://doi.org/10.3991/ijet.v15i11.13275

Omar, L. I., & Salih, A. A. (2024). Systematic review of English/Arabic machine translation postediting: Implications for AI application in translation research and pedagogy. *Informatics, 11*(2), Article 23. https://doi.org/10.3390/informatics11020023

OPEC Fund for International Development. (n.d.). *ARABTERM: The Arabic technical terminology dictionary*. Retrieved October 19, 2026, from https://opecfund.org/where-we-work/list/arabterm-the-arabic-technical-terminology-dictionary

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pollock, D., Peters, M. D. J., Tricco, A. C., Munn, Z., Jia, R. M., Alexander, L., Pieper, D., Evans, C., Godfrey, C. M., Brandão de Moraes, E., Saran, A., Campbell, F., & Khalil, H. (2026). Scoping reviews. In E. Aromataris, C. Lockwood, K. Porritt, B. Pilla, & Z. Jordan (Eds.), *JBI manual for evidence synthesis*. JBI. https://doi.org/10.46658/JBIMES-24-09

Qassem, M. (2026). *Analysis of English-Arabic translations of medical course titles in health college study plans at Saudi universities* [Preprint]. Research Square. https://doi.org/10.21203/rs.3.rs-9646127/v1

Qassem, M., & Aldaheri, M. M. (2023). Can machine translate dialogue acts: Evidence from translating dialogues from English to Arabic. *3L: Language, Linguistics, Literature, 29*(4), 63–81. https://doi.org/10.17576/3L-2023-2904-05

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, Article 39. https://doi.org/10.1186/s13643-020-01542-z

Sabtan, Y. M. N., Hussein, M. S. M., Ethelb, H., & Omar, A. (2021). An evaluation of the accuracy of the machine translation systems of social media language: Google's Arabic-English translation as an example. *International Journal of Advanced Computer Science and Applications, 12*(7), 406–415. https://doi.org/10.14569/IJACSA.2021.0120746

Samman, H. M. (2022). *Evaluating machine translation post-editing training in undergraduate translation programs: An exploratory study in Saudi Arabia* [Doctoral dissertation, University of Southampton]. University of Southampton Institutional Repository. https://eprints.soton.ac.uk/469163/

Sarsak, H. I. (2024). Developing the first digital occupational therapy dictionary worldwide: A promising mobile application model for occupational therapy service providers and recipients. *Inquiry: The Journal of Health Care Organization, Provision, and Financing, 61*, Article 00469580241282060. https://doi.org/10.1177/00469580241282060

Saudipedia. (n.d.-a). *Arabic Observatory of Translation (AOT)*. Retrieved October 20, 2026, from https://saudipedia.com/en/arabic-observatory-of-translation-aot

Saudipedia. (n.d.-b). *King Salman Global Academy for Arabic Language (KSAA)*. Retrieved October 19, 2026, from https://saudipedia.com/en/king-salman-global-academy-for-arabic-language-ksaa

Saudipedia. (n.d.-c). *The Literature, Publishing and Translation Commission*. Retrieved October 19, 2026, from https://saudipedia.com/en/article/1635/government-and-politics/commissions/the-literature-publishing-and-translation-commission

Saudipedia. (n.d.-d). *Saudi Data and Artificial Intelligence Authority (SDAIA)*. Retrieved October 19, 2026, from https://saudipedia.com/en/saudi-data-and-artificial-intelligence-authority-sdaia

Saudipedia. (n.d.-e). *Saudi Vision 2030*. Retrieved October 19, 2026, from https://saudipedia.com/en/saudi-vision-2030

Saudipedia. (n.d.-f). *Tarjim initiative*. Retrieved October 20, 2026, from https://saudipedia.com/en/tarjim-initiative

Sieny, M. E. (1989). Machine translation in Saudi Arabia. In *Proceedings of Machine Translation Summit II*. https://aclanthology.org/1989.mtsummit-1.21/

Tiedemann, J. (2012). Parallel data, tools and interfaces in OPUS. In N. Calzolari, K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 2214–2218). European Language Resources Association. https://aclanthology.org/L12-1246/

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Ziemski, M., Junczys-Dowmunt, M., & Pouliquen, B. (2016). The United Nations Parallel Corpus v1.0. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 3530–3534). European Language Resources Association. https://aclanthology.org/L16-1561/