
| RESEARCH ARTICLE

Generative Artificial Intelligence in Language and Literary Studies: A Systematic Review of Applications, Ethical Challenges, and Future Directions

John Mark Terry

Independent Researcher, United Kingdom

Corresponding Author: John Mark Terry, **E-mail:** Johnmarj2018@gmail.com

| ABSTRACT

Generative artificial intelligence (GenAI) is transforming language and literary studies by introducing new approaches to language learning, writing, translation, textual interpretation, literary analysis, and creative production. Powered by large language models and other generative technologies, these systems can generate and revise texts, summarize complex literary works, identify linguistic patterns, support multilingual communication, and facilitate interactive learning. However, their growing adoption raises important concerns regarding authorship, academic integrity, linguistic bias, cultural representation, interpretive reliability, data privacy, and the potential marginalization of less-resourced languages and literary traditions. This systematic review examines the applications, ethical challenges, and future directions of GenAI in language and literary studies. It adopts a structured literature-review approach to synthesize research on educational applications, computational linguistics, translation, literary interpretation, creative writing, and digital humanities. The review organizes the literature around five major themes: language learning and teaching, academic and creative writing, translation and multilingual communication, literary analysis and interpretation, and ethical and institutional governance. The synthesis suggests that GenAI can expand access to language-learning support, facilitate text-based research, improve writing processes, and support new forms of literary experimentation. Nevertheless, its effectiveness depends on human oversight, contextual accuracy, critical evaluation, and sensitivity to linguistic and cultural diversity. The review further identifies continuing challenges associated with fabricated information, algorithmic bias, unclear intellectual-property arrangements, overreliance on automated assistance, and inadequate representation of minority languages. It argues that GenAI should complement rather than replace human linguistic judgment, literary expertise, and creative agency. Future research should prioritize transparent evaluation frameworks, culturally responsive models, multilingual equity, responsible pedagogical integration, and interdisciplinary collaboration. The review contributes an integrated perspective for researchers, educators, writers, and institutions seeking to use generative technologies responsibly in language and literary scholarship.

| KEYWORDS

Generative artificial intelligence; language studies; literary studies; large language models; machine translation; academic integrity; digital humanities.

| ARTICLE INFORMATION

ACCEPTED: September 13, 2026 **PUBLISHED:** October 19, 2026 **DOI:** <https://doi.org/10.61424/jlls.v4i4.1101>

1. Introduction

Generative artificial intelligence (GenAI) has emerged as an influential technological development with significant implications for the humanities, particularly language and literary studies. Unlike traditional computational systems

Copyright: © 2026 The Author(s). This open access article is distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC-BY-NC-SA) License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>), published by Bluemark Publishers.

that primarily classify, retrieve, or analyze existing information, generative systems can produce new textual content, translate passages, summarize documents, simulate dialogue, and respond to prompts in natural language (Sengar, 2025). The development of large language models (LLMs), including conversational AI systems, has made these capabilities increasingly accessible to students, researchers, teachers, translators, editors, and creative writers. Consequently, GenAI is changing not only how language is studied and taught but also how literary texts are interpreted, produced, disseminated, and evaluated.

Language studies encompass linguistics, applied linguistics, sociolinguistics, discourse analysis, pragmatics, language acquisition, translation, and related disciplines concerned with the structure and use of human language. Literary studies, meanwhile, investigate literary texts through historical, cultural, aesthetic, philosophical, and theoretical perspectives. Although these fields have traditionally employed different analytical approaches, they share an interest in language as a medium of meaning, communication, identity, and cultural expression (Surbakti, 2025). GenAI creates opportunities for greater interaction between these disciplines by combining language generation with textual analysis, corpus-based investigation, translation, and computational interpretation.

One of the most prominent applications of GenAI is language education. Conversational systems can provide learners with opportunities to practice writing, reading, vocabulary, grammar, and dialogue in relatively flexible learning environments. They can generate explanations at different levels of difficulty, create examples of grammatical structures, simulate communicative situations, and offer immediate feedback on written work. These functions may be particularly useful in contexts where learners have limited access to individualized instruction or opportunities for authentic language practice (Law, 2024). However, the educational value of AI-generated feedback depends on its accuracy, suitability for learners' proficiency levels, and alignment with pedagogical objectives. Fluent responses should not automatically be interpreted as linguistically or pedagogically reliable.

GenAI also offers new possibilities for linguistic research. Researchers can use language models to explore discourse patterns, formulate hypotheses, generate preliminary annotations, compare linguistic constructions, and assist with the organization of qualitative data. In pragmatics and discourse analysis, these systems may help researchers examine conversational strategies, politeness, argumentation, and the rhetorical organization of texts. In applied linguistics, they can support the development of learning materials, writing assessment, and language-learning activities (Al-kfairy, 2024). Nevertheless, model-generated linguistic judgments require validation against authentic language data because AI systems may reproduce dominant language conventions while inadequately representing regional varieties, multilingual practices, and socially situated language use.

In literary studies, GenAI introduces both methodological opportunities and fundamental interpretive questions. Researchers can use AI-assisted tools to summarize lengthy works, compare characters and themes, identify recurring motifs, examine narrative structures, and generate preliminary interpretations for further investigation. Such applications may support exploratory analysis across large textual collections, particularly within digital humanities projects (Tirpan, 2024). Generative models can also assist writers in brainstorming narrative ideas, experimenting with genres, exploring alternative perspectives, and revising stylistic choices. These possibilities encourage new forms of collaboration between human creativity and computational systems.

However, literary interpretation involves more than identifying recurring words, themes, or narrative patterns. Literary works are embedded in historical circumstances, cultural traditions, ideological conflicts, aesthetic conventions, and authorial choices. An AI-generated interpretation may therefore be linguistically persuasive while overlooking irony, ambiguity, historical allusion, intertextuality, or culturally specific meanings. Similarly, a model may produce a plausible account of a literary work while inventing quotations, misattributing ideas, or confusing fictional details (Raitskaya, 2025). Human expertise remains essential for evaluating interpretations and distinguishing textual evidence from generated speculation.

The integration of GenAI into language and literary studies also raises important ethical questions. Academic institutions must determine how AI-assisted work should be disclosed, assessed, and credited. Researchers and publishers must consider the ownership and permitted use of AI-generated material, the provenance of training data, and the responsibilities associated with machine-assisted scholarly production. In educational settings, unrestricted dependence on generated answers may reduce opportunities for independent interpretation, sustained reading, and the development of original writing skills (Arista, 2025). At the same time, excessively restrictive policies may prevent legitimate uses of AI for accessibility, language support, and research assistance. Responsible integration therefore requires policies that distinguish between supportive assistance and the misrepresentation of intellectual work.

Another important concern is linguistic and cultural inequality. Many generative models perform more consistently in widely represented languages than in languages for which digital training data are limited. This disparity can affect translation quality, grammatical explanations, literary interpretation, and the representation of cultural knowledge. Minority languages, oral traditions, regional dialects, and locally produced literary works may be inadequately represented or misinterpreted (Esteves, 2024). Consequently, the expansion of GenAI should not be evaluated solely in terms of technical performance or productivity; it should also be assessed in relation to linguistic diversity, cultural preservation, equitable access, and the representation of different knowledge traditions.

Although research on AI in education, natural language processing, computational linguistics, and digital humanities has developed considerably, these areas are frequently examined separately. Language-learning studies may emphasize instructional effectiveness, while computational research prioritizes model performance, and literary scholarship focuses on interpretation, authorship, and aesthetics. Ethical discussions cut across all these areas but are not always integrated into analyses of practical applications (Zhang, 2024). A consolidated review can help establish connections among these research strands and clarify the distinctive implications of GenAI for language and literary scholarship.

Against this background, the present review examines the applications, ethical challenges, and future directions of GenAI in language and literary studies. It is guided by three objectives: to identify and synthesize major applications of GenAI across language education, linguistic research, translation, writing, and literary analysis; to critically examine ethical, methodological, and sociocultural challenges associated with its use; and to identify research priorities for responsible, inclusive, and academically rigorous integration (Tata, 2025). By connecting technological capabilities with the interpretive and cultural concerns of the humanities, the review seeks to provide a conceptual foundation for researchers, educators, students, translators, writers, and academic institutions navigating the expanding relationship between generative technologies and humanistic inquiry.

2. Methodology

2.1. Research Design

This study adopts a systematic literature-review approach to examine the applications, ethical challenges, and future directions of generative artificial intelligence in language and literary studies. A systematic approach is appropriate because research on GenAI is expanding across several interconnected disciplines, including applied linguistics, language education, computational linguistics, translation studies, literary criticism, creative writing, and digital humanities. The review aims to synthesize the available evidence, identify recurring themes, compare different scholarly perspectives, and establish research gaps that can inform future investigation. The methodological approach is informed by established systematic-review principles, including transparent search procedures, explicit eligibility criteria, structured evidence extraction, and thematic synthesis. Previous reviews of GenAI in language education have demonstrated the value of systematically organizing research to identify applications, methodological limitations, and unresolved questions (Application of Generative Artificial Intelligence in Language Teaching and Learning, 2024; A Systematic Review of the First Year of Publications on ChatGPT and Language Education, 2024).

2.2. Literature Search Strategy

The literature search framework encompasses publications addressing generative AI, large language models, and conversational AI in language-related and literary contexts. Relevant scholarly databases and indexing platforms for implementing the search include Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar. These sources cover education, linguistics, computer science, interdisciplinary humanities, and digital research. Searches encompass literature published from January 2017 to October 2026, with particular attention to the period following the public emergence of widely accessible generative AI systems in late 2022. Earlier publications are relevant where they establish theoretical foundations for computational linguistics, machine translation, automated writing support, and digital literary analysis.

The search strategy combines keywords related to generative technologies with terms describing language and literary applications. Representative search expressions include “generative artificial intelligence” AND “language studies”, “ChatGPT” AND “language learning”, “large language models” AND linguistics, “generative AI” AND literary analysis, “AI-assisted creative writing”, “large language models” AND translation, and “generative AI” AND “digital humanities”. Additional searches address ethical concerns using terms such as academic integrity, authorship, linguistic bias, cultural representation, copyright, privacy, and AI literacy. Search expressions should be adapted to the syntax and indexing practices of individual databases, and reference lists of relevant review articles should be examined to identify additional eligible studies.

2.3. Inclusion and Exclusion Criteria

The review focuses on scholarly publications that make a substantive contribution to understanding the role of GenAI in language or literary studies. Eligible materials include empirical research articles, systematic and scoping reviews, theoretical papers, methodological studies, and relevant institutional guidance. Studies are considered relevant when they examine language learning and teaching, linguistic analysis, writing and revision, translation, literary interpretation, creative writing, textual generation, or ethical and cultural implications in these areas. Publications discussing AI governance and academic integrity are included when their findings have direct implications for language education, literary scholarship, or humanistic research.

Studies were excluded when they focus exclusively on unrelated technical applications, employ traditional predictive AI without a meaningful connection to generative systems, or mention language and literature only incidentally. Duplicate publications, inaccessible records that cannot be assessed adequately, non-substantive commentary, and sources without sufficient methodological or conceptual detail should also be excluded from the formal synthesis. Where non-peer-reviewed research is included because of its relevance to rapidly evolving technologies, its publication status should be recorded and its claims evaluated cautiously. Institutional reports may be used to contextualize ethical and policy considerations but should be distinguished from peer-reviewed empirical evidence.

2.4. Screening and Selection of Studies

Study selection proceeded through identification, deduplication, title and abstract screening, and full-text eligibility assessment. Initially retrieved records should be screened against the review objectives, followed by an examination of potentially relevant full texts. Reasons for exclusion at the full-text stage should be documented to ensure transparency. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram can be used to report the number of records identified, duplicates removed, records screened, full texts assessed, and studies retained. The PRISMA approach supports transparent reporting of study selection, but the flow diagram must reflect the actual search and screening results rather than estimated or illustrative counts.

The available evidence for this manuscript does not include a completed database export or verified screening log. Accordingly, exact record counts, the final number of included studies, and a claim of exhaustive database coverage cannot be reported reliably here. Before the manuscript is submitted as a completed systematic review, these details should be established through execution and documentation of the search protocol. This distinction is important

because methodological transparency is necessary to distinguish a completed systematic review from a structured thematic overview.

2.5. Data Extraction and Thematic Analysis

Information extracted from eligible studies included author and publication year, research context, study design, disciplinary focus, AI technology examined, principal applications, reported benefits, limitations, and ethical implications. Where empirical evidence is available, the extraction process should also record participant characteristics, data sources, evaluation methods, and the principal outcomes reported by the authors. This structured approach allows findings from different disciplines to be compared while preserving the differences between experimental results, qualitative observations, conceptual arguments, and policy recommendations.

The extracted evidence is organized thematically to address the review objectives. Five overarching analytical categories guide the synthesis: GenAI in language learning and teaching; writing, translation, and multilingual communication; linguistic research and discourse analysis; literary interpretation, creative writing, and digital humanities; and ethical challenges and future research directions. Within these categories, recurring patterns are compared, disagreements are examined, and methodological limitations are identified. The synthesis emphasizes the distinction between demonstrated outcomes and proposed applications, avoiding the assumption that the availability of a technological capability necessarily establishes its effectiveness.

3. Findings and Discussion

3.1. Applications of Generative AI in Language Studies

3.1.1. Language Learning and Teaching

One of the most extensively investigated areas of GenAI application is language learning and teaching. Conversational AI systems can provide interactive explanations, simulated dialogues, vocabulary exercises, grammar examples, reading-comprehension questions, and writing feedback. Their conversational format enables learners to ask follow-up questions, request alternative explanations, and practice language production in different communicative situations. For example, an English-language learner may ask a system to explain the difference between formal and informal expressions, generate dialogues for a professional setting, or provide feedback on a paragraph (Matos, 2025). A learner studying French or Spanish may use similar interactions to practice sentence construction, explore vocabulary in context, and examine differences between grammatical structures.

These applications may complement conventional language instruction by offering additional opportunities for practice outside classroom hours. They may also help teachers prepare differentiated learning materials for students with different proficiency levels. For instance, a teacher can request several versions of a reading passage, develop questions that progress from literal comprehension to critical interpretation, or construct exercises focused on particular grammatical difficulties (Samala, 2025). However, generated materials require review because a model may introduce unnatural expressions, incorrect explanations, or examples that do not reflect the intended language variety.

Evidence from existing reviews supports the view that GenAI has multiple potential roles in language education. A scoping review published in *Computers and Education Open* in 2024 examined 41 publications and identified applications and challenges across language teaching and learning, while a separate systematic review published in *Computers and Education: Artificial Intelligence* synthesized 36 studies on ChatGPT and language education (Korchak, 2025). These reviews highlighted applications in self-directed learning, content generation, and teacher workflows, alongside the need for stronger empirical evidence and further evaluation of feedback quality.

The significance of these findings lies in the distinction between accessibility and learning effectiveness. Immediate responses can increase opportunities for practice, but they do not automatically produce lasting improvements in proficiency. Learners may accept incorrect feedback or develop a habit of requesting complete answers rather than working through linguistic problems independently. Effective integration therefore requires tasks that encourage learners to explain their decisions, revise their own work, verify AI suggestions, and reflect on the differences between their original responses and the system's output (Batista, 2024). Teachers remain important in establishing

learning objectives, evaluating progress, and providing the contextual guidance that automated systems cannot consistently guarantee.

3.1.2. Writing Development and Academic Communication

GenAI has become an important source of support for writing development and academic communication. Its applications include brainstorming, outlining, paragraph organization, grammar correction, vocabulary selection, stylistic revision, summarization, and feedback on coherence. In academic language instruction, these capabilities can help learners recognize recurring writing difficulties and experiment with alternative ways of expressing an argument (Bahroun, 2023). For example, a student may use AI feedback to identify unclear transitions, examine whether a paragraph supports its topic sentence, or compare a formal academic formulation with a less formal version.

AI-assisted writing may also be useful for learners writing in an additional language. Such learners may understand a subject well but experience difficulty expressing complex arguments in the conventions of academic English or another target language. A generative system can suggest revisions while explaining grammatical choices and differences in register. This can make revision more interactive than a process limited to the identification of errors by a teacher (Preiksaitis, 2023).

Nevertheless, writing quality involves more than grammatical correctness. Strong academic writing requires argument development, evidence evaluation, disciplinary awareness, appropriate citation, and an identifiable line of reasoning. If students delegate these intellectual activities to a generative system, the resulting text may be polished without demonstrating their own understanding. Similarly, AI-generated summaries may omit qualifications or distort the relationship between an author's claims and supporting evidence.

A productive pedagogical approach treats AI-generated feedback as a proposal for evaluation rather than an authoritative correction. Students can be asked to retain an initial draft, document which suggestions they accepted or rejected, and explain their revisions. Teachers can assess both the final product and the reasoning that informed its development. This approach allows GenAI to support writing improvement while preserving the importance of independent thought, drafting, and revision (Rabbani, 2025). It also encourages the development of AI literacy as part of broader academic communication skills.

3.1.3. Linguistic Analysis and Discourse Research

In linguistics, GenAI can assist with exploratory analysis of language use, discourse organization, pragmatic meaning, and rhetorical structure. Researchers may use language models to generate preliminary annotations, classify text segments, identify possible themes, compare registers, and formulate hypotheses for subsequent testing. In discourse analysis, for instance, AI-assisted coding may help organize a large collection of interview transcripts according to tentative categories such as persuasion, politeness, disagreement, or expressions of uncertainty (Wang, 2025). In sociolinguistics, models can be used to generate questions about language variation or to explore how particular expressions change across contexts.

However, these uses must be distinguished from validated linguistic measurement. A model's classification of a sentence as polite, persuasive, or offensive may depend on its training data and the assumptions embedded in its response patterns. Pragmatic meaning is often context-dependent, and the same expression may communicate different intentions across relationships, communities, and situations (Ning, 2024). AI-generated annotations can therefore introduce systematic errors if researchers accept them without examining the original data.

Another concern is the difference between generated language and naturally occurring language. A model can produce convincing examples of a dialect or conversational style without accurately representing how speakers use that variety in everyday life. Such examples are not substitutes for authentic corpora, participant observation, interviews, or ethically collected language data. In studies of multilingual communication, code-switching, and

regional language practices, researchers must pay particular attention to whether model-generated explanations reflect the experiences of the communities being studied.

The most defensible role for GenAI in linguistic research is therefore that of a research assistant whose outputs require verification. Researchers can compare AI-generated annotations with human coding, measure agreement, examine disagreements, and document the model and prompts used. Such practices can improve analytical efficiency while preserving methodological accountability (García-López, 2025). GenAI may be especially useful in exploratory stages, but claims about actual language behavior should remain grounded in appropriate linguistic evidence.

3.2. Generative AI in Translation and Multilingual Communication

3.2.1. Machine-Assisted Translation and Localization

Generative language models offer new possibilities for translation by producing context-sensitive drafts, explaining alternative word choices, adjusting tone, and assisting with localization. Beyond direct word-for-word conversion, these systems can help users compare idiomatic expressions, adapt written communication to different audiences, and identify ambiguities in source texts (Al-Kfairy, 2024). In professional and educational settings, such features may facilitate the translation of correspondence, learning materials, research summaries, and informational content.

For example, an initial translation of an institutional announcement can be revised for a formal academic audience, while a literary passage may be examined through several alternative translations to explore differences in imagery and tone. Researchers may also use AI-generated translations as preliminary drafts before conducting a detailed human review (Al Naqbi, 2024). Such practices can reduce the effort involved in producing initial versions and support multilingual communication where professional translation resources are limited.

However, translation quality cannot be assessed solely through fluency. An output may sound natural while omitting a qualification, altering a relationship between characters, or replacing a culturally specific concept with an inaccurate approximation. These problems are particularly consequential in literary translation, where rhythm, imagery, ambiguity, voice, and stylistic variation may contribute directly to meaning (Hughes, 2021). In legal, medical, or other high-stakes communication, unverified translation may also have consequences beyond stylistic quality.

Human translators remain essential for resolving ambiguity, interpreting cultural references, maintaining authorial voice, and making contextually justified decisions. A responsible workflow uses AI-generated text as a draft, compares it against the source, and subjects important passages to human review. Evaluation should consider semantic accuracy, cultural appropriateness, terminology, register, and consistency rather than fluency alone.

3.2.2. Linguistic Diversity and Low-Resource Languages

The development of GenAI raises important questions about the representation of linguistic diversity. Languages with extensive digital resources and strong representation in training data may receive more reliable support than languages with limited digitized text, standardized orthographies, or publicly available corpora (Khan, 2024). This disparity may affect translation, grammar explanations, vocabulary generation, and the interpretation of culturally specific expressions.

Low-resource languages often include varieties whose knowledge is transmitted through oral traditions, community interaction, and locally produced texts rather than extensive digital archives. If generative systems inadequately represent these sources, they may generate inaccurate expressions, standardize dialectal differences, or misrepresent cultural concepts (Yan, 2024). Such limitations are relevant to African languages, Indigenous languages, and other linguistic communities whose digital resources may be unevenly developed.

GenAI can nevertheless contribute to language documentation and revitalization when used collaboratively. Researchers and community language practitioners may employ carefully validated tools to organize texts, develop

draft learning materials, or explore translation alternatives. Yet community participation is essential because linguistic knowledge cannot be reduced to automated output (Kirova, 2023). Decisions about which materials may be digitized, how cultural knowledge is represented, and who may access resulting resources should respect the rights and preferences of the relevant communities.

Future development should therefore prioritize multilingual evaluation, community-governed datasets, appropriate representation of dialects, and tools that permit speakers to correct model outputs. Success should be measured not only by performance in widely used languages but also by the ability to support linguistic plurality without erasing local distinctions (Sohail, 2023). This position aligns with UNESCO's emphasis on inclusion and the protection of linguistic and cultural diversity in the development and use of GenAI.

3.3. Generative AI in Literary Studies and Creative Writing

3.3.1. Literary Interpretation and Textual Analysis

GenAI offers new ways to explore literary texts by generating summaries, identifying recurring themes, comparing characters, examining narrative perspectives, and proposing possible interpretations. In teaching, AI-assisted discussion can help students begin analyzing unfamiliar works or generate questions for seminar discussion (Adlani, 2025). In research, models can assist with preliminary exploration of a large collection of texts, enabling researchers to identify passages that may warrant closer reading.

For example, an initial examination of several novels might use AI to suggest recurring representations of migration, social class, gender, memory, or environmental change. A researcher could then examine the relevant passages, compare them with historical and theoretical scholarship, and determine whether the proposed pattern is supported by the texts (Hagendorff, 2024). In this process, AI-generated interpretations function as hypotheses rather than findings that can be accepted without further analysis.

The main limitation is that literary meaning depends on context, form, and interpretive judgment. Irony may contradict a narrator's literal statements; an unreliable narrator may deliberately mislead readers; and a symbol may change meaning across different historical or cultural settings. A generative model may overlook these features or provide an interpretation that appears convincing but lacks textual evidence (Budhwar, 2023). It may also invent quotations or misattribute passages, making source verification indispensable.

Close reading remains central because it connects interpretation to specific textual details, narrative structures, linguistic choices, and aesthetic effects. AI-assisted analysis is most valuable when it stimulates alternative interpretations that researchers subsequently test against the original work. Students and researchers should be encouraged to cite the actual passages supporting their arguments and distinguish between a model's suggestions and conclusions established through scholarly analysis (Ooi, 2025).

3.3.2. Creative Writing and Literary Experimentation

GenAI can support creative writing by helping authors brainstorm ideas, develop characters, explore alternative narrative structures, experiment with dialogue, and revise passages. Writers may request several possible openings for a story, examine how a scene changes when narrated from another character's perspective, or compare stylistic choices in different drafts (Ding, 2024). These activities can help writers overcome initial uncertainty and make the revision process more exploratory.

In educational settings, AI-generated writing can also provide material for critical discussion. Students may compare human-written and AI-generated narratives, examine how characterization develops, or identify differences in imagery, pacing, and voice. Rather than treating generated work as an unquestioned example of quality, instructors can ask students to evaluate its coherence, originality, emotional effects, and dependence on familiar conventions.

Nevertheless, generative writing raises questions about creative agency and authorship. A text produced through extensive automated generation may obscure the extent of the human writer's contribution. Models may also reproduce recognizable patterns from their training data, resulting in formulaic characterization, predictable plots, or stylistic imitation (Munaye, 2025). The existence of a novel combination of words does not necessarily establish meaningful originality, and the legal and ethical status of particular uses of generated text may depend on applicable law, licensing, and platform terms.

A responsible approach preserves the writer's decision-making role. Authors can use AI for brainstorming or feedback while maintaining control over narrative purpose, revision, and final selection. Where required by a publisher, institution, or assignment policy, AI assistance should be disclosed accurately. Research on creativity should also distinguish between perceived usefulness, the novelty of generated outputs, and the quality of work produced through human–AI collaboration. These are related but different outcomes (Sengar, 2025).

3.3.3. Digital Humanities and Large-Scale Literary Research

Digital humanities combines computational methods with humanistic questions to investigate cultural texts, historical documents, and literary collections. GenAI may extend this work by supporting preliminary classification, summarization, thematic exploration, and the generation of research questions from large collections (Surbakti, 2025). It can potentially help scholars navigate extensive archives and identify material for closer examination, particularly when used alongside established methods such as corpus analysis, bibliographic research, and digital text annotation.

For instance, researchers studying representations of urban life across a collection of twentieth-century novels might use AI to identify candidate passages involving housing, migration, public space, or industrial change. Those passages can then be examined through close reading and compared with historical evidence (Law, 2024). The combination of computational exploration and human interpretation may help researchers formulate questions that would otherwise be difficult to identify quickly.

However, large-scale analysis introduces methodological risks. Model outputs may change with prompts, model versions, or sampling conditions, making exact replication difficult. A model's summary of a collection may also overemphasize prominent themes while overlooking minority perspectives or subtle formal features. Where copyright or access restrictions apply, researchers must ensure that the use of texts and generated outputs complies with relevant legal, institutional, and licensing requirements (AI-kfairy, 2024).

The methodological value of GenAI in digital humanities therefore depends on transparent documentation and appropriate validation. Researchers should record model versions, prompts, text-selection procedures, and verification methods where feasible. They should also explain how automated suggestions were converted into analytical claims and acknowledge limitations in the available corpus (Tirpan, 2024). This approach helps preserve the interpretive depth of literary scholarship while making computational assistance more transparent and reproducible.

3.4. Ethical Challenges of Generative AI in Language and Literary Studies

3.4.1. Academic Integrity, Authorship, and Originality

Academic integrity represents one of the most significant challenges associated with GenAI in language and literary studies. Generative systems can produce essays, literary interpretations, research summaries, and creative texts that resemble conventional academic writing. Although these capabilities can support legitimate learning and research activities, they complicate the distinction between assistance and the substitution of a student's or researcher's intellectual contribution (Raitskaya, 2025). The central issue is not simply whether AI has been used, but how it has been used and whether the resulting work accurately represents the individual's contribution.

In language education, students may rely on AI to generate complete assignments rather than use it to improve their writing or develop their analytical skills. In literary scholarship, researchers may use generated interpretations

without independently examining the primary text or consulting relevant critical literature. These practices risk weakening the relationship between academic assessment and demonstrated competence (Arista, 2025). They may also produce arguments that appear sophisticated but contain unsupported interpretations, fabricated quotations, or inaccurate references.

Authorship becomes particularly complex in creative writing and scholarly publication. A human author may use AI for brainstorming, editing, translation, or substantial text generation, with each activity involving a different level of intellectual contribution. Consequently, a single rule treating all AI use as equivalent may be inadequate. Institutions and publishers should establish clear policies specifying acceptable assistance, disclosure expectations, and responsibility for the accuracy and originality of submitted work.

A constructive approach emphasizes transparency and demonstrable intellectual engagement. Students may be asked to submit drafts, explain interpretive choices, verify sources, or document how AI suggestions influenced their revisions. Researchers should independently validate generated claims and comply with relevant publication policies (Esteves, 2024). Ultimately, responsibility for submitted work remains with the human author, irrespective of the extent of technological assistance.

3.4.2. Linguistic Bias, Cultural Representation, and Algorithmic Limitations

Generative models learn patterns from training data that may reflect unequal representation of languages, communities, cultural traditions, and social perspectives. Consequently, their outputs can reproduce stereotypes, privilege dominant linguistic conventions, or misrepresent groups whose experiences are inadequately represented in available data (Zhang, 2024). These concerns are especially important in sociolinguistics, postcolonial literary studies, Indigenous literary scholarship, and research examining multilingual societies.

For example, an AI system might treat a standardized variety of a language as inherently more appropriate than a regional variety or interpret a culturally specific expression according to assumptions associated with a different linguistic community. In literary interpretation, such biases may influence how characters, social relationships, religious traditions, and historical experiences are described (Tata, 2025). An apparently neutral summary may therefore reproduce dominant cultural perspectives while overlooking marginalized voices.

These limitations do not mean that every AI-generated response is necessarily biased. Rather, they indicate that outputs require critical examination in relation to their context and purpose. Researchers should evaluate how models represent different language varieties, cultural traditions, and social groups (Matos, 2025). Where appropriate, community perspectives and specialist knowledge should inform the interpretation of culturally sensitive material.

Improving linguistic and cultural representation requires more than expanding the quantity of training data. It also requires attention to data quality, consent, provenance, representational balance, and the involvement of relevant communities in evaluation. Such measures are particularly important when GenAI is used to document endangered languages, interpret oral traditions, or develop educational materials for linguistically diverse populations.

3.4.3. Hallucinations, Reliability, and Research Validity

Generative AI systems can produce statements that appear credible but are inaccurate or unsupported. These errors, commonly described as hallucinations, may include fabricated references, incorrect quotations, misleading summaries, invented linguistic rules, and false historical details. In language and literary research, such errors can undermine the validity of an argument because scholarship depends on accurate engagement with primary texts, verifiable sources, and established evidence (Samala, 2025).

For example, a model asked to explain a literary passage may attribute a quotation to the wrong character or interpret a theme without sufficient textual support. Similarly, an AI-generated bibliography may contain

nonexistent publications or incorrect bibliographic details. A researcher who accepts such outputs without verification risks introducing errors into manuscripts, classroom materials, or published analyses.

The risk is heightened when models are used to summarize large volumes of literature. Summaries may omit methodological limitations, overlook contradictory findings, or present tentative conclusions as established facts (Korchak, 2025). Fluent language can make these weaknesses difficult to recognize, particularly for users who lack expertise in the subject.

Reliability should therefore be established through independent verification rather than inferred from the confidence or coherence of an answer. Researchers should check quotations against original texts, verify bibliographic records through reliable scholarly sources, and compare model-generated interpretations with relevant evidence (Batista, 2024). For empirical studies, automated outputs should be evaluated against suitable reference standards, including expert judgments or validated datasets where appropriate.

Transparent reporting is also essential. Researchers should describe the role of GenAI in data collection, analysis, translation, or writing when that role is methodologically relevant. Such documentation allows readers to understand how conclusions were developed and evaluate the extent to which automated assistance may have influenced the results.

3.4.4. Copyright, Privacy, and Intellectual Property

Copyright and intellectual-property questions arise when GenAI systems are trained on, reproduce, transform, or generate material derived from existing literary and linguistic works. Authors, translators, publishers, and researchers may have concerns about whether protected works were used in training, whether generated outputs reproduce recognizable elements of existing texts, and how responsibility should be allocated when disputes arise (Bahroun, 2023). The legal position varies across jurisdictions and depends on the facts, applicable legislation, licensing arrangements, and relevant court decisions.

These concerns are particularly important in literary studies because literary works are often protected by copyright and may contain distinctive expressions, characters, or stylistic elements. Researchers and students should distinguish between discussing a literary work, generating a critical analysis, reproducing protected expression, and using a model to imitate a particular text (Preiksaitis, 2023). These activities raise different questions and should not be treated as legally or ethically identical.

Privacy presents another challenge. Researchers may inadvertently disclose confidential interviews, unpublished manuscripts, peer-review reports, student assignments, or identifiable information when entering material into external AI systems. The consequences depend on the platform's data-handling practices, institutional requirements, and the sensitivity of the information. UNESCO's guidance emphasizes the importance of protecting privacy and establishing appropriate institutional safeguards for the use of GenAI in education and research (Rabbani, 2025).

Institutions should consequently establish procedures for handling sensitive data, assessing approved tools, and protecting unpublished or confidential material. Researchers should review relevant terms of service and data policies before uploading research materials. They should also respect copyright, licensing, confidentiality agreements, and institutional research-ethics requirements (Wang, 2025). These precautions can help ensure that the practical advantages of GenAI do not undermine the rights of authors, participants, students, or research collaborators.

3.5. Future Directions for Research and Practice

3.5.1. Human-Centred and Responsible Integration

Future integration of GenAI should prioritize human judgment, critical engagement, and disciplinary expertise. In language education, AI tools should complement teacher feedback and structured learning rather than replace

opportunities for communication, independent practice, and reflection (Ning, 2024). In literary studies, generated interpretations should encourage further inquiry rather than become substitutes for close reading, theoretical engagement, and evidence-based criticism.

Educational institutions can promote this approach by developing guidelines that specify appropriate uses of GenAI across different tasks. For instance, an institution may permit brainstorming and grammar feedback while requiring students to produce their own interpretations and document substantial AI assistance (García-López, 2025). Assessment methods can include oral explanations, reflective commentary, annotated drafts, and source-based analysis to establish whether students understand the work they submit.

Researchers should similarly distinguish between tasks suitable for automation and those requiring expert judgment. Preliminary text classification, draft translation, and organization of research material may benefit from AI assistance, whereas culturally sensitive interpretation, validation of claims, and decisions about scholarly significance require careful human oversight.

Future empirical studies should investigate how different forms of AI assistance affect learning, writing development, interpretive reasoning, and research quality. Longitudinal research would be particularly useful for determining whether repeated AI use strengthens learners' independent abilities or encourages excessive dependence (Al-Kfairy, 2024). Such studies should examine differences across educational levels, language proficiency, disciplinary contexts, and access conditions rather than assuming that all users experience the technology in the same way.

3.5.2. Multilingual Equity and Culturally Responsive AI

A second priority is the development of more inclusive systems that represent diverse languages, dialects, literary traditions, and cultural perspectives. Current applications should be evaluated for their ability to support linguistic communities beyond those most strongly represented in digital training resources (Al Naqbi, 2024). This includes examining translation accuracy, the quality of grammatical explanations, sensitivity to regional variation, and the interpretation of culturally specific texts.

Collaborative approaches involving linguists, literary scholars, translators, local educators, and community representatives can help establish more appropriate evaluation criteria. In language documentation, for example, speakers and community researchers should participate in decisions about which materials are collected, digitized, interpreted, and made accessible. Such participation can reduce the risk of treating community knowledge merely as material for technological development (Hughes, 2021).

Research should also examine whether AI-assisted translation and writing tools improve access to education and scholarship in underrepresented languages. This requires attention to affordability, connectivity, digital literacy, and the availability of locally appropriate resources (Khan, 2024). Without these considerations, the benefits of GenAI may be distributed unevenly, reinforcing existing inequalities between institutions, linguistic communities, and regions.

Culturally responsive development should not seek to eliminate linguistic variation in favor of standardized output. Instead, it should support the legitimate diversity of language use and recognize that meaning is shaped by social relationships, historical experience, and cultural context. This objective is consistent with UNESCO's emphasis on equitable, human-centred approaches that protect linguistic and cultural diversity (Yan, 2024).

3.5.3. Methodological Innovation and Interdisciplinary Collaboration

GenAI creates opportunities for collaboration between linguistics, literary criticism, computer science, education, translation studies, and digital humanities. Future research can investigate how language models perform when analyzing metaphor, irony, narrative voice, discourse relations, multilingual interaction, and historically situated

literary texts (Kirova, 2023). Such work should compare model-generated analyses with human expert judgments and examine the conditions under which automated suggestions are useful, misleading, or incomplete.

Methodological innovation should be accompanied by rigorous evaluation. Studies should use clearly defined research questions, transparent sampling procedures, appropriate comparison groups where relevant, and explicit measures of performance. When investigating literary interpretation, researchers might compare the quality of AI-generated and human-generated analyses using criteria such as textual evidence, contextual accuracy, interpretive depth, and recognition of ambiguity (Sohail, 2023). In language learning, researchers could examine changes in writing quality, grammatical accuracy, learner confidence, and independent performance over time.

Reproducibility also deserves greater attention. Researchers should document model versions, prompts, relevant settings, data-selection decisions, and verification procedures to the extent possible. Because commercial systems may change over time, identical prompts may not consistently produce identical responses. Researchers should acknowledge this limitation and preserve appropriate records of outputs when permitted by ethical, legal, and institutional requirements (Adlani, 2025).

Interdisciplinary collaboration is especially important because technical performance alone cannot determine whether a system produces meaningful linguistic or literary knowledge. Computer scientists can contribute evaluation methods and model expertise, while linguists and literary scholars can identify contextual, cultural, and interpretive requirements (Hagendorff, 2024). Educators can assess pedagogical value, and ethicists and legal specialists can help address questions of privacy, authorship, and accountability. Such collaboration can support more robust frameworks for evaluating GenAI across different humanistic applications.

3.5.4. Research Gaps and Emerging Priorities

The reviewed themes suggest several priorities for future investigation. First, more longitudinal and comparative research is needed to establish the sustained effects of GenAI on language acquisition, independent writing, and critical interpretation. Second, the evidence base should extend beyond short-term perceptions of usefulness to examine measurable learning outcomes and the quality of scholarly work (Budhwar, 2023). Third, research should address underrepresented languages and literary traditions, particularly where digital resources and opportunities for technological participation remain limited.

Fourth, further work is needed on the relationship between GenAI and literary creativity. Researchers should distinguish between the novelty of machine-generated text, the quality of human–AI collaboration, and the development of an individual writer's creative abilities. Fifth, interdisciplinary studies should investigate the reliability of automated literary interpretation, especially for culturally specific texts, complex narrative structures, and historical materials (Ooi, 2025). Sixth, empirical research should examine how different institutional policies affect responsible use, disclosure practices, and the development of AI literacy.

Finally, researchers should investigate the environmental and infrastructural costs associated with large-scale AI use, alongside questions of access and institutional capacity. The adoption of GenAI is not solely a pedagogical or interpretive decision; it also involves technological dependence, resource allocation, and broader questions about the sustainability of digital research practices (Ding, 2024).

These priorities point toward an approach in which technological innovation is evaluated through disciplinary standards and humanistic values (Munaye, 2025). The future of GenAI in language and literary studies will depend not simply on the increasing sophistication of models but on whether their use strengthens independent inquiry, respects cultural diversity, supports equitable participation, and produces verifiable scholarly contributions.

4. Conclusion

Generative artificial intelligence is reshaping language and literary studies by expanding opportunities for language learning, writing development, translation, linguistic analysis, literary interpretation, creative experimentation, and

digital humanities research. Its capacity to generate and transform text can support individualized learning, accelerate preliminary research tasks, facilitate multilingual communication, and encourage new forms of interaction between human creativity and computational technologies. Nevertheless, these opportunities are accompanied by substantial challenges involving academic integrity, authorship, linguistic and cultural bias, fabricated information, copyright, privacy, and the potential weakening of independent intellectual engagement. The review highlights that the usefulness of GenAI depends on the context in which it is applied, the quality of its outputs, and the extent to which human expertise informs their evaluation. In language education, responsible integration should strengthen learners' communicative competence and critical thinking rather than encourage dependence on automated responses. In literary studies, AI-generated interpretations should complement close reading, historical contextualization, and theoretically informed analysis rather than replace them. Future research should prioritize longitudinal evaluation, transparent methodologies, culturally responsive model development, multilingual equity, and interdisciplinary collaboration. Institutions should establish clear policies for AI disclosure, assessment, data protection, and responsible scholarly practice, while educators and researchers should develop the critical AI literacy necessary to evaluate generated content. Importantly, a genuinely inclusive approach must recognize the needs of underrepresented languages and literary traditions and involve relevant communities in the development and assessment of AI-supported resources. Ultimately, GenAI should be understood as a supplementary intellectual tool whose contribution must be judged by its capacity to enhance, rather than diminish, human agency, linguistic diversity, interpretive depth, and scholarly integrity. Its long-term value in language and literary studies will depend on a balanced relationship between technological capability, ethical responsibility, and the enduring importance of human language, creativity, and cultural understanding.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers

Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

References

- [1] Adlani, H., Ridwan, R., & Laili, M. I. (2025). Artificial intelligence in Arabic language learning: A systematic literature review of educational trends and future directions. *Al-Muyassar: Journal of Arabic Education*, 4(2).
- [2] Al Naqbi, H., Bahroun, Z., & Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. *Sustainability*, 16(3), 1166.
- [3] Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024). A Systematic Review and Analysis of Ethical Challenges of Generative Ai: An Interdisciplinary Perspective. Available at SSRN 4833030.
- [4] Al-Kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024, August). Ethical challenges and solutions of generative AI: An interdisciplinary perspective. In *Informatics* (Vol. 11, No. 3, p. 58). MDPI.
- [5] Arista, A., Ismail, M. A., Shuib, L., & Herawan, T. (2025). A Systematic Literature Review on Ethical Framework for Adoption of Generative Artificial Intelligence. *Malaysian Journal of Computer Science*, 38, 108-128.
- [6] Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, 15(17), 12983.
- [7] Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, 15(11), 676.
- [8] Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., ... & Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, 33(3), 606-659.
- [9] Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*, 29(11), 14151-14203.
- [10] Esteves, R. M. A. (2024). *Applications, Challenges, and Ethical Implications of Generative AI: A Systematic Review* (Master's thesis, Universidade NOVA de Lisboa (Portugal)).

- [11] García-López, I. M., & Trujillo-Liñán, L. (2025, June). Ethical and regulatory challenges of Generative AI in education: a systematic review. In *Frontiers in Education* (Vol. 10, p. 1565938). Frontiers Media SA.
- [12] Hagendorff, T. (2024). Mapping the Ethics of Generative AI: A Comprehensive Scoping Review: T. Hagendorff. *Minds and Machines*, 34(4), 39.
- [13] Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks-enabled human-artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. *Frontiers in artificial intelligence*, 4, 604234.
- [14] Khan, N., Khan, Z., Koubaa, A., Khan, M. K., & Salleh, R. B. (2024). Global insights and the impact of generative AI-ChatGPT on multidisciplinary: a systematic review and bibliometric analysis. *Connection science*, 36(1), 2353630.
- [15] Kirova, V. D., Ku, C. S., Laracy, J. R., & Marlowe, T. J. (2023). The ethics of artificial intelligence in the era of generative AI. *Journal of Systemics, Cybernetics and Informatics*, 21(4), 42-50.
- [16] Korchak, A. E., Patarakin, Y. D., & Costley, J. (2025). Exploring the use of generative artificial intelligence by university students: a systematic literature review. *Вестник Российского университета дружбы народов. Серия: Информатизация образования*, 22(1), 37-57.
- [17] Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174.
- [18] Matos, T., Santos, W., Zdravetski, E., Coelho, P. J., Pires, I. M., & Madeira, F. (2025). A systematic review of artificial intelligence applications in education: Emerging trends and challenges. *Decision Analytics Journal*, 15, 100571.
- [19] Munaye, Y. Y., Admass, W., Belayneh, Y., Molla, A., & Asmare, M. (2025). ChatGPT in education: A systematic review on opportunities, challenges, and future directions. *Algorithms*, 18(6), 352.
- [20] Ning, Y., Teixayavong, S., Shang, Y., Savulescu, J., Nagaraj, V., Miao, D., ... & Liu, N. (2024). Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. *The Lancet Digital Health*, 6(11), e848-e856.
- [21] Ooi, K. B., Tan, G. W. H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., ... & Wong, L. W. (2025). The potential of generative artificial intelligence across disciplines: Perspectives and future directions. *Journal of Computer Information Systems*, 65(1), 76-107.
- [22] Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review. *JMIR medical education*, 9(1), e48785.
- [23] Rabbani, S. A., El-Tanani, M., Sharma, S., Rabbani, S. S., El-Tanani, Y., Kumar, R., & Saini, M. (2025). Generative artificial intelligence in healthcare: applications, implementation challenges, and future directions. *BioMedInformatics*, 5(3), 37.
- [24] Raitskaya, L., & Tikhonova, E. (2025). The 2025 landscape of generative AI in scholarly writing and publishing: a scoping review of uses and ethical approaches. *Journal of language and Education*.
- [25] Samala, A. D., Rawas, S., Wang, T., Reed, J. M., Kim, J., Howard, N. J., & Ertz, M. (2025). Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects. *Education and Information Technologies*, 30(3), 3239-3278.
- [26] Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2025). Generative artificial intelligence: a systematic review and applications. *Multimedia tools and applications*, 84(21), 23661-23700.
- [27] Sohail, S. S., Farhat, F., Himeur, Y., Nadeem, M., Madsen, D. Ø., Singh, Y., ... & Mansoor, W. (2023). Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions. *Journal of King Saud University Computer and Information Sciences*, 35(8), 101675.
- [28] Surbakti, F. P. S. (2025). Systematic literature review on generative AI: ethical challenges and opportunities. *International Journal of Advanced Computer Science and Applications*, 16(5), 307-315.
- [29] Tata, T. S., & Heri, H. S. (2025). Navigating the Generative AI Revolution in Education: A Systematic Review of Applications, Ethical Considerations, and Future Directions. *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, 14(2).
- [30] Tirpan, E. C. (2024). The ethical issues in generative artificial intelligence: A systematic review. *Business & Management Studies: An International Journal*, 12(4), 729-747.
- [31] Wang, Y., Zhang, T., Yao, L., & Seedhouse, P. (2025). A scoping review of empirical studies on generative artificial intelligence in language education. *Innovation in Language Learning and Teaching*, 1-28.
- [32] Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., ... & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112.
- [33] Zhang, X., Zhang, P., Shen, Y., Liu, M., Wang, Q., Gašević, D., & Fan, Y. (2024). A systematic literature review of empirical research on applying generative artificial intelligence in education. *Frontiers of Digital Education*, 1(3), 223-245.