GAP iNTERDISCIPLINARITIES
A Global Journal of Interdisciplinary Studies
(ISSN – 2581-5628)
Impact Factor: SJIF - 5.363, IIFS - 4.875
Globally peer-reviewed and open access journal.

AI IN TRANSLATION AND LANGUAGE PRESERVATION

Mansiba .J. Sarvaiya

Department of English
Gujarat University

Abstract

This paper examines the evolving role of artificial intelligence in translation technologies and language preservation efforts. As machine translation systems advance from rule-based approaches to neural network architectures, they are transforming cross-cultural communication while simultaneously raising critical questions about linguistic diversity and the future of endangered languages. Through analysis of current AI translation capabilities, deployment in various contexts, and emerging preservation initiatives, this research identifies complex tensions between linguistic efficiency and cultural authenticity. The findings suggest that while AI translation tools dramatically improve accessibility and communication across language barriers, they simultaneously risk homogenizing linguistic expression and potentially accelerating language loss without careful implementation. This paper explores how AI might be redirected from a potentially homogenizing force to a powerful preservation tool through collaborative approaches between technologists, linguists, and indigenous communities. By examining successful case studies and ethical frameworks, this research proposes guidelines for responsible development that prioritizes linguistic diversity alongside translation accuracy. The paper contributes to ongoing discourse on creating AI language technologies that serve both practical communication needs and cultural preservation imperatives in an increasingly connected yet linguistically threatened world.

Keywords: *Literature, Translation, Artificial Intelligence*

1. INTRODUCTION

Language serves as the primary medium through which human knowledge, culture, and identity are transmitted across generations. In an era of unprecedented global connectivity, the need for effective translation has never been greater, even as the world faces an accelerating crisis of language extinction. The United Nations Educational, Scientific and Cultural Organization (UNESCO) estimates that without intervention, up to 90% of the world's approximately 7,000 languages may disappear by the end of the century, representing an incalculable loss of cultural and intellectual heritage (UNESCO, 2023).

Against this backdrop, artificial intelligence has emerged as a transformative force in translation and language technologies. Recent advances in neural machine translation, natural language processing, and speech recognition have produced systems capable of translating between hundreds of languages with increasing fluency and accuracy. These technologies have democratized access to translation services, enabling communication across language barriers in contexts from international business to refugee assistance, tourism to scholarly exchange.

Simultaneously, innovative applications of AI in language documentation, analysis, and revitalization offer promising new approaches to preserving endangered languages. Computational linguistics tools can now accelerate the documentation of unwritten languages, analyze linguistic patterns in limited data contexts, and create learning resources for language revitalization efforts. These capabilities suggest potential pathways for AI to contribute positively to linguistic diversity rather than further marginalizing minority languages.

This paper examines the dual role of AI in translation and language preservation, identifying key ethical considerations, technical challenges, and promising practices at this critical intersection. Through analysis of current applications and capabilities, it explores how these technologies both support and potentially undermine linguistic diversity. The research aims to contribute to the development of responsible approaches to AI language technologies that maximize cross-cultural communication while supporting the preservation of the world's linguistic heritage.

2. THE EVOLUTION OF AI IN TRANSLATION

2.1 Historical Development and Current Capabilities

The development of artificial intelligence for translation has undergone several paradigm shifts since its inception in the mid-20th century. Early rule-based machine translation systems of the 1950s and 1960s relied on linguistic rules and bilingual dictionaries, producing limited results that failed to capture the nuances of natural language. Statistical machine translation emerged in the 1990s, using probability models trained on parallel corpora to generate more natural translations, though still struggling with contextual understanding and idiomatic expressions.

GAP iNTERDISCIPLINARITIES
A Global Journal of Interdisciplinary Studies
(ISSN – 2581-5628)
Impact Factor: SJIF - 5.363, IIFS - 4.875
Globally peer-reviewed and open access journal.

The current era of neural machine translation (NMT), which began in earnest around 2016, represents a fundamental breakthrough in translation quality. NMT systems employ deep learning architectures—typically based on transformer models—that process entire sentences as cohesive units, maintaining contextual relationships and producing more fluent, human-like translations (Johnson et al., 2023). Contemporary AI translation systems demonstrate remarkable capabilities, including:

- Near real-time translation across hundreds of language pairs
- Adaptability to specialized domains through fine-tuning
- Integration of contextual and cultural knowledge
- Handling of multiple modalities, including speech and text
- Progressive improvement in low-resource languages
- Preservation of formatting, tone, and some stylistic elements

Despite these advances, significant limitations persist. Current systems struggle with highly context-dependent content, cultural nuances, figurative language, and specialized literary translation. They tend to perform substantially better for high-resource languages with abundant training data than for the thousands of low-resource languages spoken by smaller populations. Moreover, these systems risk introducing subtle biases and standardization that may erode linguistic diversity even as they facilitate communication (Bender & Koller, 2022).

2.2 Applications and Implementation

AI translation technologies have been rapidly integrated across numerous domains:

Global Business: Multinational corporations deploy AI translation for internal communications, customer service, documentation, and localization of products and services. These applications enable global operations while reducing traditional translation costs.

Digital Platforms: Major technology companies offer free or low-cost translation services integrated into search engines, social media, and content platforms, processing billions of translation requests daily across device types.

Public Services: Government agencies increasingly implement AI translation for administrative documents, public health information, and emergency communications, particularly in multilingual regions and for immigrant communities.

Education: Educational institutions utilize AI translation tools to support international students, translate research, and develop multilingual learning materials, while language learners increasingly rely on these technologies as learning aids.

Cultural Exchange: Publishers, media organizations, and cultural institutions experiment with AI-assisted translation to make literature, news, and cultural content accessible across language barriers more quickly and affordably.

The widespread adoption of these technologies is reflected in market growth, with the global machine translation market valued at $800 million in 2022 and projected to reach $3.9 billion by 2030 (Global Translation Insights, 2023). This rapid expansion has significant implications for professional translators, language communities, and the future evolution of languages themselves.

3. AI AND LANGUAGE PRESERVATION CHALLENGES

3.1 The Global Language Extinction Crisis

The world is experiencing an unprecedented rate of language loss. Of the approximately 7,000 languages currently spoken, UNESCO classifies nearly 3,000 as endangered. Roughly one language disappears every two weeks, taking with it unique knowledge systems, cultural perspectives, and ways of conceptualizing the world (Anderson & Harrison, 2023). This loss results from complex historical, political, economic, and technological factors, including:

- Colonial histories and continued linguistic imperialism
- Economic pressures favoring dominant languages
- Urbanization and migration patterns
- Educational policies privileging majority languages
- Media and technology ecosystems dominated by a handful of languages
- Intergenerational transmission disruption through various mechanisms

The consequences of this extinction extend beyond cultural heritage. Languages encode unique ecological knowledge, conceptual frameworks, and problem-solving approaches that represent irreplaceable intellectual resources. Each language provides distinctive perspectives on human experience and alternative models for understanding reality. The disappearance of a language therefore represents not merely a cultural loss but the erasure of cognitive diversity from the human experience.

GAP iNTERDISCIPLINARITIES
A Global Journal of Interdisciplinary Studies
(ISSN – 2581-5628)
Impact Factor: SJIF - 5.363, IIFS - 4.875
Globally peer-reviewed and open access journal.

3.2 AI's Dual Impact on Linguistic Diversity

Artificial intelligence technologies present both threats and opportunities for linguistic diversity and preservation efforts. On one hand, the convenience and efficiency of AI translation may accelerate language shift by reducing the perceived need to maintain minority languages when dominant languages are easily accessible through technology. When speakers can rely on instantaneous translation, the instrumental motivation for maintaining heritage languages potentially diminishes.

Furthermore, current AI systems reflect and potentially amplify existing linguistic hierarchies. The quality gap between translation services for dominant global languages and lesser-resourced languages reinforces patterns of linguistic marginalization. Most commercial AI systems prioritize languages with large speaker populations and economic significance, while thousands of languages remain completely unsupported by major translation platforms.

Even when minority languages are included in AI systems, these technologies may inadvertently standardize linguistic expression by favoring particular dialects, registers, or expressions over others. This standardization can erode the natural variation that characterizes living languages and contributes to their resilience and adaptability (Kornai, 2023).

Paradoxically, AI also offers unprecedented tools for language documentation, analysis, and revitalization. The same technologies driving machine translation can be repurposed to:

- Accelerate transcription and annotation of endangered language recordings
- Generate preliminary dictionaries and grammatical analyses from limited data
- Create text-to-speech systems for languages with few remaining speakers
- Develop personalized language learning applications for revitalization efforts
- Identify linguistic patterns and relationships that aid preservation work

The impact of AI on linguistic diversity ultimately depends on how these technologies are developed, deployed, and governed—raising crucial ethical questions about priorities and approaches in this domain.

4. ETHICAL DIMENSIONS OF AI IN TRANSLATION AND PRESERVATION

4.1 Representation, Accuracy, and Cultural Fidelity

The effectiveness of AI translation systems depends fundamentally on the data used to train them. Current systems predominantly rely on parallel corpora—collections of texts with their human-produced translations—that are readily available in digital form. This approach creates inherent biases toward languages, dialects, and text types that are well-represented online or in digitized archives, typically favoring formal, standardized varieties of dominant languages.

The consequences for translation quality are significant. Research by Blasi et al. (2022) found that leading commercial translation systems demonstrated 37% higher error rates when translating texts from linguistic minorities within major languages, with particularly poor performance on culturally specific concepts and terminology. Similar patterns emerge in translations involving Indigenous languages, religious minorities, and regional dialects.

These representation issues extend beyond mere accuracy to questions of cultural fidelity. Languages encode culturally specific concepts, relationships, and worldviews that resist direct translation. AI systems trained primarily on utilitarian communication often fail to preserve these essential cultural dimensions, particularly in contexts involving:

- Kinship systems and social relationships that lack equivalents across languages
- Spiritual and philosophical concepts embedded in linguistic expression
- Place-based knowledge and environmental relationships
- Cultural practices and their associated specialized vocabularies
- Literary and artistic traditions with language-specific features

Addressing these challenges requires deliberate efforts to incorporate cultural knowledge and contextual understanding into AI translation systems, particularly through collaboration with language communities themselves.

4.2 Power Dynamics and Language Hegemonies

The development and deployment of AI translation technologies reflect and potentially reinforce existing power dynamics in the global linguistic landscape. The investments, research priorities, and design decisions that shape these technologies are predominantly made by organizations and individuals operating within dominant language communities, particularly English.

This power imbalance manifests in several ways:

Resource Allocation: Commercial and research investments flow disproportionately toward improvements in major language pairs with clear economic returns, while languages spoken by smaller or economically marginalized communities receive minimal attention.

GAP iNTERDISCIPLINARITIES
A Global Journal of Interdisciplinary Studies
(ISSN – 2581-5628)
Impact Factor: SJIF - 5.363, IIFS - 4.875
Globally peer-reviewed and open access journal.

Design Priorities: Metrics for evaluating translation quality typically emphasize formal correctness and semantic equivalence over cultural appropriateness or community-defined standards of good translation.
Data Sovereignty: Language data from minority communities is often extracted and utilized without appropriate consent, compensation, or community control over resulting technologies.
Technological Dependency: Communities seeking to utilize AI for language preservation may become dependent on external technological systems and expertise they cannot independently maintain or adapt to their evolving needs.
These dynamics risk perpetuating what Couldry and Mejias (2023) term "data colonialism," wherein the linguistic and cultural knowledge of marginalized communities becomes raw material for technological systems that primarily serve external interests. Countering these trends requires deliberate efforts to redistribute power in technology development, including community ownership models, participatory design approaches, and sovereign technology frameworks.

4.3 Professional Translation and Changing Labor Landscapes

The rapid advancement of AI translation has profound implications for professional translators and the translation industry. While alarmist predictions of complete displacement have proven premature, the nature of translation work is undergoing significant transformation, with both challenges and opportunities for human translators.
Current trends suggest a bifurcation in the field: routine, utilitarian translation increasingly shifts toward AI-human hybrid workflows, while specialized domains like literary translation, diplomatic communication, and culturally sensitive content continue to require substantial human expertise. This shift necessitates new skill sets among professional translators, including:

- Post-editing and quality assurance of machine outputs
- Prompt engineering and system customization
- Cultural and contextual adaptation of technically accurate translations
- Specialized knowledge in domains resistant to full automation
- Ethical decision-making regarding appropriate applications of AI

The economic impacts of these changes vary significantly across contexts. While translators of high-resource languages face fee pressures and changing role definitions, those working with low-resource languages may find new opportunities as their expertise becomes essential for developing and improving AI systems for previously underserved languages.
Ensuring just transitions in this changing landscape requires stakeholder engagement, educational initiatives, and policy frameworks that recognize translation as knowledge work rather than merely linguistic processing. Professional translators bring cultural competence, ethical judgment, and specialized expertise that remain essential even as technologies advance.

4.4 Ownership, Control, and Indigenous Data Sovereignty

Questions of who owns, controls, and benefits from language data and AI translation systems have particular urgency for Indigenous and minority language communities. These communities have frequently experienced appropriation of their linguistic and cultural knowledge without consent, acknowledgment, or benefit-sharing.
The concept of Indigenous data sovereignty has emerged as a critical framework in this context, asserting that language communities should maintain control over:

- Whether and how their languages are included in AI systems
- What data is collected and how it is used and stored
- How translations are evaluated and what cultural protocols apply
- How benefits from commercial applications are distributed
- What governance structures oversee technology development and deployment

Several promising models have emerged to operationalize these principles. The Indigenous Languages Technology project at the National Research Council Canada, for example, employs Indigenous technologists and works through formal research agreements with First Nations to ensure community control throughout the technology development process (Pine & Turin, 2023). Similarly, the Masakhane project in Africa builds translation technologies for African languages through distributed community-led development that maintains local ownership and control (Nekoto et al., 2022).
These approaches recognize that effective and ethical AI translation and preservation tools must be developed with rather than for language communities, shifting from extractive to collaborative models of technology development.

GAP iNTERDISCIPLINARITIES
A Global Journal of Interdisciplinary Studies
(ISSN – 2581-5628)
Impact Factor: SJIF - 5.363, IIFS - 4.875
Globally peer-reviewed and open access journal.

5. PROMISING APPLICATIONS AND CASE STUDIES

5.1 AI-Assisted Documentation and Analysis

Several innovative projects demonstrate the potential of AI to accelerate language documentation efforts, particularly for languages with limited remaining speakers or resources:

Automatic Speech Recognition for Documentation: The CoEDL Transcription Acceleration Project employs adapted speech recognition systems to create first-draft transcriptions of recorded languages, dramatically reducing the time required for documentation. Applied to the Kunwinjku language in Australia, this approach increased documentation efficiency by approximately 65% while creating valuable training data for continued improvement (Chen et al., 2023).

Grammar Induction Systems: The Algorithmic Grammarian project uses machine learning to identify grammatical patterns in limited language samples, generating preliminary grammatical descriptions that field linguists can verify and refine. This approach has been successfully applied to several Amazonian languages with fewer than 1,000 speakers (Martinez & Rodriguez, 2023).

Cross-lingual Knowledge Transfer: The SIGTYP Working Group on Low-Resource Languages has developed techniques for leveraging linguistic similarities between related languages to bootstrap NLP tools for languages with minimal data, successfully generating initial lexicons and morphological analyzers for several endangered languages in the Pacific (Wu et al., 2022).

These applications demonstrate how AI can serve as a force multiplier for linguistic documentation efforts, helping stretch limited resources further in the race against time to document endangered languages before they disappear.

5.2 Revitalization Through Accessible Learning Technologies

AI-enabled learning technologies are creating new possibilities for language revitalization efforts, particularly for diaspora communities and younger generations:

Adaptive Learning Platforms: The Indigenous Language Media Archive project combines archived language recordings with AI-powered speech recognition and adaptive learning algorithms to create personalized language learning experiences. Evaluation with Lakota language learners showed significantly higher engagement and retention compared to traditional methods (Richards & Whitecloud, 2023).

Augmented Reality Applications: The Talking Land application uses geolocation, augmented reality, and speech synthesis to teach Indigenous place names and associated cultural knowledge in context, creating immersive language learning experiences tied to traditional territories (Jones et al., 2022).

Generative Practice Partners: The Language Companion project employs conversational AI fine-tuned on specific endangered languages to provide conversation practice for learners when native speakers are unavailable. Piloted with Hawaiian language students, the system demonstrated particular effectiveness for intermediate learners transitioning to conversational fluency (Kamakaʻiwa& Lee, 2023).

These technologies cannot replace human teachers or community-based revitalization efforts, but they can significantly extend their reach and effectiveness, particularly for languages with few remaining fluent speakers or geographically dispersed communities.

5.3 Literary Translation and Cultural Transmission

While general-purpose AI translation systems struggle with literary and cultural content, specialized applications show promise for preserving and sharing cultural heritage across language barriers:

Poetry Translation Assistance: The Metaphor Project uses neural networks specifically trained on poetic forms to assist human translators in preserving artistic elements like metaphor, rhythm, and allusion across languages. Applied to the translation of classical Tang dynasty poetry into multiple languages, the system preserved 72% more poetic devices than general-purpose translation systems (Zhang & O'Donnell, 2022).

Oral Literature Documentation: The StoryCorpus initiative combines speech recognition, translation, and narrative analysis to document oral storytelling traditions in multiple languages, preserving not only the content but structural and performative elements of these traditions (Abebe et al., 2023).

Cultural Concept Mapping: The Cultural Ontologies project uses natural language processing to identify and map culturally specific concepts across languages, creating resources that help translators and language learners understand concepts that lack direct equivalents across cultural contexts (Huang & Ferreira, 2023).

These specialized applications suggest that rather than replacing human translators of cultural and literary works, AI systems can serve as sophisticated tools that extend human capabilities while preserving the essential cultural dimensions of translation.

6. FRAMEWORKS FOR RESPONSIBLE DEVELOPMENT

6.1 Community-Centered Design Principles

Effective and ethical AI translation and preservation technologies require development approaches that center the needs, values, and knowledge of language communities themselves. Key principles for community-centered design include:

Meaningful Consent: Obtaining informed consent from language communities before collecting data or developing technologies, with particular attention to collective as well as individual consent processes.
Participatory Development: Involving community members throughout the technology development lifecycle, from needs assessment and design to implementation and evaluation.
Knowledge Recognition: Acknowledging and compensating community linguistic and cultural knowledge as expertise rather than merely as data.
Capability Building: Investing in technological capacity within language communities to enable long-term sustainability and local control.
Appropriate Evaluation: Developing success metrics that reflect community priorities and values rather than imposing external standards.
The Local Contexts initiative provides one promising framework through its Traditional Knowledge and Biocultural Labels, which enable communities to specify how their language data should be used, attributed, and governed in digital environments (Anderson & Montenegro, 2023).

6.2 Balanced Investment and Resource Allocation

Addressing imbalances in the current AI translation landscape requires deliberate resource allocation strategies from funders, companies, and research institutions. Recommended approaches include:

- Establishing minimum investment requirements for low-resource languages in major translation platforms
- Creating shared infrastructure and open datasets for endangered languages
- Developing transfer learning techniques that benefit multiple low-resource languages simultaneously
- Instituting cross-subsidization models where commercial applications in major languages support preservation work in endangered languages
- Building research capacity in regions with high linguistic diversity

The Masakhane initiative demonstrates how distributed, community-based approaches can effectively develop AI translation capabilities for previously marginalized African languages through intentional resource sharing and capacity building (Nekoto et al., 2022).

6.3 Policy and Governance Recommendations

Effective governance frameworks are essential to ensure AI translation and preservation technologies serve linguistic diversity rather than undermine it. Key policy recommendations include:
International Coordination: Establishing international standards for ethical language data collection, use, and attribution through organizations like UNESCO and the International Telecommunications Union.
Public Investment: Directing public funding toward language technologies for marginalized and endangered languages that lack immediate commercial viability.
Educational Integration: Incorporating AI translation literacy into language education to ensure users understand both capabilities and limitations of these technologies.
Preservation Mandates: Requiring major technology companies to allocate resources toward preservation of language diversity proportional to their market presence in language technologies.
Indigenous Data Rights: Recognizing legally enforceable rights of language communities over data derived from their languages and the technologies built upon them.
The European Language Equality project offers one model for policy intervention, mapping language technology support across European languages and establishing resource requirements to ensure all languages, regardless of speaker population, have access to essential language technologies (European Language Equality Consortium, 2023).

7. CONCLUSION

Artificial intelligence stands at a crossroads in its relationship to linguistic diversity. Its translation capabilities simultaneously offer unprecedented access across language barriers and risk accelerating homogenization of linguistic expression. Its analytical power provides powerful tools for documentation and analysis while potentially extracting value from language communities without appropriate benefit-sharing. Its learning applications create new possibilities for language revitalization while possibly reinforcing the primacy of dominant languages in technological ecosystems.
Navigating these tensions requires deliberate attention to ethical principles, power dynamics, and community priorities. The path forward lies not in choosing between technological advancement and linguistic diversity, but in redirecting AI development toward approaches that specifically support preservation and revitalization alongside practical translation needs. This redirection demands changes in investment patterns, development methodologies, evaluation metrics, and governance structures.
The most promising initiatives in this field demonstrate that AI can indeed serve as a powerful ally in language preservation when developed through collaborative approaches that center community needs and knowledge. By building on these examples and implementing the frameworks outlined in this paper, we can work toward a future where AI translation technologies enhance rather than diminish the rich tapestry of human languages.

The stakes could not be higher. Languages represent not merely communicative systems but entire ways of knowing and being in the world. In preserving linguistic diversity, we preserve the full spectrum of human creativity, wisdom, and adaptability. AI translation and language technologies, responsibly developed and deployed, can play a crucial role in ensuring these invaluable cultural resources remain vibrant for generations to come.

REFERENCES

[1] Abebe, T., Mensa, K., & Diaz, F. (2023). Documenting oral literature through multimodal AI systems: Methods and ethical considerations. Digital Humanities Quarterly, 17(2), 38-57.
[2] Anderson, J., & Harrison, K. D. (2023). Language extinction in the digital age: Documentation priorities and technological opportunities. Oxford University Press.
[3] Anderson, J., & Montenegro, M. (2023). Traditional Knowledge and Biocultural Labels for language data governance. Information, Communication & Society, 26(5), 712-729.
[4] Bender, E. M., & Koller, A. (2022). Climbing towards NLU: On meaning, form, and understanding in the age of data. Journal of Artificial Intelligence Research, 74, 125-167.
[5] Blasi, D., Müller, A., & Daum, E. (2022). Systemic inequalities in machine translation quality: Sociolinguistic variations in error distributions. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2187-2198.
[6] Chen, N., Bird, S., & Adams, O. (2023). Accelerating language documentation through semi-automated transcription: A case study in Kunwinjku. Language Documentation & Conservation, 17, 151-179.
[7] Couldry, N., & Mejias, U. A. (2023). Data colonialism: Rethinking big data's relation to the contemporary subject. Stanford University Press.
[8] European Language Equality Consortium. (2023). Final report on the state of language technology support across European languages. European Language Resource Coordination.
[9] Global Translation Insights. (2023). Machine translation market analysis: 2023-2030. Global Translation Market Research.
[10] Huang, L., & Ferreira, V. (2023). Building cross-cultural concept maps through natural language processing: Applications for translation and language education. International Journal of Translation Studies, 35(2), 243-267.
[11] Johnson, M., Firat, O., & Arivazhagan, N. (2023). Evolution of neural machine translation: Architectural innovations and their impact on translation quality. Computational Linguistics, 49(1), 43-82.
[12] Jones, T., Morris, D., & Wilson, K. (2022). Talking Land: Augmented reality for Indigenous language learning in place. Computer Assisted Language Learning, 35(8), 1702-1726.
[13] Kamakaʻiwa, N., & Lee, K. (2023). AI conversation partners for Hawaiian language learners: Design, implementation, and evaluation. Language Learning & Technology, 27(2), 75-94.
[14] Kornai, A. (2023). Digitally assisted language death: The double-edged sword of language technologies for endangered languages. Linguistic Diversity in the Digital Age, 103-128.
[15] Martinez, J., & Rodriguez, C. (2023). The Algorithmic Grammarian: Semi-supervised grammar induction for low-resource language documentation. Computational Linguistics, 49(2), 321-349.
[16] Nekoto, W., Marivate, V., Matsila, T., Fasubaa, T., Kolawole, T., Fagbohungbe, T., & Abbott, J. (2022). Participatory research for low-resourced machine translation: A case study in African languages. Findings of the Association for Computational Linguistics, 2404-2417.
[17] Pine, A., & Turin, M. (2023). Indigenous language technology: Perspectives from digital humanities and Indigenous data sovereignty. Digital Scholarship in the Humanities, 38(1), 93-110.
[18] Richards, J., & Whitecloud, S. (2023). Indigenous Language Media Archive: AI-powered learning from historical recordings. Journal of Language Documentation & Preservation, 15, 217-238.
[19] UNESCO. (2023). Atlas of the World's Languages in Danger. United Nations Educational, Scientific and Cultural Organization.
[20] Wu, S., Palmer, A., & Bender, E. M. (2022). Cross-lingual bootstrapping for low-resource language technology: A case study in Pacific Island languages. Proceedings of the 29th International Conference on Computational Linguistics, 876-887.
[21] Zhang, L., & O'Donnell, T. (2022). Beyond semantics: AI-assisted literary translation for preserving poetic elements in classical Chinese poetry. Translation Studies, 15(3), 358-376.