

Endangered Languages, Digital Preservation And Linguistic Human Rights In The Age Of AI Translation

¹DR. SHIKHA MISHRA

¹Assistant Professor

¹SDP College for Women, Ludhiana, Punjab

Abstract—Today's modern world is known from name of technology, AI, digitalization, globalization etc, but we should not to be forget about our roots. Indian roots are in our culture, religion, God, social and moral values, morality, and spiritualism etc. Our society is going to be change daily with very fast speed. Westernization, globalization and modernization, brightness of world, luxurious life attract the human mind. These all made our life changed. If we critically analyze these changes, we come to know some are good and some are bad for us. Modernizations make some cultural things to be in endangered situations that are sometimes lost due to our changed mind setup. But in other aspect, Modern world of AI, computerized society help to save things.

Artificial intelligence enhances endangered language preservation by digitizing oral histories, automating translations for low-resource languages, and building phonetic speech synthesis. By bridging communication gaps, AI creates digital ecosystems that empower indigenous communities to protect their linguistic heritage for future generations. Languages are essential for establishing social cohesiveness, transferring information, making social and defining cultural identity. Despite this, urbanization and globalization are threatening to wipe out a number of languages worldwide. New technologies, daily new inventions make some positive and negative aspects for languages of world. In this paper we study endangered languages, and their preservation by various Artificial Intelligence tools.

Index Terms—Endangered Languages, Digital Preservation, AI Tools

I. Introduction

Earth is one of the beautiful creation of God. God creates plants, birds, animals, human beings and nature at all. God gives brain to human being. Due to brain human being not only develops oneself but also develop surroundings. Man sometimes uses his mind critically for humanity. AI, Digitalization, Technology created by humans mind. These all become boss of human mind.

Against this technology and 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.

II. Languages

Language is any formal system of gestures, signs, sounds, and symbols used or conceived as a means of communicating thought. It is a tool for our intellectual and emotional expression. UNESCO has recognized India as one of the most linguistically diverse countries. As per census 2011, more than 19,500 languages or dialects are spoken in India as mother tongues while only 121 languages are spoken by 10,000 or more people. This shows the quantum of endangered languages in India. A number of steps have been taken at the national and international level to conserve the endangered languages but much more needs to be done for giving due protection to the vulnerable languages.

Human beings have the power of speech, by which one can share own ideas to others. Language is a highly structured, symbolic system of communication used to convey meaning through speech, writing, or gestures. It enables humans to express complex thoughts, share cultural identities, and discuss things not physically present. It consists of vocabulary, grammar, and syntax. Normal as well as abnormal persons all have their own language and own way of languages.

Regional, state, national and international languages are of different types. Some languages are used by maximum population while some are not used by everyone.

English is one of the international languages. English is maximum used by maximum no. of people of world, while regional languages, state languages and tribal languages are not known and spoken by everyone.

III. Endangered languages

An endangered language is one at risk of falling out of use as its speakers die or shift to speaking more dominant languages. Language endangerment is a severe concern, especially for regions with deep indigenous roots. The loss of a language represents the loss of invaluable cultural knowledge, including oral histories, environmental traditions, and unique community perspective. Intergenerational disruption, globalization, migration and lack of institutional societal support make a language endangered.

UNESCO provides a classification system in its 'Atlas of Endangered languages' as:

- (a) **Vulnerable** – most children speak the language, but it may be restricted to certain domains (e.g., home)
- (b) **Definitely endangered** – children no longer learn the language as a 'mother tongue' in the home;
- (c) **Severely endangered** – language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves;
- (d) **Critically endangered** – the youngest speakers are grandparents and older, and they speak the language partially and infrequently;

(e) **Extinct** – there are no speakers left.

IV. Endangered languages in India

Globally, around 3,000 languages are considered endangered, with experts warning that a significant portion could vanish by the end of the century. India is multicultural and multi lingual country. India is incredibly diverse linguistically, with **122 major languages** and **1,599 other languages** recorded, as well as thousands of dialects. While the country has no single national language, the Indian Constitution officially recognizes **22 scheduled languages** for administration and official purposes. There are regional, state, national and many international languages spoken by people of India.

India has the highest number of endangered languages in the world, with nearly 200 indigenous tongues at risk of dying out, according to UNESCO's Atlas of the World's Languages in Danger. Over 40 of these are officially classified as "critically endangered" as they are spoken by shrinking, isolated tribal communities

Indian Languages like Kanashi, Balti, Gahri, Jad, Great Andamanese, Shampen, Birhor and Asuri etc. are some of the endangered languages.

V. Reasons for Endangeredness

Languages become endangered primarily due to a breakdown in intergenerational transmission, which occurs when parents stop teaching their native language to their children. According to UNECO Survey report over 50% of the world's languages are currently losing speakers, putting them on a path toward extinction.

There are cultural, economical, political, historical physical and demographical reasons make languages endangered.

¹ Cultural and Economic Shifts

Globalization: The expansion of global trade, education and digital media elevates major languages like English, making them necessary for commerce, technology, and entertainment at world level.

Economic Incentives: Communities used frequently dominant languages to secure employment and improve quality of life

Urbanization and Migration: Moving from isolated rural villages to cities scatters speech communities and forces migrants to adopt the dominant city language

Social Stigmatization: Native tongues are sometimes associated with lower social status or rural poverty, prompting parents to push children toward more prestigious national languages

² Political and Historical Pressures

Colonialism: Historically, colonial communities aggressively imposed alien languages while devaluing or completely dismantling indigenous language systems.

Overt Repression: Governments have actively banned regional languages from public spaces, media, and school systems to force national unity

School System Assimilation: Historical residential or boarding school systems physically punished children for using indigenous speech, causing lasting intergenerational trauma and sudden language drops

Government Apathy: A lack of state funding for native-language radio, text materials, public libraries, or primary education leaves minor languages out of formal life

³ Physical and Demographical Crises

Natural Catastrophes: Events like earthquakes, tsunamis, or famines can completely wipe out small, geographically isolated language communities.

War and Genocide: Violent conflicts and targeted ethnic cleansing physically destroy the population of speakers, or force survivors to hide their identity by dropping their language.

Intermarriage: When people from different linguistic communities marry, the household usually defaults to a single dominant language, failing to pass down the minor one.

VI. Language preservation

Digital preservation is the active management of digital materials to ensure they remain accessible, usable, and authentic over time. It combats risks like media degradation, software obsolescence, and hardware failure, making it essential for libraries, archives, and individuals.

Digital preservation of language utilizes technology—such as AI, audio-video archiving, and mobile apps—to document, revitalize, and store endangered languages. It ensures linguistic diversity survives by making dictionaries, grammar, and oral histories globally accessible to both native speakers and learners

Artificial Intelligence accelerates language preservation by enabling communities to document, digitize, and revitalize endangered or oral languages. Through automated speech-to-text transcription, real-time translation, and dynamic digital archives, AI helps safeguard linguistic heritage that was once at risk of fading away

AI translation bridges the digital divide but poses threats to Linguistic Human Rights (LHR)—the fundamental right to communicate, access information, and receive justice in one's mother tongue. While AI democratizes access, it risks imposing hegemonic languages, erasing cultural nuances, and creating dangerous inaccuracies in critical sectors like law and healthcare. Artificial intelligence reshapes global communication, the concept of **Linguistic Human Rights (LHR)** has entered a complex, transformative

era. Language is not merely a tool for exchanging data; it is intrinsically tied to human dignity, cultural identity, and self-determination. While AI-driven translation and Large Language Models (LLMs) offer unprecedented accessibility, they simultaneously introduce critical challenges to LHR.

VII. Use of Technologies and Methods for Digital Preservation of Languages

Nowadays, there are many technologies and variety methods used to preserve language:

AI & Machine Learning: Artificial intelligence is utilized to process vast amounts of unstructured voice and text data, helping create automated translation models and pronunciation checkers for minority dialects.

Breaking Down Barriers: AI live translation tools and multimodal Large Language Models enable instantaneous communication, drastically improving access to digital services and global information

Digital Archives: Multimedia databases and repositories house audio recordings, dictionaries, folklore, and syntax to ensure languages are accessible, redisplayed, and preserved long-term

Mobile Applications: User-friendly educational apps bring gamified language learning to younger generations, bridging physical and generational gaps.

VIII. Conclusion

Artificial intelligence serves as a double-edged sword for linguistic diversity. While it offers unprecedented tools to document and revitalize endangered languages, it also risks homogenizing global communication. Responsible, community-led AI implementation is essential to ensure technological progress preserves, rather than erases, unique cultural identities. 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.

References

- [1] Sarvaiya, J.M (2025). AI in translation and language preservation. GAP, Vol –VIII, Special issue March 2025. pp.49-55.
- [2] JPine, A., & Turin, M. (2023). Indigenous language technology: Perspectives from digital humanities and Indigenous data sovereignty. *Digital Scholarship in the Humanities*, 38(1), 93-110.
- [3] Richards, J., & Whitecloud, S. (2023). Indigenous Language Media Archive: AI-powered learning from historical recordings. *Journal of Language Documentation & Preservation*, 15, 217-238.

- [4] UNESCO. (2023). Atlas of the World's Languages in Danger. United Nations Educational, Scientific and Cultural Organization.
- [5] 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.
- [6] 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.