GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Reviving Endangered Indian Languages through AI-Powered Translation Tools

Dr. Devina Krishna
Assistant Professor
Department of English
Patna Women's College (Autonomous)
devina.eng@patnawomenscollege.in

ABSTRACT

Recent advancement of Artificial Intelligence lends a new lease of life to vulnerable languages and helps preserve them for future generations. It provides access to digital services such as translation applications in Indian languages, creating hope in humans against language loss and safeguarding precious linguistic riches. The present paper sheds light on the impact of artificial intelligence in bringing out digital services in Indian languages to reduce language loss. The work employs a qualitative methodology framework focusing on a survey method to determine the types of technological tools used for translation. And their usage in creating open-source datasets in local languages and finding how vulnerable languages can be stored and used for future generations through content creation. Further, it aims to showcase the influence of technological giants like Google Translate, Bhashini, Anuvadini and other applications that support Indian linguistic diversity and lead to the preservation of diversity by creating linguistic storehouses. In addition, the paper explores copyright issues related to AI-generated translations, as the translations are produced by machines rather than humans. In sum, the study showcases the positive side of AI-driven translation applications that can influence the survival of vulnerable languages, leading to a diverse digital ecosystem and the issues of the authority of the final products constructed by the new technology.

Keywords: digital, preservation, technology, translation, vulnerable

Introduction

We live in an epoch where language and technology go hand in hand and chances are slim to segregate these. Both are the expressions of human creativity. Language was considered a dialect with an army and a navy and later there was a transformation of the prior sense of the

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

definition. It has been said that in modern world language will cease to exist without technology. Technology plays a great role in being hand in glove with it. The present world offers a variety of scope for languages to exist.

India is an abode to many languages comprising of scheduled, non-scheduled, and lesser-known languages. Many of the languages fall under the 'endangered category' representing distinct levels of endangerment. On one hand, we notice that languages are ways for preserving history and heritage of one's nation in a simplified manner. But if these languages face extinction, then that leads to the loss of many aspects in worldly scenario. Therefore, linguists, language scientists, and language enthusiasts deal with the issue of language loss and devise means to preserve the vulnerable codes. One way is the intervention of technology through Artificial Intelligence (AI) that is being used to revive endangered Indian languages through machine learning algorithm that can transcribe and translate oral histories and stories from endangered language speakers. AI- powered technology allows for the preservation of valuable cultural knowledge that might be vanished otherwise. AI technology helps through ways such as: Speech Recognition, Text-Analysis, Images and Audio Processing. It serves as Interactive Language Learning Platforms, provides Speech Recognition and Pronunciation Practice thereby strengthening the foundation for the protection of Indian linguistic heritage. Humans are the biggest protectors of human languages; however, Artificial Intelligence can renew and preserve linguistic codes.

In the domain of law, copyrights are ascribed to a legal/natural person instead of a non-human with no legal protection. The issue of IP ownership of AI generated works needs acute concern and has been addressed in the present work.

Objective

The objective of the present study is to discover the prolific role of the buzzing Artificial Intelligence in enhancing language preservation through its contributions to translation mechanism. It endeavours to showcase Artificial Intelligence's role in revitalizing endangered Indian languages through translation tools. In addition, the paper explores copyright issues related to AI-generated translations, as the translations are formed by computers rather than humans. In sum, the study showcases the constructive side of AI-driven translation applications

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

that can impact the existence of vulnerable languages, leading to a diverse digital ecosystem and the issues of the authority of the ultimate products created by the novel technology.

Review of Literature

Current methods of translations are often limited to a small amount of people, however, by incorporating additional translations into applications, more people can utilize these languages to learn more about them.

people, however, by incorporating additional translations into applications, more people can utilize these lan-guages to learn more about them.

AI-powered translation tools have the potential to aid in the preservation and revitalization of endangered Indian languages. AI-powered tools can assist in documenting and preserving endangered languages through automated transcription, translation, and analysis (Kulkarni et al., 2020).

Machine learning systems in the form of algorithms can help identify and classifying linguistic patterns, simplifying language documentation (Gupta et al., 2019). Moreover, AI-driven translation tools can increase the accessibility of endangered languages by providing real-time translations, promoting their use in day-to-day life (Srivastava et al., 2020).

AI-driven language learning platforms can engross younger generations in learning endangered languages, ensuring their maintenance (Mishra et al., 2019). It is a fact that the dearth and quality issues impede the expansion of AI-powered translation tools for endangered languages (Jha et al., 2020). Cultural and linguistic nuances may be lost in AI-generated translations, necessitating human lapses (Kumar et al., 2018).

Previous works on AI-translations and copyright issues stated that the term ‘authorship’ should be redefined to include both human and non-human authors (Hristov, 2017). However, there is a need for more research in the domains of technology, law and language whereas the present study seeks to fill up the gaps which the earlier studies could not fulfil.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Hypothesis

The present work presents the following research hypotheses:

Primary hypothesis

AI-led translation tools seek to revive endangered Indian linguistic codes thereby making huge impact upon users.

Ancillary hypothesis:

Severe copyright issues exist related to AI-generated translations

Methodology

The present work incorporates the use of qualitative approach mingled with an exploratory research design. To conduct the research, a questionnaire was designed at the onset of the research journey to find the effect of various AI-powered translation tools that impact and make a difference in the linguistic life of the users. The questionnaire was further circulated among young undergraduate, post graduate students, teachers, and translators. The respondents were users of different translation tools and applications used for the purpose of translation and language learning. The study included thematic analysis of the interviews that were conducted with the informants. Also, the study assessed the effectiveness of AI-powered translation tools in preserving the languages and sought to identify challenges and opportunities in using AI for language preservation.

Data

The data collection process for this study will involve a systematic search for relevant primary and secondary data sources.

Primary data sources:

For the proposed research, the survey analysis method is being used to gather first-hand information from the AI translation tools' users. A questionnaire on the topic was circulated through the instant messaging platform "WhatsApp" and data was collected in the form of the generated responses. The survey provided a platform for users to share their experiences and challenges while using the tools in question. The survey was conducted to assess language usage and attitudes towards AI-powered translation tools.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Secondary data sources:

Academic journals and relevant web sources were being sourced for the present study.

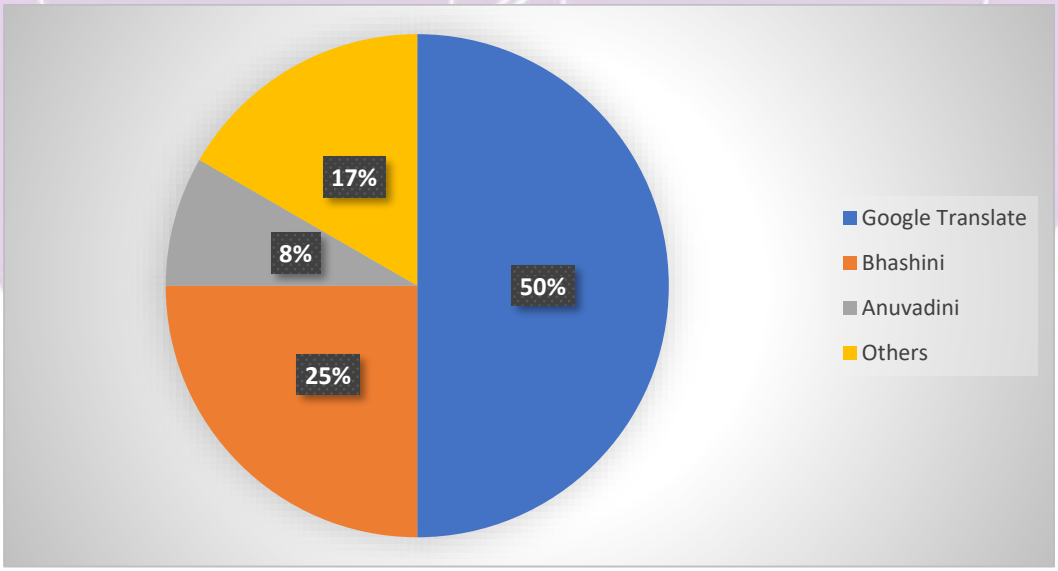
The proposed research effectively combined primary data from user surveys and secondary data from academic sources to explore the experiences of AI translation tool users. This dual approach yielded comprehensive insights into the challenges and benefits associated with these tools, ultimately enhancing understanding and guiding future developments in AI translation technology.

Discussion

The purpose of this study is to identify the online translation tools use among the students of higher education and to measure the impact of such tools in their lives. The results of questionnaire reveal the experience related to intensity and purpose of Google Translate, the most frequently used, other tools and translation procedures. This study involved 40 college going undergraduate students who filled up the questionnaire with diverse responses. Besides Google Translate, Bhashini app supports translation in 22 scheduled languages.

The following were the results of the questionnaire:

1. Which is your frequently used translation tool?



The current survey revealed the most frequently used translation tools among respondents. The results show that:

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Google Translate is the clear leader, with 50% of respondents relying on it for their translation needs.

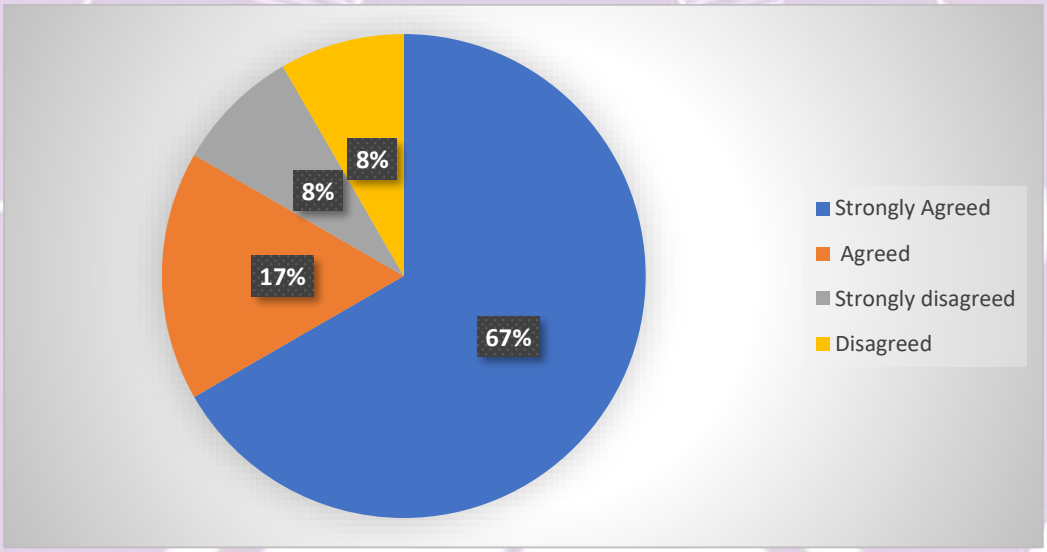
Bhashini is the second most popular choice, used by 25% of respondents.

Anuvadini is used by a smaller but still significant 8% of respondents.

A notable 17% of respondents use other translation tools not listed here, indicating a diverse range of preferences.

These results suggested that Google Translate was the dominant player in the translation tool market, but there is still room for other tools like Bhashini and Anuvadini to carve out their niche. The "others" category also highlights the existence of alternative tools that cater to specific needs or preferences.

2. Are you comfortable and content when you use the translation tools?



The survey assessing user comfort and contentment with translation tools revealed overwhelmingly positive results. Respondents expressed their level of agreement with the statement "I am comfortable and content when using translation tools" as follows:

An impressive 67% of respondents strongly agreed, indicating a high level of satisfaction and confidence in using translation tools.

A further 17% agreed, showing a significant majority of users (84% total) are comfortable and content with translation tools.

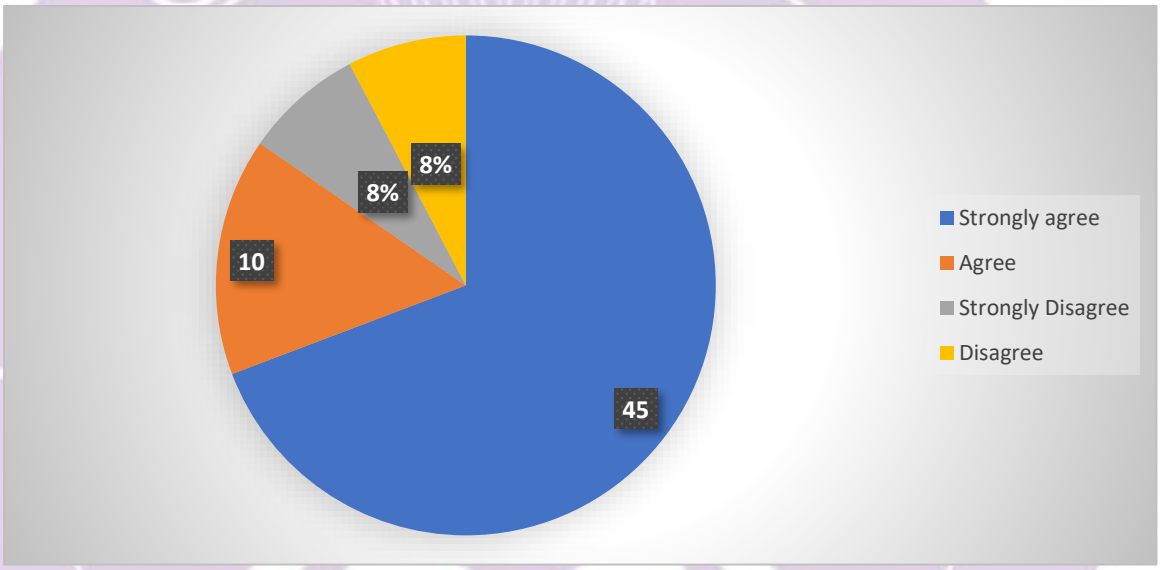
GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

A small but notable 8% of respondents slightly agreed, indicating some reservations or limitations in their satisfaction.

Only 8% of respondents disagreed, suggesting a minority of users may have had negative experiences or concerns with translation tools.

Overall, these results suggested that the vast majority of users are comfortable and content with translation tools, highlighting their effectiveness and user-friendliness.

3. Is it easy to learn a new language by using translations apps like Google Translate to translate words from our native language?



Here's a description for the given data:

Effectiveness of Translation Apps in Language Learning

The survey explored the ease of learning a new language using translation apps like Google Translate. Respondents expressed their opinions on the statement "It is easy to learn a new language by using translation apps to translate words from our native language" as follows:

A significant 45% of respondents strongly agree, indicating a strong belief that translation apps are a valuable tool for language learning.

An additional 10% agree, bringing the total percentage of positive responses to 55%. This suggests that more than half of the respondents find translation apps helpful in learning a new language.

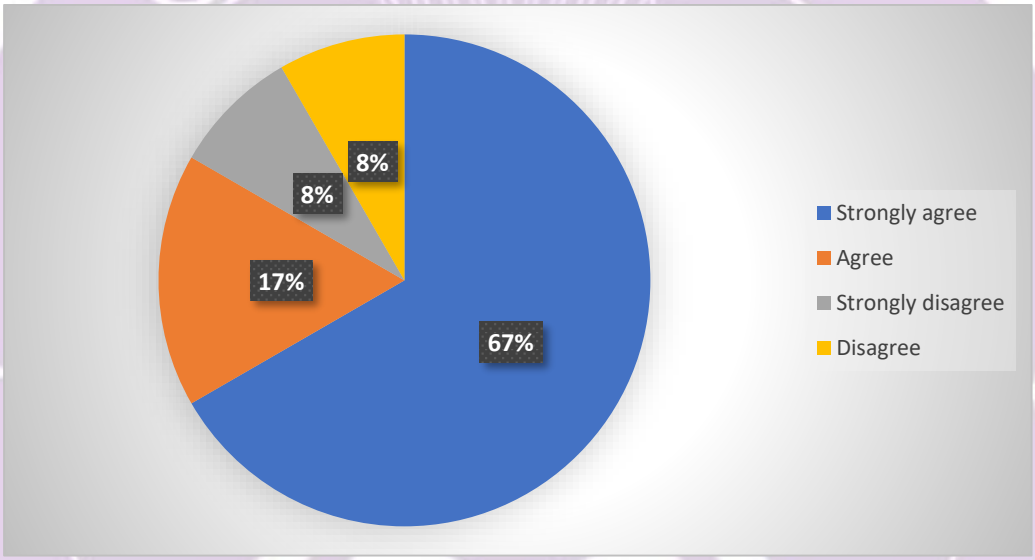
GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

On the other hand, 8% of respondents strongly disagree, indicating a strong conviction that translation apps are not effective for language learning.

A further 8% disagree, bringing the total percentage of negative responses to 16%. This suggests that a minority of respondents do not find translation apps useful for language learning.

Overall, these results suggest that a majority of users believe that translation apps like Google Translate can be a helpful tool in learning a new language, although a notable minority disagree.

4. Are there more advantages than disadvantages to translation application services?



A survey assessed users' opinions on the benefits and drawbacks of translation application services. Respondents were asked if they believe there are more advantages than disadvantages to using these services. The results show:

An overwhelming 67% of respondents strongly agree, indicating a strong conviction that the benefits of translation application services far outweigh the drawbacks.

A further 17% agree, bringing the total percentage of positive responses to 84%. This suggests that most users believe the advantages of translation application services significantly outweigh the disadvantages.

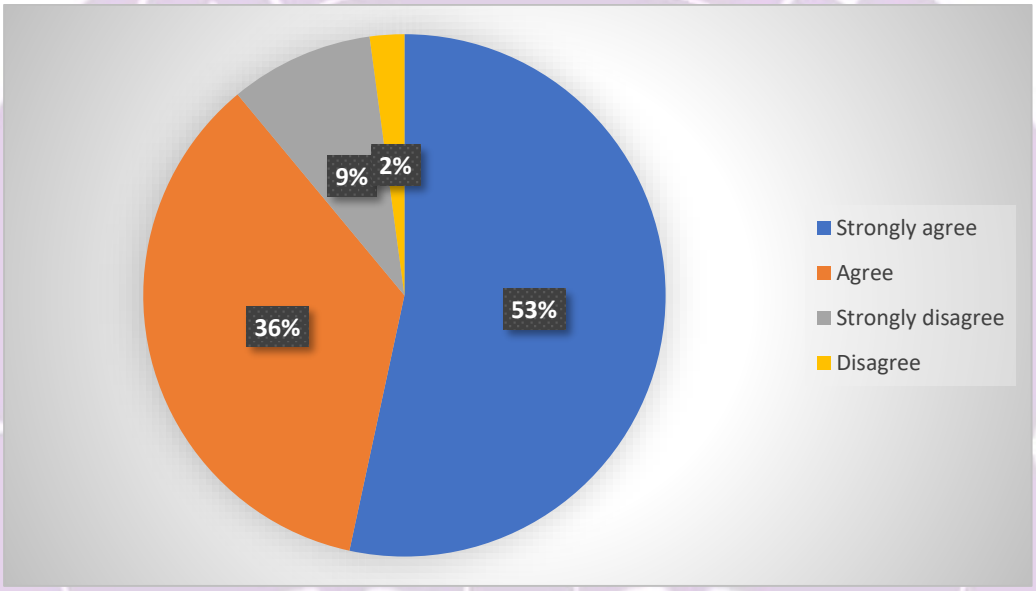
In contrast, only 8% of respondents strongly disagree, indicating a strong belief that the disadvantages outweigh the advantages.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

An additional 8% disagree, bringing the total percentage of negative responses to 16%. This suggests that a small minority of users believe the disadvantages of translation application services are more significant than the advantages.

Overall, these results indicate a strong consensus among users that translation application services offer more benefits than drawbacks, highlighting their perceived value and usefulness.

5. Do translation applications help in reviving endangered languages through usage and preservation?



A survey investigated the potential of translation applications in helping to revive endangered languages. Respondents were asked if they believe translation applications can aid in the usage and preservation of endangered languages. The results show:

A majority 53% of respondents strongly agree, indicating a firm belief that translation applications can play a significant role in reviving endangered languages.

A further 36% agree, bringing the total percentage of positive responses to 89%. Further it suggests that almost 9% strongly believe translation applications do not play some role in revival of codes whereas 2% disagree to the fact that the tools safeguard languages.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Platforms and initiatives utilizing AI-powered translation tools for endangered Indian languages:

Various platforms and initiatives have been designed that utilize AI-powered translation tools for endangered Indian languages. A few of these include:

1. AI4Bharat's Indian Language Translation Platform: This platform develops AI-powered translation tools for endangered Indian languages
2. The Endangered Languages Project (ELP) India: This platform also leads to safeguarding endangered codes through AI-powered translation tools.
3. Bhasha's Indian Language Translation Platform: Another platform of this kind develops AI-led tools for Indian languages
4. The People's Linguistic Survey of India (PLSI): This is a platform that aims to employ AI- led translation tools for documentation.
5. Google's Endangered Languages Project: This application supports language preservation and documentation through the AI-powered tools of translation
6. Microsoft's Azure Cognitive Services: This platform offers machine-learning translation for various Indian languages, including endangered ones

Endangered Indian languages supported by AI-powered translation tools

Santali, one of the vulnerable languages is supported by AI-powered translation tools. Google Translate, one of the leading AI-led translation tools, supports Santali language translation (text and speech). Another platform AI4Bharat's Indian Language Translation Platform include Santali language support for machine translation. Also, the villagers in Karnataka are among thousands of speakers of diverse Indian languages producing speech data for tech firm Karya, which is making datasets for firms such as Microsoft and Google to use in AI models for the domains of education, health sector and various other areas as well. The government of India is working tirelessly in producing digital services and thereby crafting linguistic datasets through Bhashini aimed at creating open-source datasets in regional languages for producing AI tools and applications.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Language, law, and technology

On discussing the interface between technology and law, copyrights issues related to AI-driven translations are complex and evolving. A few of the key points can be kept in mind and addressed:

1. Proprietorship: The first thing to question is that who owns the rights to AI-generated translations? It can be the creator of the AI-tool, the user who inputs the text, or someone else?
2. Imitative works: The next thing to consider is whether AI-generated translations could be regarded as derived works of the original text, or novel creations.
3. Authorizing: The kind of authorizations that are required to use and distribute AI-generated translations.
4. Correctness and liability: To whom should we consider responsible for errors or inaccuracies in AI-generated translations?
5. Inventiveness and creativity: The question is that whether AI-generated translations be considered original and creative works, or are they simply mechanical reproductions.
6. Impartial use: The question is whether AI-generated translations be used under fair use provisions, or do they require authorization from the copyright holder.
7. International copyright law: It is a question to be probed as to how do different countries' copyright laws apply to AI-generated translations.

Limitations

While AI tools offer promising solutions for preserving languages, it is also essential to acknowledge their limitations. The study includes certain limitations as well that need closer introspection. Here are the possible limitations of the study:

1. Small sample size: The study banked on a small sample size as it collected responses from a limited number of participants

2. Dearth of diversity: The informants may not be assorted enough in terms of language, culture, or geographic location.

3. Self-selection bias: Only those who are interested in translation applications or language preservation may have responded to the survey. This becomes a major drawback when it comes

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

to the idea of self-selection bias and as such, this leads to a smaller number of informants taking interest in the survey.

4. Limited scope: The study only focused on translation applications and did not explore other factors that influence language preservation.

5. Assessment system of measurement: The evaluation of AI based language model can be stimulating and tough due to limited resources and assessment criteria.

6. Language limitations: The study may be biased towards languages with more developed translation applications and resources as a few languages have been included in the tools.

7. Technical limitations: The study ceases to present the limitations and potential errors of translation applications themselves.

These limitations highlight the need for further research to fully recognize and realize the role of translation applications in language preservation.

Conclusion and Implications of the Study

AI tools provide solutions for the preservation of endangered languages; it is essential to be aware of the limitations too. The significance of this research lies in its exploration of how AI-powered translation tools can preserve the vulnerable Indian linguistic codes through certain human made tools. By investigating the impact of the AI-powered translation tools on the preservation and promotion of local languages, this research aims to highlight the potential for these technologies to bridge the communication gaps and substitute inclusivity in a plurilingual society. Besides, this study aims to comprehend the role of AI-driven language translation tools in diminishing the risk of language loss by the project practice and documentation at length. Also, the study will provide a rich source of data for future linguistic researchers, language enthusiasts and language scientists. The study showcasing the pros and cons of AI-powered translation tools will deliver significant information towards the linguistic fraternity to confirm that technology within the Indian subcontinent leads to the maintenance of linguistic ecosystem.

References:

Arora, V. (2020). Artificial intelligence and language preservation: A systematic review. *Journal of Language and Linguistics*, 19(4), 1050-1071.

GYANBODH
An International Multidisciplinary Peer Reviewed Journal
ISSN:3048-9792
Volume: 2
Issue: 5
Special Issue: 1
November: 2025

Gupta, Y. (2019). Machine learning algorithms for linguistic pattern identification and language documentation. Journal of Language Modelling, 7(1), 1-25. DOI: 10.15398/jlm.v7i1.273

Hristov, K. (2017). *Artificial Intelligence and Copyright Dilemma*. University of Illinois, Chicago.

Jain, A., & Jain, S. (2020). AI-powered translation tools for Indian languages: Challenges and opportunities. Proceedings of the 2020 International Conference on Artificial Intelligence and Language Processing (AILP), 123-128.

Jha, S. K. (2020). AI-powered translation tools for endangered languages: A case study of Santali language. In Proceedings of the 17th International Conference on Natural Language Processing (pp. 345-353).

Kulkarni, A. (2020). AI-powered tools for endangered language preservation. In Proceedings of the 12th Language Resources and Evaluation Conference (pp. 3421-3428).

Kumar, A., & Sharma, D. (2019). AI-based language preservation and revitalization: A case study of the Thadou language. Language Documentation & Conservation, 13, 101-125.

Kumar, A., & Singh, P. (2019). Machine translation for Indian languages: A review. International Journal of Advanced Research in Computer Science, 10(2), 309-316.

Mishra, A. (2019). Language preservation through technology: A case study of an endangered Indian language. Language Documentation & Conservation, 13, 46-64.

Patel, H. (2020). Developing AI-powered translation tools for endangered Indian languages (Master's thesis). Indian Institute of Technology, Kanpur.

Rehman, A. (2020). *Language preservation and AI: A critical analysis*. Routledge.

Singh, R. (Ed.). (2019). *Language and technology: Perspectives from India*. Cambridge University Press.

UNESCO. (2019). Language preservation and AI: A report on the use of artificial intelligence in language preservation efforts. UNESCO Institute for Statistics.