Journal of Digital Sociohumanities
ISSN: 3032-7865 **Vol. 2 No. 2 2025**
Website : https://jds.unand.ac.id
Email : jds@hum.unand.ac.id | digitalsociohumanities@hum.unand.ac.id

Impact of AI- Powered Translation Tools: Upholding Indian Linguistic Diversity

Devina Krishna[1], Vadini Gupta[2], Kiran Kumari[3], Kumari Nidhi[4]
Patna Women's Collage, Patna, India

Article information

DOI : 10.25077/jds.2.2.70-80.2025
Correspondence : krishna.devina@gmail.com

Submission Track

Submission : March 24, 2025
Final Review : May 29, 2025
Accepted : May 30, 2025

Keywords

AI-powered translation, linguistic diversity, communication, linguistic loss, endangered languages, digital archiving

Abstract

This paper aims to investigate the transforming potential of AI-driven translating tools in safeguarding and advancing India's great linguistic variety. By allowing their translation into generally spoken tongues, technologies including Bhashini, Anuvadini, Google Translate, and DeepL are becoming increasingly effective tools for bridging linguistic barriers and reducing the loss of threatened languages. By doing this, these instruments become quite important in preserving cultural identities and guaranteeing the survival of linguistic legacy. This study emphasizes how AI translation tools not only improve efficient communication among many language communities but also greatly help to digital archiving of languages for next generations, so increasing their accessibility and understandable value to more general audiences. This paper shows that AI-driven translating tools not only serve as means of linguistic preservation but also as essential resources for ensuring the continuous relevance and survival of India's several languages in the digital era by combining a mixed-methods approach—integrating both quantitative survey data collected via WhatsApp and Instagram and qualitative insights drawn from academic articles, web sources. These results have broad ramifications since they highlight the pressing need of legislators, technologists, and cultural players to fund the creation of inclusive, accurate, and sensitive AI language tools. Moreover, the research implies that ensuring that minority and threatened languages are meaningfully reflected in the digital sphere will depend on cooperative efforts among local communities, linguists, and technology developers. This study creates opportunities for future innovation aiming at promoting linguistic equity and cultural sustainability in India and beyond by advancing the debate on the ethical and cultural obligations of artificial intelligence development.

INTRODUCTION

India is globally recognized as one of the most linguistically diverse nations, where language not only serves as a means of communication but is deeply intertwined with cultural identity, traditional knowledge, and community heritage. According to the 2011 Census of India, the country is home to 121 major languages and 270 minor languages, with 22 Scheduled Languages officially recognized in the Eighth Schedule of the Indian Constitution (Shivakumar Jolad & Aayush Agarwal, 2024). These Scheduled Languages account for 96.71% of the population's mother tongues, while the remaining 99 languages are categorized as non-scheduled and spoken by smaller communities. Additionally, Ethnologue (2024) lists 424 living languages across the Indian subcontinent, affirming India's position as a cradle of linguistic plurality.

India's languages originate from six major language families: Indo-Aryan, Dravidian, Tibeto-Burman, Austroasiatic, Tai-Kadai, and Great Andamanese (Abbi, 1999). Among these, the Indo-Aryan languages constitute the largest group, spoken by approximately two-thirds of the population, covering regions from Bangladesh in the east to Pakistan in the west and Sri Lanka in the south to Nepal in the north (Shackle, 2001). The Dravidian language family comprises around 70 languages, predominantly spoken in southern India, with over 215 million speakers across India, Pakistan, and Sri Lanka (Shackle,

2001). The Tibeto-Burman languages, a subset of the broader Sino-Tibetan language family, are spoken along the Himalayan belt, stretching from Jammu and Kashmir in the west to Assam, Indo-China, and parts of eastern China (Peterson, 2015). The Austro-Asiatic languages include indigenous tribal languages spoken in central and eastern India, such as Munda languages like Santali, Mundari, and Ho, which have distinct phonetic and grammatical structures. The Tai-Kadai language family, earlier considered part of the Sino-Tibetan group, is now recognized as a distinct family spoken primarily in parts of Northeast India, Southern China, and Southeast Asia, with a history spanning approximately two thousand years (Peterson, 2015) . Lastly, the Great Andamanese languages, nearly extinct today, were once spoken by ten different tribal communities in the Andaman Islands. These linguistic groups highlight India's deep-rooted cultural and historical diversity (Breton, 1997).

Despite this richness, India has experienced the extinction of over 220 languages in the past 50 years due to various socio political and economic factors such as globalization, urban migration, standardization of education, and the dominance of majority languages in media and administration (Ghose et al, 2024). This alarming trend threatens not only the medium of communication but also the cultural identity, indigenous wisdom, and oral traditions of numerous ethnic groups (Romaine, 2015). Language is a vessel of collective memory, community ethos, and intergenerational values; its loss is akin to the erasure of intangible heritage (Gupta, 2009; Ghose et al., 2024).

In this context, the role of digital technology and Artificial Intelligence (AI) has become increasingly significant in reversing language loss and promoting linguistic resilience (Ray et al., 2024; Wang, 2024). Recent advancements in AI, particularly in Natural Language Processing (NLP), Machine Learning (ML), Neural Machine Translation (NMT), and Deep Learning (DL), are being deployed to create language translation tools that facilitate real-time translation and improve over time through data feedback mechanisms. These tools offer promising solutions for the documentation, conservation, and revitalization of endangered and lesser-used languages (Kavitha et al., 2023; Wang, 2024) .

Google Translate, Anuvadini, and Bhashini—a government-led project started under the National Language Translation Mission (NLTM)—are notable examples of these tools. In order to provide AI-powered translation services in all 22 Scheduled Indian languages, Bhashini is a strategic intervention that aims to eventually expand to regional dialects and indigenous languages (Sinha, 2009). Bhashini aims to increase the accessibility and inclusivity of the internet and digital governance through open-source models, cooperation with research institutions, and crowdsourcing of linguistic data (Nag, 2024). In addition to facilitating communication across India's wide range of languages, these translation tools are essential for public service delivery, education, media localization, and digital archiving (Vijayalakshmi et al., 2024; Zaki & Ahmed, 2024).

With broad ramifications for social justice, communication, and the sharing of knowledge, artificial intelligence (AI) has emerged as a disruptive force in contemporary society. While addressing ethical issues, Warwick's introductory work examines basic AI concepts such as machine learning, robotics, and consciousness (Warwick, 2011). There are advantages and disadvantages to the quick development of AI-powered language technologies, especially for non-native English speakers who might encounter linguistic obstacles (Park, 2024). Though accuracy, bias mitigation, and ethical considerations continue to be obstacles, AI-powered translation systems hold promise for overcoming linguistic barriers and advancing communication equity (Zaki & Ahmed, 2024). Globally, AI's social effects differ greatly depending on the geographic and cultural setting, and low- and middle-income nations may be more susceptible to adverse effects and less likely to gain from new developments in AI (Hagerty & Rubinov, 2019). These studies highlight the necessity of ongoing research and development to guarantee that AI technologies mitigate potential risks and foster inclusivity.

Artificial Intelligence (AI) has the potential to preserve and revitalize endangered languages, according to recent research. Unprecedented opportunities to record, examine, and preserve linguistic

diversity are presented by AI technologies (Kavitha et al., 2023). Although AI-driven language analysis improves the accuracy of documentation, digital archives combined with AI-based tools can produce extensive repositories for managing and storing linguistic data (Dayanand et al., 2023). In order to potentially recreate lost languages and create educational models for communities that are at risk, artificial intelligence (AI) systems can be trained to identify and examine patterns in understudied languages (Ermolova et al., 2024). Technologies for speech recognition, machine learning, and natural language processing hold promise for automating transcription, linguistic analysis, and data collection (Ray et al., 2024). The lack of data for rare languages, linguistic bias in AI systems, and ethical issues like privacy and cultural sensitivity are still problems, though (Ermolova et al., 2024; Ray et al., 2024). AI offers a promising way to preserve linguistic heritage in spite of these obstacles.

Real-time transcription and translation across multiple languages, with an emphasis on Indian languages, can be facilitated by AI-powered systems (Ravindra et al., 2024). In order to possibly recreate lost languages and create educational models for small ethnic groups, these technologies can assist in identifying and analyzing linguistic patterns of understudied or endangered languages (Ermolova et al., 2024). Nevertheless, issues still exist, such as the dispersed nature of data in South Asian languages (Arora et al., 2022) and linguistic bias in AI systems that could worsen linguistic and social inequality. Only a small portion of the world's 7000+ languages are represented in current NLP applications, despite the promise of language technologies to promote multilingualism (Joshi et al., 2020). In order to ensure inclusive communication and preserve cultural heritage, it is imperative that these disparities be addressed and that linguistic diversity be given priority in AI development.

Increased use of AI technologies is associated with higher engagement metrics, and they have an impact on communication styles and self-presentation on social media platforms (Zulfiqar et al., 2024). Chatbots for customer support, predictive analytics for data-driven decision-making, and personalized content curation are all made possible by AI in marketing (Aggarwal et al., 2024). Although ethical and privacy concerns still exist, AI tools on social media platforms enhance user interaction, personalize experiences, and optimize communication strategies (Chaudhari & Bhangale, 2024). AI algorithms may contribute to the spread of false information and echo chambers, even though they improve user experiences by providing personalized content. The efficacy of AI in changing the digital landscape was demonstrated by the 96.6% accuracy of a suggested methodological framework for evaluating AI-driven systems in social media (Niveditha, 2023).

Even though there is a lot of research on using AI for language translation and preservation, some important areas are still missing. First, most studies focus on how the tools work but not on how people actually use them or how they affect society. Second, while AI tools support many major languages, they often ignore local dialects and tribal languages, making the technology less inclusive. Third, AI still struggles to understand the meaning behind slang, idioms, and cultural expressions, which are important for real communication. Lastly, there isn't enough research on how these tools affect daily life—like how they help with education, language use, or saving cultural traditions. More real-world studies are needed to see how useful these tools really are for protecting endangered languages.

With little research and little in-depth analysis, the subject of AI-translation tools is still largely unexplored. This new field offers a novel and creative area of research that is full of promise but still needs more academic focus. More thorough research is urgently needed to comprehend the effects, difficulties, and potential of AI as it continues to transform language translation. The goal of the current study is to examine and comprehend how AI-powered language translation programs, like DeepL, Google Translator, Bhashini, and Anuvadini, affect the preservation and advancement of Indian languages. Additionally, this study guarantees that all languages have a role in the digital future, regardless of their prevalence, making it more dependable and accessible to everyone. Additionally, it assesses how accurate and promising these tools are at deciphering languages with such linguistic variations. In order to ensure the revival and

revitalization of languages, this study also examines how these AI-driven tools can encourage the use of Indian languages in a variety of contexts, including communication, media, and education.

This research explores two central questions: How do AI-powered translation tools affect the preservation and promotion of regional languages in India, and is there scope for AI-driven language tools to resurrect and revitalize endangered languages? By examining the role of artificial intelligence in linguistic preservation, the study seeks to understand both the current impact of these technologies on India's rich tapestry of regional languages and their potential to support the survival and revitalization of languages that are at risk of disappearing.

METHODS

This study examined the effects of AI-powered translation tools on regional language promotion and preservation in India by combining quantitative and qualitative techniques. A structured survey that was sent to active users of AI translation tools via Instagram and WhatsApp comprised the quantitative component. The majority of the 86 responses (72.1%) were from students, offering a helpful overview of a demographic that is actively using technology. In order to provide quantifiable insights into these tools' effects on linguistic diversity, the survey collected information on usage frequency, perceived advantages, difficulties, and general experiences.

A semi-structured interview with a Bhashini Division expert was used to gather qualitative data for the study in order to enhance the analysis. The interview focused on the Bhashini app's operational advantages and disadvantages as well as its wider function in language preservation. In order to give the primary findings more context and support, secondary data from scholarly publications, reports, and reliable websites were also examined.

In order to determine usage patterns, compute frequencies, and display trends in charts and graphs, the gathered survey data was examined using descriptive statistical techniques. Thematic analysis was used to identify recurrent themes and professional opinions about the practical, ethical, and cultural ramifications of AI-powered translation from the qualitative interview data. By using a combination of methods, the study was able to record both quantifiable trends and complex viewpoints, offering a comprehensive picture of how AI translation tools can help stop language loss and support the survival of regional and endangered languages in India.

RESULT

Investigating the prevalence and usage trends of AI-powered translation tools in India is the first step in the quantitative analysis. The information seeks to evaluate the degree of accuracy these technologies offer when translating between languages as well as their effect on maintaining India's rich linguistic diversity.

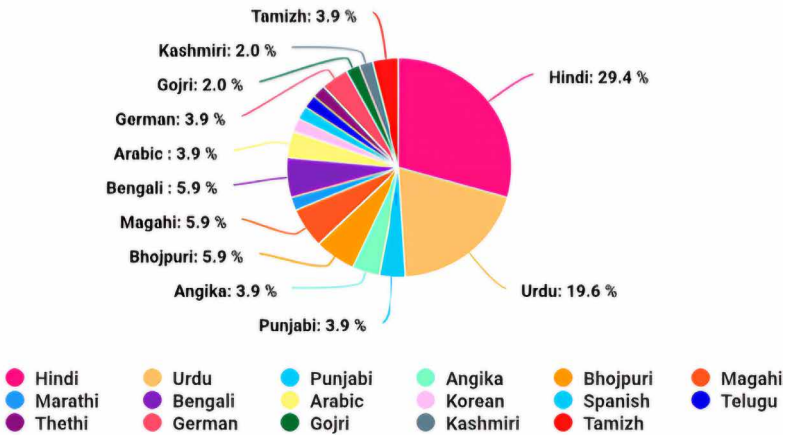

Figure 1. Participant's mother tongue

The distribution of participants' mother tongues is shown in Figure 1, underscoring the wide linguistic scope of the study. The chart shows that the majority of respondents, or 29.4% of the sample as a whole, identified Hindi as their mother tongue, with Urdu coming in second at 19.6%. With significant regional representation, other noteworthy language groups include Bengali (5.9%), Bhojpuri (5.9%), and Magahi (5.9%). Kashmiri and Gojri made up 2.0% of the sample, while smaller groups such as Arabic, German, Tamizh, Angika, and Punjabi each made up 3.9%. This combination of major Indian languages and lesser-spoken or minority languages highlights the importance of assessing how AI translation tools can promote inclusivity and cultural preservation across diverse mother-tongue communities and illustrates the wide-ranging linguistic diversity captured in the study.

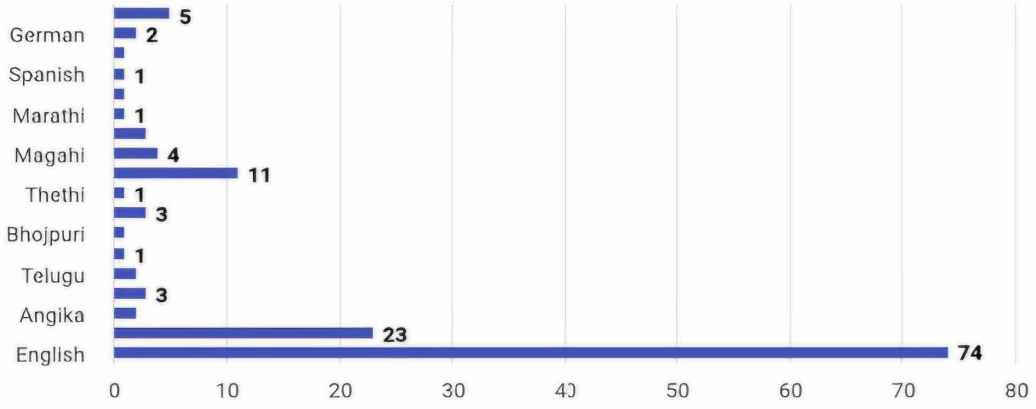

Figure 2. Another languaged spoke by the participants

The emphasis in Figure 2 switches to the other languages that participants speak in addition to their mother tongue. With 74 participants reporting proficiency, the data shows that English is by far the most frequently spoken additional language, highlighting its function as a bridging or global language within the group. With 23 participants speaking it, Angika is the next most common additional language. Thethi (11), German (5), Magahi (4), Bhojpuri (3), Telugu (3), and smaller numbers for Spanish and Marathi follow. The participants' remarkable multilingualism is demonstrated by this distribution, which supports the study's emphasis on how AI translation tools may improve interlanguage communication and promote cultural inclusivity.

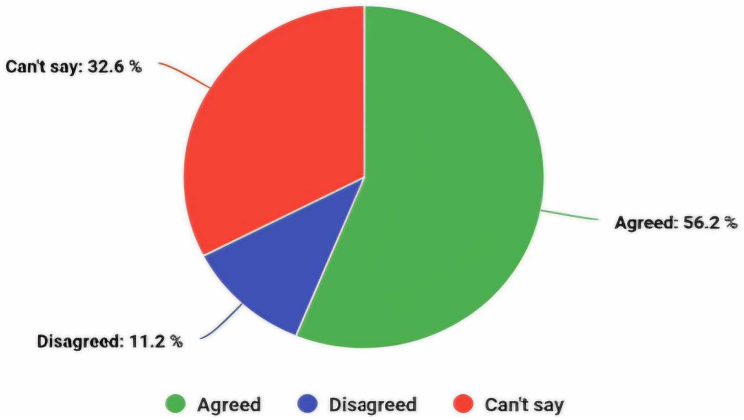

Figure 3. Participants perceptions toward AI-power translation tools for translating regional langguage

Participants' perceptions of how well AI-powered translation tools translate regional languages are examined in Figure 3. The majority, 56.2%, expressed confidence in the current capabilities of the technology by agreeing that these tools are effective. However, 11.2% disagreed, indicating skepticism or discontent with the accuracy provided, while 32.6% answered "Can't say," indicating uncertainty or insufficient exposure to the tools. In order to increase user awareness, education, and trust in AI translation technologies, these findings imply that although a large percentage of users have a favorable opinion, a significant portion are still unsure.

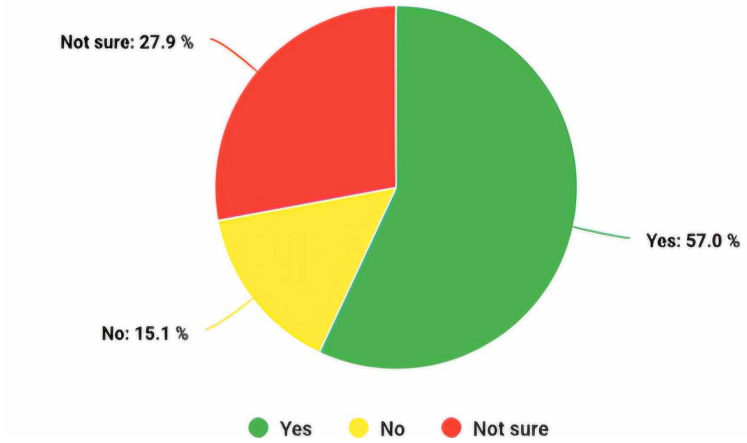

Figure 4. Participants' perceptions of the availability of regional language content as a result of AI translation

Participants' perceptions of the availability of regional language content as a result of AI translation tools are examined in Figure 4. The fact that 57.0% of respondents selected "Yes" indicates that these tools are, in fact, helping to increase linguistic representation online. On the other hand, 27.9% said they were "Not sure," indicating reluctance or ignorance about changes in the availability of content, and 15.1% said they had not noticed such an increase. This trend suggests that even though more than half of the participants see improvements, there is still a sizeable portion that is either unaware of these developments or does not have the necessary knowledge, which may indicate areas for outreach or awareness campaigns.

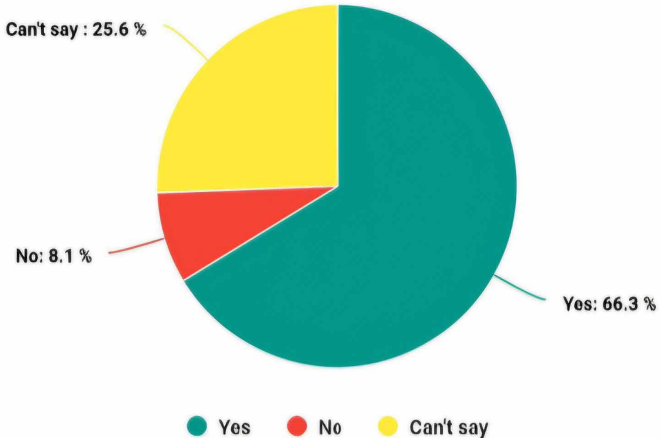

Figure 5. Participants' opinions on AI-powered translation tools fo reviving endangered languages

Participants' opinions on whether AI-powered translation tools have the potential to revive endangered languages are finally shown in Figure 5. A resounding 66.3% of respondents said "Yes," indicating hope for AI's potential contribution to the preservation of endangered languages. In contrast, 25.6% of respondents chose "Can't say," indicating uncertainty or insufficient knowledge to make a decision, and 8.1% chose "No," expressing doubts about the viability or efficacy of such efforts. The overall results show a high degree of optimism regarding the revolutionary potential of AI translation technologies, despite the percentage of respondents who expressed uncertainty. This suggests that further research, awareness campaigns, and proven results could boost confidence in this field even more.

Furthermore, an in-depth interview with an expert from the Bhashini Division of the Ministry of Electronics and Information Technology allowed further qualitative study. This specialist gave insightful analysis of the strategic objectives, technical procedures, and larger vision behind the Bhashini project, as well as information on how national development of AI-powered translating tools is underlined. Important topics, including the integration of all scheduled Indian languages, the future inclusion of regional and endangered languages, the function of crowdsourced data in building strong language models, and the

rigorous validation mechanisms in place to ensure translation accuracy and quality, were discussed in the interview. This qualitative element gave the quantitative results more context and depth, so providing a better knowledge of the technical possibilities as well as the socio-cultural relevance of artificial intelligence translating tools in India's changing digital scene.

An expert from the Ministry of Electronics and Information Technology's Bhashini Division described the app's lofty objectives, saying it is a national endeavor to preserve India's rich linguistic legacy rather than merely a translation tool. The expert underlined that the app actively works to preserve and strengthen India's languages while facilitating smooth translation across the nation's diverse linguistic landscape. The expert stated in the interview that the app is designed to "*include all scheduled languages for translation into various other languages*," highlighting a vision based on three main themes: language preservation, linguistic expansion, and inclusive translation.

In order to ensure that no community is left behind in the digital age, the expert went on to say that the app's immediate objective is "*to provide translation for all scheduled Indian languages*." In addition to providing technical functionality, Bhashini promotes cultural inclusivity by incorporating all scheduled languages into its extensive translation framework. According to the expert, this inclusivity is essential because it provides people all over India with useful communication tools that honor and reflect the country's linguistic and cultural diversity.

The expert also emphasized that the app's long-term goals go beyond simple translation; by promoting regional variations and underrepresented dialects, which are frequently in danger of being marginalized or extinct, it hopes to promote linguistic expansion. In order to prevent the digital ecosystem from favoring dominant languages over smaller, more vulnerable ones, the expert stated, "*We are working towards incorporating lesser-spoken languages and dialects in future phases*."

The expert framed language preservation as both a cultural priority and a national imperative, highlighting the close connection between it and Bhashini's broader mission. Bhashini is a unique combination of innovation and cultural stewardship, integrating preservation efforts directly into its technological framework to ensure that India's diverse and expansive linguistic mosaic endures well into the digital era.

The expert noted that "*at the moment, it offers translations for all scheduled languages and aims to expand to include regional languages from India*," addressing the urgent need for greater linguistic inclusion. This planned expansion is a strategic commitment to capturing the diverse array of regional dialects and local languages that are the foundation of India's cultural identity, rather than just a technical improvement. The developers are committed to improving the app's features to give regional and lesser-known languages a prominent and respected online presence because they are well aware that many of these languages are underrepresented in digital spaces.

The expert further mentioned that "*the company's long-term goal is to incorporate all regional languages and to revitalize endangered languages in India*." This innovative strategy makes Bhashini more than just a translation tool; it turns into an essential tool for preserving the country's linguistic legacy. The urgent need to ensure that India's distinctive and delicate linguistic legacy is passed down to future generations is met by Bhashini's prioritization of language preservation in addition to digital innovation. By doing this, the app hopes to establish a digital ecosystem that celebrates and preserves linguistic diversity, reinforcing the nation's rich cultural heritage in a world growing more interconnected by the day.

DISCUSSION

The main goal of this study was to assess how well AI-powered translation tools work to revive India's endangered languages. It also aimed to evaluate how accurate the translations produced by these

tools were in terms of context. The study came to the conclusion that platforms like Google Translate and Bhashini have a great deal of potential to help revive endangered Indian languages (Ermolova et al., 2024; Kavitha et al., 2023; Nanduri & Bonsignore, 2023). The Bhashini app's primary objective is to incorporate all Indian languages into its translation framework. After completing this fundamental task, the app will concentrate on initiatives to revive endangered languages (Dayanand et al., 2023). This demonstrates how crucial AI-powered translation technologies are anticipated to be in maintaining and protecting India's linguistic diversity in the years to come.

Since these tools are still in the early stages of development, their main goal is still to guarantee that all scheduled and regional Indian languages are fully included. After this integration is finished, the resources will be in a better position to focus on reviving endangered languages, and they might even help bring extinct ones back to life. Furthermore, by providing translations for regional content posted on social media, these tools greatly expand the content's audience among non-native speakers, assisting in their assimilation into the cultural mainstream. Although it is true that these tools currently struggle to provide translations that are contextually accurate, it is crucial to remember that they are still in their infancy and that ongoing advancements and improvements are anticipated in the future.

This research found important trends about the effects of AI-powered translation tools on various age groups and user backgrounds by analyzing primary data collected through a thorough questionnaire. Students made up a large portion of the sample (72.1%), with the majority of responders (59.3%) being between the ages of 21 and 25. The high level of engagement among younger, tech-savvy people—who frequently adopt digital innovations like AI-driven language technologies early on—is reflected in this demographic distribution.

One of the survey's main conclusions was that 61.6% of respondents think AI translation tools actively support regional language preservation by increasing their visibility and accessibility online. The availability and reach of regional content online has also increased noticeably, according to 57% of respondents, which suggests that these tools are contributing to the amplification of linguistic diversity across media platforms. The study also reveals that most respondents think these tools are useful, with 48.8% explicitly stating that the translations were appropriate for the context and satisfied user needs.

The interview with Expert 1 added vital qualitative depth to these survey findings, bolstering the study's hypotheses even more. In order to ensure inclusivity across the nation's linguistic spectrum, the expert described the methodical integration of all 22 scheduled Indian languages into platforms such as the Bhashini app. The expert also explained how crowdsourcing methods are being used to create extensive and reliable language datasets, with input from a variety of Indian speakers enhancing the platform's linguistic capabilities. The expert stressed the significance of stringent validation and quality assurance procedures, which are intended to sustain high standards of translation accuracy and dependability as the system develops further.

When combined, the research's conclusions highlight the fact that AI-powered translation tools are much more than simply practical communication tools; they are becoming important tools for advancing digital inclusivity, fostering cultural representation, and preserving India's rich linguistic legacy. In an increasingly interconnected world, these technologies are helping to counteract the marginalization of less dominant languages by increasing the visibility, accessibility, and usability of regional and minority languages in digital spaces. This ensures that diverse linguistic traditions continue to be alive and relevant.

Importantly, the study indicates that although existing tools have certain drawbacks, especially in areas like cultural sensitivity and contextual nuance, further development of these tools has great potential for the future. These platforms have the potential to significantly impact the revitalization of languages that are in danger of declining or going extinct as they grow to include not only scheduled languages but also endangered and lesser-known regional dialects. By doing this, AI translation tools have the

potential to be vital allies in the larger effort to protect and enhance India's distinctive linguistic mosaic, guaranteeing that it will continue to exist for many generations to come as a living, breathing part of the country's collective identity in addition to being a cultural treasure.

The research's timeframe imposes certain constraints that may limit the ability to generalise the findings. Also, each language has a unique idiomatic expression, slang, and culturally embedded meanings. This limits AI-powered translation tools from capturing the complexities of language. With varied linguistic diversity capturing languages across the country, some languages, especially those from the Western regions, remain underexplored. One limitation of this study is that the analysis of the interview conducted with expert 1, which was intended to be performed using the NVivo app, could not be completed due to time constraints.

In light of the dominance of the languages spoken by the majority or dominant group, this research is significant because it examines how AI-powered translation tools can preserve and improve the rich linguistic diversity. This study attempts to demonstrate how these AI-powered translation tools can close communication gaps and promote inclusivity in a multilingual environment by examining how they affect regional language promotion and preservation. This study also seeks to comprehend how AI-powered language translation tools can reduce the likelihood of language loss. Also, knowing the advantages and difficulties of AI translation tools will help technologists, educators, and policymakers make sure that technological advancements help preserve India's linguistic legacy.

Additionally, the study had a number of shortcomings that should be noted. Advanced qualitative analysis tools like NVivo, which could have offered a more in-depth analysis of the interview data, were not used due to time constraints. Furthermore, some parts of India, especially those linked to languages that are less well-documented or seriously endangered, are still underrepresented in AI language datasets. These discrepancies limit how broadly the study's conclusions can be applied and emphasize the necessity of a more inclusive, varied research methodology in subsequent studies.

Future studies should focus on a wider range of languages, particularly those that are endangered or do not have a strong enough online presence. Involving linguists, cultural historians, and local communities in the design, development, and validation of AI models would guarantee greater cultural and contextual fidelity in addition to improving the technical accuracy of translations. Advanced qualitative and ethnographic techniques could enhance future research by providing more in-depth understandings of user experiences, community sentiments, and the wider sociocultural ramifications of these technologies.

This study concludes by highlighting the revolutionary potential of AI-powered translation tools in conserving and reviving India's linguistic legacy. Current tools are difficult to use, especially when it comes to obtaining contextual accuracy and full language coverage, but their growing popularity and intentional, inclusive development initiatives make them essential tools in the global fight against language loss. In order to guarantee that every language, regardless of how few speakers it has, finds a safe and permanent place in the digital future, it is critical that advancements in AI translation be driven by inclusive policies, moral behavior, and cultural sensitivity.

CONCLUSION

This study set out to explore the role and effectiveness of AI-powered translation tools in the revitalisation of endangered languages in India, with a specific focus on tools such as Bhashini and Google Translate. The research reveals that while these technologies are still in their developmental stages, they already demonstrate significant potential in supporting linguistic diversity, digital inclusion, and cultural preservation. The study's findings, supported by both primary data and expert insights, indicate that AI translation tools are increasingly recognised by users as valuable resources for preserving and promoting regional languages in a rapidly digitising society. One of the central observations emerging from the data is that a substantial proportion of users—particularly students and young adults—perceive these tools

as instrumental in the preservation of their linguistic heritage. This not only reflects growing awareness of language endangerment but also suggests a shift in how regional languages are being consumed and valued in the digital realm. The translation of regional content, especially on social media platforms, is helping to bring once-marginalised languages into mainstream discourse, thus increasing their visibility and relevance among younger audiences. Moreover, the tools' ability to facilitate communication across linguistic barriers is fostering a greater sense of inclusivity in a multilingual society. Insights drawn from expert interviews further reinforce the idea that AI-powered translation tools are being developed with a long-term, strategic vision. The planned integration of all 22 scheduled languages, along with the use of crowdsourcing and validation frameworks, highlights the commitment to ensuring accuracy, quality, and representation. However, the findings also underscore some critical limitations that currently hinder the full potential of these tools. AI systems still struggle to grasp the cultural nuances, idiomatic expressions, and localised slang that are intrinsic to every language. This gap is particularly significant in the context of India's vast linguistic diversity, where each language carries deeply rooted socio-cultural contexts that are not easily captured by automated systems.

REFERENCES

Aggarwal, D., Sharma, D., & Saxena, A. B. (2024). Exploring the role of AI for enhancement of social media marketing. *Journal of Media Culture and Communication*, *45*, 1–11. https://doi.org/10.55529/jmcc.45.1.11

Arora, A., Farris, A., Basu, S., & Kolichala, S. (2022). Computational historical linguistics and language diversity in South Asia. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2203.12524

Chaudhari, V., & Bhangale, S. K. (2024). Enhancing engagement and communication using artificial intelligence on social media. *International Journal of Education Modern Management Applied Science & Social Science*, *06*(04(I)), 35–39. https://doi.org/10.62823/ijemmasss/6.4(i).6947

Dayanand, N. S. A., Devi, N. M. U., & Kumar, N. R. (2023). Digital archives and preservation techniques for revitalising endangered languages. *Shanlax International Journal of English*, *12*(S1-Dec), 174–182. https://doi.org/10.34293/rtdh.v12is1-dec.133

Ermolova, T. V., Savitskaya, N. V., Dedova, O. V., & Guzova, A. V. (2024). The potential of artificial intelligence in preserving linguistic diversity for future generations. *Science for Education Today*, *14*(1), 184–198. https://doi.org/10.15293/2658-6762.2401.09

Ghose, A., Bharadwaj, S., & Ali, S. M. A. (2024). Linguistic Genocide of minority and mother tongue Languages: Unravelling international implications on Indian laws through a critical discourse. *Journal of Asian and African Studies*. https://doi.org/10.1177/00219096241243058

Gupta, P. (2009). Endangered Languages: Some concerns. *Economics and Political Weekly*, *XLIV*(32). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2123899

Hagerty, A., & Rubinov, I. (2019, July 18). *Global AI Ethics: A review of the social impacts and ethical implications of artificial intelligence*. arXiv.org. https://arxiv.org/abs/1907.07892

Jolad, S., & Agarwal, A. (2024). Mapping India's linguistic diversity and exclusion in the Indian Census. In *Practices of Digital Humanities in India*. Routledge. https://doi.org/10.4324/9781003325239

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. https://doi.org/10.18653/v1/2020.acl-main.560

Kavitha, R., Fathma, S. S., & Vinitha, V. (2023a). AI: a catalyst for language Preservation in the digital era. *Shanlax International Journal of English*, *12*(S1-Dec), 102–106. https://doi.org/10.34293/rtdh.v12is1-dec.39

Kavitha, R., Fathma, S. S., & Vinitha, V. (2023b). AI: a catalyst for language Preservation in the digital era. *Shanlax International Journal of English*, *12*(S1-Dec), 102–106. https://doi.org/10.34293/rtdh.v12is1-dec.39

Kavitha, R., Fathma, S. S., & Vinitha, V. (2023c). AI: a catalyst for language Preservation in the digital era. *Shanlax International Journal of English*, *12*(S1-Dec), 102–106. https://doi.org/10.34293/rtdh.v12is1-dec.39

Korom, F. J., & Abbi, A. (1999). Languages of tribal and Indigenous peoples of India: the Ethnic space. *Journal of the American Oriental Society*, *119*(2), 354. https://doi.org/10.2307/606144

Nanduri, D. K., & Bonsignore, E. M. (2023). Revitalizing Endangered Languages: AI-powered language learning as a catalyst for language appreciation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.09394

Niveditha, K. P. (2023). AI in Social Media: Navigating the Balance between User Engagement and Misinformation. *International Journal of Innovative Research in Computer and Communication Engineering*, *12*(07), 5250–5257. https://doi.org/10.15680/ijircce.2024.1207107

Park, S. (2024). AI chatbots and linguistic injustice. *Journal of Universal Language*, *25*(1), 99–119. https://doi.org/10.22425/jul.2024.25.1.99

Peterson, J. (2015). Introduction – advances in the study of Munda languages. *Journal of South Asian Languages and Linguistics*, *2*(2), 149–162. https://doi.org/10.1515/jsall-2015-0008

Ravindra, N. V., N, N. D. K., & Raj, N. P. V. (2024). A Systematic survey of Multilingual speech transcription and Translation. *International Journal of Advanced Research in Science Communication and Technology*, 517–523. https://doi.org/10.48175/ijarsct-15369

Ray, N. S. (2024a). Exploring the role of artificial intelligence in language documentation and endangered language preservation. *Tuijin Jishu/Journal of Propulsion Technology*, *45*(02), 1630–1639. https://doi.org/10.52783/tjjpt.v45.i02.6122

Ray, N. S. (2024b). Exploring the role of artificial intelligence in language documentation and endangered language preservation. *Tuijin Jishu/Journal of Propulsion Technology*, *45*(02), 1630–1639. https://doi.org/10.52783/tjjpt.v45.i02.6122

Romaine, S. (2015). The global extinction of languages and its consequences for cultural diversity. In *Multilingual education* (pp. 31–46). https://doi.org/10.1007/978-3-319-10455-3_2

Shackle, C. (2001). Speakers of South Asian languages. In *Cambridge University Press eBooks* (pp. 227–243). https://doi.org/10.1017/cbo9780511667121.016

Shapiro, M. C., & Breton, R. J. (1999). Atlas of the Languages and Ethnic Communities of South Asia. *Journal of the American Oriental Society*, *119*(3), 495. https://doi.org/10.2307/605944

Sinha, R. M. K. (2009). *Indian National Translation Mission: Need for Integrating Human-Machine Translation*. ACL Anthology. https://aclanthology.org/2009.mtsummit-government.6/

Vijayalakshmi, B., Deshpande, K. A., & Krishan, N. (2024). Use of AI to Augment Multilingual Content in Cyberspace for Development – An Indian Case Study. In *IEEE Global Humanitarian Technology Conference (GHTC)* (pp. 31–35). IEEE. https://doi.org/10.1109/ghtc62424.2024.10771552

Wang, L. (2024). Artificial intelligence's role in the realm of endangered languages: Documentation and teaching. *Applied and Computational Engineering*, *48*(1), 123–129. https://doi.org/10.54254/2755-2721/48/20241249

Warwick, K. (2013). Artificial Intelligence: The Basics. In *Routledge eBooks*. https://doi.org/10.4324/9780203802878

Zaki, M. Z., & Ahmed, U. (2024). Bridging linguistic divides: The impact of AI-powered translation systems on communication equity and inclusion. *Journal of Translation and Language Studies*, *5*(2). https://doi.org/10.48185/jtls.v5i2.1065

Zulfiqar, N., Shaikh, B. A., Shah, S. A., & Awan, A. (2024). Investigating how AI Technologies Influence Changing Social Norms and Behaviors on Social Media, Particularly in Areas like Communication and Self-Presentation. *Review of Education Administration and Law*, *7*(4), 167–183. https://doi.org/10.47067/real.v7i4.371