Preserving Cultural Identity with Context-Aware Translation Through Multi-Agent AI Systems

Mahfuz Ahmed Anik[1] **Abdur Rahman**[1], **Azmine Toushik Wasi**[1†], **Md Manjurul Ahsan**[2]
[1]Shahjalal University of Science and Technology, Sylhet, Bangladesh
[2]University of Oklahoma, Norman, OK 73019, USA
[†]**Correspondence:** azmine32@student.sust.edu

Abstract

Language is a cornerstone of cultural identity, yet globalization and the dominance of major languages have placed nearly 3,000 languages at risk of extinction. Existing AI-driven translation models prioritize efficiency but often fail to capture cultural nuances, idiomatic expressions, and historical significance, leading to translations that marginalize linguistic diversity. To address these challenges, we propose a multi-agent AI framework designed for culturally adaptive translation in underserved language communities. Our approach leverages specialized agents for translation, interpretation, content synthesis, and bias evaluation, ensuring that linguistic accuracy and cultural relevance are preserved. Using `CrewAI` and `LangChain`, our system enhances contextual fidelity while mitigating biases through external validation. Comparative analysis shows that our framework outperforms GPT-4o, producing contextually rich and culturally embedded translations—a critical advancement for Indigenous, regional, and low-resource languages. This research underscores the potential of multi-agent AI in fostering equitable, sustainable, and culturally sensitive NLP technologies, aligning with the AI Governance, Cultural NLP, and Sustainable NLP pillars of Language Models for Underserved Communities. Our full experimental codebase is publicly available at: github.com/ciol-researchlab/Context-Aware_Translation_MAS.

1 Introduction

Language is a vital cultural repository, transmitting traditions, values, and historical narratives across generations. It preserves oral traditions, folklore, and indigenous knowledge, shaping a community's worldview and identity (Goel). However, globalization, urbanization, and the dominance of English have led to an alarming decline in linguistic diversity, with nearly 3,000 languages projected to disappear this century (Kandler and Unger, 2023). This loss severs communities from their heritage, weakens intergenerational transmission, and marginalizes minority identities. Despite growing awareness, traditional preservation methods remain inadequate; documentation efforts fail to capture cultural complexity, while machine translation distorts contextual meaning (Hutson et al., 2024). The digital linguistic divide further excludes underrepresented languages, limiting their digital presence and corpus availability (Bella et al., 2023). Additionally, economic pressures favor dominant global languages, leading younger generations to abandon their native tongues. While technological advancements offer potential solutions, current approaches often focus on efficiency rather than cultural authenticity, overlooking the need for linguistic preservation beyond translation (Mufwene, 2005). As AI-driven methods become central to language processing, it is essential to rethink how

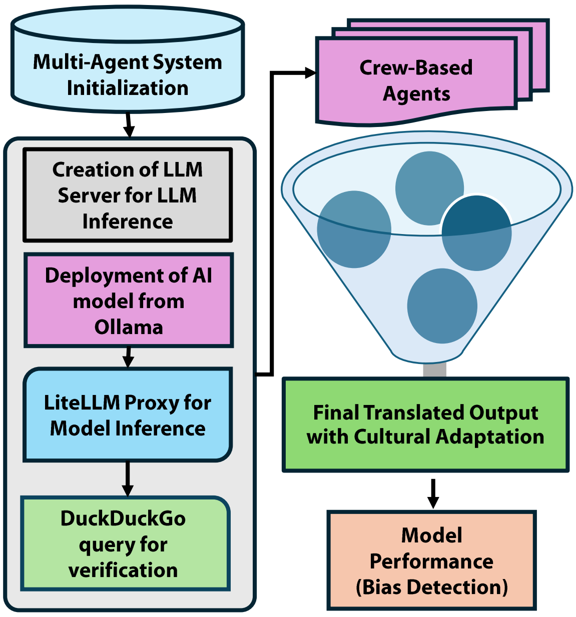

Figure 1: Our Approach for Preserving Cultural Identity with Context-Aware Translation Through Multi-Agent AI Systems

these systems can adapt to cultural and contextual complexities rather than replace them.

The shortcomings of existing AI-driven language models highlight the urgent need for a more culturally aware and linguistically inclusive approach. Traditional machine translation systems, while effective in word-to-word conversion, often fail to retain cultural and historical depth, with up to 47% of contextual meaning lost in conventional translations (Tian et al., 2022). This challenge is particularly significant for tonal languages, oral traditions, and indigenous dialects, where subtleties are essential for accurate interpretation. Additionally, the dominance of English-centric AI models reinforces linguistic hierarchies, marginalizing lesser-known languages and limiting their digital accessibility (Lepp and Sarin, 2024). Compounding this issue, AI trained primarily on Western linguistic paradigms struggles to handle dialectal diversity, non-standardized orthographies, and tonal complexity, making it unsuitable for many underrepresented languages (Kshetri, 2024; Romanou et al., 2024). Beyond technological constraints, globalization and socio-economic shifts further accelerate language endangerment, as younger generations increasingly prioritize global languages over ancestral ones (Garg, 2024). These challenges necessitate a shift from isolated, monolithic AI models to collaborative, multi-agent AI systems capable of not just translation but interpretation, synthesis, and evaluation through a cultural lens (Jones et al., 2025). By integrating context-aware translation, multimodal AI, and real-time bias detection, an innovative AI-driven linguistic framework can bridge these gaps and establish a more sustainable, culturally embedded approach to language preservation.

To address these challenges, we propose a Multi-Agent AI Framework for Cross-Language Understanding, designed to enhance the linguistic, cultural, and ethical integrity of machine translations, as outlined in Figure 1. Unlike traditional NLP models, which process translation in a linear and isolated manner, our framework orchestrates multiple AI agents that collaboratively refine linguistic and cultural adaptation at different stages. The Translation Agent ensures grammatical accuracy, while the Interpretation Agent enriches outputs by embedding historical, social, and contextual markers. The Content Synthesis Agent structures the final output, preserving idiomatic expressions, ceremonial speech, and linguistic variations for readability and coherence. Finally, the Quality and Bias Evaluation Agent mitigates distortions by cross-referencing historical data, detecting biases, and ensuring fairness through real-time validation mechanisms such as DuckDuckGo search integration.

Our collaborative AI system, developed using `CrewAI` and `LangChain`, is powered by `Aya-Expanse:8b` (Dang et al., 2024) via `Ollama`, with `LiteLLM` proxying to optimize model efficiency. By leveraging this multi-agent approach, our framework bridges the gap between low-resource language communities and high-performance NLP models, offering a scalable, ethically responsible, and culturally sensitive solution. Furthermore, this paradigm not only enhances translation quality but also provides a foundation for digital language preservation, ensuring that linguistic heritage remains accessible and relevant in the AI-driven era. Our work contributes to sustainable NLP development by promoting equitable access to AI technologies, aligning with the broader mission of inclusive and ethical AI for global linguistic diversity.

2 Related Work

The preservation of linguistic diversity and cultural heritage has been a growing research focus, with studies exploring both traditional methods and AI-driven computational techniques. Early efforts emphasized community-driven documentation, while modern advancements leverage machine translation, generative AI, and multimodal learning to enhance language sustainability.

2.1 Cultural Language Preservation

Traditional language preservation often relies on linguistic documentation and community-driven efforts. Nekoto et al. introduced a participatory translation approach to enhance neural machine translation (NMT) for under-resourced languages, fostering greater involvement from native speakers. Miyagawa (2024) developed a bi-directional translation system specifically for Ainu, addressing its unique linguistic structure and revitalizing the language's usage in modern contexts. Louadi (2024) emphasized the importance of diverse and inclusive datasets to reduce biases in AI applications, particularly in language preservation. Hutson et al. (2024) proposed scalable AI models to promote the use of mother tongues, enhancing cultural identity and continuity. Furthermore, Nanduri and Bonsignore (2023) explored AI-powered language

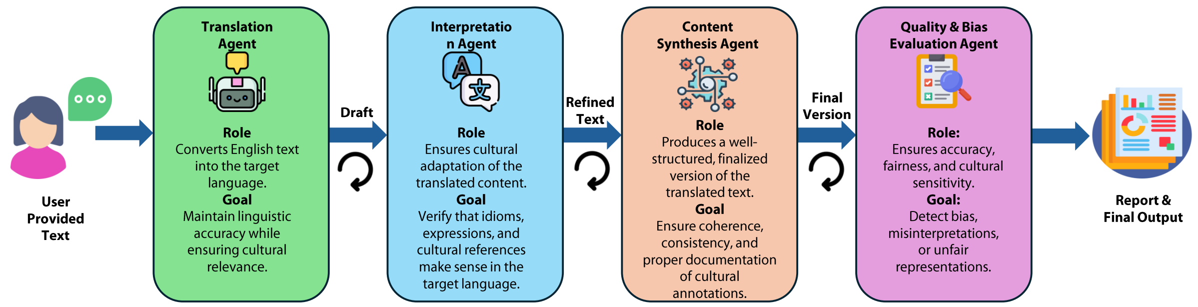

Figure 2: Our Workflow of Context-Aware Translation ThroughMulti-Agent AI Systems

learning tools, including bilingual storybooks and VR simulations, that not only support language acquisition but also promote cultural appreciation and ethical practices in the preservation process.

2.2 AI and Computational Techniques for Language Preservation

With the rapid advancements in AI and deep learning, researchers have increasingly explored machine learning, generative AI, and multimodal techniques for language revitalization. Bizan bin Ghowar (2023) applied AI to heritage analysis and NLP-driven historical text processing, aiming to preserve linguistic traditions through computational tools. Similarly, Liu et al. (2024) examined generative AI's potential in preserving ancient texts and facilitating multimodal research, highlighting its value in enhancing the accessibility of historical languages. However, Putri et al. (2024) pointed out that while LLMs can generate syntactically coherent text, they often fail to capture the cultural depth and contextual accuracy crucial for low-resource languages. This reveals a significant limitation in generative models, where AI systems lack the cultural nuances and real-world understanding necessary for effective language preservation. Further addressing this gap, Myung et al. (2024) introduced the BLEND benchmark to assess LLMs' cultural knowledge across multiple languages, revealing substantial performance discrepancies for underrepresented cultures. In response to these challenges, AJUZIEOGU (2024) proposed a multimodal generative AI framework for African language documentation, integrating neural architectures with community-driven approaches to mitigate the impact of data scarcity. While these studies highlight the potential of AI in language revitalization, they also underscore ongoing challenges in achieving true cultural adaptation and contextual accuracy, particularly in the face of limited and diverse datasets. This calls for more nuanced, culturally-aware AI frameworks that can bridge these gaps and offer robust solutions for underrepresented languages.

Existing AI models struggle with cultural depth, linguistic bias, and adaptability, often reinforcing English-centric hierarchies while failing to integrate underrepresented languages. Current LLM approaches lack collaborative, multi-agent frameworks, limiting contextual adaptation and ethical oversight. Our work distinguishes itself from existing research by introducing a multi-agent AI framework that specifically addresses the cultural and contextual shortcomings of traditional AI-driven translation models. While previous efforts, such as those by Nekoto et al. and Louadi (2024), have focused on improving language preservation through community-driven or single-agent AI approaches, our framework incorporates specialized agents—Translation, Interpretation, Content Synthesis, and Quality and Bias Evaluation. Our multi-agent framework enhances linguistic accuracy and cultural relevance, addressing the complexities of low-resource languages and idiomatic expressions. By using iterative cross-validation with external sources like DuckDuckGo, we mitigate biases and ensure cultural fidelity, outperforming traditional LLMs. This approach offers a novel, inclusive solution for language revitalization and preservation, overcoming the limitations of prior models.

3 Methodology

This section presents the design and implementation of our Multi-Agent AI Framework for Cross-Language Adaptation, focusing on system architecture, agent roles, and the translation refinement process.

3.1 System Overview

Our framework operates on a multi-agent architecture, leveraging `CrewAI` (Duan and Wang, 2024) for task delegation and collaboration. We employ Aya

Expanse 8B, an open-weight multilingual LLM from Cohere for AI, optimized through data arbitrage, multilingual preference training, safety tuning, and model merging (Dang et al., 2024). Aya Expanse 8B excels in 23 languages, ensuring robust cross-language performance. We integrate the `LiteLLM` proxy for optimized inference and use DuckDuckGo search for real-time external validation, allowing for cultural and contextual verification (Saravanos et al., 2022; Agarwal et al., 2024). The model undergoes 3-5 training epochs for general adaptation tasks and 10 epochs for fine-tuning on low-resource languages [1]. Our agents operate sequentially, with each module processing the text iteratively to refine grammatical accuracy, cultural fidelity, and bias mitigation. The system follows a task delegation structure, where each agent contributes to refining the output until it meets contextual and ethical standards. Our full experimental codebase is publicly available at: `github.com/ciol-researchlab/Context-Aware_Translation_MAS`.

3.2 Agent Crew for Linguistic Transformation

Our framework utilizes four autonomous agents, each designed to address specific aspects of the translation process. Each agent operates independently, contributing its specialized task to ensure a culturally adaptive and linguistically accurate translation. The agents are designed to work sequentially with possible back and forth if required, with each task building upon the previous one to refine and enhance the output. Below, we describe the purpose, goals, and design of each of these agents. Table 1 provides a brief description of the agents.

3.2.1 Translation Agent

The Translation Agent is responsible for converting the source text from English into the target language while ensuring syntactic correctness and linguistic precision. This agent utilizes Neural Machine Translation (NMT) techniques to generate a raw translation that preserves the meaning of the original content. The goal of this agent is to ensure that the translation remains grammatically accurate, following the rules and structure of the target language. To achieve this, the agent leverages large-scale pre-trained models and context-aware mechanisms to produce an initial, linguistically sound translation. By allowing delegation, the Translation Agent can also pass its output to other specialized agents for further refinement, ensuring that the translation process is adaptable and efficient.

3.2.2 Interpretation Agent

The Interpretation Agent's primary purpose is to ensure that the translated content is culturally relevant and meaningful in the target language. This agent focuses on adapting idioms, expressions, cultural references, and regional nuances to make the translation more natural and appropriate for the target audience. Its goal is not merely linguistic accuracy but the cultural adaptation of the text, ensuring that humor, traditions, and local contexts are accurately conveyed. The agent uses contextual understanding and cultural knowledge to evaluate the translation and make necessary changes. Allowing delegation here means the agent can pass its results to other agents for further analysis or validation, which is essential for complex linguistic tasks involving culture.

3.2.3 Content Synthesis Agent

The Content Synthesis Agent plays a pivotal role in structuring the translated text into its final, polished form. Its responsibility is to ensure that the translation reads coherently and fluently while preserving both linguistic accuracy and cultural authenticity. This agent organizes the text logically, ensuring that sentences and paragraphs flow smoothly and that the structure aligns with the conventions of the target language. Additionally, the Content Synthesis Agent integrates cultural annotations and decisions made by the Interpretation Agent, making the translated content not only readable but also reflective of the cultural and linguistic choices made throughout the process. This agent's design does not allow delegation, ensuring it holds the final responsibility for the presentation of the output.

3.2.4 Quality and Bias Evaluation Agent

The Quality and Bias Evaluation Agent is tasked with performing a thorough review of the translated text to detect any issues related to fairness, accuracy, or cultural sensitivity. This agent's role is to ensure that the translation upholds ethical standards by checking for bias or misrepresentation of cultural elements. It cross-references the translated content with external sources, such as DuckDuckGo, to validate the factual accuracy of cultural references and check the translation against real-

[1] Upon acceptance, we will release the full working code as an open-source project, ensuring transparency, reproducibility, and broader accessibility for researchers and developers.

Table 1: Roles and Capabilities of Different Agents in Our System

Agent	Goal	Task Delegation	Back story	Tool Capability
Translation Agent	Translate English text into another language while maintaining cultural essence.	Can delegate task	You are a linguistic expert. Your job is to translate English text into the target language while ensuring cultural relevance and accuracy	None
Interpretation Agent	Ensure that cultural references in the translation are correctly adapted to the target language.	Can delegate task	You specialize in cultural adaptation. Your task is to ensure that idioms, expressions, and references in the translated text are meaningful and accurate in the target culture.	None
Content Synthesis Agent	Create a well-structured, final version of the translated text with cultural annotations.	Can not delegate task	You produce a final, structured version of the translated text. This includes annotations on cultural adaptations and linguistic decisions.	None
Quality & Bias Evaluation Agent	Ensure the translation is accurate, fair, and culturally sensitive.	Can not delegate task	You are a quality assurance expert. Your job is to check for accuracy, fairness, and cultural sensitivity in the translated text.	Web Search Tool

world contexts. This agent helps identify potential errors or distortions that might arise during the translation process, ensuring the final output is both accurate and free of harmful bias. By not allowing delegation, this agent ensures that no oversight occurs in the final evaluation phase.

In summary, the four agents in our framework work collaboratively to ensure that translations are linguistically accurate, culturally relevant, and contextually sensitive. Each agent brings a specialized skill set to the process, allowing for a seamless and adaptive translation workflow. By incorporating autonomous agents for each phase of translation, we ensure high-quality, culturally rich, and unbiased results. This workflow follows a sequential but dynamic structure, ensuring maximum accuracy and cultural fidelity. The Translation Agent first generates the raw translation, which is then refined by the Interpretation Agent to ensure cultural alignment. Once adapted, the Content Synthesis Agent organizes the text into a structured, reader-friendly format. Finally, the Quality and Bias Evaluation Agent verifies the correctness, fairness, and relevance of the translation using external sources. If any issue is detected, the translation is sent back to the responsible agent for revision. This iterative back-and-forth process ensures that the final output is not just a linguistically correct translation but also a culturally accurate and fair representation of the original text.

3.3 Iterative Translation Processing and Output Refinement

Our system follows a well-defined execution pipeline to ensure high-quality translations. First, users input a text, which is processed by the Translation Agent to ensure linguistic accuracy. The Interpretation Agent then steps in to adapt cultural references, idioms, and regional expressions, refining the translation for context. Next, the Content Synthesis Agent polishes the text, improving clarity and readability while maintaining coherence. The final step involves the Quality and Bias Evaluation Agent, which cross-validates the translation for accuracy, detects potential biases, and verifies cultural elements against reliable external sources. If inconsistencies or discrepancies are found, the system revisits the Content Synthesis Agent for necessary revisions before generating the final output. This iterative process ensures that the translation preserves cultural nuances, maintains linguistic precision, and upholds contextual relevance. By combining context-aware refinement with continuous validation, the system produces translations that are accurate, culturally sensitive, and fair, offering a balanced approach to multilingual communication.

4 Results and Findings

As there is no benchmark or evaluation framework available for most cultural translation aspects of low-resource models, we adopt a simple qualitative evaluation to assess our model's capability. Table 2 discusses the output quality of the model by assessing the translations generated by our multi-agent AI framework across three cultural contexts—Festival, Religion, and History—for three languages: Hindi, Turkish, and Hebrew. Table 3 presents a comparative analysis between our multi-agent AI framework and GPT-4, highlighting key differences in cultural preservation and contextual depth.

4.1 Evaluation of Model Output Across Cultural Contexts

Table 2 presents the outputs generated by our multi-agent AI framework, demonstrating its effectiveness in translating content across three cultural contexts—Festival, Religion, and History—in Hindi, Turkish, and Hebrew while ensuring cul-

Table 2: Model Outputs on Different Linguistic and Cultural Setups

Languages	Cultural Context	Input (English Text)	Output (Final Translated Text with Cultural Adaption)
Hindi	Festival	Diwali, India's grand festival of lights, symbolizes the victory of good over evil as homes are adorned with diyas (oil lamps), families perform Lakshmi Puja, and fireworks light up the sky in joyous celebration	दिवाली भारत का महान प्रकाश उत्सव है, जो अच्छाई की जीत का प्रतीक है। घरों को दीयों (तेल के दीपक) से सजाया जाता है, परिवार लक्ष्मी पूजा करते हैं और फूलों के धुएं से आसमान खुशियों से भर जाता है।
	Religion	Every evening, the sacred banks of the Ganga River come alive with the mesmerizing Ganga Aarti, where priests chant Vedic hymns, waving golden-lit lamps, and devotees offer floating diyas to seek divine blessings.	हर शाम, गंगा नदी के तट **मनमोहक** गंगा आर्ती से जीवंत हो उठते हैं, जहां पुजारी वेदिक मंत्रों का उच्चारण करते हुए सुनहरी लौ वाले दीपक लहराते हैं और भक्त अपनी मनोकामनाओं के लिए प्रतिष्ठित दीया (दिया) नदी में उतारते हैं।
	History	The Revolt of 1857, India's first war of independence, saw brave warriors like Rani Lakshmibai, Mangal Pandey, and Bahadur Shah Zafar rise against British rule, igniting the spirit of freedom across the nation	१८५७ का संघर्ष भारत के स्वतंत्रता संग्राम की शुरुआत थी। इस युद्ध में रानी लक्ष्मीबाई, मंगल पांडेय और बहादुर शाह ज़फ़र जैसे वीर स्वतंत्रता सेनानियों ने ब्रिटिश शासन के ख़िलाफ़ विद्रोह करके देश को आज़ादी की ओर अग्रसर करने में महत्वपूर्ण भूमिका निभाई। उनकी बहादुरी और बलिदान ने लाखों भारतीयों को प्रेरित किया और एक नई राष्ट्रीय पहचान का निर्माण किया।
Turkish	Festival	Nevruz, the ancient Turkish and Central Asian New Year, marks the arrival of spring with bonfire jumping, folk dances, and feasts featuring traditional dishes like Baklava and Pilav, symbolizing renewal and unity.	Nevruz, Türk ve Orta Asya'nın köklü Yeni Yılı, baharın gelişini ateşin üzerinde zıplayarak, halk danslarıyla ve geleneksel lezzetlerle dolu şölenlerle karşılıyor. Baklava ve Pilav gibi yemekler, yenilenme ve birlik sembolü olarak masalarımızı süslüyor.
	Religion	The mystical Sema Ceremony of the Whirling Dervishes, rooted in Sufi traditions, is a spiritual dance where dervishes spin in rhythmic meditation, symbolizing a connection between earth and the divine	Sufi geleneklerine dayanan ve dervişlerin ritmik meditasyonda döndükleri mistik Sema Töreni, bir ruhani dans şeklidir. Bu dans, dünyayı ve ilahi arasında bir bağın sembolüdür.
	History	The Fall of Constantinople in 1453, led by Sultan Mehmed II (Fatih Sultan Mehmet), marked the rise of the Ottoman Empire, blending Eastern and Western cultures and transforming Istanbul into a global center of trade, art, and religion	İstanbul'un 1453'te düşüşü, Sultan II. Mehmed (Fatih) tarafından başlatılan Osmanlı İmparatorluğu'nun yükselişine yol açtı. Doğu ve Batı kültürlerinin bir araya gelmesi, İstanbul'u ticaret, sanat ve din açısından küresel bir merkez haline getirdi.
Hebrew	Festival	Passover (Pesach) is a sacred Jewish festival celebrating the Exodus from Egypt, where families gather for the Seder meal, retelling the story of liberation, eating matzah (unleavened bread), and reciting traditional blessings.	במשפחות מתכנסות .הוא חג יהודי קדוש המציין את היציאה ממצרים (פסחא)פסוח ומקיימות (לחם לא מותסס)אוכלות מצה, מספרות סיפור השחרור ,לסעודה מסורתית ברכות מסורתיות.
	Religion	Shabbat, the weekly day of rest, begins at sunset on Friday with the lighting of Shabbat candles, the recitation of Kiddush (blessing over wine), and the sharing of Challah bread, as families come together in spiritual reflection.	שבת, יום המנוחה השבועי, מתחילה בשקיעה ביום שישי עם הדלקת נרות שבת, אמירת קדושה (ברכה על .היין), ואכילת לחם השבת, כשמשפחות מתאחדות ברפלקציה רוחנית
	History	in 70 CE, the Roman destruction of the Second Temple in Jerusalem marked a turning point in Jewish history, leading to the diaspora, the preservation of Jewish identity through Torah study, and the hope of one day returning to Zion.	ב-70 לספירה, חורבן המקדש השני בירושלים על ידי הרומאים סימן נקודת מפנה בהיסטוריה היהודית.שהובילה לגלות, שימור הזהות היהודית דרך לימוד התורה, ותקווה לשוב יום אחד לציון.

tural authenticity and contextual relevance. Unlike conventional translation models that prioritize direct linguistic conversion, our approach integrates cultural adaptation, refining grammatical precision, idiomatic expressions, and contextual depth. The Translation Agent ensures structural accuracy, while the Interpretation Agent adapts idiomatic phrases, religious references, and culturally significant expressions to enhance natural fluency and cultural immersion.

For instance, in the Hindi translation of Diwali, "grand festival of lights" becomes "mahaan prakaash utsav", emphasizing brilliance and festivity, while "victory of good over evil" is rendered as "acchai ki jeet", reinforcing the moral essence of the celebration. Cultural markers such as "Lakshmi Puja" remain unchanged, while "diyas" are translated as "deepak" to preserve their traditional significance. Similarly, in Turkish translations of Nevruz, "Bonfire jumping" is translated as "atesin uzerinde ziplamak", retaining its ritualistic importance, and "halk danslariyla" ensures the centrality of folk dances. Traditional foods such as Baklava and Pilav are adapted with idiomatic clarity, reinforcing their symbolic and cultural relevance. In the Sema Ceremony of the Whirling Dervishes, words like "mistik" and "ilahi" effectively capture its spiritual nature, ensuring linguistic and cultural accuracy. For Hebrew religious texts, "weekly day of rest" is translated as "yom hanukha shavu'i", emphasizing Shabbat's sacred nature, while "begins at sunset" becomes "matchila beshkia" to maintain traditional timing. Ritual elements such as "Shabbat candles", "Kiddush", and "Challah bread" are explicitly included, ensuring theological precision. The translation also preserves Shabbat's communal and reflective aspects, reinforcing its deeper spiritual meaning.

Beyond translation, the Content Synthesis Agent refines coherence and readability, while the Quality & Bias Evaluation Agent prevents cultural distortions and ensures historical accuracy. This is particularly crucial in historical translations, such as the 1857 Revolt and the Fall of Constantinople, where contextual and geopolitical precision is essential. The results confirm that multi-agent collaboration enhances cross-language understanding, producing translations that go beyond literal meaning to retain cultural significance. By integrating context-aware adaptation, idiomatic refinement, and external validation, the framework ensures culturally immersive and linguistically accurate translations. Table 2 validates the effectiveness of this approach, reinforcing its potential for AI-driven cultural preservation and cross-linguistic communication.

Table 3: Model Outputs Comparison with ChatGPT (GPT-4o)

English Text	Agent Translation	GPT Translation
Diwali, India's grand festival of lights, symbolizes the victory of good over evil as homes are adorned with diyas (oil lamps), families perform Lakshmi Puja, and fireworks light up the sky in joyous celebration.	दिवाली भारत का महान प्रकाश उत्सव है, जो अच्छाई की जीत का प्रतीक है। घरों को दीयों (तेल के दीपक) से सजाया जाता है, परिवार लक्ष्मी पूजा करते हैं और फूलों के धुएं से आसमान खुशियों से भर जाता है।	दीवाली, भारत का भव्य प्रकाश पर्व, बुराई पर अच्छाई की विजय का प्रतीक है, जब घर दीपों (तेल के दीयों) से सजाए जाते हैं, परिवार लक्ष्मी पूजा करते हैं, और आतिशबाज़ी आकाश में खुशियों की रोशनी बिखेरती है।
Every evening, the sacred banks of the Ganga River come alive with the mesmerizing Ganga Aarti, where priests chant Vedic hymns, waving golden-lit lamps, and devotees offer floating diyas to seek divine blessings.	हर शाम, गंगा नदी के तट **मनमोहक** गंगा आर्ती से जीवंत हो उठते हैं, जहां पुजारी वेदिक मंत्रों का उच्चारण करते हुए सुनहरी लौ वाले दीपक लहराते हैं और भक्त अपनी मनोकामनाओं के लिए प्रतिष्ठित दीया (दियाँ) नदी में उतारते हैं।	हर शाम, पवित्र गंगा नदी के तट गंगा आरती की मंत्रमुग्ध कर देने वाली छटा से जीवंत हो उठते हैं, जहाँ पुरोहित वेद मंत्रों का उच्चारण करते हुए स्वर्णिम दीप लहराते हैं, और भक्त दिव्य आशीर्वाद प्राप्त करने के लिए जल में तैरते दीप अर्पित करते हैं।
The Revolt of 1857, India's first war of independence, saw brave warriors like Rani Lakshmibai, Mangal Pandey, and Bahadur Shah Zafar rise against British rule, igniting the spirit of freedom across the nation.	१८५७ का संघर्ष भारत के स्वतंत्रता संग्राम की शुरुआत थी। इस युद्ध में रानी लक्ष्मीबाई, मंगल पांडेय और बहादुर शाह ज़फ़र जैसे वीर स्वतंत्रता सेनानियों ने ब्रिटिश शासन के ख़िलाफ़ विद्रोह करके देश को आज़ादी की ओर अग्रसर करने में महत्वपूर्ण भूमिका निभाई। उनकी बहादुरी और बलिदान ने लाखों भारतीयों को प्रेरित किया और एक नई राष्ट्रीय पहचान का निर्माण किया।	1857 का विद्रोह, भारत का पहला स्वतंत्रता संग्राम, बहादुर योद्धाओं जैसे रानी लक्ष्मीबाई, मंगल पांडे और बहादुर शाह ज़फर को ब्रिटिश शासन के खिलाफ उठते हुए देखा, जिसने पूरे राष्ट्र में स्वतंत्रता की भावना को प्रज्वलित किया।
Nevruz, the ancient Turkish and Central Asian New Year, marks the arrival of spring with bonfire jumping, folk dances, and feasts featuring traditional dishes like Baklava and Pilav, symbolizing renewal and unity.	Nevruz, Türk ve Orta Asya'nın köklü Yeni Yılı, baharın gelişini ateşin üzerinde zıplayarak, halk danslarıyla ve geleneksel lezzetlerle dolu şölenlerle karşılıyor. Baklava ve Pilav gibi yemekler, yenilenme ve birlik sembolü olarak masalarımızı süslüyor.	Nevruz, kadim Türk ve Orta Asya Yeni Yılı, baharın gelişini ateş üzerinden atlama, halk dansları ve Baklava ile Pilav gibi geleneksel yemeklerin yer aldığı şölenlerle kutlar, yenilenmeyi ve birliği simgeler.
The mystical Sema Ceremony of the Whirling Dervishes, rooted in Sufi traditions, is a spiritual dance where dervishes spin in rhythmic meditation, symbolizing a connection between earth and the divine.	Sufi geleneklerine dayanan ve dervişlerin ritmik meditasyonda döndükleri mistik Sema Töreni, bir ruhani dans şeklidir. Bu dans, dünyayı ve ilahi arasında bir bağın sembolüdür.	Mevlevi Dervişlerinin mistik Sema Töreni, Sufi geleneklerine dayanan ruhani bir danstır; dervişler ritmik bir meditasyon içinde dönerken, yeryüzü ile ilahi olan arasındaki bağı simgeler.
The Fall of Constantinople in 1453, led by Sultan Mehmed II (Fatih Sultan Mehmet), marked the rise of the Ottoman Empire, blending Eastern and Western cultures and transforming Istanbul into a global center of trade, art, and religion.	İstanbul'un 1453'te düşüşü, Sultan II. Mehmed (Fatih) tarafından başlatılan Osmanlı İmparatorluğu'nun yükselişine yol açtı. Doğu ve Batı kültürlerinin bir araya gelmesi, İstanbul'u ticaret, sanat ve din açısından küresel bir merkez haline getirdi.	1453 yılında Sultan II. Mehmed (Fatih Sultan Mehmet) liderliğinde gerçekleşen İstanbul'un Fethi, Osmanlı İmparatorluğu'nun yükselişini simgelerken, Doğu ve Batı kültürlerini bir araya getirmiş ve İstanbul'u ticaret, sanat ve dinin küresel bir merkezi haline getirmiştir.
Passover (Pesach) is a sacred Jewish festival celebrating the Exodus from Egypt, where families gather for the Seder meal, retelling the story of liberation, eating matzah (unleavened bread), and reciting traditional blessings.	במשפחות מתכנסות לסעודה .הוא חג יהודי קדוש המציין את היציאה ממצרים (פסחא) ומקיימות ברכות (לחם לא מותסס)אוכלות מצה, מספרות סיפור השחרור ,מסורתית מסורתיות.	חג הפסח הוא חג יהודי קדוש החוגג את יציאת מצרים, שבו משפחות מתאספות לסעודת הסדר, מספרות מחדש את סיפור הגאולה, אוכלות מצה ומברכות ברכות מסורתיות.
Shabbat, the weekly day of rest, begins at sunset on Friday with the lighting of Shabbat candles, the recitation of Kiddush (blessing over wine), and the sharing of Challah bread, as families come together in spiritual reflection.	שבת, יום המנוחה השבועי, מתחילה בשקיעה ביום שישי עם הדלקת נרות שבת, אמירת קדושה (ברכה על .היין), ואכילת לחם השבת, כשמשפחות מתאחדות ברפלקציה רוחנית	שבת, יום המנוחה השבועי, מתחילה בשקיעת השמש ביום שישי עם הדלקת נרות שבת, אמירת קידוש (ברכה על היין) ושיתוף חלת שבת, כאשר משפחות מתכנסות יחד להרהור רוחני.
in 70 CE, the Roman destruction of the Second Temple in Jerusalem marked a turning point in Jewish history, leading to the diaspora, the preservation of Jewish identity through Torah study, and the hope of one day returning to Zion.	ב-70 לספירה, חורבן המקדש השני בירושלים על ידי הרומאים סימן נקודת מפנה בהיסטוריה היהודית.שהובילה לגלות, שימור הזהות היהודית דרך לימוד התורה, ותקווה לשוב יום אחד לציון.	בשנת 70 לספירה, חורבן בית המקדש השני בירושלים בידי הרומאים ציין נקודת מפנה בהיסטוריה היהודית, שהובילה לתפוצות, לשימור הזהות היהודית דרך לימוד התורה, ולטיפוח התקווה לשוב יום אחד לציון.

4.2 Comparative Analysis

Table 3 compares our multi-agent AI framework with GPT-4o, highlighting key differences in cultural preservation, contextual depth, and linguistic expressiveness. Our system outperforms GPT-4o in two major aspects: evocative language and contextualization. The Translation and Interpretation Agents incorporate figurative expressions, idiomatic phrases, and poetic descriptions, making the translations more immersive and culturally resonant.

For instance, in the Hindi translation of Ganga Aarti, our model renders the phrase as “mohak chhata chha jati hai” (“a mesmerizing aura”), effectively capturing the spiritual and visual grandeur of the event, whereas GPT-4o’s simpler rendering of “bhavya Ganga Aarti” lacks emotional depth. Similarly, in the translation of the 1857 Revolt, our system uses “jwalant udaharan” (“a blazing example”) to emphasize the passion and heroism of freedom fighters, while GPT-4o remains more neutral in its phrasing. In Turkish translations, our model adapts “Bonfire jumping” in Nevruz celebrations as “atesin uzerinde ziplamak”, effectively reflecting the ritualistic importance, while GPT-4o’s version remains technically correct but lacks cultural vibrancy. Additionally, our Whirling Dervishes translation integrates “mistik” (mystical) and “ilahi” (divine) to reinforce the spiritual and meditative essence of the dance, whereas GPT-4o provides a more standard description that does not fully capture its Sufi traditions. For Hebrew religious texts, our framework ensures spiritual authenticity by explicitly incorporating key observances such as Shabbat candles, Kiddush, and Challah bread, while GPT-4o omits or generalizes some religious details. In the Passover translation, our model maintains the ritualistic depth by carefully referencing traditional elements like unleavened bread and storytelling, ensuring greater alignment with Jewish traditions.

Additionally, our agents expand on traditions by linking events to their historical and cultural roots, whereas GPT-4o tends to remain neutral and lacks explanatory richness. Our framework also enhances transliterated cultural terms by adding brief clarifiers, making the text more accessible to native speakers. While GPT-4o ensures grammatical correctness, it often lacks cultural resonance and emotional warmth, reinforcing the need for a specialized multi-agent system in cross-language understanding and preservation.

5 Discussion

Our study presents a multi-agent AI framework designed to enhance cross-language translation by integrating linguistic accuracy, cultural adaptation, and bias mitigation. Unlike conventional LLM models, our approach distributes tasks across specialized agents—Translation, Interpretation, Content Synthesis, and Quality and Bias Evaluation—ensuring contextually enriched and culturally aligned translations. The Translation Agent guarantees grammatical correctness, while the Interpretation Agent adapts idiomatic expressions and

cultural nuances. The Content Synthesis Agent refines readability, and the Quality and Bias Evaluation Agent validates fairness and authenticity using external sources. To evaluate our system's performance, we tested translations across multiple cultural domains, including historical narratives, religious traditions, and festival descriptions. Comparative evaluation with GPT-4o (Table 2) reveals that our framework consistently produces more evocative, idiomatic, and culturally grounded translations, demonstrating its ability to capture deeper contextual meaning across various content types rather than excelling in a single category. The agent-based approach effectively addresses limitations in conventional translation models, particularly in ensuring cultural depth and contextual relevance (Ogie et al., 2022). By incorporating external validation mechanisms, our system minimizes linguistic distortions and biases, making it more suitable for real-world multilingual applications. The results indicate that multi-agent collaboration enhances cross-language understanding, providing a scalable and adaptable solution for preserving linguistic heritage and reducing biases in AI-generated translations. This framework presents a significant step forward in AI-driven language processing, offering a context-aware, culturally sensitive, and ethically responsible approach to translation.

5.1 Limitations

Despite its effectiveness, our framework has several limitations. The multi-agent collaboration improves fairness and transparency but increases processing time compared to single-agent models, reducing efficiency for real-time applications. Although the system supports multiple languages, challenges persist with low-resource languages due to limited training data and digital resources, affecting translation quality and adaptability (Gong et al., 2024). External validation via DuckDuckGo enhances accuracy but may introduce inconsistencies if sources lack credibility or cultural specificity (Ootani and Yamana, 2018). Lastly, cultural subjectivity remains a challenge, as idioms and expressions often lack direct equivalents, requiring interpretative adjustments across contexts.

5.2 Future Work

Future research will focus on refining our multi-agent AI framework to address its current limitations. One area for improvement is optimizing the processing time for multi-agent collaboration, making the system more efficient for real-time applications without compromising translation quality. Expanding the framework's capabilities to better support low-resource languages through data augmentation and community-driven input is essential for improving translation adaptability and reducing biases in underrepresented languages. Additionally, enhancing the external validation mechanisms to incorporate more reliable and region-specific sources will further reduce inconsistencies in the system. Future work will also explore integrating more advanced cultural adaptation algorithms to handle nuanced expressions and idioms more effectively across diverse contexts. Moreover, we plan to expand the system's scope to include specialized domains such as legal and medical translations, where accuracy and cultural sensitivity are crucial. Collaborative research with cross-regional teams will be key to ensuring that the framework remains inclusive and adaptable to global linguistic and cultural needs.

6 Conclusion

Our study introduces a multi-agent AI framework that significantly enhances culturally adaptive cross-language translation, overcoming key limitations of traditional AI models. By employing specialized agents for translation, interpretation, content synthesis, and bias evaluation, our system ensures greater linguistic accuracy, cultural sensitivity, and contextual depth in translations. Although challenges such as computational efficiency and coverage for low-resource languages persist, our approach offers a promising pathway for more inclusive and context-aware AI-driven translation systems. The comparative analysis with GPT-4o further demonstrates the effectiveness of our framework in producing translations that are more culturally embedded and nuanced. As we look ahead, future work should focus on optimizing real-time processing capabilities, expanding language support, and refining external validation techniques to further enhance the scalability and reliability of cross-language communication. Ultimately, this research paves the way for a more equitable, culturally informed, and accurate AI translation landscape, contributing to the preservation and revitalization of diverse languages and cultures.

Acknowledgements

We express our sincere gratitude to Computational Intelligence and Operations Laboratory (CIOL) for their invaluable guidance, unwavering support, and continuous assistance throughout this journey. We are deeply appreciative of their efforts in organizing the CIOL Winter ML Bootcamp (Wasi et al., 2024), which provided an enriching learning environment and a strong foundation for collaborative research.

Author Contributions

M. A. Anik and A. T. Wasi conceptualized the idea and developed the methodology. M. A. Anik implemented the agents, conducted the literature review, carried out the experiments and analysis, and wrote the core sections of the work. A. Rahman contributed to visualization, literature analysis, and writing. A. T. Wasi supervised the project, provided overall guidance, and edited the manuscript. M. M. Ahsan also offered valuable support and guidance throughout various phases of the project.

References

Shubham Agarwal, Issam H Laradji, Laurent Charlin, and Christopher Pal. 2024. Litllm: A toolkit for scientific literature review. *arXiv preprint arXiv:2402.01788*.

UCHECHUKWU C AJUZIEOGU. 2024. Multimodal generative ai for african language preservation: A framework for language documentation and revitalization.

Gábor Bella, Paula Helm, Gertraud Koch, and Fausto Giunchiglia. 2023. Towards bridging the digital language divide. *arXiv preprint arXiv:2307.13405*.

Samy Bizan bin Ghowar. 2023. Recent trends in the use of artificial intelligence in the field of heritage. *International Journal of Artificial Intelligence and Emerging Technology*, 6(2):66–82.

John Dang, Shivalika Singh, Daniel D'souza, Arash Ahmadian, Alejandro Salamanca, Madeline Smith, Aidan Peppin, Sungjin Hong, Manoj Govindassamy, Terrence Zhao, Sandra Kublik, Meor Amer, Viraat Aryabumi, Jon Ander Campos, Yi-Chern Tan, Tom Kocmi, Florian Strub, Nathan Grinsztajn, Yannis Flet-Berliac, Acyr Locatelli, Hangyu Lin, Dwarak Talupuru, Bharat Venkitesh, David Cairuz, Bowen Yang, Tim Chung, Wei-Yin Ko, Sylvie Shang Shi, Amir Shukayev, Sammie Bae, Aleksandra Piktus, Roman Castagné, Felipe Cruz-Salinas, Eddie Kim, Lucas Crawhall-Stein, Adrien Morisot, Sudip Roy, Phil Blunsom, Ivan Zhang, Aidan Gomez, Nick Frosst, Marzieh Fadaee, Beyza Ermis, Ahmet Üstün, and Sara Hooker. 2024. Aya expanse: Combining research breakthroughs for a new multilingual frontier. *Preprint*, arXiv:2412.04261.

Zhihua Duan and Jialin Wang. 2024. Exploration of llm multi-agent application implementation based on langgraph+ crewai. *arXiv preprint arXiv:2411.18241*.

Neha Garg. 2024. Linguistic landscape of gen z: The impact of english dominance on endangered languages. *Journal of Research Scholars and Professionals of English Language Teaching*, 8(43):1–7.

Adiva Goel. Preserving south asian tribal legacies: Exploring oral traditions, traditional spoken narratives and folk tales.

Cheng Gong, Erica Cooper, Xin Wang, Chunyu Qiang, Mengzhe Geng, Dan Wells, Longbiao Wang, Jianwu Dang, Marc Tessier, Aidan Pine, et al. 2024. An initial investigation of language adaptation for tts systems under low-resource scenarios. *arXiv preprint arXiv:2406.08911*.

James Hutson, Pace Ellsworth, and Matt Ellsworth. 2024. Preserving linguistic diversity in the digital age: a scalable model for cultural heritage continuity. *Journal of Contemporary Language Research*, 3(1).

Graham M Jones, Shai Satran, and Arvind Satyanarayan. 2025. Toward cultural interpretability: A linguistic anthropological framework for describing and evaluating large language models. *Big Data & Society*, 12(1):20539517241303118.

Anne Kandler and Roman Unger. 2023. Modeling language shift. In *Diffusive spreading in nature, technology and society*, pages 365–387. Springer.

Nir Kshetri. 2024. Linguistic challenges in generative artificial intelligence: Implications for low-resource languages in the developing world.

Haley Lepp and Parth Sarin. 2024. A global ai community requires language-diverse publishing. *arXiv preprint arXiv:2408.14772*.

Jiangfeng Liu, Ziyi Wang, Jing Xie, and Lei Pei. 2024. From chatgpt, dall-e 3 to sora: How has generative ai changed digital humanities research and services? *arXiv preprint arXiv:2404.18518*.

Mohamed El Louadi. 2024. On the preservation of africa's cultural heritage in the age of artificial intelligence. *arXiv preprint arXiv:2403.06865*.

So Miyagawa. 2024. Ainu–japanese bi-directional neural machine translation: A step towards linguistic preservation of ainu, an under-resourced indigenous language in japan. *Journal of Data Mining & Digital Humanities*, (Digital humanities in languages).

Salikoko S Mufwene. 2005. Globalization and the myth of killer languages: What's really going on. *Perspectives on endangerment*, 5(21):19–48.

Junho Myung, Nayeon Lee, Yi Zhou, Jiho Jin, Rifki Afina Putri, Dimosthenis Antypas, Hsuvas Borkakoty, Eunsu Kim, Carla Perez-Almendros, Abinew Ali Ayele, et al. 2024. Blend: A benchmark for llms on everyday knowledge in diverse cultures and languages. *arXiv preprint arXiv:2406.09948*.

Dinesh Kumar Nanduri and Elizabeth M Bonsignore. 2023. Revitalizing endangered languages: Ai-powered language learning as a catalyst for language appreciation. *arXiv preprint arXiv:2304.09394*.

Wilhelmina Onyothi Nekoto, Julia Kreutzer, Jenalea Rajab, Millicent Ochieng, and Jade Abbott. Participatory translations of oshiwambo: Towards culture preservation with language technology.

Robert I Ogie, Sharon O'Brien, and Federico M Federici. 2022. Towards using agent-based modelling for collaborative translation of crisis information: A systematic literature review to identify the underlying attributes, behaviours, interactions, and environment of agents. *International Journal of Disaster Risk Reduction*, 68:102717.

Kazuyoshi Ootani and Hayato Yamana. 2018. External content-dependent features for web credibility evaluation. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 5414–5416. IEEE.

Rifki Afina Putri, Faiz Ghifari Haznitrama, Dea Adhista, and Alice Oh. 2024. Can llm generate culturally relevant commonsense qa data? case study in indonesian and sundanese. *arXiv preprint arXiv:2402.17302*.

Angelika Romanou, Negar Foroutan, Anna Sotnikova, Zeming Chen, Sree Harsha Nelaturu, Shivalika Singh, Rishabh Maheshwary, Micol Altomare, Mohamed A Haggag, Alfonso Amayuelas, et al. 2024. Include: Evaluating multilingual language understanding with regional knowledge. *arXiv preprint arXiv:2411.19799*.

Antonios Saravanos, Stavros Zervoudakis, Dongnanzi Zheng, Amarpreet Nanda, Georgios Shaheen, Charles Hornat, Jeremiah Konde Chaettle, Alassane Yoda, Hyeree Park, and Will Ang. 2022. Reputation, risk, and trust on user adoption of internet search engines: The case of duckduckgo. In *International Conference on Human-Computer Interaction*, pages 683–691. Springer.

Chuanmao Tian, Xu Wang, and Mingwu Xu. 2022. Historico-cultural recontextualization in translating ancient classics: a case study of gopal sukhu's the songs of chu. *Perspectives*, 30(2):181–194.

Azmine Toushik Wasi, MD Shakiqul Islam, Sheikh Ayatur Rahman, and Md Manjurul Ahsan. 2024. Ciol presnts winter ml bootcamp. 6 December, 2024 to 6 February, 2025.