

Preserving Indigenous Culture and Sovereignty in Large Language Models.

Srishti Gupta^{a*}, Taranum Bano^b, Yu-Che Chen^b, Edouardo Zendejas^b, Sarah Krafka^b, Chun-Hua Tsai^b

^aUniversity at Albany, SUNY, Albany, New York, United States, sgupta4@albany.edu

^bUniversity of Nebraska Omaha, Omaha, Nebraska, United States, tbano@unomaha.edu, ychen@unomaha.edu, ezendejas@unomaha.edu, skrafka@unomaha.edu, chunhuatsai@unomaha.edu

Publishing history: Submitted*: e.g 01 January 2021, Revised*: e.g 01 June 2021, Accepted*: e.g 01 July 2021, Published*: e.g 01 August 2021

(Revised date: In the case of multiple rounds of reviews, we need both revised dates).

Abstract. The advancement of large language models (LLMs) presents a significant opportunity to shape public understanding of indigenous histories and knowledge domains. Despite notable achievements, their applicability to marginalized data sources like indigenous languages and cultures remains uncertain, raising concerns about Indigenous data governance and sovereignty. Historical mistrust rooted in colonialism now extending into contemporary technology systems further complicates responsible AI deployment in these contexts. This highlights the need for ethical frameworks that center indigenous rights and self-determination. This paper presents a study involving multiple tribal communities in the Midwest region of the United States. Through an iterative co-design process, we developed TribalLLM, an LLM-based chatbot designed to preserve tribal culture, history, and sovereignty. We employed the Retrieval-Augmented Generation (RAG) method to integrate tribally approved data into the LLM, enabling culturally grounded responses while maintaining a locally controlled database that reflects indigenous data sovereignty principles. The system was evaluated through online interviews with ten participants representing diverse tribal communities. Our findings reveal several key themes: the importance of AI responses being grounded in tribally authored knowledge; the imperative of maintaining Indigenous sovereignty through community-led empowerment; strategies for embedding oral traditions, language, and cultural values within LLMs; and the relationship between AI and Indigenous peoples' right to self-determination. Participants emphasized that trust in the tool is inseparable from trust in the research team, highlighting the relational nature of technology adoption in Indigenous contexts. We conclude with design implications for culturally sensitive LLM development, including sovereignty-first data practices, multimodal outputs honoring oral and visual communication preferences, transparent data provenance, and sustained community engagement. This work advances both AI scholarship and digital government research by offering a concrete pathway for embedding Indigenous data sovereignty and cultural continuity into public-facing AI systems.

Keywords. Indigenous sovereignty, large language models, cultural preservation, retrieval-augmented generation, co-design, tribal communities, AI design

1. Introduction

The proficiency of advanced Large Language Models (LLMs) raises critical questions about their ability to accurately represent and process data from sources that are traditionally underrepresented or marginalized, with a particular emphasis on indigenous languages and cultures (Kasneci et al., 2023; Stap and Araabi, 2023). This situation has heightened concerns over data misuse and sovereignty issues, leading many indigenous communities to oppose using their data and languages for training AI models (Carroll et al., 2019; Chandran, 2023). These apprehensions reflect a broader, long-standing institutionalized mistrust towards dominant

power entities. Historically, this distrust originated with ancient colonizers and has now extended to contemporary power structures in the realm of information and technology (Shiri et al., 2022). This scenario highlights the importance of recognizing and respecting data sovereignty as a fundamental right of indigenous communities. It is crucial to address these issues by establishing ethical guidelines and frameworks that prioritize the rights and perspectives of indigenous peoples in the development and application of AI technologies, ensuring their cultural heritage is preserved and respected.

With the advancement of AI, recent research has explored applying these technologies to interpret and convey cultural heritage, culture, and values (Spennemann, 2023; Holguin, 2023). These developments offer a unique opportunity to preserve or internalize human culture and language within these AI systems (Colas et al., 2022), addressing a crucial research area in the preservation of Indigenous culture and linguistic diversity (Roy, 2015; Fraisse et al., 2019; Park et al., 2022). The integration of LLMs in capturing and disseminating indigenous knowledge represents a critical intersection of technology, cultural preservation, and Indigenous sovereignty. It raises important questions about how these advanced computational models can be sensitively and accurately adapted to represent the complexities and nuances of indigenous languages and cultural narratives.

Scholarship has begun to explore how Indigenous culture can be incorporated into AI design (Gallamaso, 2024; Munn, 2024; Vamanu, 2012). Building on these works, our study develops a culturally sensitive large language model (LLM) and conducts an exploratory evaluation of its effectiveness, with the broader goal of advancing approaches to more inclusive and respectful AI design. Specifically, we investigate the research question: ***How can LLM tools be designed to strengthen and preserve indigenous culture and sovereignty?*** To address this, we conducted an online interview study with participants from multiple tribal communities across the Midwest region of the United States. Our research team has been engaged with these communities through ongoing meetings and discussions on the role of AI in tribal contexts [citations removed for peer review]. This study represents one strand of that broader collaboration, focusing on how AI might support cultural preservation and uphold tribal sovereignty. Through interviews, we sought to understand how Indigenous community members perceive our prototype tool, TribalLLM, and to learn what they view as most important for cultural preservation and how they envision AI contributing to sovereignty and cultural continuity.

Our findings show how indigenous sovereignty is essential for the protection, preservation, and flourishing of indigenous cultures through several key themes: the significance of AI-generated responses being relevant to reservation lands, the imperative of maintaining Indigenous sovereignty and empowering stakeholders through educational programs, strategies for presenting culture and embedding Indigenous data within LLMs, and the relationship between AI technologies and Indigenous peoples' right to self-determination. These themes illuminate the complex interplay between advanced AI tools and the cultural, social, and political contexts of Indigenous communities, highlighting the importance of a nuanced and respectful approach to the integration of these technologies.

Our work contributes to the Digital Government Society and AI literature by foregrounding tribal culture, sovereignty, data governance, and trust as central design concerns for public-facing AI systems. By empirically examining the co-design and evaluation of a culturally sensitive LLM with Indigenous communities, this research extends digital government discussions of equity, participation, and legitimacy beyond service delivery to include the stewardship of cultural and linguistic knowledge. It demonstrates how AI systems can function not only as technical artifacts but also as governance instruments that shape power relations between Indigenous communities and dominant institutions. In doing so, this paper advances AI scholarship by operationalizing Indigenous data sovereignty within LLM design and evaluation, offering a concrete methodological and conceptual pathway for embedding community values, self-determination, and cultural continuity into AI systems intended for public and governmental contexts.

2. Related Works

The relationship between Indigenous cultures, technology, and AI design is deeply shaped by historical, cultural, and epistemological contexts. The connection between Indigenous peoples and land is profound and influences their interaction with digital technologies. Contrary to the common view of cyberspace as placeless, scholars argue that it is inherently land-based due to its material infrastructure (Caranto Morford and Ansloos, 2021). This perspective foregrounds the socio-political implications of technology for Indigenous communities and highlights the importance of digital sovereignty and cyber-justice (Duarte, 2017; Caranto Morford and

Ansloos, 2021). Recognizing the territorial nature of cyberspace allows us to better understand how Indigenous communities engage with technology and why culturally grounded approaches to digital innovation are essential.

Digitization, digital archives, and storytelling have become key strategies for preserving Indigenous Cultural Heritage (Shiri et al., 2022; Unru et al., 2022). At the same time, they highlight the need to respect tribal rights and perspectives in decolonizing data practices (Roy, 2015). Yet, data can also marginalize Indigenous peoples by erasing traditions or reframing histories through colonial lenses (Carroll et al., 2019). The principle of Indigenous data sovereignty seeks to counter these issues by giving communities control over their data (Kukutai and Taylor, 2016; Caranto Morford and Ansloos, 2021), but its implementation remains challenged by colonial legacies embedded in modern information systems. As AI technologies advance, ensuring they do not reproduce these harms requires collaborative, culturally sensitive design (Carroll et al., 2019). Recent applied studies echo this dual potential. (Qureshi et al., 2025), for example, document how arts-based research and the Indigenous method of *yarning* were used in South Australia to digitally preserve narratives and techniques, while also educating artists about Indigenous Cultural and Intellectual Property (ICIP). Similarly, (Goulding et al., 2020) analyze two case studies from Aotearoa New Zealand and frame digital technologies as a “double-edged sword” for cultural sustainability as they can revitalize language and cultural knowledge but also raise issues of ownership, control, and consultation.

Traditionally, AI systems have reflected Western, upper-class-centric worldviews (Ghaith, 2024). Designing AI for Indigenous communities, however, requires incorporating local epistemologies to build inclusive and culturally relevant systems. Long-term collaborations with Indigenous groups emphasize moving beyond mere recording of knowledge toward reinterpretation, re-contextualization, and community empowerment (Winschiers-Theophilus et al., 2019). This transcultural engagement enables blended frameworks that acknowledge cultural diversity rather than isolating or homogenizing traditions. Nevertheless, challenges persist, especially in the Global South, where issues such as data scarcity, ethical concerns, limited technological literacy, and skepticism toward AI undermine trust (Anuyah et al., 2023). Similar dynamics emerge in Indigenous contexts in the United States, emphasizing the importance of careful co-design.

Recent advances in LLMs such as ChatGPT have further raised both opportunities and risks for Indigenous knowledge systems. These models demonstrate remarkable capacity for synthesizing information, multilingual interaction, and complex reasoning (Kasneci et al., 2023; Susnjak, 2023), with applications ranging from healthcare to translation. Early explorations have also examined how LLMs represent Indigenous knowledge, testing whether their outputs align with community interpretations (Holguin, 2023). Yet, deploying such technologies for Indigenous contexts requires sensitivity to cultural and social frameworks (Barnosky et al., 2022; Lewis et al., 2020a). Techniques like Retrieval Augmentation Generation (RAG) (Lewis et al., 2020b; Baek et al., 2023; Gao et al., 2023) make it possible to combine LLMs with compact tribal datasets, extending their use to culturally specific domains such as healthcare and emergency management (Kruse et al., 2016; Tsai et al., 2023; Potts et al., 2021). Building effective systems in these contexts necessitates strategies like culturally sensitive data collection, transfer learning across linguistic and cultural divides, and attention to cultural variation in knowledge and communication (Hershovich et al., 2022). Ultimately, co-design approaches that emphasize reinterpretation, community empowerment, and respect for Indigenous epistemologies (Smith et al., 2021; Burnette et al., 2014) are critical to ensuring that AI systems are both trustworthy and aligned with community needs.

2.1. Digitization and Tribal Culture Preservation

The connection between indigenous cultures and land is profound and significantly influences their interaction with technology. Contrary to the popular notion that cyberspace is devoid of territoriality, scholars argue that it has a land-based nature due to its internet infrastructure (Caranto Morford and Ansloos, 2021). This perspective highlights the physical realities and socio-political implications of digital technologies for indigenous communities and their ancestral lands, emphasizing the need to consider indigenous digital sovereignty and cyber-justice (Duarte, 2017; Caranto Morford and Ansloos, 2021). By recognizing the territorial aspect of cyberspace, we can better understand the complexities surrounding indigenous cultures and their engagement with technology.

The use of digital archives and storytelling methods has become increasingly important in preserving Indigenous Cultural Heritage (Shiri et al., 2022; Unru et al., 2022), while also emphasizing the need for respect and

recognition of tribal peoples' rights and perspectives in decolonizing data and cultural preservation (Roy, 2015). However, data can also be used to marginalize indigenous peoples by erasing their traditional ways of life and recontextualizing their histories (Carroll et al., 2019). The concept of 'Indigenous data sovereignty' seeks to address this issue by maintaining control over the data of Indigenous communities, but it is challenged by the legacy of colonialism within modern information systems (Kukutai and Taylor, 2016; Caranto Morford and Ansloos, 2021). As AI technologies advance, it is crucial to consider these complexities and ensure that they do not perpetuate the marginalization of indigenous communities, highlighting the need for a collaborative and culturally sensitive approach in developing AI solutions (Carroll et al., 2019).

In line with these concerns, recent work by (Qureshi et al., 2025) provides an applied account of preservation through digital practice. Using arts-based research (ABR) and the Indigenous method of *yarning*, the authors document workshops and video-documented interviews with artists at an Indigenous art centre in South Australia to digitally preserve narratives, techniques (e.g., felting), and symbolic knowledge, while educating artists about Indigenous Cultural and Intellectual Property (ICIP) and ethical digitization. Their findings show how digital tools (recording, archiving, platform sharing) can both empower communities to safeguard and transmit cultural heritage and introduce risks (digital divide, appropriation, commodification). Similarly, (Goulding et al., 2020) examine two case studies from Aotearoa New Zealand and frame digital technologies as a "double-edged sword" for Indigenous cultural sustainability. While digitization projects and storytelling apps can revitalize language and cultural knowledge, they also surface challenges of ownership, control, consultation, and cultural homogenization. These findings underscore the importance of Indigenous authority in decisions around digital representation to ensure technologies strengthen rather than diminish cultural heritage.

2.2. Cultural Considerations in AI

Traditionally, AI systems have been designed with a Western, upper-class-centric focus (Ghaith, 2024). Developing AI for Indigenous communities, however, requires integrating local epistemologies to create more inclusive and culturally relevant technologies that mitigate bias. Long-term collaborations with Indigenous groups worldwide show the need to move beyond simply recording knowledge toward reinterpretation, recontextualization, and community empowerment (Winschiers-Theophilus et al., 2019). This approach calls for transcultural engagement—creating blended frameworks rather than treating cultures in isolation.

Applying this framework to AI design poses specific challenges. A review of cultural considerations in the Global South identified issues such as scarce localized datasets, ethical concerns in data use, limited technological literacy, and skepticism toward AI that undermines trust (Anuyah et al., 2023). Similar challenges also surfaced in our work with Indigenous communities in the Midwestern United States.

Our study examines how LLMs can be adapted to capture and share Indigenous knowledge. Prior research underscores key strategies: culturally sensitive data collection and annotation, use of transfer learning to bridge linguistic and cultural differences, and attention to cultural variation in linguistic forms, knowledge, values, and communication (Hershcovich et al., 2022). These strategies are essential for building culturally aware NLP and translation systems.

2.3. LLMs and Retrieval Augmentation Generation (RAG)

The introduction of LLMs like ChatGPT has marked a significant advancement in natural language understanding and conversational capabilities, approaching the nuances of human language and interaction (Kasneci et al., 2023). These models are adept at generating accurate and coherent responses to complex queries, necessitating an amalgamation of information synthesis, analytical processing, and practical application (Kasneci et al., 2023). Trained on extensive textual datasets, these models capture the complexities and nuances of human language, enabling multilingual interactions and broadening their applicability across various sectors, including customer service, content creation, healthcare, and language translation (Susnjak, 2023). LLMs have also been used to explore indigenous knowledge, employing different prompts to test if ChatGPT aligns with interpretations from indigenous communities (Holguin, 2023).

The deployment of emerging technologies like LLMs for indigenous communities requires keen awareness and integration of their distinct cultural and social contexts (Barnosky et al., 2022). Incorporating indigenous knowledge systems into the discourse surrounding AI's societal role can pave the way for innovative

methodologies tailored to unique challenges (Lewis et al., 2020a). Techniques like Retrieval Augmentation Generation (RAG) enable specialized text generation by combining pre-trained LLMs with non-parametric memory (Lewis et al., 2020b; Baek et al., 2023; Gao et al., 2023), allowing for the integration of compact tribal datasets into LLMs and enhancing their applicability to specific domains like healthcare and emergency management (Kruse et al., 2016; Tsai et al., 2023; Potts et al., 2021; Kasneci et al., 2023). To develop culturally relevant AI technologies, it is essential to adopt a nuanced approach that emphasizes reinterpretation, re-contextualization, and community empowerment (Winschiers-Theophilus et al., 2019), addressing key issues like data scarcity, ethical concerns, and limited technological literacy (Anuyah et al., 2023). This requires trans-cultural engagement, understanding of cultural variations in linguistic forms and values (Hershovich et al., 2022), and co-design approaches that prioritize collaborative frameworks respecting Indigenous values and epistemologies (Smith et al., 2021; Burnette et al., 2014), ultimately promoting the development of AI systems aligned with community needs and values.

3. Methods

3.1. System Development and Pilot Testing

We developed *TribalLLM* as illustrated in Figure 1. The primary aim of creating TribalLLM was to explore a data sovereignty framework. This framework involves using a tribally recognized database maintained locally, in contrast to the direct transmission of data to a cloud-based API (Gupta et al., 2024). Specifically, our proposed framework allows the indigenous community to self-maintain a database, submitting only necessary and tribally sanctioned data to the LLM service, which we used the ChatGPT 3.5 model. This approach is intended to safeguard data sovereignty and governance. Additionally, this design offers greater flexibility to incorporate functions and features based on findings from our experiments. This structure aligns with the principles of indigenous data autonomy and reflects a tailored approach to integrating advanced technological solutions with indigenous communities' specific needs and values.

This project involved the manual curation of a comprehensive PDF document featuring content from the official websites of three indigenous tribes: the Omaha Tribe of Nebraska, the Sac and Fox Tribe, and the Chitimacha Tribe of Louisiana. Leveraging the LangChain framework, we integrated this data into a local database, enabling our system to match user prompts with relevant context data and generate responses using the ChatGPT API. To ensure the cultural sensitivity and usability of our system, we showcased it at the National Congress of American Indians (NCAI) 2023 in Louisiana, soliciting feedback from attendees. This feedback was instrumental in refining the system's interface and conceptual approach, with improvements including the addition of icons, sample prompts, and layout adjustments. The iterative process of development and user feedback is essential for creating a tool that effectively harnesses AI capabilities while addressing the unique needs and preferences of indigenous communities. In our formal interview study, the chatbot artifact served as a catalyst for discussion with participants, rather than a means of evaluating response quality; therefore, we did not solicit ratings or comments on the AI responses. Validation of the chatbot's responses was conducted collaboratively during sessions, with participants reviewing outputs and providing corrections or contextual feedback where responses diverged from their tribal knowledge. Additionally, we acknowledge that the dynamic nature of Gen-AI models may lead to variable responses, and while embedding tribal context can enhance response relevance, it does not guarantee comprehensive coverage of all questions based on the embedded dataset. This limitation shows the importance of ongoing development and refinement to ensure the chatbot's effectiveness in supporting indigenous communities.

3.2. Data Collection and Analysis

To further understand the perceptions and experiences associated with using TribalLLM, we conducted an online interview study involving 10 participants from various tribes. This study aimed to gather insights into how TribalLLM was received and utilized by individuals from different indigenous backgrounds, providing valuable feedback for further development and refinement of the tool. The demographic data is shown in Table 1. Participants were recruited by a tribal collaborator (one of the co-authors) who had personal connections with these tribes. She assisted in disseminating information about the interview concerning a newly developed AI tool and recruited individuals interested in participating in the project. Although 12 participants were initially scheduled, ten ultimately completed the one-hour online interview via Zoom. The relatively small sample size reflects the inherent difficulty of recruiting from specific indigenous tribal populations, which are limited in

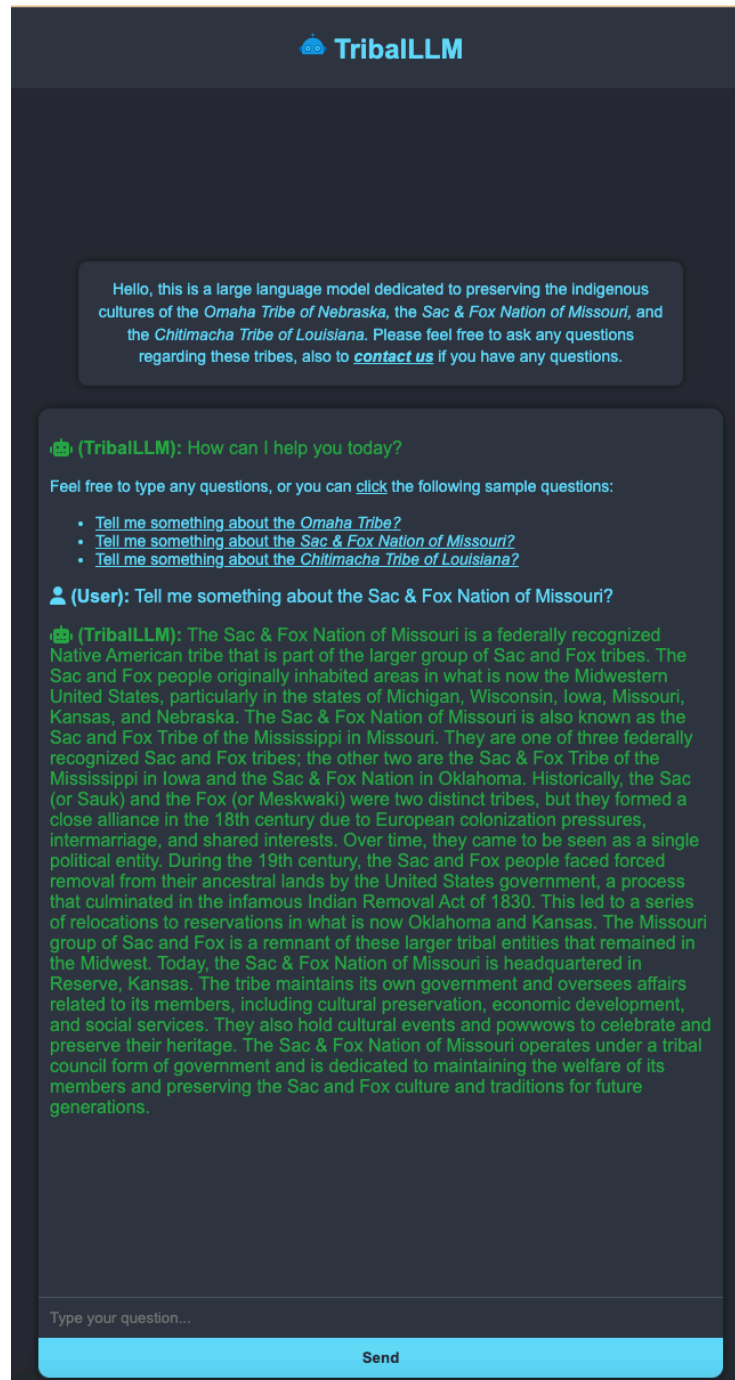


Fig. 1 – This figure displays a screenshot of the TribalLLM chatbot, which integrates data from multiple tribal nations into a large language model. Users can input various prompts to elicit responses from the application, demonstrating the chatbot’s capacity to access and process the embedded tribal data.

size and access; we recruited as many participants as feasibly possible given our available resources and community connections (Caine, 2016).

The online interview process commenced with a brief demonstration of TribalLLM, facilitated through screen sharing. Additionally, we provided participants with a link, enabling them to submit prompts and interact with the app independently. However, most participants preferred to rely on our guidance and demonstration rather than engaging directly with the app. In cases where participants did not interact directly, a researcher entered prompts into the chatbot on their behalf, and the generated responses were reviewed and discussed together with the participant in real time. The initial set of sample prompts included queries like

Tab. 1 – Participant Demographics

ID	Gender	Tribe	Occupation	Age
S-1	F	Muscogee Creek	Retired	~80
S-2	F	Yankton/Santee	Doctoral Student	~30
S-3	F	Cheyenne Arapaho	Retired	~70
S-4	F	Sicangu Lakota	Adjunct Professor	~65
S-5	M	Omaha	Student/Marine Reserves	~20
S-6	M	Navajo	Director of Health Disparity and Health Equity	~43
S-7	F	Kickapoo	Graduate Student	
S-8	F	Cheyenne Arapaho	University Academic Advisor	~50
S-9	F	Winnebago	Retired	~60
S-10	F	Omaha	Cultural Director of a school district	~70

'Tell me about the [target] tribe' and 'What should I do if a natural disaster impacts me?' These prompts were strategically chosen to align with the dual objectives of presenting tribal information and specific governance scenarios to users, thereby garnering their feedback and insights. The interview protocol encompassed various segments, including 1) General Background, 2) Interaction with the Chatbot, and 3) Specific Questions designed to elicit requirements for governance and tribal culture preservation. Sample questions included: 'What is your experience with chatbots like ChatGPT?', 'Do you consider the chatbot to accurately describe your tribe's cultures, histories, and government?', and 'If we want to design this chatbot for outside people to understand your tribe, what information do we need to have?'. 2-3 research team members facilitated the interview for each participant in every session.

Qualitative data from all interviews were collected using Zoom's cloud recording service. We employed an inductive, qualitative thematic analysis to open-code the transcribed recordings of the semi-structured interviews, identifying emergent themes (Braun and Clarke, 2006). Specifically, we followed Braun and Clarke's six-phase thematic analysis framework, using an inductive approach where themes were derived directly from the data rather than being imposed by a pre-existing theoretical framework. Two authors of this paper then engaged in discussions and iterations to develop key and high-level themes related to AI chatbots, culture, sovereignty, and user experience. Following data collection, each researcher thoroughly reviewed the dataset to gain an initial understanding of its contents. Frequent meetings were held during the first two weeks, approximately once every three or four days, to discuss impressions of the dataset and insights from the co-design process.

In planning and executing this study, we were acutely aware of the ethical considerations required when engaging with vulnerable populations. Collaboration was established with two faculty members at our university, both of whom are American Indians with direct ties to tribal councils and community members. Their connections were invaluable in recruiting participants for the study. We present demographic data, including gender, age, and occupation, in Table 1; all such data were collected voluntarily. We meticulously removed all identifiable personal information in our commitment to uphold participant privacy. All aspects of the research received Institutional Review Board (IRB) approval, and official resolutions from tribal governments were in place before any research activities commenced.

4. Findings

4.1. *Sovereignty, Representation, and Empowerment*

4.1.1. Native Voice in Sovereignty

Tribal sovereignty holds significant importance for all tribal members. However, the chatbot's responses to questions about the tribe lacked specific information emphasizing sovereignty. Participants noted that while the chatbot acknowledged involvement in preserving cultural heritage, it fell short of explicitly mentioning aspects like administering their own nation, overseeing rules, and managing jurisdiction, especially within the context of reservations. There is a desire for the chatbot to convey a more comprehensive understanding of tribal sovereignty and governance.

“If it’s just coming up like Wikipedia American history kind of things, you know. Because that’s the one thing that as Native Americans, we’re trying to rewrite that history. Our own voices and things. I wish it would be more specific to Native voices rather than pulling like the regular GPT from everywhere and just getting that generic information, you know.” S-8

As mentioned by S2-8 in the quote above, a fundamental issue contributing to the inaccurate representation of tribal sovereignty in chatbot responses is that much of the available Native American information on the internet is authored by white Americans, leading to a loss of the authentic voice of native authors. To address this, the chatbot should prioritize cultural sensitivity, accuracy, and representation of the perspectives and narratives actively advocated and reclaimed by Native Americans in the rewriting of their history. Respecting tribal sovereignty involves engaging in government-to-government conversations and interactions, allowing each tribe to decide what aspects of their culture they want to showcase and discuss. There is a need for an appropriate and respectful approach that aligns with the principles of tribal sovereignty.

4.1.2. Representation and Misrecognition

Participants also linked these issues of voice to broader struggles of representation in society. Native Americans continue to encounter racism, stereotyping, and dismissal of their knowledge systems, which are often based in oral history and not recognized as “scholarly” by dominant institutions.

“It was a problem for me in my master’s program, because I wanted to write about natives, and they would come back at me and say, this isn’t scholarly research. I would say there is none. What do you want me to do? ... I finally got them to the point where they would accept our tribal knowledge as part of what I was writing about. But it was a fight.” S-4

S-4’s experience illustrates how Indigenous knowledge is often devalued within mainstream academia, echoing broader cultural ignorance and exclusion. Other participants shared similar experiences of teasing, stereotyping, or cultural dismissal in predominantly white environments. These stories highlight the urgency of resisting reductive portrayals and affirming Native identity on Native terms.

Within this context, participants saw potential for chatbots not only to educate about tribal sovereignty but also to counter stereotypes and expand cultural understanding. By amplifying Native voices, such tools could both strengthen cultural identity within communities and promote broader acceptance and respect outside them - particularly for younger generations, who may see AI as a pathway to learning, empowerment, and belonging.

4.2. Trust and Relationships in Technology Adoption

4.2.1. Building Trust with Communities

Researchers discovered intriguing insights into tribal communities’ views on trusting the chatbot. While we learned that participants were concerned about the source of information fed into the chatbot, we also discovered that it was crucial for tribal communities to build trust with the research and development team of the project to trust the chatbot.

“Trust has to be built it just that’s where it’s not the program. Don’t sell the program. Sell yourself to the people that you are trustworthy. And when I’m gone, the people after me will be equally well trained and trustworthy and to keep that personal contact going, otherwise, it will fail.” S-1

The participant emphasizes the need for researchers or individuals to establish personal connections and prove their trustworthiness over time. The participant mentions that there might be doubt or reluctance initially, but building trust is crucial for the success of any program or collaboration. This stems from the historical distress and skepticism among Native communities when engaging with government entities or researchers. They mentioned that researchers often seem obstructive or fail to deliver on their promises, leading to reservations about collaborations. Furthermore, the participants suggests that the process of building

trust involves personal contact, establishing relationships, and sharing information over time. They express a preference for in-person interactions, such as sitting down, sharing a meal, and getting acquainted, as these methods contribute to creating deeper and more meaningful connections.

In addition to researchers establishing trust with the tribal community, it was crucial for tribal communities to be aware that the information provided to the chatbot was crafted by and for native individuals. The process involves gradually instilling trust in the technology by training tribal members, starting with human input to showcase the chatbot's capabilities. Participants suggested incorporating interactive exercises and scenarios to acquaint users with the technology, recognizing that technology can be intricate. This highlights the significance of a personalized approach, addressing specific tribal concerns and conducting training sessions to guarantee ease of use and comprehension.

"A lot of this technology is very confusing. And you show them concrete examples that they can use with the chatbot. It's just gonna have to be ease of use slowly easing them in using them." S-4

4.2.2. Trust in Data and Sources

While most participants found the chatbot response to be accurate, they also pointed out some missing information in certain responses and were curious to know the source of the information.

"I don't know if this chatbot does that. Checks like, okay, who is this written? By what kind of knowledge do they have? What kind of education do they have in the background? That's something that I think I don't know." S-7

S-7 expressed curiosity about the source of information and emphasized the importance of ensuring that the chatbot disseminates correct information that respects tribal culture. The participant acknowledges the challenge of gaining trust within Native communities, citing historical trauma and a general distrust of the government and authority figures. To encourage the use of the AI chatbot, the suggested approach is to present it as a tool for providing aid. By highlighting its role in aiding governance, reducing workload, and addressing communication challenges, S-7 believes this framing could make the technology more acceptable to community members with reservations rooted in historical and systemic issues. The ultimate goal is to position the technology as a helpful resource rather than reinforcing negative perceptions associated with government involvement.

Considering the diversity among Native American tribes, each with its unique culture, ceremonies, and heritage, it is imperative not to generalize or stereotype all Native Americans into a single category. Participants emphasize that the chatbot should recognize and respect these differences. To achieve this, connecting with specific tribal websites is suggested. These websites often contain information about tribal authors, news articles, and publications by native doctors and Ph.D. holders. A way to verify if the information source is authentic or not. Participants also suggested paying close attention to author names since many native authors tend to place the name of their tribe after their last name. However, the authentic information from native authors available on the internet and individual tribal websites is very scarce and not enough to train and develop an effective LLM chatbot. Researchers learned that this is because of lack of written documentation of tribal knowledge.

"We didn't write down a lot of our knowledge. So a lot of what we know is only shared by our side, not by the other side, because so much has been taken and stolen and our religion was turned into the hippie movement. So we have a lot of problems with people. Which is why I think what you guys are doing is great, because if you come out of tribe respectfully and with knowledge, they will accept you. If you come at them calling them Brule Sioux. They're gonna say, Okay, these are White people who don't understand who we are and are not gonna try." S-4

S-4 highlights the historical issue of their knowledge being taken, stolen, and misrepresented, particularly with their religion being associated with the hippie movement. The term "Brule Sioux," given by white Americans, is commonly used on the internet despite being derogatory to native tribes. Much of the information is transmitted through oral tradition from elders to younger generations, representing the cultural norm for preserving knowledge and customs. Participants recommend researchers obtain and record tribal knowledge by directly engaging with elders, tribal historians, teachers, and agencies to ensure accurate representation.

4.3. Cultural Preservation and Indigenous Knowledge

4.3.1. Oral Traditions and Storytelling

Understanding the tribal culture is a challenging endeavor due to its multifaceted nature. There are diverse elements and interpretations embedded in the cultural fabric. For instance, instead of being centered around a single festival, tribal ceremonies and traditions unfold throughout the year, with each season introducing its unique practices. From winter to summer, spring to fall, each season carries its distinct set of rituals. And these rituals are different for different families and communities. Cultural narratives vary significantly between the eastern and western sides of the nation, and traditions in the north differ from those in the south.

Traditionally tribal culture is passed down to generations orally by elders of the tribe. In many native communities, information is often absorbed through oral traditions due to the deep-rooted practice of storytelling. This approach led to misunderstandings with non-native individuals who, upon their arrival, failed to engage in active listening. The clash arose because there was a lack of comprehension regarding attempts to convey messages verbally. Reading wasn't a prevalent practice, and the written word held less significance compared to spoken words, believed to emanate from the heart. This cultural difference resulted in a constant struggle, as non-native individuals prioritized conveying information through written means, expecting strict adherence to documented information. However, thought processes in these communities often prioritize the spoken word as a more heartfelt and genuine form of communication.

Many of our native people, we learn more by listening, because we're such storytellers. We listen, which is something that I think we came at odds with the non-native people when they started coming into this country because they didn't listen... the written word didn't mean too much to us as the spoken word... because that's what we felt was coming from your heart. S-3

Hence, a significant portion of tribal cultural knowledge is transmitted orally through storytelling, with elders playing a crucial role in passing this wisdom to the younger generation. The data generation is a collaborative effort. Direct communication with elders within the family is highly valued as a rich source of cultural knowledge, offering guidance on various aspects of tribal traditions and language. While contemporary resources like books and other media allow the younger generation to delve into indigenous culture, participants preferred storytelling over reading extensive and potentially dry written material. Moreover, there is great respect for elders and tribal leaders in the native culture. Elders are given priority and served first in community functions or gatherings. The cultural code emphasizes the importance of respecting and prioritizing the elders, such as children, and those who may be physically impaired or not fully developed.

Yeah. I've gone to stomp dances and other events, and when I walk in the door with my brothers and sisters... the young girls and boys immediately come with a plate of food. They get you a chair. Make sure you're comfortable. Ask if you need water. S-1

Preserving tribal culture and customs in the AI chatbot is seen as a valuable means of passing down this heritage, especially with many Native Americans living off reservations and grandchildren residing away from their grandparents. Participants emphasized the importance of incorporating oral traditions into the chatbot's responses, highlighting that Native Americans tend to prefer visual and oral modes of communication over written knowledge. To effectively capture and convey tribal traditions, utilizing generative AI that accommodates these preferences is considered beneficial.

Cause I also put into the chat, What is the culture of Navajo people? And it did a pretty good interpretation... the only thing I don't like is that it's just one big, huge paragraph. At least for a Navajo person, it's very visual. And so I think, for just native people in general, it's more kind of a visual thinking versus literally reading. S-6

4.3.2. Language Preservation

Tribal languages play a crucial role in their culture but are gradually diminishing. English dominates official tribal governance and meetings. Participants acknowledge that native languages are not given priority in these contexts, and the convenience of using English often prevails, especially in a capitalistic society. However,

there is a desire to reverse this trend and operate in native languages, emphasizing the importance of creating a priority for using and preserving their languages within the context of sovereign nations.

The chatbot was considered a crucial tool for addressing language preservation issues. Participants emphasized the need to integrate language into the chatbot, with some suggesting a voice component as an ideal way to preserve the language and respect oral traditions. By incorporating tribal languages into various aspects of daily life, services, or infrastructure, a more immersive and effective language preservation effort can be achieved. If the AI chatbot offers a platform or resources for language revitalization efforts, it would be warmly welcomed by the community. Providing accessibility and support for language-related activities within the chatbot could motivate community members to engage, explore its features, and potentially increase their willingness to learn and interact with the technology.

“For you know those who are more English speaking they might not necessarily have to have all the instructions in, say Dakota, but it would be nice you know, culturally comforting. At least you say something, start off a sentence, or some sort of like, relating to the context a word or two, or a phrase, or something that is in Dakota, or something like that.” S-2

Participants suggested that language revitalization efforts should be tailored to each tribe’s unique language situation. They recommend incorporating cultural elements into instructions or sentences to offer a sense of cultural comfort, especially for those who may prefer it. The suggestion includes providing language options for users. For example, in emergency notifications, even if the entire response cannot be provided in different languages, including a brief alert message in the tribal language. This is particularly important for elders who may primarily speak the tribal language and have limited proficiency in English.

5. Discussion

This paper presents an online interview study with tribal community members in the Midwest of the United States, examining perspectives on the role of LLMs in preserving Indigenous culture and values. While this study is part of a broader ongoing co-design research agenda, here we focus specifically on interview findings that contribute valuable perspectives for the future development of LLMs, particularly in the domain of sovereignty and cultural preservation.

First, we emphasize the importance of prioritizing sovereignty not only in data handling but also in how AI systems represent Indigenous perspectives. Participants expressed concern that generic AI outputs reproduce dominant non-Native narratives, often erasing sovereignty and self-governance. For design, this means that training corpora and retrieval pipelines must explicitly center Indigenous-authored sources and allow tribes to determine what knowledge is represented, under what terms, and in what form. Respect for sovereignty in AI design requires not only technical safeguards (e.g., data governance protocols) but also epistemic ones: ensuring Indigenous knowledge systems are treated as authoritative rather than subordinated to Western notions of “scholarly” legitimacy (Wang et al., 2024). Second, our study showcases an artifact that employs a data embedding approach to preserve Indigenous culture, data, and values within LLMs, offering design implications for maintaining data sovereignty. This method illustrates a practical application of embedding Indigenous perspectives in AI technology while respecting their autonomy and cultural significance. However, our findings highlight that technical embedding alone is insufficient; adoption depends on relational trust. Systems should be accompanied by sustained partnerships, training opportunities for tribal members, and clear provenance indicators (e.g., showing the source and authorship of responses). In practice, this calls for designing AI that is transparent about its sources and gradually introduced through community-led demonstrations, so that trust in the system is built alongside trust in the people behind it.

5.1. Indigenous Knowledge, Land, and Data Sovereignty

Our findings emphasize the deep connection between Indigenous knowledge and the lands of tribal communities, demonstrating the importance of ensuring AI technologies serve these cultural and local needs. Indigenous knowledge is inherently tied to specific lands, embodying cultural narratives, practices, and values uniquely suited to each community’s history and context (Caranto Morford and Ansloos, 2021). Participants showed sensitivity to the dynamics between their sovereign rights and interactions with other governing entities, including state and federal bodies. They highlighted the potential for LLMs to support tribal businesses,

resources, and education, demonstrating how these tools could bolster sovereignty by equipping Indigenous communities with tailored, empowering resources (Tsai et al., 2023). To achieve this, AI must be designed to resist homogenization across tribes, enabling customization for local languages, traditions, and governance structures. Design frameworks should treat diversity and specificity not as a technical challenge to be minimized but as a cultural imperative to be honored.

An essential component of sovereignty is Indigenous Data Sovereignty, where Indigenous data, traditionally created within communities and often transmitted orally, is now digital. The National Congress of American Indians (NCAI) asserts that tribes have long served as data creators and stewards, with data embedded in cultural practices and communal roles (NCAI, 2018). However, as LLMs become more prominent, Indigenous data remains underrepresented online, raising concerns about authenticity and perpetuating misinformation. There is a critical need for tribal members to maintain control and ownership over their data in AI contexts, particularly in training and fine-tuning models. Since much Indigenous knowledge is preserved orally and relationally, designers must consider multimodal AI systems, incorporating audio, video, and storytelling formats that better reflect Indigenous epistemologies (Vamanu, 2012). A voice-enabled chatbot, for example, could help preserve oral traditions and languages while supporting intergenerational learning for youth living away from reservations.

One promising approach involves incorporating tribal members as citizen scientists, empowering them to become data creators and owners within the LLM development pipeline (Vadapalli et al., 2024). Our study introduces a framework that uses the Retrieval-Augmented Generation (RAG) method, which builds a local database of curated datasets to enhance response accuracy based on specific cultural contexts. While this approach moves towards Indigenous data sovereignty, our findings indicate that further development is necessary to ensure precise representation of Indigenous knowledge in AI models. Preserving Indigenous culture, values, and language in digital formats presents unique challenges. Efforts to create corpora (Dutta, 2019), library archives (Roy, 2015), and digital interfaces (Shiri et al., 2022) have met with complex historical and power dynamics that can undermine Indigenous communities' data sovereignty. A significant consideration has been the physical placement of data storage, where some propose locating servers on tribal lands rather than external institutions. Our approach reflects a land-based perspective on technology, respecting the ties between Indigenous sovereignty, land, and cultural knowledge (Ginsburg, 2016). By focusing on Indigenous perspectives in the digital preservation and dissemination of knowledge, this work highlights the need for AI systems that honor the cultural and sovereign rights of Indigenous communities.

5.2. Accountability, Empowerment, and Future Directions

In the context of preserving Indigenous knowledge and values, transparency and collaboration are crucial when working with tribal communities, involving clear communication about stakeholders and data processing, as well as continuous dialogue between community members, researchers, and government officials (Ghaith, 2024). Beyond transparency, AI design must incorporate ongoing community review mechanisms, ensuring that systems evolve with cultural needs and remain accountable to tribal authorities. This includes ethical reviews of what knowledge should or should not be digitized, and mechanisms to prevent reductive or stereotypical portrayals (Belarde-Lewis, 2011). Furthermore, achieving accountability in these systems requires a bi-directional relationship where technical performance feeds back into the governance rules established between sovereign entities. This ensures that AI-enabled tools remain responsible to the specific public values and administrative needs of the tribal communities they serve (Chen et al., 2025). Our research team's long-term project with tribes has demonstrated the value of a mutual learning process, where researchers strive to understand Indigenous perceptions, values, and knowledge while sharing scientific findings to empower them, exemplifying a commitment to inclusive design and respect for Indigenous autonomy (Ghaith, 2024).

By incorporating Indigenous perspectives in state-of-the-art AI technologies and prioritizing human expertise in evaluations, we can ensure that AI systems are continuously refined and adapted to maintain ethical standards and cultural sensitivity. Ultimately, these systems should not only preserve cultural knowledge but also serve as tools for empowerment, resisting misrepresentation, supporting sovereignty, and strengthening identity across generations.

Table 2 synthesizes the key design principles emerging from our findings, mapping participant insights to concrete implications for LLM design in Indigenous contexts.

Tab. 2 – Design Implications for LLMs in Indigenous Communities

Design Principle	Finding	Implication for LLM Design
Sovereignty-first	Generic AI erases Native voice	Center Indigenous-authored sources
Relational trust	Trust in tool = trust in team	Community-led onboarding
Oral/visual preference	Written responses feel alienating	Multimodal outputs
Data provenance	Participants questioned info sources	Show authorship/source in responses

A significant limitation of this research is the power imbalance in the AI co-design process, which arises from the unintentional lack of technical knowledge about AI among indigenous tribe participants, despite efforts to design collaborative iterations and foster long-term relationships that respect their autonomy and cultural significance. As participants' trust in the chatbot is closely tied to their trust in the research team, establishing trust with tribal communities is paramount, but it is also important to address the power imbalance by educating them about the fundamental workings of AI, including data collection and utilization. To overcome this challenge, future work aims to design comprehensive participatory design studies that empower tribal communities to collaborate more effectively, providing informed feedback and design ideas for the iterative improvement and evaluation of TribalLLM, ultimately enabling the LLM chatbot to better support tribal communities in protecting and reinforcing their sovereignty and culture.

6. Conclusion

This paper has presented an iterative co-design study with tribal communities in the Midwest United States, providing valuable insights into integrating LLMs in indigenous contexts. The study began by introducing ChatGPT to tribal members, revealing its lack of cultural sensitivity and inadequate representation of tribal culture and values. Addressing this, the Retrieval-Augmented Generation (RAG) method was utilized, enabling the incorporation of tribally authorized data into LLMs. This approach ensured the chatbot's functionality, cultural authenticity, and respectfulness, aligning with indigenous data sovereignty principles.

Our findings extend beyond technicalities to explore AI's ethical and cultural implications in the context of indigenous knowledge and history. This research contributes significantly to the discourse on LLMs, especially regarding cultural preservation. It underscores the need for data sovereignty and respectful handling of culturally sensitive information. Our co-design approach with tribal communities has set a precedent in HCI research, demonstrating the importance of inclusive design, collaboration, and mutual respect. These partnerships ensure that the developed technologies are advanced and functional, culturally sensitive, and reflective of the communities they serve. This collaborative model advocates for a deep engagement in technology design, paving the way for future research and development in HCI that is both inclusive and respectful of indigenous voices and perspectives.

Contributor Statement

- **Srishti Gupta:** Conceptualization, Investigation, Data Curation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing.
- **Taranum Bano:** Resources, Investigation, Writing – Original Draft, Writing – Review & Editing.
- **Yu-Che Chen:** Resources, Investigation, Writing – Review & Editing.
- **Edouardo Zendejas:** Resources, Investigation.
- **Sarah Krafka:** Resources, Investigation.
- **Chun-Hua Tsai:** Supervision, Conceptualization, Methodology, Validation, Writing – Original Draft, Writing – Review & Editing.

Use of AI

AI tools (such as GenAI assistance for Latex writing in Overleaf) were used in certain places to correct and improve grammar and sentence structures.

Conflict of Interest

There is no conflict of interest.

References

- Anuyah, O., Wan, R., Adejoro, C., Yeh, T., Metoyer, R., and Badillo-Urquiola, K. (2023). Cultural considerations in ai systems for the global south: A systematic review. In *Proceedings of the 4th African Human Computer Interaction Conference*, pages 125–134.
- Baek, J., Jeong, S., Kang, M., Park, J. C., and Hwang, S. J. (2023). Knowledge-augmented language model verification. *arXiv preprint arXiv:2310.12836*.
- Barnosky, J. T., Lauland, A., Miro, M. E., Balagna, J., Ecola, L., Kim, S., Kolb, C., Leuschner, K. J., Mitch, I., Parker, A. M., Payne, L. A., Price, C. C., Reese, T., Regan, L., Resetar, S. A., Schnaubelt, C. M., Steratore, R., and Sukkamp, K. M. (2022). Streamlining Emergency Management: Issues, Impacts, and Options for Improvement. Technical report, RAND Corporation.
- Belarde-Lewis, M. (2011). Sharing the private in public: Indigenous cultural property and online media. In *Proceedings of the 2011 iConference*, pages 16–24. ACM New York, NY, USA.
- Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.
- Burnette, C. E., Sanders, S., Butcher, H. K., and Rand, J. T. (2014). A Toolkit for Ethical and Culturally Sensitive Research: An Application with Indigenous Communities. *Ethics and Social Welfare*, 8(4):364–382. Publisher: Routledge _eprint: <https://doi.org/10.1080/17496535.2014.885987>.
- Caine, K. (2016). Local standards for sample size at chi. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 981–992.
- Caranto Morford, A. and Ansloos, J. (2021). Indigenous sovereignty in digital territory: a qualitative study on land-based relations with# nativetwitter. *AlterNative: An International Journal of Indigenous Peoples*, 17(2):293–305.
- Carroll, S. R., Rodriguez-Lonebear, D., and Martinez, A. (2019). Indigenous data governance: strategies from united states native nations. *Data Science Journal*, 18.
- Chandran, R. (2023). Indigenous groups fear culture distortion as ai learns their languages. *The Japan Times*.
- Chen, Y.-C., Tsai, C.-H., and Zendejas, E. (2025). Multi-level collaborative governance framework for designing accountable ai systems for emergency management. In *Conference on Digital Government Research*, volume 26.
- Colas, C., Karch, T., Moulin-Frier, C., and Oudeyer, P.-Y. (2022). Language and culture internalization for human-like autotelic ai. *Nature Machine Intelligence*, 4(12):1068–1076.
- Duarte, M. E. (2017). *Network sovereignty: Building the internet across Indian country*. University of Washington Press.
- Dutta, U. (2019). Digital preservation of indigenous culture and narratives from the global south: In search of an approach. *Humanities*, 8(2):68.
- Fraisse, A., Zhang, Z., Zhai, A., Jenn, R., Fisher Fishkin, S., Zweigenbaum, P., Favier, L., and Mustafa El Hadi, W. (2019). A sustainable and open access knowledge organization model to preserve cultural heritage and language diversity. *Information*, 10(10):303.
- Gallamaso, J. D. (2024). Indigenous culture in the age of artificial intelligence: Challenges and moral responsibility. *Ignatian International Journal for Multidisciplinary Research*, 2(7):824–837.
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., and Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.
- Ghaith, K. (2024). Ai integration in cultural heritage conservation—ethical considerations and the human imperative. *International Journal of Emerging and Disruptive Innovation in Education: VISIONARIUM*, 2(1):6.
- Ginsburg, F. (2016). Indigenous media from u-matic to youtube: Media sovereignty in the digital age. *Sociologia & Antropologia*, 6:581–599.
- Goulding, A., Campbell-Meier, J., and Sylvester, A. (2020). Indigenous cultural sustainability in a digital world: Two case studies from aotearoa new zealand. In *International Conference on Information*, pages 66–75. Springer.
- Gupta, S., Chen, Y.-C., and Tsai, C. (2024). Utilizing large language models in tribal emergency management. In *Companion Proceedings of the 29th International Conference on Intelligent User Interfaces*, pages 1–6.
- Herscovich, D., Frank, S., Lent, H., de Lhoneux, M., Abdou, M., Brandl, S., Bugliarello, E., Piqueras, L. C., Chalkidis, I., Cui, R., et al. (2022). Challenges and strategies in cross-cultural nlp. *arXiv preprint arXiv:2203.10020*.

-
- Holguin, B. E. (2023). *An Indigenous Archaeological Perspective on the Use of Artificial Intelligence in Reconstructing Chumash History*. PhD thesis, University of California, Santa Barbara.
- Kasneji, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., et al. (2023). Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103:102274.
- Kruse, C. S., Bouffard, S., Dougherty, M., and Parro, J. S. (2016). Telemedicine use in rural native american communities in the era of the aca: a systematic literature review. *Journal of medical systems*, 40:1–9.
- Kukutai, T. and Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. ANU press.
- Lewis, J. E., Abdilla, A., Arista, N., Baker, K., Benesiinaabandan, S., Brown, M., Cheung, M., Coleman, M., Cordes, A., Davison, J., Duncan, K., Garzon, S., Harrell, D. F., Jones, P.-L., Kealiikanakaoleohaililani, K., Kelleher, M., Kite, S., Lagon, O., Leigh, J., Levesque, M., Mahelona, K., Moses, C., Nahuewai, I. I., Noe, K., Olson, D., Parker Jones, ., Running Wolf, C., Running Wolf, M., Silva, M., Fragnito, S., and Whaanga, H. (2020a). Indigenous Protocol and Artificial Intelligence Position Paper. Place: Honolulu, HI Publisher: Indigenous Protocol and Artificial Intelligence Working Group and the Canadian Institute for Advanced Research.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., et al. (2020b). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.
- Munn, L. (2024). The five tests: designing and evaluating ai according to indigenous māori principles. *AI & SOCIETY*, 39(4):1673–1681.
- NCAI (2018). Resolution kan-18-011: Support of us indigenous data sovereignty and inclusion of tribes in the development of tribal data governance principles.
- Park, N., Regenbrecht, H., Duncan, S., Mills, S., Lindeman, R. W., Pantidi, N., and Whaanga, H. (2022). Mixed reality co-design for indigenous culture preservation & continuation. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 149–157. IEEE.
- Potts, C., Ennis, E., Bond, R., Mulvenna, M., McTear, M. F., Boyd, K., Broderick, T., Malcolm, M., Kuosmanen, L., Nieminen, H., et al. (2021). Chatbots to support mental wellbeing of people living in rural areas: Can user groups contribute to co-design? *Journal of Technology in Behavioral Science*, 6:652–665.
- Qureshi, A., Wilson, D., and Sarantou, M. (2025). Indigenous communities re-interpreting and preserving cultural heritage through narratives while navigating the digital age. In *Digital Indigenous Cultural Heritage*, pages 121–143. Springer.
- Roy, L. (2015). Indigenous cultural heritage preservation: A review essay with ideas for the future. *IFLA journal*, 41(3):192–203.
- Shiri, A., Howard, D., and Farnel, S. (2022). Indigenous digital storytelling: digital interfaces supporting cultural heritage preservation and access. *International Information & Library Review*, 54(2):93–114.
- Smith, R. C., Winschiers-Theophilus, H., Loi, D., de Paula, R. A., Kambunga, A. P., Samuel, M. M., and Zaman, T. (2021). Decolonizing design practices: Towards pluriversality. In *Extended abstracts of the 2021 CHI conference on human factors in computing systems*, pages 1–5.
- Spennemann, D. H. (2023). Chatgpt and the generation of digitally born “knowledge”: How does a generative ai language model interpret cultural heritage values? *Knowledge*, 3(3):480–512.
- Stap, D. and Araabi, A. (2023). Chatgpt is not a good indigenous translator. In *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)*, pages 163–167.
- Susnjak, T. (2023). Applying bert and chatgpt for sentiment analysis of lyme disease in scientific literature. *arXiv preprint arXiv:2302.06474*.
- Tsai, C.-H., Rayi, P., Kadire, S., Wang, Y.-F., Krafka, S., Zendejas, E., and Chen, Y.-C. (2023). Co-design disaster management chatbot with indigenous communities. In *20th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, volume 2023, page 1.
- Unru, S. A., Zhodzikov, A. V., et al. (2022). The role of digitalisation in preserving the cultural heritage of indigenous peoples. *European Proceedings of Social and Behavioural Sciences*.
- Vadapalli, J., Gupta, S., Karki, B., and Tsai, C.-H. (2024). Incorporating citizen-generated data into large language models. In *Proceedings of the 25th Annual International Conference on Digital Government Research*, pages 1023–1025.
- Vamanu, I. (2012). Dimensions of indigenous heritage discourse and the use of technologies of display in an ethnographic exhibition: a case study of the” song for the horse nation” exhibition (national museum of the american indian, new york). In *Proceedings of the 2012 iConference*, pages 641–643.
- Wang, Y.-F., Chen, Y.-C., Huang, Y.-C., Redwing, C., and Tsai, C.-H. (2024). Tribal knowledge cocreation in generative artificial intelligence systems. In *Proceedings of the 25th Annual International Conference on Digital Government Research*, pages 637–644.
- Winschiers-Theophilus, H., Zaman, T., and Stanley, C. (2019). A classification of cultural engagements in community technology design: introducing a transcultural approach. *Ai & Society*, 34:419–435.