Designing a writing assistant for the Nheengatu Language

Desenvolvendo um assistente de escrita para a língua nheengatu

Claudio Pinhanez, Raquel Cordeiro, João Paulo Bento, Julio Nogima, Enrico Santarelli

participatory design, indigenous languages, digital literacy, co-creation, writing tools

The recent expansion of internet access in Indigenous villages and communities across the Amazon has opened new opportunities for learning, communication, and cultural exchange. While many young people are bilingual, the predominance of Portuguese in digital environments poses risks to bilingual education and to the preservation of Indigenous languages. This paper presents the development of a digital tool designed to support the written use of Indigenous languages – specifically Nheengatu – within online contexts. The goal is to encourage greater use of Nheengatu on the internet by enabling young people to produce digital content in their own language and to strengthen their familiarity with it in everyday digital practices. We describe the design and co-creation process of these tools, developed in collaboration with Indigenous teachers and students through participatory workshops, usability evaluations, and community interviews. The paper discusses the first learnings from this process and outlines ongoing efforts to expand the initiative to other Indigenous languages and educational settings.

design participativo, línguas indígenas, letramento digital, cocriação, ferramentas de escrita

A recente expansão do acesso à internet em aldeias e comunidades indígenas em toda a Amazônia abriu novas oportunidades para aprendizagem, comunicação e intercâmbio cultural. Embora muitos jovens sejam bilíngues, a predominância do português nos ambientes digitais representa riscos para a educação bilíngue e para a preservação das línguas indígenas. Este artigo apresenta o desenvolvimento de uma ferramenta digital projetada para apoiar o uso escrito de línguas indígenas – especificamente o nheengatu – em contextos on-line. O objetivo é incentivar o uso mais frequente do nheengatu na internet, possibilitando que jovens produzam conteúdos digitais em sua própria língua e fortaleçam sua familiaridade com ela nas práticas digitais cotidianas. Descrevemos o processo de design e de cocriação dessas ferramentas, desenvolvido em colaboração com professores e estudantes indígenas por meio de oficinas participativas, avaliações de usabilidade e entrevistas com a comunidade. O artigo discute os primeiros aprendizados desse processo e delineia os esforços em andamento para ampliar a iniciativa para outras línguas indígenas e contextos educacionais.

1 Introduction

The growing expansion of internet access in Indigenous villages and communities across the Amazon, driven by satellite connectivity technologies, has significantly broadened access to knowledge and opportunities – particularly among adolescents and young people. While many of these individuals are bilingual, the near-exclusive use of Portuguese in digital spaces poses a risk to bilingual education and may contribute to the gradual weakening of the use of their Indigenous language. This concern extends beyond the Amazon, affecting Indigenous peoples across diverse regions, cultures, and linguistic families. Although native languages are often spoken in daily life, their limited presence in online environments raises concerns about intergenerational transmission and long-term vitality.

A key factor contributing to this situation is the lack of appropriate digital writing tools for Indigenous languages – such as specialized keyboards, spellcheckers, autocomplete systems, and digital dictionaries. These tools, which are widely available for Portuguese and other dominant languages, enhance the ease and quality of writing while reducing the effort required to produce digital texts. While most children first learn to read and write using paper-based media, digital literacy rapidly becomes essential as they transition to communication through word processors, email, and platforms such as WhatsApp. For Indigenous languages to remain vibrant and relevant, it is crucial that these digital literacy practices also occur within the Indigenous languages themselves.

To address this need, we developed a digital writing assistant designed to simplify and foster the use of Indigenous languages among community members – particularly teenagers and young adults. The development process combines multiple component tools, which are prototyped, integrated, and tested with Indigenous speakers in authentic educational and communicative contexts. We initiated the prototype in collaboration with students and teachers, conducted a usability inspection, and performed community interviews to understand local needs and practices. Currently, the project is evolving through co-creation workshops with Indigenous students and educators who are actively shaping the tool's design and functionality.

This article presents the context and foundations of this project, outlining previous work that motivated the research, followed by a detailed account of the co-design process and the main findings so far. We conclude by reflecting on the lessons learned, methodological challenges, and future directions for technology-supported language revitalization and intercultural design.

2 Strengthening Brazilian indigenous languages

Brazil is home to a remarkable linguistic diversity. There are 391 Indigenous ethnic groups in the country, and 295 Indigenous languages are spoken. Between 2010 and 2022, the number of speakers increased from 293,853 to 433,980, however, due to the growing dominance of Portuguese, the proportion of Indigenous people who reported speaking an Indigenous

language decreased from 37.35% to 28.51% (Instituto Brasileiro de Geografia e Estatística [IBGE], 2022). This alarming scenario underscores the urgency of developing effective strategies for the preservation and revitalization of Indigenous languages.

As highlighted in a recent publication by the United Nations Educational, Scientific and Cultural Organization (UNESCO), technologies such as Artificial Intelligence (AI), Natural Language Processing (NLP), and automated speech recognition can play a critical role in language revitalization efforts (Llanes-Ortiz, 2023). In this context, the *Prolind* project[1] has been working toward strengthening Brazilian Indigenous languages through AI. The main goal of this initiative is to explore the development of technological tools which support Indigenous communities in documenting, preserving, and revitalizing their languages.

1 https://c4ai.inova.usp.br/research.html#research

The project team comprises a multidisciplinary group of researchers, professors, professionals, students, and interns from the University of São Paulo (USP), São Paulo State University (UNESP), and IBM Research Brazil. Since the project's inception in 2022, the team has been in contact with several Indigenous communities in Brazil, including the Guarani Mbya, Guarani Kaiowá, Guarani Nhandewa, Tupi, Terena, Baré, Wassu, Tukano, Pankararu, Zoé, and Mehinako peoples. The project also engages with key organizations working with Indigenous communities, such as the Ministry of Indigenous Peoples, the Inter-American Development Bank (IDB), the National Foundation of Indigenous Peoples (FUNAI), the Plurinational Union of Indigenous Students (UPEI), the Socio-Environmental Institute (ISA), and other NGOs.

Between August 2024 and July 2025, our work has focused on the Nheengatu language, spoken by approximately 13,000 people across three different areas of the Amazon region and parts of Northeastern Brazil. Nheengatu is used by various ethnic groups, including cases where it has been adopted by peoples who lost their original languages, such as the Baré people (Epps & Stenzel, 2013). Since mid-2023, the team has engaged with multiple leaders and groups connected to the Nheengatu language across Brazil. These collaborations have taken place alongside the development of preliminary machine translation models (to and from Brazilian Portuguese) and early prototypes of a writing assistant designed for Nheengatu.

The focus of this paper is the *Yẽgatu*[2] *Digital* project, which explores the use of AI-based tools in Indigenous middle and high schools serving Baré communities near São Gabriel da Cachoeira, one of the most linguistically diverse areas in the Amazon. This project was developed in partnership with FOIRN[3] (Federation of the Indigenous Organizations of the Rio Negro), one of the largest Indigneous organizations in Brazil.

2 Yẽgatu is *Nheengatu* written according to the orthography used in the Upper Rio Negro region.

3 https://foirn.org.br/

3 AI-based tools for indigenous language

There has been considerable work, especially in the last five years, exploring how Artificial Intelligence (AI) technologies can be applied to support endangered and Indigenous languages. Kuhn et al. (2020) discussed multiple

initiatives from the Indigenous Languages Technology (ILT) project at the National Research Council of Canada, including the construction of corpora for several languages, annotation tools, speech recognition systems, and read-along audiobooks. Neubig et al. (2020) summarized a 2019 workshop on the state of technology for language documentation, while Mager et al. (2023) provided an in-depth discussion on the challenges and common approaches to developing machine translation systems for Indigenous languages in the Americas. C. Pinhanez et al. (2023) also presented several applications of NLP technologies in scenarios involving endangered languages.

Our work on developing digital writing assistants began with the identification of a need expressed initially by Guarani Mbya speakers (Pinhanez et al., 2024) and has since been centered on Nheengatu-speaking communities. The first prototype of the app was developed in partnership with undergraduate students from Insper, an engineering college in São Paulo, through a special collaboration program (Figure 1).

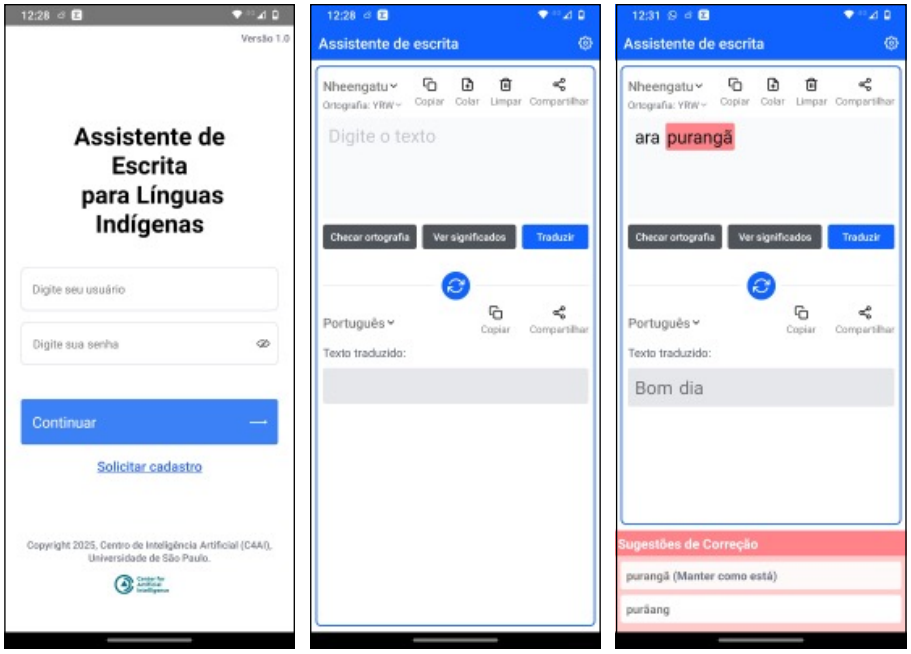

Figure 1 Initial prototype showing the login screen, translation tool, and spell-checking features of the writing assistant.

The main tools we have been building include:

- **Word Dictionary**: provides access to words, meanings, and translations based on approximate search.
- **Word Completion**: suggests words to complete a partially typed entry.
- **Next-Word Prediction**: suggests the next word in a partially typed sentence.
- **Spell Checker**: suggests corrections for words in partially or fully typed sentences.

- **Translator**: Translates words and sentences to and from the Indigenous language and another language, or between different orthographies of the same language (technical details of this tool were described in the previous section).

In many ways, these are typical writing-support tools like those available in mainstream word processors, social media platforms, and other applications for widely spoken languages. However, such systems are usually developed by large, specialized teams with substantial resources and ongoing technical support – conditions rarely available for endangered Indigenous languages. This limitation has prompted us to develop alternative strategies for building these tools under constraints of limited data and technical resources.

As part of the development process, we conducted a **usability inspection** of the current prototype version to identify design issues and generate recommendations for improvement. The method used was a **heuristic evaluation**, in which an expert reviews the interface according to Nielsen's ten usability heuristics[4] (Nielsen, 1994). Each usability issue identified was classified by severity:

- **Low**: minor confusion without hindering task completion.
- **Medium**: causes user frustration or task errors.
- **High**: prevents task completion, leads to significant misunderstanding, or causes loss of important information.

In total, **20 usability issues** were identified: 8 of low severity, 8 of medium, and 4 of high (Table 1). The most frequently violated heuristic was *recognition rather than recall*, cited six times. Other commonly affected heuristics included *visibility of system status*, *consistency and standards*, and *error prevention* (each cited four times). *Match between the system and the real world*, *flexibility and efficiency of use*, *aesthetic and minimalist design*, and *help and documentation* were each cited three times. *User control and freedom* was cited twice, and *help users recognize, diagnose, and recover from errors* once.

4 1) Visibility of system status;
2) Match between system and the real world;
3) User control and freedom;
4) Consistency and standards;
5) Error prevention;
6) Recognition rather than recall;
7) Flexibility and efficiency of use;
8) Aesthetic and minimalist design;
9) Help users recognize, diagnose, and recover from errors;
10) Help and documentation.

Table 1 Usability issues identified during the heuristic evaluation.

Issue	Severity	Heuristic	Recommendations
App icon does not reflect the visual identity	Low	6	• Co-create a visual identity and app icon with students.
No option to remember username and password	Medium	5, 6, 10	• Add a "Remember me" option. • Ask whether the user wants to save the password on the device. • Add a "Forgot my password" link.
Error message provides no guidance for correction	Low	9	• Include guidance such as: "Check if you typed it correctly. Pay attention to extra spaces and uppercase/lowercase (Caps Lock)."
"Check spelling" function is active even with no text	High	4,5	• Show an immediate message indicating text is required. • Disable the button when no text is entered.

Table 1 Usability issues identified during the heuristic evaluation. (continued)

Issue	Severity	Heuristic	Recommendations
"See meanings" button active without content	Low	5	• Disable the button when there is no content to analyze.
Translation requires multiple taps	Medium	7	• Perform the translation and close the keyboard with a single tap.
Orthography abbreviations are difficult to understand	Medium	2	• Automatically select orthography based on user login. • Include a sample word for each orthography. • Add an explanation or hint in parentheses.
Lack of guidance after clicking "See meanings"	Medium	1, 6	• Automatically display the meaning of the first word. • Visually indicate clickable words.
"See meanings" button does not reflect the highlighted word	Low	1, 6	• Change button state to reflect when the function is active.
Meaning area is limited and hard to navigate	Medium	4, 8	• Reorganize layout to show meaning below the text, replacing the translation area temporarily with the selected word.
Reduced space for dictionary interaction	Medium	8	• Optimize button spacing (copy, paste, clear, share). • Reduce text box height. • Remove the share button from the input area.
Little indication that more content is scrollable	High	4	• Add a scroll bar, visual cut-off, or increase visible area.
Special characters bar occupies too much space	Low	8	• Redesign to fit compactly in a single row.
"Copied" message disappears too quickly	Medium	1, 6	• Increase display time to ensure readability.
Long-press selection of words is not supported	Low	1, 2, 4	• Allow long-press selection consistent with other apps.
Paste icon does not follow recognizable standards	Low	2	• Replace with standard copy/paste icons.
No help or support channel available	Low	10	• Add a help icon with a basic guide. • Add a link for reporting issues or suggestions.
Large text size breaks the top menu layout	High	5, 7	• Adjust layout to accommodate larger text sizes.
Menu terms cause confusion	High	4, 10	• Use clearer or more consistent terms. • Add a quick explanation in a help button.
Settings are not saved after closing the app	Medium	3, 6, 7	• Save and persist user settings.

For each identified issue, we proposed recommendations and created a new **interactive prototype in Figma** to illustrate the suggested improvements. These usability evaluations not only informed technical improvements to the prototype but also set the stage for a deeper collaborative process with Indigenous communities. The next phase of our work focused on co-creating the writing assistant together with Baré educators and students, ensuring that its design and functionality truly reflected local linguistic practices, educational needs, and cultural values.

4 Co-creating a writing assistant

This project is grounded in the principles established by the Los Pinos Declaration (2020), which defines the key elements guiding all actions for Indigenous languages, particularly within the context of the UNESCO International Decade of Indigenous Languages (2022–2032). Among these principles, one of the most fundamental is the commitment to work with and for Indigenous communities from the very beginning of any initiative.

Following these principles, our team has committed to not only collaborating with Indigenous communities but also critically reflecting on the ethical implications of AI research practices. This includes questioning conventional data collection, algorithmic modeling, and validation methods that often overlook the agency, consent, and cultural specificity of the communities involved.

Following Indigenous protocols, the engagement process with the Baré communities began in April 2024 with a presentation during a local organization's assembly in which the community approved the idea of establishing a partnership and authorized the researchers and technicians to begin work within the Indigenous territory. In September of 2024 we visited two communities who had volunteered to host the project and, together with their leadership and school teachers, established the main tenants of the project. In May 2025, we signed a partnership between the University of São Paulo (USP), the Federation of Indigenous Organizations of the Rio Negro (FOIRN), and the two Baré communities to launch a weekly workshop program using a hybrid digital classroom. As accorded with the communities, the program aimed to support the use of the writing assistant in local middle and high schools, using tablets and satellite-based Internet connections. Initially, the activities took place in two schools located deep within the Amazon forest, with plans to extend the program to other schools with similar conditions.

Our goal has been to co-design and co-develop AI-based tools for Nheengatu in close collaboration with these communities, studying their adoption in real-life contexts. Through this process, we aim to contribute to the ongoing efforts of Indigenous peoples to ensure that their languages remain alive – spoken and written by both younger and older generations, not only in daily communication but also increasingly within the digital realm.

Additionally, the program seeks to explore, together with teachers and local content creators, how such tools can foster the production of more diverse and higher-quality digital materials written in Indigenous languages. By doing so, the initiative aspires to bridge traditional linguistic practices and contemporary digital expression.

4.1 Working with teachers to refine orthography and language use in the app

We consider the current design, content, and capabilities of our tools as only an initial step – a foundation to enable early community use and jump-start the co-development process. We expect that further linguistic data and

community input will be necessary to make the tools useful across a wider variety of contexts and communication needs.

Since mid-April 2025, we have been collaborating remotely with two educators from the Tabocal and Juruti communities to refine the writing assistant and better support the written use of their Indigenous languages. This collaboration has followed a participatory and iterative approach, organized around three main activities:

- **Dictionary review and word meaning verification** – we supported and observed the teachers' process of reviewing dictionary entries and correcting word meanings. This task was essential for ensuring linguistic accuracy and cultural appropriateness within the app, as well as aligning orthographic variants and preferred spellings used in the communities.
- **Text creation within the app** – the teachers were invited to compose short texts directly in the writing assistant. This hands-on activity allowed us to identify real-world usability challenges and gather detailed feedback on interface design, word suggestions, and error correction features. Their input guided concrete interface improvements and new functionality priorities.
- **Content filtering for educational contexts** – in a subsequent testing round, the educators helped identify terms and expressions that might be inappropriate for school contexts. This feedback supported the development of culturally sensitive content filters, ensuring that the tool remains aligned with the communities' pedagogical and ethical values.

This phase has been crucial in aligning the app's linguistic resources with the communities' actual language use and orthographic conventions, while also strengthening trust and ownership among local educators who will play a key role in promoting the tool's continued adoption and improvement.

4.2 Digital classrooms

As part of the partnership agreement, in **July 2025**, we deployed two fully equipped **digital classrooms**, one in each community school, to support the ongoing co-creation activities. The equipment was donated by a private donor and approved by **FOIRN** and **USP**. Each classroom includes **20 student tablets**, **one laptop for the teacher**, **one laptop to control the classroom system**, **a 40-inch monitor**, **a video-conferencing camera**, **a Starlink Internet system**, and a **250 W solar panel kit** to power and charge all devices (Figure 2).

The digital classrooms were designed to enable **hybrid workshops**, integrating Indigenous teachers and students with remote researchers. This setup allows for continuous collaboration even across vast geographical distances within the Amazon region. The primary objective was to provide an **educational infrastructure** that supports teaching and learning in **Nheengatu**, facilitating reading, writing, and the creation of digital materials in the language.

Figure 2 Students using tablets, with the wall-mounted monitor and camera that enable online collaborative workshops.

The digital tools and initial prototypes of the writing assistants developed by our team were **embedded into this platform**, allowing the classrooms to serve as both **learning environments and co-design spaces**. Through these hybrid interactions, teachers and students can explore new ways of engaging with written Nheengatu while contributing directly to the tool's iterative design and development process.

4.3 Interviews with the community

During the field trip to install the digital classrooms, we also took the opportunity to talk with members of both communities to better understand their needs and everyday experiences with digital tools and the Nheengatu language. Over the course of a week, we interviewed 25 people, including teachers and students aged 11 to 20, from Juruti and Tabocal.

With teachers, we focused on the digital tools they use, the main challenges they face in the classroom, and what could help strengthen the teaching of Nheengatu in online environments. With students, we explored their language practices – whether they speak Nheengatu from childhood, how and when they use it, and what kind of content they consume or produce online.

Most students use Facebook Messenger as their primary communication platform, rather than WhatsApp or Instagram, as it connects directly with their social circles on Facebook. Girls tend to spend more time watching short videos on Kwai, while boys often play Free Fire, sometimes in groups where they communicate in Portuguese. This pattern suggests an opportunity to use games as informal spaces for practicing Nheengatu.

Although most students can read and write Nheengatu, many struggle with pronunciation and express embarrassment or fear of making mistakes

when speaking. Technical limitations also affect writing: most mobile phones lack the necessary diacritics, such as the tilde, making it difficult to spell words correctly. Teachers confirmed these challenges and noted that, in class, students often learn by first forming sentences in Portuguese and then translating them into Nheengatu.

Teachers also mentioned the need for structured teaching materials to guide substitute instructors and support classroom continuity. Some expressed concerns about children's exposure to online content and the risk of excessive gaming, emphasizing the importance of training for both teachers and parents on safe and educational uses of technology.

From these conversations, several opportunities emerged for the next stages of co-design: integrating the app with Messenger, adding an audio feature to support pronunciation, developing a keyboard with Nheengatu-specific diacritics, and creating game-based learning experiences to encourage practice in informal settings. The results also point to the need for digital literacy initiatives to ensure technology becomes an ally in preserving and strengthening the use of Nheengatu.

4.4 Co-creation workshops

The co-creation workshops for the writing assistant began in August 2025, with three weekly 90-minute sessions: one with middle school students from Juruti (around 20 participants), one with upper middle school students from Tabocal (around 15 participants), and one with high school students from Tabocal (around 40 participants). Due to high interest, the high school group was divided into three rotating subgroups. After the first month, adjustments were made to improve participation: five new 7th-grade students joined the Tabocal middle school group, and the large high school cohort was divided into two weekly sessions. The program currently runs four workshops per week – two for middle school and two for high school students – including recent graduates who asked to remain involved.

Workshops are conducted via Google Meet, using the digital classroom equipment installed earlier in the year (Figure 3). The classroom monitor displays the remote facilitators, while the camera and microphone capture the group's audio and video. Screen sharing is used for presentations, and clear communication protocols have been established, including muting microphones when not speaking. This hybrid setup allows for dynamic interaction between students, teachers, and researchers, despite the distance between the communities and the university team.

The first collective activity was to co-create the app's name, encouraging students to take ownership of the tool while practicing digital writing. We adapted the Crazy 8s ideation method, in which participants generate eight ideas in quick succession. To fit the online format, we created a shared Google Slides document with eight slides per student. However, the original one-minute-per-idea pace proved too fast for participants still adapting to the digital tools, so the activity time was extended.

Figure 3 Teachers, students, and facilitators collaborating in a hybrid session to test and discuss improvements to the writing assistant.

The second step involved building a collective mind map in Miro. The facilitator transcribed students' contributions while they explained the meaning of each term in Nheengatu. To deepen the reflections, we asked each student to write four words or phrases related to "assistant," "writing," "languages," and "Indigenous." This change led to richer responses, as Nheengatu often conveys ideas through short expressions rather than isolated words. The resulting map offered important insights: many students described the assistant as something that would "teach," revealing a perception that required clarification – that the tool is meant to support writing, not replace teachers.

After the ideation sessions, we merged all contributions into a single mind map, grouping similar words and translating them into Nheengatu (Figure 4). The translated list – reviewed by teachers and students – was then transferred to a Google Form for community voting (Figure 5). The form included the 23 name proposals, each shown with its translation into Portuguese (in parentheses). Students voted during class, and the link was also shared on WhatsApp, allowing others in the community who were not attending the workshops to participate in choosing the final name.

In total, we received 96 responses. The most voted names were Pinimasa Yẽgatu Rupi ("Writing in Yẽgatu"), with 37 votes, followed by Baré Tayega ("Language of the Barés"), with 31 votes. After announcing the results in all workshops, we updated the app to display the community-chosen name, allowing participants to immediately see their ideas reflected in the tool they helped create (Figure 5).

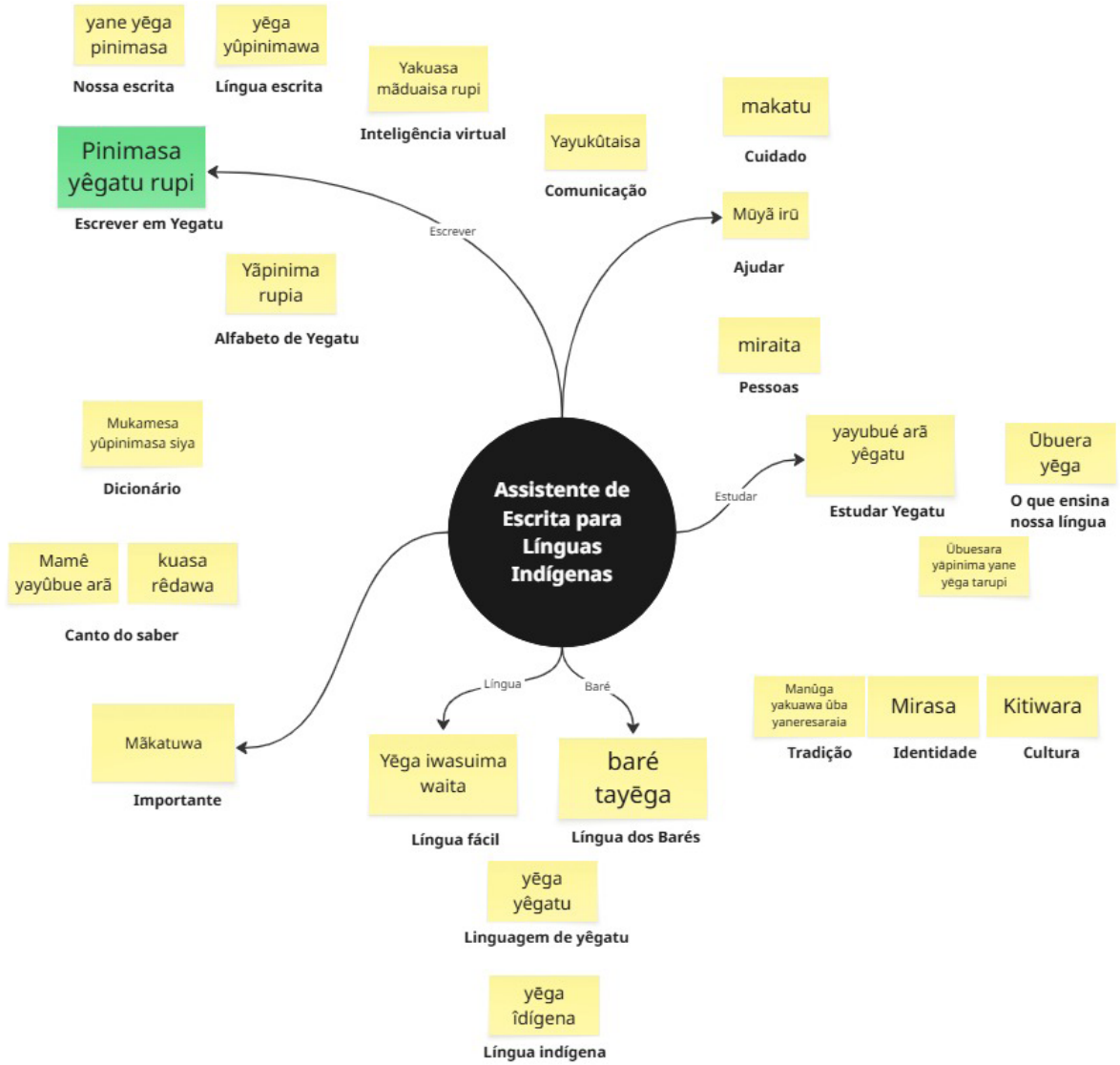

Figure 4 Ideas generated by students during a participatory activity to select a name for the writing assistant.

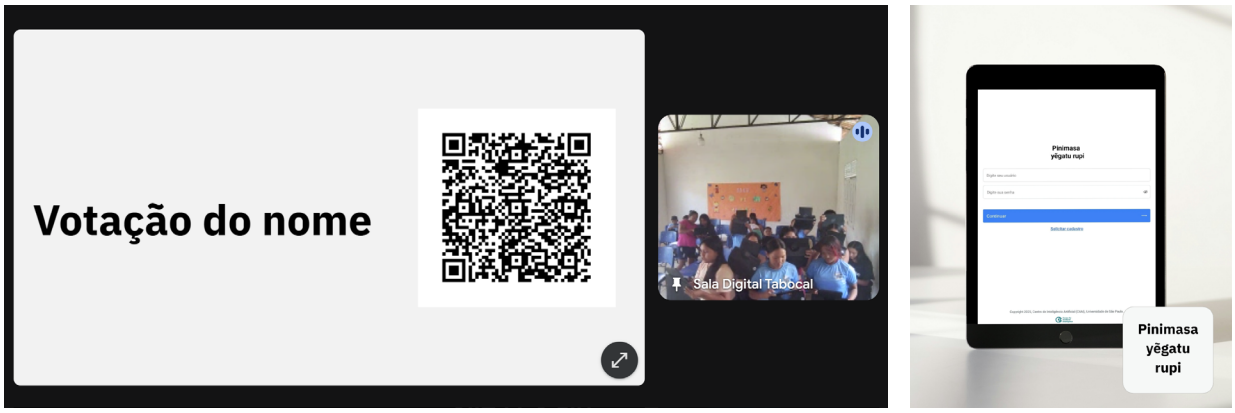

Figure 5 Google Forms QR code used for students to vote on their preferred app name, and the updated interface showing the selected name.

5 Conclusion

The weekly workshops and the ongoing co-creation of the Yẽgatu Digital app continue to evolve. The next phase focuses on observing writing classes in Nheengatu to better understand students' main challenges and how the app can effectively support them. The remote workshops have already proven valuable, making it possible to maintain an intense, continuous weekly collaboration that has brought significant mutual learning.

The creation of a "third space" for hybrid collaboration – neither entirely local nor external – has emerged as a promising model for inclusive and participatory design that could be adapted to other remote or underserved educational contexts. However, it is still too early to fully assess the success of this model, as the co-design phase is still in progress. It also remains to be seen how this hybrid space will evolve as it becomes more integrated into routine school activities and more familiar to both local and remote participants. Nevertheless, this approach already shows potential as a technology-supported co-design methodology, fostering agency among young learners by linking digital literacy with cultural identity and offering pathways for sustained engagement rather than short-term interventions.

Despite these achievements, the distance between participants has presented challenges. At times, the remote facilitator could not clearly hear or see classroom interactions, making it difficult to follow spontaneous conversations between students and teachers. Some students were shy and hesitant to express their difficulties when asked, which might have been easier to observe during in-person sessions. Furthermore, since classroom communication often occurs in Nheengatu – a language not spoken by the facilitator – some nuances may have been lost or misinterpreted in translation.

Co-design processes have the potential to empower Indigenous youth by providing opportunities for creative and linguistic self-expression. Iterative, culture-centered design approaches have been shown to foster empowering learning experiences (Edwards-Vandenhoek, 2018). In our case, this empowerment became visible when students actively participated in naming the digital assistant and expressing themselves through writing and visual activities. Teachers also reported beginning to use the new technology for broader educational purposes – showing videos in other classes, enrolling in online courses, and joining remote meetings. These practices demonstrate how, when introduced through culturally grounded co-design, technology can extend beyond the initial project to strengthen educational opportunities and digital agency across the community.

The workshops are scheduled to continue through the end of 2025, when the project will undergo a joint evaluation by all partners. The results will guide improvements and planning for 2026. At this stage, FOIRN has already expressed interest in expanding the digital classroom model to ten additional Baré schools, as well as potentially adapting it to schools working with two other Indigenous languages. As this plan advances, we expect to identify methods to streamline the organization, decision-making, and deployment processes – such as simplified approval workflows, collective

training on the use of digital classrooms, and systems for peer support and knowledge exchange.

Ultimately, Yẽgatu Digital represents an emerging model of how technology, co-creation, and Indigenous knowledge can converge to support language revitalization, digital inclusion, and intercultural education.

Acknowledgment

Most and foremost, we acknowledge the vast amounts of knowledge and wisdom we have generously received from the people we worked with in the Indigenous communities, from their leaderships, and from people of NGOs who work with them: the teachers and students of the Tabocal dos Pereiras e Juruti communities, in particular the teachers and local leaders Miguel Alemão and Reinaldo Barreto.

This work has been partially funded by the Center for Artificial Intelligence (C4AI-USP), with support by the FAPESP (São Paulo Research Foundation, grant #201907665-4) and by the IBM Corporation.

References

Edwards-Vandenhoek, S. (2018). 'Over there, in the future': The transformative agency of place-based design education in remote Aboriginal communities. *International Journal of Art & Design Education*, *37*(4), 622–637. https://doi.org/10.1111/jade.12209

Epps, P., & Stenzel, K. (2013). *Upper Rio Negro: Cultural and linguistic interaction in northwestern Amazonia*. Museu Nacional. http://www.museunacional.ufrj.br/ppgas

Instituto Brasileiro de Geografia e Estatística. (2022). *Censo demográfico 2022*. https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html

Kuhn, R., Davis, F., Désilets, A., Joanis, E., Kazantseva, A., Knowles, R., Littell, P., Lothian, D., Pine, A., Running Wolf, C., Santos, E., Stewart, D., Boulianne, G., Gupta, V., Maracle Owennatékha, B., Martin, A., Cox, C., Junker, M.-O., Sammons, O., & Souter, H. (2020). The Indigenous Languages Technology project at NRC Canada: An empowerment-oriented approach to developing language software. *Proceedings of the 28th International Conference on Computational Linguistics*, 5866–5878. https://doi.org/10.18653/v1/2020.coling-main.516

Llanes-Ortiz, G. (2023). *Digital initiatives for indigenous languages*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000387186

Mager, M., Bhatnagar, R., Neubig, G., Vu, N. T., & Kann, K. (2023). Neural machine translation for the indigenous languages of the Americas: An introduction. *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas* (*AmericasNLP*), 109–133. https://doi.org/10.18653/v1/2023.americasnlp-1.13

Neubig, G., Rijhwani, S., Palmer, A., MacKenzie, J., Cruz, H., Li, X., Lee, M., Chaudhary, A., Gessler, L., Abney, S., Hayati, S. A., Anastasopoulos, A., Zamaraeva, O., Prud'hommeaux, E., Child, J., Child, S., Knowles, R., Moeller, S., Micher, J., … Littell, P. (2020). A summary of the first workshop on language technology for language documentation and revitalization. *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages* (SLTU) and *Collaboration and Computing for Under-Resourced Languages* (CCURL), 342–351.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 152–158. https://doi.org/10.1145/191666.191729

Pinhanez, C., Cavalin, P., Vasconcelos, M., & Nogima, J. (2023). Balancing social impact, opportunities, and ethical constraints of using AI in the documentation and vitalization of indigenous languages. *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 6174–6182. https://doi.org/10.24963/ijcai.2023/685

Pinhanez, C., Cavalin, P., Storto, L., Finbow, T., Cobbinah, A., Nogima, J., Vasconcelos, M., Domingues, P., Mizukami, P. de S., Grell, N., Gongora, M., & Gonçalves, I. (2024). Harnessing the power of artificial intelligence to vitalize endangered indigenous languages: Technologies and experiences. https://doi.org/10.48550/arXiv.2407.12620

United Nations Educational, Scientific and Cultural Organization. (2020). *Los Pinos Declaration [Chapoltepek] – Making a decade of action for indigenous languages*. https://unesdoc.unesco.org/ark:/48223/pf0000374030

About the authors

Claudio Pinhanez
claudio@pinhanez.com
IBM Research Brazil
São Paulo, SP

Raquel Cordeiro
raquelcorreacordeiro@gmail.com
University of São Paulo
São Paulo, SP

João Paulo Bento
joaopaulofernandes.bento@gmail.com
University of São Paulo
São Paulo, SP

Julio Nogima
jnogima@br.ibm.com
IBM Research Brazil
São Paulo, SP

Enrico Santarelli
enricosantarelli@hotmail.com
IBM Research Brazil
São Paulo, SP

Submission date/*Artigo recebido em*: 18/11/2025
Approvement date/*Artigo aprovado em*: 8/12/2025