

ARTICLE

When AI is not the solution: *Towards integral co-design processes*[†]

Sofía Trejo, Hamlet Antonio García Zúñiga, Ivan Vladimir Meza Ruiz, and Gaby León Ortiz

Introducción

Artificial Intelligence (AI) is currently promoted as a tool for the development of numerous knowledge areas and as a possible solution to several global issues, from gender inequality to climate change (OCDE 2019; Vinuesa et al. 2020). However, data requirements, computing infrastructures, and access to specialized knowledge have led to these systems being developed, to a large extent, in the private sector or in collaboration with it (Trejo 2022). The concentration of AI development on these actors has led to the production of technology that emulates the ideologies, practices, and issues of the contexts and groups that develop them. In particular, the hegemonic AI systems we have today reproduce colonialism and extractivism (Ricaurte 2019, 2022; Milan and Treré 2019), in addition to operating under a power centralization logic, where its profits concentrate in a few hands (Crawford and Joler 2018). Therefore, it is common for AI systems not to align with the context and needs of the populations on which they operate, which are generally not included in the various stages of their life cycle, from design to operation.

This project is an exploratory effort to develop technologies that move away from the practices mentioned above, which serve as tools for liberation while promoting at all stages of their development: gender perspective, collaboration, co-design, shared benefits, digital autonomy, and data sovereignty. To this end, we sought to create an AI system and a co-design methodology that would help solve a real problem located in a specific time and context in collaboration with the population on which the technology would act. In particular, we sought to co-design an AI-based solution with indigenous language interpreters to address the lack of data and statistics regarding their labor situation in Mexico, taking into account their

[†] The research team wishes to express their gratitude to all the staff of CEPIADET who took part in the design and implementation of the project, as well as to the Human Rights Ombudsman of the People of Oaxaca and their personnel for facilitating their time and facilities. No less important is the gratitude we extend to the <A+> Alliance which sponsored the project; to the f<A+i>r Network, which provided us with constant support during its development and; to Paola Ricaurte and Tatiana Revilla for helping design the workshops. We also thank the group of interpreters who took part in the workshops, whose knowledge and experience were essential to the project: Angélica Morelos Gómez, Donald Rivera Lezema, Ernesto Cortés Venegas, Flor Silvestre de Jesús Ramírez, José Antonio Orozco Gutiérrez, Jovita Cortés Venegas, Juan Ruiz Roque, Virginia Jiménez Antonio, Gaby León Ortiz, and Gabriela García García.

experiences and situated knowledge¹.

The research problem was selected because of the urgent and great need in Mexico to guarantee access to justice for Indigenous Peoples, who, on many occasions, cannot exercise their right (individual and collective) to get assisted in criminal trials and proceedings by interpreters with knowledge of their language and culture (Diario Oficial de la Federación 2003). Due, in part, to the gap between the law and the resources (financial, institutional, and human) ²needed to guarantee the right to interpretation. Hence, indigenous language interpretation is fundamental to access justice in the country, particularly for those who speak any of these languages.

The development of an AI-based solution, in particular a conversational agent (CA), seemed reasonable given the scalability of this technology, that could allow for the creation of independent agents that could be used, controlled, and monitored autonomously by various groups of interpreters to collect data needed to map the status of indigenous language interpretation in Mexico, providing more tools to guide decision-making and allowing interpreters to have agency and influence in public policy related to interpretation and access to justice. However, during the research process, we identified conditions that made an AI-based solution as the one proposed unsuitable, for the time being. From the technical point of view, an AI solution has to satisfy different conditions to be developed, such as data availability, the correspondence between the technological level and the scale of the problem, and the availability of an infrastructure to host the solution. This project identified a new set of conditions related to the user population and its organizational stage, which are fundamental for an AI solution to align with some of the principles that this project seeks to promote, namely, co-design, digital autonomy, and data sovereignty.

This article documents the research process followed to co-design the proposed AI solution and assess the feasibility of its implementation. To do so, we will begin by contextualizing interpretation and access to justice for Indigenous Peoples in Mexico, to then present the design proposal for the conversational agent and the project's co-design methodology.

Overview of the interpretation of indigenous languages in the legal field in Mexico

According to the most recent General Census of Population and Housing of the National Institute of Statistics and Geography (INEGI 2020), there are currently 7,364,645 people over three years of age who speak any of the 364 linguistic variants of the indigenous languages recognized in legal terms in Mexico; which is equivalent to 6.1% of the countrys population in that age range. Of this amount, 51.4% are

1. These items are discussed at length in an extended version of the article (in Spanish), available at https://turing.iimas.unam.mx/fair/docs/paper_extended_es.pdf.

2. For an in-depth discussion, we recommend the synthesis of the National Human Rights Commission recommendation published in *Diario Oficial de la Federación* (Official Journal of the Mexican Federation) on December 7, 2021.

women (3,783,447), while 48.6% are men (3,581,198). The five federal entities with the greatest number of indigenous language speakers are Oaxaca, Chiapas, Yucatán, Guerrero, and Hidalgo; this corresponds, to a great extent, with the southeast of the country, which coincides with several factors of exclusion, such as poverty and poor access to education, among others..

The context of interpretation and access to justice for Indigenous Peoples in Mexico is varied and complex. It involves various agents, such as interpreters, community representatives (sometimes), the state justice system (ministers, public defenders, judges), as well as circumstances of a very different nature, from the political to the social spheres, through academia, for example. Despite this, the discussion on the subject can be considered relatively new.

Linguistic assistance or mediation in trials was used in the colonial period, initially in a practical sense since few indigenous people knew Spanish (Moreno et al. 2002; Nansen Daz and Martínez 2003; Nesvig 2012; Cunill 2018). However, before the approval of the *Ley General de Derechos Lingüísticos de los Pueblos Indígenas* (General Law on Linguistic Rights of Indigenous Peoples, LGDLPI, by its acronym in Spanish), and its publication in the *Diario Oficial de la Federación* (Official Journal of the Mexican Federation) on March 13, 2003, indigenous languages interpretation had not been seriously contemplated as a State responsibility, or as part of an organized infrastructure involving public institutions and the three levels of government (federal, local and municipal).

Legal framework

To understand the history of indigenous language interpretation in the legal field in Mexico, one must consider the changes that the legislation has undergone at different levels. For example, in 2001, the *Constitución Política de los Estados Unidos Mexicanos* (the Political Constitution of the United Mexican States) got amended to include various legal figures and institutions in law enforcement:

In all trials and proceedings in which indigenous persons are a part, either individually or collectively, their uses, customs, and cultural specificities will be accounted for; consequently, they have the right to be assisted by interpreters and defenders with knowledge of their language and culture at all times.

(*Constitución Política de los Estados Unidos Mexicanos*, Article 2,
paragraph A, Section VIII)

In this respect, the Bilingual Public Prosecutors Offices ³ and the Specialized Unit for the Attention of Indigenous Affairs were also created (Diario Oficial de la Federación 2003). Thus, before 2003, the enforcement and administration of justice in Mexico showed a marked and constant inequity when dealing with indigenous language speakers. In particular, the culture and traditions of the defendants were

3. in Michoacán on 2nd October 2002

ignored, and those in charge of translation or interpretation did not have the necessary training. In short, the legal framework existed, but not the conditions to ensure its enforcement.

The LGDLPI emerged as an institutional way of meeting the long-standing demands for inclusion and rights that had been risen, at various times and in different spaces, by diverse Indigenous Peoples living in Mexico. Undoubtedly, at a national level, two events prompted the social reflection that concluded in the discussion, approval, and publication of the LGDLPI: the criticism of the commemoration of the 500th anniversary of the struggle for the defense of American territories against the European invasion, as well as the uprising of the Zapatista Army of National Liberation (January 1, 1994). At the international level, it is clear that certain circumstances made it possible to rationalize the complex process between colonized societies and a legal framework, which comprehended both States and democratic forms of representation and participation; such is the case of the *International Labor Organization Convention 169* (1989), the *Universal Declaration of Linguistic Rights* (1996), the *Durban Declaration* (2001), the *United Nations Declaration on the Rights of Indigenous Peoples* (2007) and the *American Declaration on the Rights of Indigenous Peoples* (2016), to cite just a few examples.

The LGDLPI is broad in nature: it confers a particular status to the indigenous languages spoken in the country, by classifying them as national languages with the same validity as Spanish (articles 7 and 9 of the LGDLPI); it grants linguistic rights to part of the population; and it sets forth the creation and functions of the National Institute of Indigenous Languages (INALI, by its acronym in Spanish), setting it as a decentralized and sectorized body within the Ministry of Public Education. In addition, the LGDLPI establishes responsibilities for government institutions caring for indigenous communities; and it establishes guidelines for the recognition, protection, and promotion of the preservation, development, and use of indigenous languages.

The above-mentioned legal framework includes the training and professionalization of indigenous languages interpreters, as observed in item (d) of Article 14 of the LGDLPI which establishes the creation and oversight of "certification and accreditation programs for bilingual technicians and professionals," and proposes "promoting the training of specialists on the subject." From a governmental perspective, this is seen as a way to strengthen, maintain and preserve linguistic diversity in Mexico, as stipulated in the *Programa de Revitalización, Fortalecimiento y Desarrollo de las Lenguas Indígenas Nacionales* (Program for the Revitalization, Strengthening, and Development of National Indigenous Languages) (Diario Oficial de la Federación 2010). In particular, the Revitalization Program identifies as priority lines of action: justice administration and enforcement with respect for the internal normative systems, based on uses and customs; and guaranteeing the assistance of interpreters, translators and defenders with knowledge of indigenous languages and cultures.

Interpretation and access to justice

The following section gives an idea of the obstacles to guaranteeing access to justice for Indigenous Peoples in Mexico and the role that interpretation plays in this process. Based on the 2006 National Consultation conducted by the National Commission for the Development of Indigenous Peoples (CDI), the trials and legal proceedings carried out by civil, criminal, judicial, agrarian, and commercial courts, at the federal and state levels, did not have interpreters or translators for indigenous languages:

For the period 2006–2007, there were only 14 bilingual defenders who knew the cultures and spoke the Tsotsil, Tseltal, Zapotec of the Isthmus, Tarahumara, Nahuatl, Mayan, Huichol, Yaqui, Chontal of Tabasco, Chinanteco, Purépecha, Zapotec of the valley, and Zoque languages in the Federal Public Defender's Office system.

(Diario Oficial de la Federación 2010, p. 31)

In 2012, the situation did not change much. For example, in Oaxaca (the entity with the greatest numerical and structural linguistic diversity), according to CDI, about 91% of the indigenous inmates in detention centers were not assisted by an interpreter during their proceedings (Sabia 2015), even though, the *Plan Estatal de Desarrollo 2010–2016* (State Development Plan 2010–2016) established that government instances must be trained to “guarantee the effective access of indigenous people to the jurisdiction of the State” (Sabia 2015, p. 54). Something similar is presented in the *Informe del diagnóstico sobre el acceso a la justicia para los indígenas en México* (Report of the diagnosis on access to justice for indigenous people in Mexico) (Alto Comisionado de las Naciones Unidas para los Derechos Humanos 2007), which indicates that only three of 562 indigenous people in detention had anthropological, cultural, or linguistic support during their proceedings, which in many cases constitute the first step to determine the assignment of interpreters.

Certification programs

The establishment of strategies, along with educational and standardization programs are central to the national infrastructure required to ensure access to indigenous language interpretation. In particular, INALI's model for accreditation and certification of translation and interpretation of indigenous languages is the responsibility of the Office of Accreditation, Certification, and Training, which mission is the professionalization of interpreters and translators and to generate conditions of accessibility and equality for indigenous language speakers. This model is based on a diagnosis of justice enforcement and administration in Mexico, conducted by INALI between 2006 and 2008. This study found insufficient coordination between different government levels to establish a body of professional interpreters and translators. In particular, the right to be assisted by an interpreter during trials was largely unknown; there was a considerable absence of interpreters and translators in public prosecutors agencies; it was not possible to identify the language of accused people; and unqualified indigenous language speakers served as interpreters and

translators. In addition, the control of payment procedures for interpretation and translation work was identified as an important issue.

The design and implementation of a coordinated and structured program interpretation within the legal system in Mexico is a recent development. Hence, the assessment of its quality is still inconsistent. In what follows, we provide some partial information on this program. In 2007, 14 bilingual interpreters were identified. In 2014 this number rose to 555, as shown in Table 1. It is worth mentioning that the languages considered in this study were those which for socio-historical reasons are best known or correspond to those with the greatest number of speakers, such as Zapotec, Mayan, and Nahuatl.

Table 1. Overview of legal interpretation in indigenous languages in 2014

Binary Gender	Certifications	Accreditations	TOTAL
Women	26	163	189
Men	66	300	366
TOTAL	92	463	555

Source: Authors’ compilation based on Sabia 2015.

Despite the success of professionalization efforts, it is still common to find flaws in the training and exercise of indigenous language interpretation. This shows an implementation gap of the *Plan Nacional de Desarrollo 2012-2018* (National Development Plan 2012-2018) (Diario Oficial de la Federación 2013) which, among other things, intended to abate impunity. In particular, this document specifies that it is essential to consolidate the education, training, updating, and specialization in topics related to interpretation and access to justice for Indigenous Peoples of those who work for Mexico’s legal system

Interpreter’s experiences

Despite the advances in legislation and the establishment of an apparatus to guide the training and practice of interpreters within the legal system, interpreters still face numerous obstacles and barriers. In this section, we present some of these issues, related to working conditions and the lack of access to formal work.

In general, no guidance is provided concerning the formalization of interpretation contracts. It is worth mentioning that it is common for the first contact with an interpreter to happen by invitation, which is seldom formal. Hence, there is little alignment with labor rights and virtually no access to social benefits (such as pensions). In addition, interpretation services tend to be requested for urgent matters, so interpreters labor conditions are far from ideal. For example, interpreters often work without technical facilities and under difficult circumstances, such as not having time to study the case for which they will interpret or not having access to

interpretation booths in courts.

In addition, society's perception and expectations of the work of an interpreter are disconnected from reality. In fact, due to a lack of understanding about the role of interpreters, defendants, witnesses, etc. constantly assume that interpreters can settle conflicts. Furthermore, is often assumed that interpreters are obliged to sympathize with the defendants and, therefore, must protect them and do everything possible to exonerate them.

In addition, there is no follow-up on interpreters and their work. This means that the experience and certification level of those interpreters who are part of the official register of indigenous language interpreters is unknown. In addition, the balance between language variants with interpreters and those without has been neglected. This has led to the assignment of interpreters in the legal field based on acquaintances and not on official records; in other words, on many occasions, the search for interpreters is done through friendships.

Something of serious concern, which discourages the collaboration of interpreters with the legal system, is that no assistance is provided when payments do not correspond to the work performed. In particular, the justice system has fixed rates, based on time and UMAs (acronym in Spanish for the *Unidad de Medida y Actualización* [Measurement and Updating Unit]⁴). So, for example, if the appointment to interpret during a proceeding is at a specific time, but the proceeding starts late (for example, two hours late) and the total time of the hearing is only 10 minutes, the payment corresponds only to the time spent on the hearing.

Finally, it is important not to lose sight of other seldom-mentioned aspects, namely, training, and professionalization. In particular, specialized programs in line with the National Program of Indigenous Languages Interpreters and Translators (acronym in Spanish PANITLI) at the undergraduate or graduate level are practically nonexistent. In addition, neither government or academic institutions promote the periodic update of interpretation programs. This situation has changed somewhat with the recent creation of a graduate program in interpretation at the Benito Juárez of Oaxaca Autonomous University. Besides this, it is necessary to strengthen and standardize certifications and official accreditations related to interpretation. This could help with the numerous and frequent criticisms towards interpreters related, for example, to their total or partial ignorance of specialized vocabulary, a product of the absence of professionalization or training. This lack of professionalization leads to interpreters not being rigorous enough and not taking advantage of the resources recommended for interpretation, such as note-taking and paraphrasing.

4. Economic reference in Mexican pesos to determine the amount of the payment of obligations and assumptions provided for in federal and state laws, as well as in legal provisions emanating therefrom.

Project proposal

As seen from the previous section, the gap between the regulatory framework and the legal practice faced by interpreters calls for strategies to strengthen their work and improve their labor conditions. In particular, it is necessary to establish clear routes on which practitioners who face the day-to-day problems of being interpreters can influence public policy and improve their situation and that of their guild. Although there are social options to think about these issues, this project proposed a solution with a technological component, based on conversational agents.

A conversational agent is an AI system that tries to solve users' needs through dialogue. This project is based on the idea that interpreters could individually communicate the issues they face in their daily work to a conversational agent. The agent could agglomerate this data (without disclosing any private information) and make this evidence available to groups of interpreters to improve decision-making and influence public policy. Tangentially, interpreters could use such agent to build collective knowledge, for example, by producing a glossary of terms relevant to interpretation in specialized areas or by sharing work tips, such as the location of courts, travel routes, etc.

Currently, conversational agents are common in modalities, such as chatbots or conversational assistants associated with devices that take orders to satisfy information needs through suggestions, provide recommendations or perform specific actions, such as playing audio. However, in most cases, there is a power asymmetry between the operator of the technology and the users. In general, conversational agents are operated by organizations for their benefit, for example, to reduce customer service costs. To our knowledge, only in a few instances, conversational agents have been used to benefit the population with whom the systems interact or upon which they act. In what follows we present some examples of such systems..

The DoNotPay chatbot was created in 2016 to dispute traffic fines (Murimi 2021). In this case, it is the chatbot collected information filled in by the user related to the reconsideration of a fine. More recently, UNESCO implemented the U-Report program ⁵ to automate social surveys through a bot that works via the mass messaging system (SMS). In this case, the system collects and aggregates data from large populations on topics of interest to UNESCO. In addition to these examples, there are reports of the use of a chatbot for consultations within a union organization (Flanagan and Walker 2021), in which incorporating information related to the organization into the design of the system helped enhance its acceptance and impact. These are some examples where power asymmetry between those who design technology and those who use it is rebalanced by designing systems to benefit the users. Our proposal is similar in many ways to the U-Report system and the union chatbot since it seeks to collect information from a group of people in a particular employment situation. However, one significant difference between these proposals and ours is that our project aimed to transfer the technology to its users

5. Program website: <https://ureport.in/about/>

and promote digital autonomy and data sovereignty.

The conversational agent proposed in this project sought to provide support to interpreters at two levels:

1. The target users of the conversational system would be interpreters of indigenous languages in Mexico who would communicate with the system in Spanish. Users would be able to access the system with the following objectives:
 - a. Access information related to their assigned activities (e.g., cases) in the style of conversational assistants.
 - b. Communicate information about a problem faced in their daily work (e.g. discrimination).
 - c. Collaboratively build collective knowledge related to the practice of interpretation
2. Secondary users of the system would be organizations or groups of interpreters who could extract cumulative global information from the conversational system. Of particular interest would be the data produced by the system, which could be used in an aggregated manner by secondary users to bring to light labor problems and serve as a basis for designing strategies and programs and improving public policies on matters related to interpretation.

In this respect, the two identified roles imposed a challenge to designing a solution that could benefit both the interpreter and their groups or organizations.

The development of the system was conceived in three stages:

1. **Stage 1:** Conceptual design of the system and co-design methodology and project feasibility assessment.
2. **Stage 2:** Development of a prototype of the conversational agent and evaluation of the co-design methodology.
3. **Stage 3:** Deployment of the CA.

This text corresponds to Stage 1, where the relationships between the different actors (particularly between the interpreters and the interpreters organizations or groups) were studied to assess the feasibility of continuing with the following stages of the project. An essential part of the research project was conducting workshops with interpreters to incorporate their experiences, opinions, needs, and expectations into design and methodology ⁶.

The project's goals were to promote digital autonomy and data sovereignty by transferring the technology to groups or organizations of interpreters so that they could operate and monitor the CA and have the necessary tools to make use of the

6. The results of these workshops are presented in the *Workshops* section.

data gathered by the system. An important part of the following stages of the project would center on providing the resources needed (human, knowledge, and infrastructure) to transfer the technology, including CA management training, data management training, and infrastructure development. It would be important for these tasks to align with the *CARE Principles* for indigenous data management (see Principles and Ethical Considerations subsection) and to be based on clear collaboration agreements with the various groups involved.

Methodology

Complementary to the exploration of the development of the CA, the team sought to generate a work methodology to balance the power relations between the different groups involved in the project and to align the work with the principles of gender perspective, collaboration, co-design, shared benefits, digital autonomy, and data sovereignty. Central to the methodology was promoting the interpreters' agency, control, and access to the research process and the products derived from it. This part of the project was based on and aligned with proposals for the ethical development of research projects created by indigenous groups and researchers (Assembly of First Nations 2009, s.f. Carroll et al. 2022; Carroll et al. 2021), in addition to taking into consideration feminist practices of listening and situated knowledge.

Part of the methodology consisted of transversally incorporating gender perspective into all stages of the project's development. To this end, both the research team and the workshops sought to achieve gender equality. In addition, a workshop on *Gender* was designed to compile the expectations, barriers, and experiences of the interpreters depending on whether they self-identified as men or women⁷. This workshop aimed to serve as the basis for incorporating gender perspective into the design of the system, allowing for the collection of data on issues such as gender equality and gender-based violence.

The first tool produced by the research team as part of the co-design methodology was a *Research Protocol and Collaboration Agreements Format*⁸. The format was used to guide the elaboration of a project-specific version of the document. The format in reference begins by establishing the general considerations for filling out the document, namely: using non-exclusive language, accessibility to the group with which one will collaborate; considering the expressions used by the culture and social organization of the groups involved; and ensuring the document is self-contained (the inclusion of a glossary is suggested). In addition to sections on general aspects of the project, the document emphasizes the inclusion of the group with which the project will collaborate (in the case of this project: indigenous language interpreters) integrally and explicitly in all stages of the design and development of the project, including the definition of goals and objectives, the identification of expected benefits, and the interpretation of the results. Furthermore, the document

7. The limitations of the binary classification of gender are acknowledged, but this was the selected segmentation, given the demographics of the people attending the workshop.

8. This format can be accessed at: <https://turing.iimas.unam.mx/fair>

suggests providing the following information: collaboration agreements, project beneficiaries, conflicts of interest, ownership, and authorship of the products derived from the research, as well as the licenses that will be used for the publication and dissemination of data, articles, reports, etc.

The following subsection presents relevant aspects of the *Preliminary Version of the Research Protocol and Collaboration Agreements* of this project, highlighting the difficulties and challenges encountered in the process of developing it.

Research protocol and collaboration agreements

The *Research Protocol and Collaboration Agreements Format* served as the basis for the elaboration of a *Preliminary Version of the Research Protocol and Collaboration Agreements (Preliminary Version)* which established the collaboration and co-design guidelines for the project. This document would be discussed and agreed upon with the Indigenous Professional Center for Counseling, Defense, and Translation⁹ (acronym in Spanish CEPIADET), the indigenous organization with which this project collaborated that, among other topics, specializes in indigenous language interpretation. In addition to the dialogue with CEPIADET, the *Preliminary Version* would be discussed with the group of interpreters who participated in the workshops, all of whom work with CEPIADET. It is important to emphasize that not all of the elements contained in the document would be agreed upon collaboratively, given that, for example, the research products and the licenses applicable to them had been established previously by the <A+> Alliance, which sponsored the project.

The *Preliminary Version* was sent to the interpreters along with the invitation to attend the workshops; the invitation contained detailed information regarding the project and the proposal for its development. Originally one of the workshops would focus on discussing the *Preliminary Version*. As mentioned, a similar discussion would be held with CEPIADET. The final version of the document would be based on these two dialogues. However, during the development of the project, it became evident that an *Indigenous Data Sovereignty* workshop would be more suitable. So a decision was made to replace the workshop centered on the *Preliminary Version* with the latter. This change was motivated by the idea that to reach agreements with the various actors involved, it would be important to balance the knowledge regarding topics such as the collection, management, access, control, and use of data. After the necessary foundations had been laid, it would be possible for all parties to be in a better position to negotiate the collaboration agreements and guidelines for the development of the following stages of the project. In addition, given the nature of the project (focused on developing an AI system), the workshops provided a space to reflect on data sovereignty and its relationship to the rights of Indigenous Peoples, particularly autonomy and self-determination. The idea for these conversations was to serve as a starting point to design strategies to transfer the technology to groups of interpreters and promote digital self-determination in subsequent stages of the

9. CEPIADET's website: <https://cepiadet.org>

project.

CARE Principles and ethical considerations

The project subscribes to and is based on the *CARE Principles* for Indigenous Data Sovereignty (Carroll et al. 2020) developed by the Global Indigenous Data Alliance (GIDA). These principles aim to serve as a tool for Indigenous Peoples to have greater power, agency, and control over the application and use of their data, as well as over their knowledge, in a way that collectively benefits them and aligns with their visions, needs, interests, and cultures. For the principles to have a real impact on the project and all its processes they were linked to the project’s ethical considerations. In other words, the ethical considerations establish strategies and mechanisms to implement the *CARE Principles* and ensure their compliance.

In the following subsections, we will present the central ideas behind each of the *CARE Principles*: Collective benefit, Authority to Control, Responsibility, and Ethics. **These principles are formulated within the project framework and are an adaptation of the works** (Research Data Alliance International Indigenous Data Sovereignty Interest Group 2019a, 2019b). After briefly introducing each of the principles, we will present some of the strategies followed to implement them in the project.

Authority to control

Indigenous communities’ rights and interests over their data and knowledge are recognized and their authority to control this data will be strengthened.

The following ethical considerations were established to implement this principle within the project. First, the project was conceived and designed as a tool to promote indigenous data sovereignty. This sovereignty consists of supporting indigenous communities to have greater agency, control, ownership, and access to their data. In particular, the *Indigenous Data Sovereignty* workshop, held as part of the project, served as a space to discuss and reflect on these matters (the workshop results are presented in the *Workshops* section). In addition, the research protocol established that the data and products derived from the present research should be accessible to everyone who collaborated on the project. To ensure accessibility an extended version of this research article in Spanish was produced¹⁰, with the intention that all the people who collaborated on the project could read the work and give feedback on it. In addition, the project established that the results derived from this research should be available to all of our collaborators, to receive suggestions and comments, which would be (and have been) incorporated into the work with the aim that the results produced include, as much as possible, the voices and views of all the people involved with the project.

10. To access this document visit https://turing.iimas.unam.mx/fair/docs/paper_extended_es.pdf

In addition, the *Preliminary Version* established clear data collection and data management procedures. Furthermore, the annexes of the document included all the forms and letters of consent that people invited to attend the workshops would have to sign to ensure their voluntary participation and the protection of their rights, particularly their data rights¹¹. Finally, the *Preliminary Version* was attached to the invitations to attend the workshops so invitees could have detailed information on all aspects of the project and ensure that their participation followed the principle of informed consent.

Responsibility

The project accepts the responsibility to communicate how data will be used to support the self-determination and collective benefit of Indigenous Peoples. It also recognizes that accountability requires significant and available evidence so that the benefits and efforts accrue to Indigenous Peoples.

One fundamental consideration for implementing this principle was the elaboration of the *Preliminary Version*, which establishes, among other things, the responsibilities, roles, and expected benefits for all the groups and individuals involved in the project. In addition, the document contains information on the provisions taken to ensure the project was developed observing the principles of transparency and accountability, including contact channels, and details on the funds available and how they were budgeted.

To ensure responsibility concerning indigenous languages and views, the project designed the workshops as a space to collect the experiences, opinions, needs, and aspirations of the interpreters to incorporate them in the design of the CA and into the research. In addition, to make the voice of the interpreters visible, their experiences were compiled in the workshops, and are an essential part of this work (subsections entitled *Perspectives*¹²). Furthermore, a media piece in collaboration with the interpreters was produced. The piece consists of a web page¹³ in which, in addition to information about the project, there are nine videos in different indigenous languages (subtitled in Spanish and English) in which interpreters share, in their languages, their experiences and views working as interpreters. It is worth noting that part of the project's budget was allocated to pay for the participation of each interpreter in the workshops (in addition to providing food and transportation funds) and for the creation of the media piece.

Ethics

All processes will contemplate power and resource asymmetries and how these affect Indigenous Rights and Human Rights. The process will include the representation

11. Before starting the workshops, an explanation was given on the need to collect the letters of consent. The handling and uses that would be given to the data collected were also clarified, and the procedures for exercising the rights over personal data were explained.

12. An extended version of these sections can be found in the Extended Version of this article

13. <http://turing.iimas.unam.mx/fair>

of relevant indigenous communities.

Our project is anchored on the principles of collaboration and co-design with indigenous language interpreters. Indigenous language interpreters collaborated (at different moments) with the conceptualization, design, and development of this work. Furthermore, an indigenous representative, Gaby Ortiz León, is part of the research team and an integral part of the development of the project.

Finally, one of the central concerns of the project was to balance power relations, particularly between the interpreters and the research team. To tackle existing asymmetries the project developed workshops on *Conversational Agents* and *Indigenous Data Sovereignty*. The main objectives of these workshops were: 1) to provide participants with the essential knowledge to understand and discuss the project; 2) to serve as a space for reflection and dialogue that would allow to incorporate the interpreters' views, experiences, and needs into the design of the CA and the methodology.

Collective benefits

The data ecosystem will be designed and operated to benefit First Nations and Indigenous Peoples.

The project was designed to promote digital autonomy and indigenous data sovereignty by providing interpreters with the knowledge and tools needed to use AI technology to collectively improve their labor conditions. Moreover, the project aimed to provide groups of interpreters with the expertise, technical support, and infrastructures required for them to autonomously (with the support of the research team) operate, monitor, and maintain the AI systems. To do so, all processes would have to be based on collaboration agreements between the research team and these actors.

Workshops

The project's intention and justification, as well as the context of the problem to be addressed, led to a three-workshop proposal: one focused on conversational agents, another linked to the sovereignty of indigenous data, and an additional one related to gender. The workshops were held in the José Mujica Cordano Auditorium of the Human Rights Ombudsman's Office of the People of Oaxaca, in the city of Oaxaca of Juárez, in October 2022. In the following subsections, we present the main results of each workshop¹⁴.

14. The extended version of the article (in Spanish) addresses the insights and discussions at the workshops in greater detail. This document can be accessed at: https://turing.iimas.unam.mx/fair/docs/paper_extended_es.pdf.

Conversational agents

The objective of this workshop was to provide interpreters with a contextualized understanding of conversational agents and to explore possibilities to align the design of the system with the context and the situated experiences of the interpreters. In particular, the design proposal for the CA contemplated the possibility for users to collaboratively generate knowledge and tools with the system's aid. For example, the CA might ask users, while interacting with them, to contribute terms to a glossary. Subsequently, the CA could ensure that interpreters evaluated and commented on the proposed terms or definitions (again, as part of user interactions). Finally, the information collected and corroborated by the users could be analyzed by the system administrators to be added to the glossary. Thus, the CA could act as an intermediary for interpreters to build resources collaboratively.

For a system like the one proposed to be viable, it would need to meet two conditions:

1. The users should actively collaborate in the production of the resources and tools through their recurrent interactions with the system.
2. The group of interpreters should possess a structure that would allow transferring the technology to them, so they could operate, manage and monitor the system and make use of the data generated.

Therefore, the proposed system design heavily depended on the existence of an organizational structure within the group of interpreters to operate and maintain the technology. Part of the objectives of this workshop was to evaluate these conditions.

During the workshop, participants reflected on artificial and conversational agents. The following insights stand out from these conversations. Interpreters have a close relationship with technology, and their status as digital citizens gives them a perspective on technological possibilities. In particular, interpreters are familiar with the use of smartphones and know conversational agents such as *Alexa* or *Siri*. In addition, participants emphasized that smartphones are an essential piece of equipment for their work as interpreters, not only as a means of communication but as a replacement for computer equipment.

Another aspect discussed during the workshop was the nature of data, and how for many systems the users provide the data, while those who implement the technology are the owners and beneficiaries of the information. This matter was discussed further in the workshop on *Indigenous Data Sovereignty*. In addition, the lack of co-responsibility in the operation of intelligent systems was discussed, along with the problems AI-based systems could generate, such as black box systems; reproduction of biases of various kinds; the disparity between who implements the systems and the end-users; the lack of mechanisms to challenge AI decisions; and the lack of clarity regarding authorship and ownership of products generated by AI.

With the background on how intelligent systems work and their limitations, the dynamic turned to design conversations to link the behavior of the agent with users expectations and needs. To do so, the workshop facilitators specified the projects objective and asked participants to script a conversation to request an interpretation job. The participants organized themselves into groups and developed scripts with different conceptualization levels. An issue that emerged during the exercise was the payment for their work as interpreters.

Perspectives¹⁵

As a result of the workshop, the interpreters identified the following functionalities they would like the agent to have:

- Ask for work schedule information. For example, where is my next interpretation job?
- Share work experiences with the agent. For example, reporting discrimination.
- Accessing formats to create contracts.
- Collect payment information. For example, tracking payments.
- Consult practical information related to their work as interpreters. For example, the location of penitentiaries and detention centers.
- Consult legal information, in particular regarding updates on laws and regulations.

During the workshop, the research team learned that the interpreters expectations do not include taking an active part in the development of the technology or in the production of collective knowledge. Furthermore, it was not possible to identify a structure within the group of interpreters or their organizations that could assume, for the time being, the roles required to prototype the system. It is important to highlight that, without this organizational structure or the intention of the interpreters to actively engage with the system it would not be possible to generate resources collaboratively or transfer the technology to the group of interpreters, at least not within the timeframe set to develop the project.

Indigenous data sovereignty

The workshop had two main objectives: 1) to serve as a space for reflection on indigenous data and its relationship to the rights, needs, and aspirations of Indigenous Peoples, and 2) to explore mechanisms and strategies to align the project with indigenous data sovereignty. The workshop development followed an *Indigenous Data Sovereignty Manual* that was provided to each participant at the beginning

15. This section reports the opinions and views of the people who attended the workshop. The information is presented in a summarized manner to facilitate readability.

of the activities. In what follows we present the main topics addressed within the workshop and the insights generated during its execution.

Before starting, it is important to emphasize that we understand indigenous data to be information or knowledge in digital format or that it can be digitized (for example, books, textiles, stories) that refer to or come from indigenous people or peoples. Therefore, indigenous data includes information on indigenous individuals, groups, ways of life, culture, language, territories, and resources.

Indigenous data

The workshop began by presenting key elements related to indigenous data. All this information was included in the workshop’s manual. In this subsection we present the main ideas discussed and insights from these discussions.

Central to any study regarding indigenous and non-indigenous data are power structures and how they permeate the various processes associated with data: its conceptualization and design, collection, interpretation, validation, ownership, access, and uses. It is important to emphasize that all these processes strongly depend on the context (place, time, people involved, etc.). Thus, one of the most relevant phenomena to consider in the case of Indigenous Peoples is colonization.

Colonization resulted in the devaluation of existing indigenous data systems and their replacement by external systems imposed by colonizers, particularly in the government and the public spheres (Pool 2016). This power asymmetry continues to this day, as Indigenous Peoples live in countries with colonial pasts, in which data systems are those inherited from colonizers (Pool 2016). In this respect, Indigenous Peoples continue to experience colonization in the digital spaces of the countries where they live.

It is important to emphasize that these power asymmetries are reproduced and often amplified with data cycles since the ideas and views of the dominant groups influence data design and determine, for example, what information is collected, what uses will be given to this information and who has access to it. This data, produced by dominant views, is used to measure progress, development, and gaps and to guide public policy. Finally, these measurements (such as indexes or indicators of poverty, health, education, etc.) are used to inform the design and collection of new data. These data cycles have been particularly detrimental to Indigenous Peoples. Given that in various spheres, such as in public discourse, these populations tend to be represented in terms of disadvantage, dysfunction, difference, disparity, and deprivation (Walter 2016).

To contextualize the previous discussion at the workshop and in Mexico, a press release published by the National Institute of Statistics and Geography on August 8, 2022, entitled *Estadísticas a Propósito del Día Internacional de los Pueblos Indígenas* [Statistics on the International Day of the World’s Indigenous Peoples] (Instituto Nacional

de Estadística y Geografía 2022) was studied after a presentation on indigenous data. The discussion was opened with workshop attendees, using this example as a basis, to share their experiences and opinions regarding the views, participation, collection, uses, access, and possible prejudices reproduced in indigenous data. We present the results of these conversations in the next subsection.

Perspectives¹⁶

According to the experience of workshop participants, on numerous occasions, indigenous data, particularly the one produced by the government, is not useful for these populations. This is reflected in various problematic aspects, such as access to justice, health, and education, three issues that were addressed in detail by the people who participated in the workshop. In particular, it was reported that existing data related to interpretation and access to justice is not useful for indigenous organizations or defenders. For example, there are official figures on the percentage of incarcerated indigenous people who were assisted defenders or interpreters with knowledge of their indigenous language. However, no more details are known about this population, such as where they are held, to which communities they belong, or the languages they speak. This example shows how indigenous data tends to reflect the interests and meet the needs of those who create it, not those of the indigenous communities. In this respect, it is the general opinion of workshop participants that should indigenous data be thought of from the perspective of the communities, it would reflect a different reality.

Data Sovereignty

After discussing indigenous data, the workshop focused on reflecting on the relationship between data and Indigenous Peoples' rights, emphasizing that Indigenous Peoples' lack of agency and control over their data limits their ability to make informed decisions and to develop policies and strategies that allow them, for example, to address local issues in alignment with their traditions, cultures, languages, and forms of organization. Therefore, for Indigenous Peoples to fully exercise their rights to autonomy and self-determination, power over indigenous data must reside within their communities (Kukutai and Taylor 2016; Snipp 2016).

In addition to the presentation on indigenous data sovereignty, the workshop explored how data referring to Indigenous Peoples produced under non-indigenous views tend to normalize a racialized reality created and recreated using these same tools. As a result, indigenous data has historically tended to be designed from the logic of non-indigenous populations. Thus, it does not align with indigenous ways of life, culture, interests, and needs. Nor does it address local issues or is used to benefit the communities from which it is collected. Furthermore, the populations whom the data refer to or from whom data were collected have no access to or control over such data. To exemplify these problems, the workshop addressed the

16. This section reports the opinions and views of the people who attended the workshop. The information is presented in a summarized manner to facilitate readability.

historical development of population censuses in Mexico, where these tensions are clearly visible.

The presentations on data sovereignty, population censuses in Mexico, and the national regulatory framework on autonomy and self-determination of Indigenous Peoples (Article 2 of the *Constitución Política de los Estados Unidos Mexicanos* [Political Constitution of the United Mexican States] and the *United Nations Declaration on the Rights of Indigenous Peoples*, particularly Articles 3 and 4) served as the basis to discuss indigenous data and data sovereignty and their relationship with the indigenous rights of autonomy and self-determination. Finally, a group exercise was conducted to reflect on possible strategies for this project to align with indigenous data sovereignty and the role interpreters could play in the process. Below we present the experiences and opinions of workshop participants related to these discussions.

Perspectives¹⁷

In relation to indigenous data, workshop participants highlighted the importance of reflecting on the criteria used, particularly by the government, to characterize indigenous people. In these conversations, indigenous languages played a fundamental role, but it was highlighted that other elements to consider could be: being born and raised in an indigenous community, knowing indigenous culture and traditions, or local history.

Concerning indigenous data sovereignty, the discussion focused on data created and collected by the government. In particular, several of the experiences shared by participants acknowledged the importance of language as a key element when thinking about indigenous data, as language is a fundamental factor in ensuring full access to education, health, and justice for indigenous people, particularly for those whose first language is not Spanish or who only speak indigenous languages. In this regard, the inclusion of language as a data variable at the national level is of utmost importance to guarantee the rights of all groups in the country as, in many cases, it is the language barrier that limits the full exercise of rights.

Finally, the following problems are reported concerning the group activity aimed at reflecting on how to align this project with indigenous data sovereignty and on the role that interpreters could play in the process: there were many doubts related to what a CA is and what its design might look like; and there was a lack of clarity concerning the role interpreters could play in the project, and the level of involvement required by the various groups (interpreters and organizations) in the construction, operation, and maintenance of the system. Thus, the time of this workshop activity was used to clarify doubts about CA and explore how the system could support the interpreters in their professional work. The results of these conversations were presented in the CA section to facilitate reading.

17. This section reports the opinions and views of the people who attended the workshop. The information is presented in a summarized manner to make it easier to read.

Gender

During the workshops, an important place was given to understanding the situated experiences of interpreters in various aspects related to their work and gender to identify elements or topics that should be incorporated into the design of the CA, such as gender discrimination or violence. This workshop was carried out in two groups, men and women, facilitated by the members of the respective gender research team. The dynamic followed a questionnaire that included the topics of interpreters' certification and training, work experience, and future aspirations.

*Women*¹⁸

Regarding the reasons for working as interpreters of indigenous languages, the first thing that is relevant to mention is that none of the participants lives exclusively from interpretation, but rather work as interpreters complementary to other jobs. Therefore, they do not see interpretation as their occupation. Instead, their primary motivations for carrying out this work are to support their countrymen, particularly people who do not speak Spanish, so that they can have greater agency and access to justice, health, etc., and to revitalize their languages and preserve their cultures.

The main barrier they have experienced for training as interpreters is economic in nature. This is mainly related to the irregularities, delays, or lack of, payment for their work as interpreters. Which makes it difficult for people to consider interpretation as their first source of income. In the workplace, the main problem experienced by the workshop participants while working as interpreters has been discrimination. However, it is important to note that the vast majority of participants consider that the discrimination they have suffered relates to them being indigenous and not their gender. Other difficulties experienced by the participants when performing their work are that their families do not want them to act as interpreters because they consider it dangerous and are concerned about their safety, insecurity problems, femicides, and the schedules in which they must work or travel. In addition to this, in many cases, performing interpreting tasks requires moving to other locations, which presents difficulties to care of children and family.

Finally, regarding labor prospects, several of the participants share an interest in being able to perform interpretation work for migrants and being able to work as interpreters in the United States. These aspirations are strongly linked to the difficulties that migrants, particularly those who do not speak Spanish or English, experience living abroad, where not being able to communicate and not knowing the institutions and laws puts them in a particularly vulnerable situation. Traveling, not only abroad but to other communities, is an important reason why several of the participants enjoy being interpreters and why they like to be able to do it. In addition, several of the participants are interested in being able to make a living from interpretation and enter the profession fully.

18. This section reports the opinions and views of the people who attended the workshop. The information is presented in a summarized manner to facilitate readability.

*Men*¹⁹

First, the facilitators explained the context of the exercise to the participants; the questionnaire was answered individually, but the questions were discussed as a group. The general feeling that was expressed by the participants was that their gender did not influence their work; however, there was an introspection that revolved around the expectations that their community has of them as men, which influenced their motives for being interpreters. They also added that one of the reasons they chose to serve as interpreters relates to the desire to help their communities of origin and the circumstances of conflict they face. An important highlight is that there is pride in being indigenous and that the participants are concerned about the lack of knowledge within indigenous communities about their rights. The barriers identified by the participants related to their work as interpreters are concentrated on: economic issues, infrastructure (there is no equipment to perform their work better), discrimination by authorities, lack of understanding of what they do as interpreters, and training itself (they do not find school offers to specialize). This last issue is important as they often do not know the topics and terms present in the law.

In addition, the participants stated that they would like to continue having the opportunity to support their community, but, above all, they would like to give talks about the justice system so that people can learn about their rights. One participant noted that the best way to help their communities in the future is to prepare and be certified as an interpreter.

The barriers they face daily to carry out their work as interpreters are: the distance they must travel from their homes to reach the courts; the time and coordination needed to perform their work as interpreters; travel expenses; problems with the financial compensation for their work; and lack of training. Discrimination is something they face in their day to day, mainly because they are indigenous, particularly if they do not pronounce Spanish well, and because, in their opinion as Indigenous People, "they must work more" and, as a man, "you cannot aspire to a salary raise or job promotion," as in their view these types of opportunities are given to women because they are women.

Discussion

The project intended to address the need to map the current state of indigenous language interpretation as a profession in Mexico and to provide tools to identify and close existing gaps between the regulatory framework and its implementation. To this end, the proposal was to create an AI system that, by interacting with interpreters, would collect data that could: serve to evaluate existing public policies on interpretation matters; help identify specific linguistic needs, such as the lack of certified interpreters for certain languages and variants; and help understand the situation of interpreters in terms of their training and the labor problems they experience. Thus,

19. This section reports the opinions and views of the people who attended the workshop. The information is presented in a summarized manner to facilitate readability.

the project explored the feasibility of designing and implementing a conversational agent that could serve as a tool so that interpreters could collaboratively contribute to the generation of data that would allow them to:

1. Bring to light the problems they face to carry on their work as interpreters.
2. Have greater agency and power in decision-making and public policy on matters related to interpretation, particularly in the justice arena.
3. Collaboratively build collective knowledge to support them in their interpretation work.

In addition to these objectives, the project created a methodology to align all stages of development with the principles promoted by this project, which are gender perspective, co-design, shared benefits, digital autonomy, and data sovereignty. Important elements of this methodology included the development of workshops with indigenous language interpreters, the elaboration of a *Research Protocol and Community Agreements*, and the establishment of strategies to implement the *CARE Principles* for the management of indigenous data.

Furthermore, to develop AI systems that align with the principles promoted by the project, it is essential for the population with which the technology will be co-designed and will act upon to take an active role at all stages of the systems life cycle, from design to deployment. In this respect, the goals and needs of the system must agree with those of the various actors involved in its development (i.e., administrator(s), developer(s), and users). In addition, to guarantee the transfer of the technology to its users or their representatives, there must exist an organizational structure within this group that can undertake the various tasks required for operating and monitoring the technology. Therefore, a key part of the project was to carry out a field assessment of these conditions.

With respect to the above, the projects research process allowed us to identify the following characteristics of the dynamics within the collaborating group of actors. First, organizations or groups of interpreters, such as CEPIADET, have an interest in mapping the landscape of interpretation in Mexico; and the data produced by the CA could be useful to them for various purposes, such as program planning or strategy design. On the other hand, statistical data is not perceived as a necessity for interpreters. This is mainly because interpretation is not currently the primary work activity of indigenous language interpreters. Instead, they perform the work as a support to their communities and to contribute to the preservation of their languages and cultures. In this respect, the interpreters main interest when interacting with the agent would not be the generation of statistical data to shed light on labor problems in their field of work or to influence public policy. They would rather have information and tools to support them in their day-to-day work as interpreters, such as glossaries of legal terms, the location of prisons, and information on transportation to those locations.

The proposed system design could accommodate the needs of both interpreters and their groups and organizations. However, it was not possible to identify an organizational structure within the group of actors with whom we collaborated that would enable us to transfer the technology, thereby limiting the possibility of aligning the project with the principles of digital autonomy and data sovereignty. In addition, interpreters expect not to take an active role in the development of the technology or in the production of collective knowledge, which represents a barrier to the principle of co-design. Therefore, the research team concluded that it is not feasible, for the time being, to develop the system under the established principles. However, the following tools were identified to support the professional work of interpreters, which are more suitable for the groups of interpreters we collaborated with, which aim at the creation of a community of interpreters and meet the needs reported in the workshops:

- The creation of spaces for direct communication and collaboration among interpreters, without the intervention of external agents, whether artificial or human. Technical options for creating these spaces include internet forums or team communication platforms.
- The use of tools for the creation of shared resources. This could be achieved through technologies such as Wikis. In particular, the project identified the creation of glossaries and linguistic mappings as areas of opportunity.

This project reached the following conclusions from the research process: the potential impact of using AI to address specific problems is real, but if the design elements are not aligned with the expectations and possible uses of the technology by the users, the impact of the technology will be null or even counterproductive for the target population. In our case, the development of the conversational agent required a community of interpreters, rather than groups that have not yet achieved cohesion. In addition, the proposal strongly required users to take an active role in co-designing the technology and constructing collective knowledge. However, these requirements were not in line with the aspirations of the interpreters who participated in the workshops. Therefore, the evaluation of the project revealed that, for the time being, deploying a system such as the one proposed would not guarantee that the objectives and principles proposed²⁰ would be met and that the system should not be developed.

We hope this project will serve as an example to reflect on the problems between the conceptual design of AI projects and the feasibility of implementing these technologies in specific contexts, respecting pre-established principles in all processes. Furthermore, we believe that papers such as this one, where it is decided not to continue with the development of AI solutions, are fundamental in discussions about technology, as they provide essential knowledge and tools to close the gap between theory and practice; this topic is not explored much in works on the subject, which tend to assume compliance with ethical principles and/or present considerable biases by reporting only those solutions that get to be developed. Finally, we hope that

20. Particularly, co-design, digital autonomy and data sovereignty.

the ideas and methodology of the project will contribute to the creation of AI systems focused on users, aimed at rebalancing the relationships of agency and power between those who develop and use technology, by promoting the gender perspective, co-design, shared benefits, digital autonomy, and data sovereignty, at all stages of development.

License

This work is licensed under a Creative Commons “Attribution 4.0 International” license.



References

- Alto Comisionado de las Naciones Unidas para los Derechos Humanos. 2007. *Informe del diagnóstico sobre el acceso a la justicia para los indígenas en México. estudio de caso en Oaxaca. proyecto de implementación de las recomendaciones derivadas del diagnóstico sobre la situación de los derechos humanos en México.*
- Assembly of First Nations. 2009. *Ethics in first nations research.* https://www.afn.ca/uploads/files/rp-research_ethics_final.pdf.
- . s.f. *First nations ethics guide on research and aboriginal traditional knowledge.* https://www.afn.ca/uploads/files/fn_ethics_guide_on_research_and_atk.pdf.
- Carroll, Stephanie Russo, Ibrahim Garba, Oscar L Figueroa-Rodriguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, et al. 2020. The care principles for indigenous data governance.
- Carroll, Stephanie Russo, Ibrahim Garba, Rebecca Plevel, Desi Small-Rodriguez, Vanessa Y Hiratsuka, Maui Hudson, and NanibaaA Garrison. 2022. Using indigenous standards to implement the care principles: setting expectations through tribal research codes. *Frontiers in Genetics*, 489.
- Carroll, Stephanie Russo, Edit Herczog, Maui Hudson, Keith Russell, and Shelley Stall. 2021. Operationalizing the care and fair principles for indigenous data futures. *Scientific Data* 8 (1): 108.
- Crawford, Kate, and Vladan Joler. 2018. Anatomy of an ai system. Retrieved September 18.
- Cunill, Caroline. 2018. Un mosaico de lenguas: los intérpretes de la audiencia de México en el siglo xvi. *Historia mexicana* 68 (1): 7–48.
- Diario Oficial de la Federación. 2003. *Unidad especializada para la atención de asuntos indígenas*, July 24, 2003.
- . 2010. *Programa de revitalización, fortalecimiento y desarrollo de las lenguas indígenas nacionales*, July 2, 2010.
- . 2013. *Plan nacional de desarrollo*, May 20, 2013.
- Flanagan, Frances, and Michael Walker. 2021. How can unions use artificial intelligence to build power? the use of ai chatbots for labour organising in the us and australia. *New Technology, Work and Employment* 36 (2): 159–176.
- INEGI. 2020. *Censo de población y vivienda 2020.* <https://www.inegi.org.mx/programas/ccpv/2020/>.
- Instituto Nacional de Estadística y Geografía. 2022. *Estadísticas a propósito del día internacional de los pueblos indígenas.* https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2022/EAP_PueblosInd22.pdf.
- Kukutai, Tahu, and John Taylor. 2016. *Indigenous data sovereignty: toward an agenda.* ANU press.
- Milan, Stefania, and Emiliano Treré. 2019. Big data from the south (s): beyond data universalism. *Television & New Media* 20 (4): 319–335.
- Moreno, Manuel Aguilar, et al. 2002. The indio ladino as a cultural mediator in the colonial society. *Estudios de cultura nahuatl* 33.
- Murimi, Renita. 2021. When humans judge other humans using machines. *MIT Computational Law Report*.
- Nansen Daz, Eréndira, and Paredes Martínez. 2003. Los intérpretes jurados como auxiliares de la administración de justicia colonial. C. Paredes Martínez y M. Terán (Coords.). *Autoridad y gobierno indígena en Michoacán*, 173–184.

- Nesvig, Martin. 2012. Spanish men, indigenous language, and informal interpreters in postcontact Mexico. *Ethnohistory* 59 (4): 739–764.
- OCDE. 2019. *Artificial intelligence in society*. <https://doi.org/https://doi.org/10.1787/eedfee77-en>. <https://www.oecd-ilibrary.org/content/publication/eedfee77-en>.
- Pool, Ian. 2016. Colonialisms and postcolonialisms fellow traveller: the collection, use and misuse of data on indigenous people. *Indigenous data sovereignty: Toward an agenda*, 57–78.
- Research Data Alliance International Indigenous Data Sovereignty Interest Group. 2019a. *Care principles for indigenous data governance*. <https://www.gida-global.org/care>.
- . 2019b. *Principios care para la gobernanza de datos indígena*. <https://www.gida-global.org/s/Principios-CREA.pdf>.
- Ricaurte, Paola. 2019. Data epistemologies, the coloniality of power, and resistance. *Television & New Media* 20 (4): 350–365.
- . 2022. Ethics for the majority world: ai and the question of violence at scale. *Media, Culture & Society* 44 (4): 726–745.
- Sabia, T. 2015. Los intérpretes de lenguas indígenas: una forma de garantizar los derechos lingüísticos y el debido proceso. *Jus Semper Loquitur*, 51–75.
- Snipp, C Matthew. 2016. What does data sovereignty imply: what does it look like. *Indigenous data sovereignty: Toward an agenda*, 39–56.
- Trejo, Sofia. 2022. Entre la técnica y la práctica. *eminist Artificial Intelligence: Towards a research-action agenda for Latin America and the Caribbean*.
- Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark, and Francesco Fuso Nerini. 2020. The role of artificial intelligence in achieving the sustainable development goals. *Nature communications* 11 (1): 233.
- Walter, Maggie. 2016. Data politics and indigenous representation in Australian statistics. *Indigenous data sovereignty: Toward an agenda* 38:79–98.