

Soberanía tecnológica y lenguas indígenas: de la dominación histórica a la era de la inteligencia artificial

Soberania tecnológica e línguas indígenas: da dominação histórica à era da inteligência artificial

Technological sovereignty and indigenous languages: from historical domination to the era of artificial intelligence

Silvia Fernández-Sabido ^{*1} y Laura Peniche-Sabido ^{†2}

¹Secretaría de Ciencia, Humanidades, Tecnología e Innovación (Secihti), Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo), Mérida, México.

²Colegio "Educativa México", Mérida, México.

Resumen

La relación entre las lenguas originarias y el desarrollo tecnológico ha estado históricamente marcada por tensiones de poder y exclusión sistémica. Este artículo analiza cómo la evolución de las tecnologías de la comunicación y la información reconfigura las relaciones de poder en las comunidades indígenas y condiciona su soberanía frente a la inteligencia artificial (IA). Mediante una revisión crítico-documental de 62 fuentes clave (67.7 % de actualidad), apoyada en herramientas de IA para el soporte en la redacción, se sistematiza una genealogía que transita desde el saber ancestral hasta el activismo digital contemporáneo. Los resultados, sustentados en taxonomías de viabilidad técnica y evolución histórica que integran los impactos reportados en la literatura, revelan un movimiento pendular en la agencia tecnológica: desde la autonomía prehispánica hasta la desposesión simbólica mediante la alfabetización latina, culminando en la actual reapropiación de herramientas digitales y en el posicionamiento frente a los desafíos del Procesamiento de Lenguaje Natural (PLN). Se concluye que la integración soberana en los Grandes Modelos de Lenguaje (LLM) requiere trascender el extractivismo de datos para establecer protocolos de colaboración horizontal y diseño participativo. La soberanía de la palabra en la era de la IA es una condición necesaria para la alineación de la tecnología con los mejores valores humanos.

Palabras clave: Lenguas indígenas. Descolonización. Inteligencia artificial. Soberanía tecnológica. Tecnologías de la comunicación.

Resumo

A relação entre as línguas originárias e o desenvolvimento tecnológico tem sido historicamente marcada por tensões de poder e exclusão sistêmica. Este artigo analisa como a evolução das tecnologias da informação e comunicação reconfigura as relações de poder nas comunidades indígenas e condiciona sua soberania diante da inteligência artificial (IA). Por meio de uma revisão crítico-documental de 62 fontes fundamentais (67,7% de atualidade), apoiada em ferramentas de IA para o suporte na redação, sistematiza-se uma genealogia que transita do saber ancestral ao ativismo digital contemporâneo. Os resultados, sustentados em taxonomias de viabilidade técnica e evolução histórica que integram os impactos relatados na literatura, revelam um movimento pendular na agência tecnológica: da autonomia pré-hispânica à despossessão simbólica mediante a alfabetização latina, culminando na atual reapropriação de ferramentas digitais e no posicionamento frente aos desafios do Processamento de Linguagem Natural (PLN). Conclui-se que a integração soberana nos Grandes Modelos de Linguagem (LLM) requer transcender o extrativismo de dados para estabelecer protocolos de colaboração horizontal e design participativo. A soberania da palavra na era da IA é, além de um direito indígena, uma condição necessária para o alinhamento da tecnologia com melhores os valores humanos.

Palavras-chave: Línguas indígenas. Descolonização. Inteligência artificial. Soberania tecnológica. Tecnologias da comunicação.

Texto livre
Linguagem e Tecnologia

DOI: 10.1590/1983-3652.2026.64143

Sección:
Artículos

Autor correspondiente:
Silvia Fernández Sabido

Editor de sección:
Hugo Heredia Ponce
Editor de maquetación:
Rayany Chaves

Recibido el:
29 de enero de 2026
Aceptado el:
22 de marzo de 2026
Publicado el:
12 de octubre de 2026

Esta obra está bajo una
licencia «CC BY 4.0».



*Email: sfernandez@centrogeo.edu.mx

†Email: penichesabidolaura@educativamexico.com.mx

Abstract

The relationship between Indigenous languages and technological development has been historically marked by power tensions and systemic exclusion. This article analyzes how the evolution of Information and Communication Technologies reconfigures power dynamics within Indigenous communities and how these dynamics condition their sovereignty in the face of Artificial Intelligence (AI). Through a critical documentary review of 62 key sources (67.7 % of which were published between 2020-2026), supported by AI tools for writing assistance, a genealogy is systematized that moves from ancestral knowledge to contemporary digital activism. The results, supported by taxonomies of technical viability and historical evolution that integrate impacts reported in the literature, reveal a pendular movement in technological agency: from pre-Hispanic autonomy to symbolic dispossession through Latin literacy, culminating in the current re-appropriation of digital tools and positioning toward the challenges of Natural Language Processing (NLP). It is concluded that the sovereign integration of these languages into Large Language Models (LLMs) requires transcending data extractivism to establish protocols for horizontal collaboration and participatory design. Sovereignty over the spoken word in the era of AI is, in addition to an Indigenous right, a necessary condition for aligning technology with the best human values.

Keywords: Indigenous languages. Decolonization. Artificial intelligence. Technological sovereignty. Communication technology.

1 Introducción

1.1 Panorama global de las lenguas indígenas

En la actualidad, la identificación de quienes pertenecen a un pueblo indígena suele basarse en su lugar de residencia, lengua, vestimenta y costumbres, así como en su situación de marginación. Sin embargo, cuando una persona de origen indígena migra de un entorno rural a uno urbano se genera un proceso de adaptación en el que se mezclan elementos de la ruralidad con las nuevas experiencias de la ciudad, reconfigurando su identidad (Corona, 2023). En este escenario surgen preguntas interesantes: ¿Se deja de ser indígena al trasladarse a una ciudad y adoptar un estilo de vida diferente?, ¿Qué diríamos si la persona deja de usar su lengua y adopta la cultura del nuevo entorno?, ¿Qué sucede si, al intentar participar en la modernidad, simplemente no encuentra su lengua en internet? La poca presencia de las lenguas originarias en el mundo digital puede convertirse en una nueva forma de exclusión que obliga a elegir entre la herencia propia y la integración tecnológica.

Hoy día, menos del 6 % de la población mundial se identifica como indígena (Ivers, 2023), lo que sugiere que cerca del 94 % de la humanidad se ha alejado de su identidad originaria en diversos grados. La condición indígena parece ser un espectro que incluye desde personas que mantienen una conexión con su territorio, costumbres y lengua, hasta quienes han perdido su herencia cultural. De aquí surge otra pregunta: ¿Se puede recuperar la identidad perdida aprendiendo todos la lengua y cultura de nuestros ancestros?

Pese a que el Atlas Mundial de las Lenguas (UNESCO, [s. d.]) reporta la pervivencia de cerca de 7 mil idiomas, su distribución y vitalidad son profundamente desiguales. Asia y África concentran la mayor diversidad lingüística, mientras que, en América Latina, países como México encabezan la lista con 68 lenguas originarias oficiales. Aunque la población indígena representa apenas el 6 % mundial, este grupo resguarda el 60 % de las lenguas del planeta y el 80 % de su biodiversidad (Ivers, 2023). Este patrimonio se encuentra en estado de vulnerabilidad; solo el 0.8 % de las lenguas documentadas son consideradas dominantes, mientras que la mayoría de las lenguas indígenas enfrentan riesgos de extinción.

1.2 Sistemas de conocimiento originarios

A pesar de que la vida de los pueblos originarios está permeada por tradiciones heredadas de la época colonial y moderna, aún se percibe un rico tapiz de saberes e idiomas ancestrales (Mager *et al.*, 2018; Makgopa, 2022) que revela su capacidad de sintetizar grandes campos de observación creando lo que ciertos autores denominan sistemas de conocimiento indígena (Ajani *et al.*, 2024; Ibáñez Blancas *et al.*, 2020). Estos sistemas abarcan los dominios de naturaleza y salud, extendiéndose hacia

tecnología, sociedad, economía, cultura, gobernanza y educación.

Las lenguas indígenas codifican filosofías sobre la identidad territorial y la reciprocidad comunitaria, manteniendo el equilibrio entre el ser y su entorno. Estudios actuales demuestran que expresiones creativas modernas en lenguas originarias siguen cimentadas en filosofías ancestrales, como el *Dalihan Na Tolu* en Indonesia (Pasaribu y Manik, 2026). Esta sabiduría se manifiesta en dimensiones prácticas; hoy día, estos pueblos conservan la habilidad de leer señales de anticipación climática, siguen interpretando signos del cuerpo ante una variedad de enfermedades y preservan formas de producción y alimentación orgánicas; resultado de miles de años de coevolución con su entorno, que presentan pocas externalidades negativas y requieren pocos insumos (Deance Bravo y Troncoso, 2017; Valladolid y Sandoval Chayña, 2007).

Aunque no todos los individuos indígenas son conscientes de ese conocimiento, encontramos su impronta en las lenguas que usan. Por ejemplo, la lengua Maya tiene nombres para todas las partes del cuerpo, externas e internas, por pequeñas que sean. Ese saber es útil para el personal médico y de partería, sea indígena o no (Janetsky, 2022; Martín Briceño y Briceño Chel, 2013). Algunos pueblos utilizan formas de comunicación únicas, como silbidos o tañido de tambores que imitan el habla, con los que transmiten mensajes a larga distancia aprovechando las propiedades bioacústicas del entorno (Meyer; Dentel y Seifart, 2012). Ejemplos de ello son las comunidades silbos gomeros de las Islas Canarias, España, o la comunidad mazateca de Oaxaca, México, que se comunican con chiflidos compartiendo mensajes complejos que viajan kilómetros a través de la accidentada geografía de sus territorios.

La tecnología, entendida como el conjunto de técnicas reunidas en un sistema y tiempo determinados que se concretan en objetos físicos (Assess Technology, s/f), ha existido desde siempre. La tecnología ancestral indígena, como los sorprendentes pigmentos nanoestructurados (Fernández-Sabido *et al.*, 2012; Yacamán *et al.*, 1996), se caracteriza por estar adaptada al ambiente y ser útil en el contexto donde se gesta. Sin embargo, cuando la tecnología proviene del exterior, los beneficios para la comunidad suelen ser marginales.

1.3 La brecha tecnológica y digital

La desconexión entre el saber ancestral y la técnica externa adquiere una nueva dimensión en la era digital. Si lenguas como el maya poseen precisiones anatómicas que la ciencia convencional apenas comienza a integrar, su ausencia en los datos que alimentan la IA es una pérdida de conocimiento especializado que reduce la diversidad del pensamiento humano.

La tecnología no es neutral; afecta el bienestar colectivo y conlleva consecuencias políticas, ya que las relaciones de poder se entrelazan de manera inadvertida en su diseño y aplicación. El algoritmo es también un texto (Araújo, 2025), es decir, una construcción cargada de intencionalidad que, en el caso de las lenguas indígenas, suele imponer una visión monocultural. Estas dinámicas ocultas pueden generar injusticias sociales, lo que subraya la necesidad de reconocer cómo las herramientas tecnológicas inciden profundamente en la cultura resultando en una brecha profunda (Alvarez Avila, 2016; Bielenberg, 1999).

En México, los estados con mayor población indígena presentan los índices más bajos de conectividad y acceso a hardware (INEGI, 2024). Quienes no emplean lenguas dominantes quedan excluidos de los beneficios de la digitalización. Aunque han surgido estrategias de Educación Intercultural Bilingüe (EIB) y marcos legales robustos, como la Ley General de Derechos Lingüísticos en México o el reconocimiento del Estado Plural en Bolivia, persiste una desactualización normativa frente a las innovaciones tecnológicas.

La lengua constituye el vehículo primordial para acceder a los saberes ancestrales. La interacción *lengua-sociedad* define todo un campo de estudio para las ciencias sociales, cuyo reto es dimensionar la pérdida epistémica que conlleva la extinción lingüística (Ibáñez Blancas *et al.*, 2020; Ríos Hernández, 2010). Ante esta amenaza, diversos actores han impulsado iniciativas para revertir el desplazamiento lingüístico, recurriendo a la tecnología para potenciar sus acciones.

1.4 Evolución histórica de la agencia tecnológica indígena

Mirar de cerca la relación entre los pueblos originarios y la tecnología, permite desmitificar la supuesta pasividad de estas comunidades frente a la innovación. Históricamente, la tecnología indígena no solo fue soberana, sino vanguardista; el desarrollo de materiales complejos como el Azul Maya (Fernández-Sabido *et al.*, 2012; Yacamán *et al.*, 1996) y la sofisticación de sus sistemas de registro e inscripción (Klein, 2001; Lenoir, 1998) demuestran una capacidad técnica autónoma donde lengua-herramienta operaban en una simbiosis epistémica. Sin embargo, la expansión colonial impuso marcos técnicos que reconfiguraron la tecnología como un instrumento de jerarquización y control (Day, 2006; Mallén Rivera, 2012), desplazando los saberes locales hacia la periferia del progreso. Pese a este desplazamiento, el siglo XXI ha sido testigo de una fase de reapropiación que se ha transformado en una estrategia de resistencia digital (Dyson; Grant y Hendriks, 2016; Francese, 2023).

Este trabajo presenta una revisión crítico-documental de los esfuerzos por integrar las lenguas y saberes de los pueblos originarios al ámbito tecnológico. El objetivo es, a partir de un análisis crítico de experiencias previas, mapear las tensiones, riesgos y oportunidades que surgen de la integración de las lenguas indígenas a la inteligencia artificial (IA). En particular, se examina su posición frente a los Grandes Modelos de Lenguaje (LLM, por sus siglas en inglés), donde la falta de representación amenaza con profundizar una brecha de invisibilidad digital. Es vital desplazar el enfoque desde la simple gestión de datos hacia procesos de intercambio humano; por ello, en este estudio se opta por el término “tecnologías de la comunicación y la información” sobre el concepto institucionalizado de “Tecnologías de la Información y la Comunicación” (TIC), reconociendo que la transmisión de saberes precede a lo técnico.

La presente investigación busca dar respuesta a la siguiente interrogante: ¿De qué manera la evolución de las tecnologías de la comunicación y la información ha reconfigurado las relaciones de poder en las comunidades indígenas, y cómo las tensiones entre la apropiación tecnológica y las estructuras de control histórico condicionan hoy su soberanía frente a la IA?

2 Metodología

La presente investigación adopta un enfoque cualitativo con un diseño documental y exploratorio, orientado a mapear la intersección entre las lenguas originarias y las tecnologías de la comunicación y la información. Se siguieron los principios generales de la declaración PRISMA para garantizar la transparencia y replicabilidad en la selección del corpus.

2.1 Estrategia de búsqueda y consolidación del corpus

El proceso de construcción del corpus se estructuró en dos etapas:

- *Recolección y curaduría*: se realizó una búsqueda en Dimensions-AI, ERIC, SciELO, Redalyc y Google Scholar. Se optó por estas fuentes debido a su alta pertinencia en lenguas originarias y contextos del Sur Global. Se utilizaron descriptores booleanos en español e inglés *tecnología AND lenguas indígenas; indigenous languages AND technology; procesamiento de lenguaje natural AND lenguas indígenas*).
- *Validación de alto impacto y analítica de red*: se realizó una fase complementaria en Web of Science (WoS) y Scopus. Para robustecer la exhaustividad de la búsqueda y descubrir conexiones latentes entre autores del Sur Global, se empleó la herramienta de analítica de citas ResearchRabbit. Esta validación facilitó la integración de estudios de vanguardia (2024-2026) y aseguró una convergencia del 85 % entre las fuentes detectadas.

2.2 Criterios de selección y análisis situado

De los 85 registros iniciales gestionados en Zotero, se aplicaron criterios de exclusión por redundancia teórica y pertinencia temática, resultando en un corpus de 62 documentos clave. La selección siguió un criterio temporal diferenciado: para el análisis de la IA y el Procesamiento de Lenguaje Natural (PLN), se aplicó un marco de vigencia (2020-2026) debido a la acelerada evolución de estas tecnologías. En contraste, para documentar la evolución histórica y la relación con las lenguas indígenas, se adoptó una perspectiva de largo alcance sin restricciones temporales. La distribución

documental del corpus es: el 49.2 % corresponde a artículos científicos, el 18.5 % a libros y capítulos especializados, y el 13.8 % a actas de congresos, complementados por informes de organismos internacionales (10.8 %) y fuentes digitales (7.7 %). El 67.7 % de las fuentes pertenece al último lustro, manteniendo una representatividad predominante de México, América Latina y África, que refuerza un enfoque de investigación situado que refleja el origen de las autoras, quienes, como descendientes de pueblos originarios, asumen un compromiso personal con la visibilización de estas lenguas en el entorno digital.

2.3 Uso de herramientas de asistencia

En cumplimiento de las políticas de transparencia, se informa que se utilizaron los modelos de lenguaje Perplexity y Gemini como asistentes de redacción para la optimización de estilo y la traducción técnica de fragmentos específicos al inglés o portugués. La supervisión crítica, la verificación de fuentes y la autoría intelectual corresponden íntegramente a las investigadoras.

3 Resultados

La Tabla 1 sistematiza la evolución de las tecnologías de la comunicación y la información en contextos indígenas, vinculando cada fase con los impactos positivos y negativos reportados. Las herramientas se categorizan según su proveniencia (interna o externa), revelando que mientras las tecnologías externas históricamente han servido como vehículos de asimilación, las iniciativas internas tienden a fortalecer la cohesión comunitaria.

Tabla 1. Evolución e impactos de las tecnologías de la comunicación y la información en contextos indígenas.

Tecnología	Proveniencia (marca temporal)	Agentes de creación y ejecución	Ejemplos	Impacto reportado en comunidades indígenas	
				Positivo	Negativo
Escritura	Interna (Pre-conquista)	Especialistas indígenas en ciencia, tecnología y artes.	Escritura en soportes variados: códigos, cerámicas, piedra.	Preservación de memorias históricas y sistemas de conocimiento (Chemla, 2005; Fernández-Sabido <i>et al.</i> , 2012; Yacamán <i>et al.</i> , 1996; Lenoir, 1998).	
Escritura	Externa (Siglo XVI)	Iglesia y escuelas	Libros impresos en papel con caracteres latinos.	Acceso a la educación y al conocimiento (Bielenberg, 1999).	Evangelización y alfabetización masiva, deslegitimación de las culturas indígenas y exclusión de las mujeres de los registros oficiales de conocimiento (Bielenberg, 1999; Chemla, 2005; Day, 2006; Garcés Velásquez, 2020; Holmes; Renn y Rheinberger, 2003; Llanes-Ortiz, 2023; Janetsky, 2022).
Navegación	Interna (Pre-conquista)	Personas artesanas, agricultoras y comerciantes indígenas.	Navegación de rutas comerciales y ciclos astronómicos.	Soberanía alimentaria y consolidación de redes de intercambio cultural (Valladolid y Sandoval Chayña, 2007; Yacamán <i>et al.</i> , 1996).	

Navegación	Externa (Siglo XV)	Potencias europeas	Conquista transoceánica		Pérdidas humanas, culturales y de territorio (Day, 2006; Gutiérrez Fonseca, 2021; Mallén Rivera, 2012).
Medios analógicos de comunicación	Externa (Siglo XIX)	Empresas productoras comerciales.	Radio, cine y televisión		Aumento de la brecha tecnológica, globalización que lleva a desaparecer tradiciones, reforzamiento de estereotipos (Cazden, 2003; Dyson; Grant y Hendriks, 2016; Iturriaga Acevedo y Reyes, 2020).
Medios analógicos de comunicación	Interna (Siglo XX)	Agentes de comunicación indígena	Radio, cine y televisión comunitarios	Difusión de información en lenguas locales, producción educativa para proteger lenguas y culturas (Mkontwana y Ndivhuwo, 2013; Gobierno de México, 2023).	
Tecnología digital	Externa (Siglo XX)	Grandes empresas tecnológicas y creadores de contenido del norte global .	Internet, computadoras, dispositivos móviles, aplicaciones.		Peligro para la diversidad lingüística, se difunden contenidos que favorecen culturas predominantes y debilitan identidades locales (Alvarez Avila, 2016; Cazden, 2003; Garcés Velásquez, 2020; Llanes-Ortiz, 2023).
Tecnología digital	Interna (Siglo XXI)	Activismo digital indígena: autorías musicales, artísticas y generadoras de contenido.	Redes de telefonía, aplicaciones de celular, plataformas digitales, contenidos en redes sociales, doblaje de películas, diseño de tipografías.	Preservación de sistemas de conocimiento indígena, revitalización de lenguas y culturas, descolonización del sistema educativo (Ajani <i>et al.</i> , 2024; Garcés Velásquez, 2020; Llanes-Ortiz, 2023; Solari Pita <i>et al.</i> , 2021).	
Inteligencia artificial	Externa (Siglo XXI)	Grandes empresas tecnológicas e instituciones de investigación con incipiente participación indígena.	Grandes modelos de lenguaje, asistentes virtuales, traductores.	Surgen algunas aplicaciones de traducción y enseñanza de lenguas originarias, se abordan problemas étnicos y sociales de los pueblos indígenas (Caswell, 2024; Tonja <i>et al.</i> , 2024).	Globalización acelerada, la brecha tecnológica se amplifica, personas indígenas relegadas a objetos de estudio y productoras de datos (Gasparotto, 2026; Joshi <i>et al.</i> , 2020; Mager <i>et al.</i> , 2018).
Inteligencia artificial	Interna (2020 - actualidad)	Comunidades, personas lingüistas y tecnólogos indígenas.	Proyectos de soberanía de datos, PLN comunitario.	Empoderamiento, autogestión y diseño culturalmente situado (Araújo, 2025; Horna-Saldaña; Perez Perez y Galeano Toro, 2025; Joshi <i>et al.</i> , 2020).	Falta de recursos, dependencia de infraestructuras externas (Mager <i>et al.</i> , 2018; Moraes <i>et al.</i> , 2025).

Fuente: Elaboración propia a partir de las fuentes revisadas.

Para comprender el contexto de vulnerabilidad en el que operan estas herramientas, en la Figura 1 se esquematizan las causas identificadas de la pérdida de las lenguas indígenas, destacando la presión de las lenguas hegemónicas y la brecha digital. Frente a este panorama, la Figura 2 ejemplifica diversas iniciativas que ilustran el alcance global del activismo digital indígena, mostrando una red de resistencia que ha ido trascendiendo las fronteras nacionales.

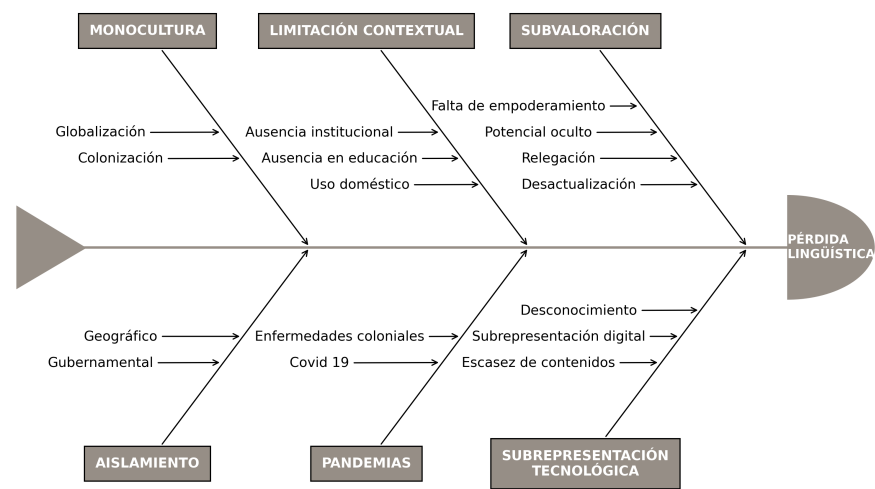


Figura 1. Esquema de las causas identificadas de la pérdida de las lenguas indígenas.
Fuente: Elaboración propia a partir de los documentos citados en el texto.

La Tabla 2 presenta un contraste entre una lengua de recursos abundantes (holandés) y una de escasos recursos (somalí). Se evidencia la disparidad en la disponibilidad de datos para el entrenamiento de algoritmos, donde variables como la presencia en Wikipedia y la funcionalidad en sistemas de traducción marcan una brecha en la representatividad digital de los idiomas y una falta de estandarización deliberada.

La Tabla 3 propone una taxonomía actualizada de las herramientas de PLN para lenguas indígenas, clasificándolas según su viabilidad técnica. Se evidencia que la mayoría de estas lenguas se encuentran en una fase de recursos digitales fragmentarios, sin embargo, las nuevas estrategias de *transfer learning* y modelos *end-to-end* están permitiendo transitar hacia modelos de IA fomentando la autonomía comunitaria.

Tabla 2. Disparidad entre una lengua con abundancia de recursos (high-resource) y otra de escasos recursos (low-resource).

Variable de comparación	Holandés (abundantes recursos)	Somalí (escasos recursos)	Impacto de la falta de estandarización (Gasparotto, 2026)
Población	29 millones	18 millones	La demografía no dicta la prioridad tecnológica, sino la hegemonía del idioma.
Presencia en Wikipedia	2 millones de artículos	5 mil artículos	Los modelos se entrenan con sesgos de representatividad masiva.
Uso en traducción	Extensivo, alta calidad	Calidad inferior	Se imponen estructuras gramaticales externas que distorsionan el sentido.
Ideología lingüística	Estandarización global	Invisibilización y fragmentación	El PLN refuerza la exclusión al no reconocer variantes.

Fuente: Elaboración propia a partir del trabajo de Mussandi y Wichert (2024).



Figura 2. Mosaico de iniciativas de activismo digital indígena.

Fuente: Elaboración propia a partir de fuentes revisadas. Imágenes tomadas de sitios oficiales (consultados en oct. 2024): Kenia (zeno.fm/radio/radio-domus-fm-kenya); Ecuador (facebook.com/groups/749828998458070); Micronesia (talking-dictionary.swarthmore.edu); México (proyecto-miyotl.web.app); Suecia (umesamiska.nu); Canada (mobile.knet.ca); México (ir.library.louisville.edu/chatino); India (commons.wikimedia.org/wiki/File:Shukla_Sorang_Sompeng.svg).

Tabla 3. Clasificación y viabilidad de las tecnologías para lenguas indígenas.

Categoría	Disponibilidad real	Estado actual
Emblemáticas (traducción, Reconocimiento Automático de Voz [RAV])	Baja funcionalidad, pero alta prioridad en investigación (Horna-Saldaña; Perez Perez y Galeano Toro, 2025).	Basadas en <i>transfer learning</i> y modelos <i>end-to-end</i> . Se afinan con pocos datos, alta relevancia cultural.
Prácticas (teclados, búsqueda)	Amplia como metodología, pero limitada como tecnología.	Utilizan modelos pre-entrenados para normalizar la escritura reduciendo la influencia de estándares externos (Gasparotto, 2026).
En espera (ortografía, gramática)	Bloqueadas por falta de texto. Dependen del éxito de la transcripción automática.	La estrategia es usar el RAV para generar los volúmenes de texto necesarios para crear correctores.
Experimentales (IA Generativa, chatbots)	Prometedoras, pero su futuro está ligado al éxito de las categorías “Emblemáticas” y “Prácticas”.	Avances, su éxito depende de la soberanía de los datos (Moraes <i>et al.</i> , 2025).

Fuente: Elaboración propia adaptada de Littell *et al.* (2018) con actualizaciones (2024-2026), a partir del trabajo de Mussandi y Wichert (2024).

4 Discusión

4.1 Genealogía tecnológica y lenguas indígenas

La Tabla 1 revela un movimiento pendular en la agencia tecnológica indígena, marcando una transición desde la soberanía ancestral hacia la imposición colonial y el posterior retorno hacia la autonomía. Mientras que en el periodo pre-conquista la tecnología era de procedencia interna, la conquista (siglo XVI) desplazó la agencia hacia el exterior, convirtiendo la escritura y los medios analógicos en herramientas de deslegitimación. El punto de inflexión hacia la recuperación de la agencia se sitúa en el surgimiento del activismo digital. En este momento, las comunidades dejan de ser sujetos pasivos para convertirse en agentes que disputan el espacio virtual. Este retorno, iniciado con la reapropiación de las redes sociales y la creación de contenidos, se extiende hoy con la IA, que es el estadio más reciente de una soberanía tecnológica activamente en construcción. Para comprender la magnitud de esta recuperación, es preciso analizar cómo los soportes tecnológicos han condicionado la producción y el control de la información.

El papel ha sido una tecnología clave tanto en la producción de conocimiento (Lenoir, 1998) como herramienta de dominación (Day, 2006). Pueblos originarios como mayas o egipcios lo usaron para registrar historias y conocimientos; sin embargo, el papel también es un dispositivo manipulable que sirve a intereses específicos (Chemla, 2005; Holmes; Renn y Rheinberger, 2003). Esta ambivalencia sugiere que el soporte material del conocimiento nunca es neutro sino sujeto a las estructuras de poder de su época; como los algoritmos actuales, que al final son textos que proyectan intencionalidades en el espacio digital (Araújo, 2025).

La escritura moderna de las lenguas originarias de América, ajustada al sistema de la fuerza colonizadora (caracteres latinos), es una tecnología cuyo efecto no se considera del todo beneficioso (Garcés Velásquez, 2020; Llanes-Ortiz, 2023). Un estudio con comunidades indígenas americanas, donde se implementaron campañas de alfabetización, concluyó que esta práctica altera la cultura que intenta preservar. Señala que el binomio iglesia-escuela ha sido el mayor promotor de la alfabetización indígena, con poco éxito fuera de estos ámbitos. Estas circunstancias han llevado a algunas comunidades a oponerse a la alfabetización, que prioriza la estructura del colonizador sobre los saberes propios (Bielenberg, 1999). Este proceso de alfabetización masiva fue una herramienta de desposesión simbólica. Al imponerse la escritura como único registro válido de la “verdad”, se invisibilizaron los saberes transmitidos por las mujeres indígenas, cuya autoridad reside tradicionalmente en la oralidad y la práctica comunitaria, tal como se observa en la discusión sobre partería de Janetsky (2022).

La hegemonía del registro escrito no fue un fenómeno aislado. Las innovaciones tecnológicas han facilitado e intensificado las dinámicas de dominación sobre los pueblos indígenas y sus territorios: la invención de grandes embarcaciones permitió los viajes de conquista transoceánicos, mientras que las armas de fuego posibilitaron el sometimiento de civilizaciones enteras (Mallén Rivera, 2012).

La expansión tecnológica transitó desde la fuerza física hacia la influencia ideológica. Los primeros medios de comunicación masiva (radio, cine y televisión) contribuyeron a formar y deformar culturas. En Alaska se ha demostrado la relación entre la desaparición de tradiciones y la introducción de los televisores (Cazden, 2003). Aunque existen producciones televisivas encaminadas a proteger lenguas y culturas, también se han usado para estereotipar a las personas indígenas (Iturriaga Acevedo y Reyes, 2020), ejerciendo una suerte de extractivismo simbólico donde la imagen del indígena se utiliza sin su consentimiento o para fines ajenos a su bienestar.

La tecnología digital (internet, dispositivos móviles y computadoras) puede ser un peligro para la diversidad lingüística, ya que difunde contenidos que favorecen a las culturas dominantes y debilitan las identidades locales (Alvarez Avila, 2016; Garcés Velásquez, 2020; Llanes-Ortiz, 2023). En este escenario, el riesgo se traslada al ámbito de los datos: el uso de lenguas indígenas en plataformas globales a menudo alimenta algoritmos externos sin que exista una retribución o protección para las personas hablantes. No obstante, la era digital también ha ofrecido una vía para la revitalización de los sistemas de conocimiento e idiomas indígenas (Ajani *et al.*, 2024), siempre que la comunidad logre transitar desde el consumo pasivo hacia la soberanía tecnológica.

Finalmente, la IA representa la frontera más reciente de esta evolución tecnológica. Al igual que ocurrió con la escritura latina o la televisión, la IA no llega como una herramienta neutra. Los Grandes

Modelos de Lenguaje (LLM) y los sistemas de traducción dependen de la recolección masiva de datos lingüísticos; sin embargo, cuando este proceso es liderado por entidades externas, se corre el riesgo de un extractivismo de datos, donde el conocimiento ancestral es procesado y monetizado sin el consentimiento o beneficio de las comunidades originarias (Mager *et al.*, 2018). Pese a este peligro, como se sugiere en la Tabla 1, la IA también abre la posibilidad de una tecnología de proveniencia interna. Si las comunidades lideran el desarrollo de sus propios modelos de Procesamiento de Lenguaje Natural (PLN), la IA puede dejar de ser una herramienta de asimilación para convertirse en un mecanismo de soberanía tecnológica que proteja y potencie la identidad indígena (Joshi *et al.*, 2020).

4.2 Causas estructurales de la erosión de las lenguas indígenas

Como se muestra en la Figura 1, la pérdida de las lenguas indígenas responde a una red de causas estructurales y desequilibrios de poder. El aislamiento geográfico y la falta de atención gubernamental limitan el acceso a servicios básicos, confinando las lenguas originarias al ámbito doméstico mientras las lenguas dominantes hegemonizan los espacios oficiales y académicos (Uekusa, 2019). A este aislamiento se suman las catástrofes demográficas; desde la colonización hasta la pandemia de Covid-19, la pérdida de hablantes de edad avanzada ha fracturado la transmisión intergeneracional de conocimientos (Gutiérrez Fonseca, 2021).

Asimismo, la subvaloración histórica ha desplazado los sistemas de conocimiento indígenas en favor de modelos europeos, un fenómeno que persiste en la actualidad (Garcés Velásquez, 2020; Makgopa, 2022). Aunque las revoluciones tecnológicas poseen un potencial democratizador, el vacío de contenidos digitales en lenguas originarias profundiza la brecha entre el desarrollo social y la adopción técnica (Deance Bravo y Troncoso, 2017; Joshi *et al.*, 2020). En la actual Cuarta Revolución Industrial (4IR), la IA amenaza con afianzar el dominio de pocos idiomas, a menos que se utilicen estos instrumentos para la emancipación mediante un enfoque multicultural y soberano (Marwala, 2023).

4.3 Activismo digital indígena frente a la hegemonía tecnológica

Frente a las dinámicas de exclusión y dominación (Figura 1), el activismo digital indígena surge como una gran fuerza de retorno hacia la soberanía. Este es un movimiento proactivo y, en gran medida, liderado por jóvenes, que busca descolonizar y democratizar la tecnología (Marmion; Obata y Troy, 2014). El fenómeno representa un ejercicio de soberanía tecnológica donde las comunidades retoman el control sobre sus propios datos, medios de comunicación y narrativas. La Figura 2 ejemplifica diversas iniciativas que ilustran este alcance a nivel global.

De manera notable, las radios comunitarias han desempeñado un papel fundamental en la difusión de información en lenguas locales, con un crecimiento constante en países como Sudáfrica (Mkontwana y Ndivhuwo, 2013) (Omusonga; Simiyu y Chesaro, 2023) y México (Gobierno de México, 2023). Asimismo, el uso estratégico de redes sociales ha permitido visibilizar movimientos de resistencia, como #FeesMustFall en el sistema educativo sudafricano (Tshifhumulo y Makhanikhe, 2021). En contextos de revitalización, el pueblo Navajo en Estados Unidos utiliza estas plataformas para promover su cultura (Reyhner; Lockard y Martin, 2024), mientras que en Ecuador se emplea Facebook para la difusión activa de la lengua quichua (Garcés Velásquez, 2020).

La creación de software y herramientas de aprendizaje autogestionadas es otro pilar del activismo. Los habitantes de la comunidad sami de Suecia personalizaban la aplicación Memrise para crear cursos de lengua (Llanes-Ortiz, 2023), mientras que en México se han desarrollado múltiples aplicaciones para la enseñanza de lenguas nacionales (Maldonado Rivera; Blanco Sánchez y Ramiro Reyes, 2018). Estos esfuerzos se complementan con el desarrollo de infraestructura tipográfica, como la creación de fuentes para el sorang-sompeng en la India (Llanes-Ortiz, 2023) y proyectos similares en Latinoamérica (Secretaría de Cultura, 2016).

En el ámbito de la expresión artística y multimedia, el activismo se manifiesta en cómics tradicionales creados en Indonesia (Llanes-Ortiz, 2023) o en la música de jóvenes que utilizan internet para distribuir rap y hip-hop en lenguas originarias (Francese, 2023). Estos esfuerzos de reapropiación técnica alcanzan hitos como el desarrollo de recursos digitales para la lengua chatino por la lingüista

Hilaria Cruz, Auten y Alfaro (2021), o el primer doblaje de cine animado al matsigenka en Perú, promoviendo valores comunitarios y representación femenina (Solari Pita *et al.*, 2021; Llanes-Ortiz, 2023).

Finalmente, el desarrollo de herramientas de productividad y conectividad refleja una búsqueda de autonomía técnica profunda. Los diccionarios parlantes de Micronesia (Llanes-Ortiz, 2023), la suite ofimática TotoOffice en México, que permitió la redacción de una tesis académica en totonaco (Deance Bravo y Troncoso, 2017), y el despliegue de redes de telefonía celular comunitaria en Canadá (Keewaytinook Mobile) y México (Tic-ac), demuestran que el activismo digital es capaz de ofrecer servicios esenciales a costos accesibles y bajo visiones propias (Dyson; Grant y Hendriks, 2016; Llanes-Ortiz, 2023).

Los casos analizados son apenas una muestra de una agencia colectiva que prepara el terreno para el desafío tecnológico más sofisticado hasta la fecha: la IA. En esta transición, la lucha por la lengua se desplaza hacia la soberanía de los datos, convirtiéndose en el eje de las nuevas formas de activismo digital. Esta evolución se alinea con las tendencias actuales en la generación de textos dinámicos mediante LLM, los cuales, como sugieren Ribeiro y Tsunoda (2025) en su reciente revisión sistemática, abren posibilidades inéditas para la producción de contenidos.

4.4 Inteligencia artificial y la soberanía de datos indígenas

El combustible principal de la IA son los datos (documentos, audios y videos) con los que se entrenan los modelos computacionales. Mientras las empresas tecnológicas los acumulan para crear aplicaciones que benefician a las lenguas dominantes, las lenguas indígenas quedan rezagadas (Joshi *et al.*, 2020; Tonja *et al.*, 2024). Esta desigualdad se hace evidente al contrastar dos idiomas con realidades digitales opuestas, tal como se ve en la Tabla 2 donde el PLN clasifica al somalí como “lengua de escasos recursos” y al holandés como “lengua de abundantes recursos” (Mussandi y Wichert, 2024).

Es imperativo precisar que la etiqueta de “escasos recursos” es un término técnico-computacional que se refiere a la carencia de datos digitales y no a una precariedad del idioma. De hecho, las comunidades indígenas cuestionan este término por su connotación peyorativa, reivindicando que sus sistemas lingüísticos son ricos, efectivos y altamente complejos.

La falta de estandarización en el desarrollo del PLN refleja lo que Gasparotto (2026) identifica como ideologías lingüísticas dañinas. Al imponer normas de escritura y estructuras gramaticales ajenas, estos sistemas invisibilizan la riqueza de los idiomas originarios. De hecho, la fuerte oralidad de estos idiomas a menudo enmascara que muchas contaron con ingeniosos sistemas de escritura que se ajustan mejor a su fonología que la versión latina, cuya falta de simplicidad ha dificultado el aprendizaje y la producción de tecnología (Joshi *et al.*, 2020).

Esta asimetría en la disponibilidad de recursos implica que las estrategias computacionales aplicadas a uno u otro grupo no son iguales pues las lenguas de escasos recursos presentan más desafíos. Moraes *et al.* (2025) señalan que incluso en lenguas con abundantes datos, los modelos de lenguaje de gran escala enfrentan dificultades de ambigüedad lingüística que se agrava en idiomas con recursos digitales limitados donde la falta de contexto puede derivar en interpretaciones erróneas.

La diversidad lingüística y los fenómenos tipológicos de las lenguas indígenas presentan desafíos fascinantes para la ciencia (González Santillan, 2021; Mager *et al.*, 2018). Salir de la cámara de resonancia tipológica, donde solo se exploran fenómenos de lenguas similares, permitirá enriquecer el campo del PLN (Joshi *et al.*, 2020; Ward, 2018). Un estudio de Littell *et al.* (2018) clasificó la viabilidad de estas herramientas en cuatro categorías; sin embargo, es preciso actualizar dicha taxonomía para reflejar los avances tecnológicos actuales (ver Tabla 3).

Tradicionalmente, la falta de presencia digital ha limitado la viabilidad de las tecnologías para lenguas indígenas. No obstante, han nacido espacios académicos como NLPAméricas (NLPA, 2021) y Africa Language Dataset Challenge (AI4D, 2019) enfocados en crear materiales de investigación para lenguas como el guaraní, el bribri y el maya, así como generar datos de calidad en regiones de Ghana, Uganda y Sudáfrica (Chiruzzo *et al.*, 2024; Siminyu *et al.*, 2021).

El desplazamiento hacia tecnologías soberanas responde a la capacidad de optimizar algoritmos sin

depender de volúmenes masivos de datos (González Santillan, 2021; Joshi *et al.*, 2020; Littell *et al.*, 2018). El paradigma técnico que lo permite es el aprendizaje por transferencia (transfer learning), que aprovecha el conocimiento de grandes arquitecturas multilingües para adaptar sistemas a lenguas indígenas mediante conjuntos de datos pequeños, pero de alta relevancia cultural.

La IA actúa como herramienta de archivo y como motor para cerrar la brecha de acceso a la información, permitiendo que los servicios digitales sean accesibles en lenguas nativas (Horna-Saldaña; Perez Perez y Galeano Toro, 2025). Si bien empresas como Google han empleado estas técnicas para incluir lenguas como el n'ko de África, el punjabi de Pakistán, el manés de Irlanda y el maya yucateco de México (Caswell, 2024), su enfoque busca la cobertura global masiva más que la autonomía local.

Este camino resulta riesgoso pues al automatizar el proceso bajo lógicas externas, se corre el peligro de minimizar la participación de las personas hablantes, resultando en productos que carecen de interpretación cultural. El estado del arte apuesta por modelos end-to-end que logran la traducción de habla a habla permitiendo que la comunidad gestione su lengua como un fenómeno vivo y esencialmente oral. La verdadera innovación es la capacidad de las comunidades para gobernar estas herramientas (Araújo, 2025). El transfer learning solo es soberano si permite que los pueblos originarios dejen de ser proveedores de datos para convertirse en diseñadores de su futuro digital.

5 Conclusiones

En este trabajo se ha trazado un panorama sobre la compleja relación entre las lenguas indígenas y la evolución de las tecnologías de la comunicación y la información. El análisis evidencia que la pérdida de vitalidad lingüística es el resultado de tensiones de poder estructurales ligadas a la colonización y globalización, que han hecho necesaria la legislación de los derechos lingüísticos. El recorrido histórico permite concluir que la procedencia de la innovación técnica determina su impacto: cuando estas tecnologías son impuestas desde el exterior, suelen intensificar dinámicas de dominación; por el contrario, cuando son apropiadas y ejercidas desde la autonomía comunitaria, se transforman en instrumentos de resistencia con resonancia global.

La llegada de la IA plantea una encrucijada epistemológica. Si los modelos de lenguaje actuales son espejos de los valores depositados en internet, la exclusión de las lenguas originarias nos priva de sistemas de conocimiento y valores esenciales para el bienestar global. Esta exclusión no debe leerse como un fenómeno exclusivo de "lo indígena", ya que representa un síntoma de cómo el entrenamiento algorítmico tiende a marginar cualquier saber no hegemónico. Es imperativo que la incorporación de estas lenguas en la IA abandone el extractivismo de datos y se dirija hacia una investigación participativa que genere productos culturalmente pertinentes. Este proceso debe evitar la homologación de "lo indígena" como un bloque uniforme; como demuestra la hegemonía del quechua sobre lenguas como el puquina, la tecnología no debe replicar lógicas de expansión que erosionen la micro-diversidad de los territorios.

Cuando todas las lenguas humanas, con su carga de sabiduría milenaria, formen parte del tejido digital, tal vez surja, como una de las grandes capacidades emergentes de la tecnología, la tan buscada alineación de la IA con los mejores valores humanos. Esta alineación podría acompañarse de una toma de conciencia global donde cada individuo rastree su ancestría indígena para reencontrarse con la lengua y el territorio, constituyendo, tal vez, la campaña de revitalización más grande y efectiva.

Se reconoce que este estudio presenta limitaciones dadas por la celeridad del desarrollo de los LLM y la necesidad de profundizar en la literatura del Sur Global en otros idiomas, como el portugués; por ello, las taxonomías de viabilidad propuestas requieren actualización constante y validación empírica en territorios de baja conectividad. Hacia el futuro, es imperativo trascender el consentimiento informado convencional para establecer protocolos de colaboración horizontal. La verdadera soberanía digital solo será posible integrando a lingüistas y tecnólogos indígenas en los equipos de diseño, otorgando a la lengua un lugar central en la construcción del futuro tecnológico.

Referencias

AI4D. *African Language Dataset Challenge*. [S. l.]: Artificial Intelligence for Development, 2019. Disponible en: <https://idl-bnc-idrc.dspacedirect.org/server/api/core/bitstreams/aaf5d39d-41bb-4367-ac65-165b2d3eeb03/content>. Acceso en: 25 enero 2026.

- AJANI, Yusuf Ayodeji; OLADOKUN, Bolaji David; OLARONGBE, Shuaib Agboola; AMAECHI, Margaret Nkechi; RABIU, Nafisa y BASHORUN, Musediq Tunji. Revitalizing Indigenous Knowledge Systems via Digital Media Technologies for Sustainability of Indigenous Languages. *Preservation, Digital Technology & Culture*, volumen 53, número 1, páginas 35-44, 2024. DOI: 10.1515/pdtc-2023-0051.
- ALVAREZ AVILA, Abelardo. *Cultura e identidad frente a la globalización*. Saarbrücken: Editorial Académica Española, 2016.
- ARAÚJO, Júlio. O algoritmo é um texto. *Texto Livre*, volumen 18, e58505, 2025. DOI: 10.1590/1983-3652.2025.58505.
- BIELENBERG, Brian. Indigenous Language Codification: Cultural Effects. *En: ANNUAL Stabilizing Indigenous Languages Symposium*. [S. l.: s. n.], 1999. páginas 1-11.
- CASWELL, Isaac. *El Traductor de Google incorpora 111 idiomas, siendo su mayor expansión hasta la fecha*. [S. l.]: Google, 2024. Disponible en: <https://blog.google/intl/es-419/actualizaciones-de-producto/informacion/el-traductor-de-google-incorpora-111-idomas-siendo-su-mayor-expansion-hasta-la-fecha/>. Acceso en: 27 jun. 2024.
- CAZDEN, Courtney Bernadine. Sustaining Indigenous Languages in Cyberspace. *En: NURTURING Native Languages*. Flagstaff: Northern Arizona University, 2003. páginas 1-7.
- CHEMLA, Karine. *History of Science, History of Text*. [S. l.]: Springer Science & Business Media, 2005.
- CHIRUZZO, Luis; DENISOV, Pavel; MOLINA-VILLEGAS, Alejandro; FERNANDEZ-SABIDO, Silvia; COTO-SOLANO, Rolando; AGÜERO-TORALES, Marvin; ALVAREZ, Aldo; CANUL-YAH, Samuel; HAU-UCÁN, Lorena; EBRAHIMI, Abteen; PUGH, Robert; ONCEVAY, Arturo; RIJHWANI, Shruti; WENSE, Katharina von der y MAGER, Manuel. Findings of the AmericasNLP 2024 Shared Task on the Creation of Educational Materials for Indigenous Languages. *En: PROCEEDINGS of the 4th Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP 2024)*. [S. l.: s. n.], 2024. páginas 224-235.
- CORONA, Ignacio. La identidad indígena como identidad urbana. Un abordaje descolonial a las crónicas de Ana Matías Rendón. *Latinoamérica. Revista de Estudios Latinoamericanos*, volumen 1, número 76, página 76, 2023. DOI: 10.22201/cialc.24486914e.2023.76.57551.
- CRUZ, Hilaria; AUTEN, Kylee y ALFARO, Sofia. *Chatino Tonal Books Project*. [S. l.: s. n.], 2021. Disponible en: <https://ir.library.louisville.edu/chatino/3>. Acceso en: 20 enero 2026.
- DAY, David. *Conquista. Una nueva historia del mundo moderno*. Barcelona: Crítica, 2006.
- DEANCE BRAVO Y TRONCOSO, Iván. TotoOffice: Experiencias interculturales en torno a la lengua y la tecnología. *En: LOS territorios discursivos en América Latina-Interculturalidad, Comunicación e Identidad*. 1. edición. Quito: Ediciones Ciespal, 2017. páginas 98-115.
- DYSON, Laurel Evelyn; GRANT, Stephen y HENDRIKS, Max. *Indigenous People and Mobile Technologies*. 1. edición. Londres: Routledge, Taylor & Francis Group, 2016. volumen 1.
- FERNÁNDEZ-SABIDO, Silvia; PALOMO-CARRILLO, Yoly; BURGOS-VILLANUEVA, Rafael y COSS, Romeo de. Comparative Study of Two Blue Pigments from the Maya Region of Yucatan. *MRS Online Proceedings Library*, volumen 1374, número 1, páginas 115-123, 2012.
- FRANCESE, Jonah. Language revitalization through Indigenous Mexican hip hop: Building towards an Indigenous hip hop futurism. *Journal of Multilingual and Multicultural Development*, volumen 45, número 1, páginas 9-21, 2023. DOI: 10.1080/01434632.2022.2143510.
- GARCÉS VELÁSQUEZ, Fernando. Las comunidades virtuales del quichua ecuatoriano: Revalorizando la lengua en un espacio apropiado. *Punto Cero*, volumen 25, número 40, páginas 9-19, 2020. Disponible en: http://www.scielo.org.bo/scielo.php?script=sci_arttext%5C&pid=S1815-02762020000100002%5C&lng=es%5C&nrm=iso. Acceso en: 25 enero 2026.

GASPAROTTO, Melissa. "Missing standardization": Identifying harmful language ideologies in natural language processing work. *Big Data & Society*, volumen 13, número 1, 2026. DOI: 10.1177/20539517251406184.

GOBIERNO DE MÉXICO. *Las radios comunitarias y su labor en el desarrollo rural*. México: Secretaría de Agricultura y Desarrollo Rural, 2023. Disponible en: <https://www.gob.mx/agricultura%7Cdgsiap/articulos/las-radios-comunitarias-y-su-labor-en-el-desarrollo-rural>. Acceso en: 20 oct. 2024.

GONZÁLEZ SANTILLAN, Diana. *CS447 Literature Review: Natural Language Processing for Indigenous Languages of the Americas*. [S. l.]: [s. n.], 2021.

GUTIÉRREZ FONSECA, Itza Nahomy. Las epidemias del México prehispánico: Un breve recorrido histórico. *Revista de Medicina y Cine*, volumen 16, páginas 237-245, 2021. DOI: 10.14201/rmc202016e237245.

HOLMES, Frederic; RENN, Jürgen y RHEINBERGER, Hans-Jörg. *Reworking the Bench: Research Notebooks in the History of Science*. [S. l.]: [s. n.], 2003.

HORNA-SALDAÑA, Cesar Jhonnatan; PEREZ PEREZ, Juan Ernesto y GALEANO TORO, Manuel Leonardo. Artificial intelligence in the preservation of native languages and bridging the information access gap for indigenous peoples. *Journal of Enabling Technologies*, volumen 19, número 1, páginas 63-75, 2025. DOI: 10.1108/JET-09-2024-0063.

IBÁÑEZ BLANCAS, Nicolás; ISCH L., Edgar; PANARIO, Daniel; GUTIERREZ, Ofelia y ZAMBRANO C., Ángela. El cambio climático y los conocimientos tradicionales, miradas desde Sudamérica. *Terra Nueva Etapa*, volumen XXXVI, número 59, 2020.

INEGI. *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2023 (ENDUTIH): Síntesis metodológica*. [S. l.]: INEGI, 2024.

ITURRIAGA ACEVEDO, Eugenia y REYES, Jaime López. *Pueblos indígenas frente al racismo mexicano*. 1. edición. [S. l.]: Contramarea Editorial - UNAM, 2020.

IVERS, Laura. *Pueblos indígenas: Panorama general*. [S. l.]: World Bank, 2023.

JANETSKY, Megan. *Estas parteras tradicionales combinan la herencia maya con la medicina occidental para salvar vidas*. [S. l.]: National Geographic, 2022. Disponible en: <https://www.nationalgeographicla.com/historia/2022/04/estas-parteras-tradicionales-combinan-la-herencia-maya-con-la-medicina-occidental-para-salvar-vidas>. Acceso en: 20 enero 2026.

JOSHI, Pratik; SANTY, Sebastin; BUDHIRAJA, Amar; BALI, Kalika y CHOUDHURY, Monojit. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. *En: PROCEEDINGS of the 58th Annual Meeting of the Association for Computational Linguistics*. [S. l.: s. n.], 2020. páginas 6282-6293. DOI: 10.18653/v1/2020.acl-main.560.

KLEIN, Ursula. Paper tools in experimental cultures. *Studies in History and Philosophy of Science Part A*, volumen 32, número 2, páginas 265-302, 2001. DOI: 10.1016/S0039-3681(01)00010-3.

LENOIR, Timothy. *Inscribing Science*. Stanford: Stanford University Press, 1998.

LITTELL, Patrick; KAZANTSEVA, Anna; KUHN, Roland; PINE, Aidan; ARPPE, Antti; COX, Christopher y JUNKER, Marie-Odile. Indigenous language technologies in Canada: Assessment, challenges, and successes. *En: PROCEEDINGS of the 27th International Conference on Computational Linguistics*. [S. l.: s. n.], 2018. páginas 2620-2632.

LLANES-ORTIZ, Genner. *Iniciativas digitales para lenguas indígenas*. [S. l.]: UNESCO/Global Voices, 2023.

MAGER, Manuel; GUTIERREZ-VASQUES, Ximena; SIERRA, Gerardo y MEZA-RUIZ, Ivan. Challenges of language technologies for the indigenous languages of the Americas. *En: PROCEEDINGS of the 27th International Conference on Computational Linguistics*. [S. l.: s. n.], 2018. páginas 55-69.

- MAKGOPA, Mokgale. Implications of the Fourth Industrial Revolution (4IR) on the Development of Indigenous Languages of South Africa: Challenges and Opportunities. *En: HANDBOOK of Research on Protecting and Managing Global Indigenous Knowledge Systems*. Hershey: IGI Global, 2022. páginas 152-161. DOI: 10.4018/978-1-7998-7492-8.ch009.
- MALDONADO RIVERA, Manuel Roberto; BLANCO SÁNCHEZ, Fernando y RAMIRO REYES, Gilberto. Las tecnologías de la información y la comunicación (TIC) fortalecen la preservación de las lenguas indígenas. *TECTZAPIC Revista de divulgación científica y tecnológica*, volumen 4, número 1, páginas 49-55, 2018.
- MALLÉN RIVERA, Carlos. La ciencia en el México colonial e independiente. *Rev. mex. de cienc. forestales*, volumen 3, número 9, páginas 03-09, feb. 2012. Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_arttext%5C&pid=S2007-11322012000100001%5C&lng=es%5C&nrm=iso. Acceso en: 25 enero 2026.
- MARMION, Doug; OBATA, Kazuko y TROY, Jakelin. *Community, identity, wellbeing: The report of the Second National Indigenous Languages Survey*. Canberra: Australian Institute of Aboriginal y Torres Strait Islander Studies (AIATSIS), 2014.
- MARTÍN BRICEÑO, Patrizia y BRICEÑO CHEL, Fidencio. *Manual de Comunicación para Médicos. Español-maya*. Mérida: Universidad Autónoma de Yucatán, 2013.
- MARWALA, Tshilidzi. The Fourth Industrial Revolution has arrived. *South African Journal of Science*, volumen 119, número 1/2, 2023. DOI: 10.17159/sajs.2023/15429.
- MEYER, Julien; DENTEL, Laure y SEIFART, Frank. A methodology for the study of rhythm in drummed forms of languages: Application to Bora Manguaré of Amazon. *En: INTERSPEECH*. [S. l.: s. n.], 2012. páginas 687-690. DOI: 10.21437/Interspeech.2012-212.
- MKONTWANA, Phinda y NDIHVUWO, Doctor Sundani. The Role of South African Community Radio Stations in Promoting South African Indigenous Languages: A Systematic Review. *International Journal of Social Science Research and Review*, volumen 6, número 10, páginas 179-189, 2013. DOI: 10.47814/ijssrr.v6i10.1536.
- MORAES, Lavínia de Carvalho; SILVÉRIO, Irene Cristina; MARQUES, Rafael Alexandre Sousa; ANAIA, Bianca de Castro; PAULA, Dandara Freitas de; FARIA, Maria Carolina Schincariol de; CLEVESTON, Iury; CORREIA, Alana de Santana y FREITAG, Raquel Meister Ko. Linguistic ambiguity analysis in large language models (LLMs). *Texto Livre*, volumen 18, e53181, 2025. DOI: 10.1590/1983-3652.2025.53181.
- MUSSANDI, Joaquim y WICHERT, Andreas. NLP Tools for African Languages: Overview. *En: PROCEEDINGS of the 16th International Conference on Computational Processing of Portuguese*. [S. l.: s. n.], 2024. páginas 73-82. Disponible en: <https://aclanthology.org/2024.propor-2.11.pdf>. Acceso en: 25 enero 2026.
- NLPA. *Thrid Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)*. [S. l.]: NLPA, 2021. Disponible en: <http://turing.iimas.unam.mx/americasnlp/index.html>. Acceso en: 25 enero 2026.
- OMUSONGA, Tony Okwach; SIMIYU, Pius y CHESARO, Daniel K. Learning indigenous languages in public primary schools in Kenya: Overview, problems and the way forward. *European Journal of Education Studies*, volumen 10, número 12, 2023. DOI: 10.46827/ejes.v10i12.5091.
- PASARIBU, Arsen Nahum y MANIK, Sondang. Thematic and linguistic representations of cultural identity in Batak Toba songs (1980–2025). *Research Journal in Advanced Humanities*, volumen 7, número 1, 2026. DOI: 10.58256/276pwn07.
- REYHNER, Jon; LOCKARD, Louise y MARTIN, Joseph. *Revitalizing Indigenous Languages Challenges and Opportunities*. Flagstaff: Northern Arizona University, 2024.
- RIBEIRO, Patrick Fernandes Rezende y TSUNODA, Denise Fukumi. Geração de textos dinâmicos e criativos com Grandes Modelos de Linguagem (LLMs): Uma revisão sistemática. *Texto Livre*, volumen 18, e60103, 2025. DOI: 10.1590/1983-3652.2025.60103.

RÍOS HERNÁNDEZ, Iván. El lenguaje: Herramienta de reconstrucción del pensamiento. *Razón y Palabra*, número 72, 2010. Disponible en: <https://www.redalyc.org/articulo.oa?id=199514906041>. Acceso en: 25 enero 2026.

SECRETARÍA DE CULTURA. *Tipografía para lenguas indígenas, herramienta de conservación*. México: Secretaría de Cultura, 2016. Disponible en: <https://www.gob.mx/cultura/prensa/tipografia-para-lenguas-indigenas-herramienta-de-conservacion?state=published>. Acceso en: 25 enero 2026.

SIMINYU, Kathleen; KALIPE, Godson; ORLIC, Davor; ABBOTT, Jade; MARIVATE, Vukosi; FRESHIA, Sackey; SIBAL, Prateek; NEUPANE, Bhanu; ADELANI, David I.; TAYLOR, Amelia; ALI, Jamiil Toure; DEGILA, Kevin; BALOGOUN, Momboladji; DIOP, Thierno Ibrahima; DAVID, Davis; FOURATI, Chayma; HADDAD, Hatem y NASKI, Malek. AI4D – African Language Program. *arXiv*, 2021. DOI: 10.48550/arXiv.2104.02516.

SOLARI PITA, Helder; DE LA TORRE TUPAYACHI, Ángela; RINCÓN GONZÁLEZ, Marcela y MUÑOZ RAMÍREZ, Alexander. Escucharnos en la pantalla grande: Osankevantite Irira / El Libro de Lila. El proceso de doblaje de películas a lenguas originarias en la comunidad Amazonía–Bajo Urubamba–Cusco – Perú. En: EXPERIENCIAS en revitalización y fortalecimiento de lenguas indígenas en Latinoamérica: Argentina–Bolivia–Chile–Ecuador–Guatemala–México–Perú. 1. edición. Cochabamba: Funproeib Andes, 2021. página 296.

TONJA, Atnafu; BALOUCHZAH, Fazlourrahman; BUTT, Sabur; KOLESNIKOVA, Olga; CEBALLOS, Hector; GELBUKH, Alexander y SOLORIO, Tamar. NLP Progress in Indigenous Latin American Languages. En: PROCEEDINGS of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S. l.: s. n.], 2024. páginas 6972-6987.

TSHIFHUMULO, Rendani y MAKHANIKHE, Tshimangadzo Justice. *Handbook of Research on Protecting and Managing Global Indigenous Knowledge Systems*. Hershey: IGI Global, 2021. DOI: 10.4018/978-1-7998-7492-8.

UEKUSA, Shinya. Disaster linguisticism: Linguistic minorities in disasters. *Language in Society*, volumen 48, número 3, páginas 353-375, 2019. DOI: 10.1017/S0047404519000150.

UNESCO. *The World Atlas of Languages*. [S. l.]: UNESCO, [s. d.] Disponible en: <https://en.wal.unesco.org/world-atlas-languages>. Acceso en: 20 oct. 2024.

VALLADOLID, Julio R. y SANDOVAL CHAYÑA, P. B. *Señas ancestrales como indicadores biológicos de alerta temprana*. [S. l.]: Programa Mundial de Alimentos, Naciones Unidas, 2007. Disponible en: <https://es.wfp.org/publicaciones/peru-senas-ancestrales-como-indicadores-biologicos-de-alerta-temprana>. Acceso en: 20 oct. 2024.

WARD, Monica. Qualitative research in less commonly taught and endangered language CALL. *Language Learning & Technology*, volumen 22, número 2, 2018. DOI: 10.64152/10125/44639.

YACAMÁN, Miguel José; RENDÓN, Luis; ARENAS, Juan y PUCHE, Mari Carmen Serra. Maya Blue Paint: An Ancient Nanostructured Material. *Science*, volumen 273, número 5272, páginas 223-225, 1996. DOI: 10.1126/science.273.5272.223.

Contribuciones de los autores

Silvia Fernández-Sabido: Conceptualización, Curación de datos, Análisis formal, Investigación, Metodología, Redacción – borrador original, Redacción – Revisión y edición; **Laura Peniche-Sabido:** Investigación, Curación de datos, Redacción – borrador original, Redacción – Revisión y edición.

Disponibilidad de datos

Los datos de investigación están disponibles en el texto del documento.

Financiación

Este trabajo fue hecho en el marco de los proyectos INMX-4974 y CBF-2025-G-368 de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) de México.