



PRISMA ODS

REVISTA MULTIDISCIPLINARIA
SOBRE DESARROLLO SOSTENIBLE

ISSN: 3072-8452

**EL SESGO DEL ALGORITMO
ANTE EL DESPLAZAMIENTO
LINGÜÍSTICO: UNA
CRÍTICA JURÍDICO-
FILOSÓFICA SOBRE LA
EXTINCIÓN DEL
INDIGENISMO EN LA ERA
DE LA IA**

*ALGORITHMIC BIAS IN THE FACE OF
LINGUISTIC DISPLACEMENT: A LEGAL-
PHILOSOPHICAL CRITIQUE OF THE
EXTINCTION OF INDIGENISM IN THE ERA
OF AI*

AUTOR

OCTAVIO RODRÍGUEZ
GONZÁLEZ
UNIVERSIDAD MICHOACANA
DE SAN NICOLÁS DE
HIDALGO
MÉXICO

doi 10.65011

Volumen 5 - Número 2
2026

El Sesgo del Algoritmo ante el Desplazamiento Lingüístico: Una Crítica Jurídico-Filosófica sobre la Extinción del Indigenismo en la Era de la IA

Algorithmic Bias in the Face of Linguistic Displacement: A Legal-Philosophical Critique of the Extinction of Indigenism in the Era of AI

Octavio Rodríguez González

octavio.rodriguez@umich.mx

<https://orcid.org/0009-0004-0816-8227>

Universidad Michoacana de San Nicolás de Hidalgo

México

Artículo recibido: 23/04/2026

Aceptado para publicación: 27/05/2026

Conflictos de Intereses: Ninguno que declarar

RESUMEN

El presente artículo analiza la crisis de las lenguas indígenas en México bajo la lente de la Inteligencia Artificial (IA), conceptualizándola no como un evento fortuito, sino como un proceso sociocultural complejo de desplazamiento lingüístico inducido por factores históricos, económicos, educativos y políticos. A través de una crítica jurídico-filosófica que articula el pensamiento de Althusser, Foucault y Sousa Santos, se argumenta que la arquitectura técnica de la IA opera como un Aparato Ideológico de Estado y un panóptico digital que invisibiliza y desplaza el Patrimonio Cultural Inmaterial de los pueblos originarios. El estudio integra datos del INEGI (2020) para evidenciar cómo el sesgo algorítmico y la escasez de corpus digitales aceleran la extinción del indigenismo, produciendo una "línea abismal" que excluye a las lenguas de baja representatividad del ecosistema de la comunicación global. Finalmente, se propone la necesidad de una transición hacia la soberanía tecnológica indígena y una justicia lingüística que desarticule el poder normalizador del algoritmo contemporáneo.

Palabras clave: inteligencia artificial, lenguas indígenas, desplazamiento lingüístico inducido, sesgo algorítmico, México

ABSTRACT

This article analyzes the crisis of Indigenous languages in Mexico within the context of Artificial Intelligence (AI), conceptualizing it not as an inertial event, but as a complex sociocultural process of induced linguistic displacement driven by historical, economic, educational, and political factors. Through a legal-philosophical critique articulating the thought of Althusser, Foucault, and Sousa Santos, it is argued that the technical architecture of AI operates as a Digital State Ideological Apparatus and a digital panopticon that invisibilizes and displaces the Intangible Cultural Heritage of original peoples. By integrating data from INEGI (2020), this study evidences how algorithmic bias and the scarcity of digital corpora accelerate the extinction of Indigenism, producing an "abyssal line" that excludes low-resource languages from the global communication ecosystem. Finally, the paper proposes a transition toward Indigenous technological sovereignty and linguistic justice aimed at dismantling the normalizing power of contemporary algorithms.

Keywords: artificial intelligence, indigenous languages, induced linguistic displacement, algorithmic bias, Mexico

INTRODUCCIÓN

En la contemporaneidad, la hegemonía de la Inteligencia Artificial (IA) no solo representa un salto cualitativo en la eficiencia técnica, sino que se constituye como un nuevo horizonte de poder que redefine la relación entre el Estado, la tecnología y los sujetos subalternos. El presente estudio aborda la problemática de las lenguas indígenas en México, no como una entidad en retiro natural, sino como víctimas de un proceso sociocultural complejo de desplazamiento lingüístico inducido por factores históricos, económicos, educativos y políticos. Este fenómeno se ve exacerbado por el despliegue de algoritmos de procesamiento de lenguaje natural que, en su arquitectura técnica, operan bajo un sesgo que invisibiliza lo no-occidental.

La tesis central de esta investigación sostiene que la IA, lejos de ser una herramienta neutra de traducción o archivo, funciona como un Aparato Ideológico de Estado y de mercado (Althusser, 1970). Bajo esta lógica, el algoritmo interpele al hablante indígena como un sujeto que debe ser "traducido" para ser reconocido, o bien, condenado a la inexistencia digital. Esta exclusión no es un fallo accidental del sistema, sino una manifestación de lo que de Sousa Santos (2014) denomina epistemicidio: una producción activa de ausencias que traza una línea abismal entre las lenguas funcionales para el capital y aquellas que son relegadas al olvido tecnológico.

Recordemos que la inteligencia artificial presenta limitaciones estructurales para la comprensión y modelización de lenguas indígenas, debido a su alta variación dialectal —frecuentemente con baja inteligibilidad mutua—, la ausencia de estandarización lingüística y la escasez de corpus digitales amplios y anotados, indispensables para su entrenamiento mediante Machine Learning; en consecuencia, los modelos basados en transformers tienden a generar salidas con errores gramaticales, interferencias del español y construcciones pragmáticamente inadecuadas; adicionalmente, la limitada representación de contextos culturales restringe la adecuada interpretación de significados situados y funciones discursivas específicas.

Desde una perspectiva jurídica, este desplazamiento colisiona frontalmente con la noción de Patrimonio Cultural Inmaterial. Mientras los marcos internacionales de la UNESCO (2003) instan a la salvaguardia de la diversidad, la realidad técnica impone una estandarización biopolítica (Foucault, 1975) donde la lengua materna es desplazada por el código binario hegemónico. Los datos del INEGI (2020), que reportan una disminución del 6.1% en la

proporción de hablantes de lenguas originarias, no son meras estadísticas; son el registro de un despojo ontológico.

Por tanto, el objetivo de este artículo es analizar, a través de una crítica jurídico-filosófica, cómo el sesgo del algoritmo acelera la extinción del indigenismo. Se busca problematizar la "política del reconocimiento" (Taylor, 1994) en la era digital, proponiendo que la verdadera protección del patrimonio lingüístico requiere una transición hacia la soberanía tecnológica que desarticule el poder normalizador de la IA contemporánea.

DESARROLLO

Estado de la cuestión

El estado de la cuestión en torno a la relación entre las lenguas originarias y la inteligencia artificial (IA) revela una tensión dialéctica donde la riqueza lingüística de México se enfrenta a una arquitectura técnica diseñada para la homogeneización. México, consolidado como uno de los epicentros de la diversidad cultural a nivel global, posee un patrimonio inmaterial vastísimo integrado por 11 familias lingüísticas, de las cuales emanan 68 agrupaciones y 364 variantes dialectales.

Esta multiplicidad no representa meramente un inventario de términos, sino una compleja red de ontologías y formas de habitar el mundo que desafían la lógica binaria y reduccionista de los modelos de lenguaje contemporáneos. Sin embargo, los datos proporcionados por el Instituto Nacional de Estadística y Geografía documentan lo que se ha denominado una "evidente extinción", donde la proporción de hablantes de lenguas indígenas ha descendido críticamente al 6.1% de la población total (INEGI, 2020, p. 2). Este declive no es un proceso biológico natural, sino un fenómeno de desplazamiento lingüístico inducido por estructuras de poder (INEGI, 2020, p. 5).

La literatura académica actual, particularmente en el campo del procesamiento de lenguaje natural (NLP), demuestra que el sesgo del algoritmo no es un error de programación, sino una manifestación de la "línea abismal" que separa las lenguas funcionales para el capital de aquellas producidas como inexistentes. Investigaciones fundamentales han señalado que los modelos de lenguaje de gran escala (LLM) operan bajo un sesgo de disponibilidad masiva, lo que condena a las lenguas indígenas a ser catalogadas como de "bajos recursos" (low-resource languages). Como sostienen Bender et al. (2021), la concentración de datos en lenguas hegemónicas genera una estratificación epistémica donde el "ser" indígena es

expulsado del ecosistema de la comunicación global (p. 615). En este sentido, la riqueza lingüística de México se encuentra en una paradoja: mientras su diversidad es celebrada retóricamente como Patrimonio Cultural Inmaterial bajo los marcos internacionales, su funcionalidad es activamente anulada en el diseño de las interfaces tecnológicas que median la realidad contemporánea (UNESCO, 2003, p. 5).

A continuación, se presenta el tabulado estadístico que sirve como base empírica para nuestra investigación. Evidenciamos como la concentración del 74.1% de los hablantes en apenas 10 lenguas demuestra que el Patrimonio Cultural Inmaterial de México se encuentra en un estado de vulnerabilidad extrema frente a la estandarización digital.

Tabla 1. Distribución de Hablantes por Lengua Indígena en México (INEGI, 2020)

<i>LENGUA INDÍGENA</i>	<i>NÚMERO DE HABLANTES</i>	<i>PORCENTAJE DEL TOTAL DE HABLANTES</i>	<i>ESTATUS DE VISIBILIDAD ALGORÍTMICA</i>
Náhuatl	1,651,958	22.4%	Alta
Maya (Maaya T'aa)	774,755	10.5%	Alta
Tzeltal	561,907	7.6%	Media
Tsotsil	550,274	7.5%	Media
Mixteco	526,593	7.1%	Media
Zapoteco	490,818	6.7%	Media
Otomí	298,861	4.1%	Baja (Riesgo de sesgo)
Totonaco	256,344	3.5%	Baja (Riesgo de sesgo)
Chol	254,715	3.5%	Baja (Riesgo de sesgo)
Mazateco	237,212	3.2%	Baja (Riesgo de sesgo)
Huasteco	168,729	2.3%	Crítica (Invisibilidad)
Mazahua	153,797	2.1%	Crítica (Invisibilidad)
Otras 56 lenguas*	1,438,682	19.5%	Zona de Epistemicidio
TOTAL	7,364,645	100%	

Fuente: Elaboración propia con datos de Instituto Nacional de Estadística y Geografía [INEGI]. (2020). Censo de Población y Vivienda 2020: Tabulados definitivos. (p. 1-12). Recuperado de <https://www.inegi.org.mx/>

Para el desarrollo de nuestro artículo, esta tabla debe interpretarse bajo la lente de la Sociología de las Ausencias de Sousa Santos. Nótese que la categoría de "Otras 56 lenguas"

representa casi una quinta parte de la población indígena, sin embargo, para un modelo de lenguaje masivo, este grupo es procesado como "ruido" estadístico.

La amalgama de las lenguas restantes constituye el núcleo más vulnerable del patrimonio inmaterial mexicano, representando un mosaico de resistencia lingüística que, pese a su diversidad, pende de un hilo demográfico y tecnológico. Este grupo incluye al P'urhépecha, cuya singularidad como lengua aislada la coloca en una posición de fragilidad extrema ante la estandarización, junto a lenguas con una presencia histórica profunda, pero con un número de hablantes que, en casos extremos como el kiliwa, el awakateko, el mocho', el ayapaneco o el ixil, apenas alcanza a unas cuantas decenas de personas, situándolas en el epicentro de la "zona de epistemicidio" (Santos, 2014, p. 52).

La lista se extiende a través de una geografía espiritual y lingüística que abarca el chinanteco, mixe, tarahumara (rarámuri), zoque, amuzgo, tojolabal, huichol (wixárika), tepehuano, chateño, popoluca, cora, cuicateco, yaqui, kanjobal, triqui, popoloca, pame, mam, chontal de Tabasco, chontal de Oaxaca, sayulteco, chuj, tepehua, kato'k, quiché, huave, pápago, caxcán, seri, pima, chichimeca jonaz, lacandón, jacalteco, o'odham, kekchí, kumiai, chocholteco, matlatzinca, tlahuica, cucapá, paipai, mame, teko, q'anjob'al, akateko y kaqchikel. Para la arquitectura de la Inteligencia Artificial, este vasto repertorio de saberes es procesado como una "ausencia activa", pues al no alcanzar los umbrales de masa crítica de datos exigidos por el mercado, estas lenguas son víctimas de un desplazamiento inducido que las excluye de la interpelación tecnológica, acelerando su tránsito hacia el silencio definitivo (Althusser, 1970, p. 28).

Lenguas con poblaciones mínimas, como el Kiliwa o el Awakateko, se encuentran en lo que Foucault definiría como la periferia del panóptico: no son dignas de ser vigiladas (codificadas) y, por tanto, son condenadas a la extinción por desuso tecnológico. Este es el corazón del desplazamiento lingüístico inducido: la creación de un sistema donde hablar una lengua fuera del "Top 10" se traduce en una exclusión absoluta de los Aparatos Ideológicos de Estado que operan mediante IA.

La desconexión aparente entre la estadística fría y la teoría crítica desaparece cuando se analiza el bilingüismo reportado, el cual alcanza un 87.2% entre la población hablante de lengua indígena (INEGI, 2020, p. 4). Este dato, lejos de ser un indicador de integración armoniosa, refleja la presión sistémica que obliga al hablante a transitar hacia el español para

poder interactuar con los Aparatos Ideológicos de Estado que hoy residen en el entorno digital (Althusser, 1970, p. 22).

Investigaciones previas sobre soberanía de datos indígenas advierten que este tránsito está mediado por un extractivismo epistémico; Kukutai y Taylor (2016) han sido enfáticos al señalar que la captura de datos lingüísticos por corporaciones globales, sin marcos de justicia o consentimiento comunitario, constituye un despojo de la propiedad intelectual y cultural (p. 15). Así, el estado de la cuestión revela que el desplazamiento inducido es el resultado de un vacío jurídico donde la IA actúa como un mecanismo de disciplina que normaliza el lenguaje y borra las huellas de la pluralidad humana (Foucault, 1975, p. 201).

Finalmente, la revisión de los antecedentes bibliográficos permite concluir que la extinción del indigenismo en la era de la IA es un proceso multifactorial donde la técnica se convierte en la última frontera del colonialismo. La importancia de la riqueza lingüística de México radica en su capacidad para ofrecer alternativas al pensamiento único; por ello, la "sociología de las ausencias" resulta esencial para comprender cómo el algoritmo produce activamente el desplazamiento de los pueblos originarios (Santos, 2014, p. 45).

Marco Teórico y Discusión Epistémica

Para consolidar una genealogía crítica que sitúe a la Inteligencia Artificial (IA) no como un evento técnico aislado, sino como la culminación de un proceso histórico de control sobre el lenguaje, el conocimiento y la identidad, es imperativo estructurar un análisis interdisciplinario. Este andamiaje se construye desde la intersección de la semiótica estructuralista, la gramática generativa, la sociología del control y la epistemología del sur, estableciendo que la infraestructura algorítmica contemporánea actúa como un tamiz que purifica la realidad para hacerla procesable, eliminando en el camino cualquier rastro de alteridad ontológica indígena. La técnica, en este sentido, no es un receptáculo neutral, sino un despliegue de la hegemonía cultural que reconfigura las asimetrías coloniales en el plano del código binario.

La base de este planteamiento reside en comprender que la lengua no es un mero instrumento de comunicación o una acumulación de datos intercambiables, sino la estructura misma que posibilita el pensamiento y la existencia política. Al respecto, Saussure (1916/1945) postuló que "la lengua es una forma y no una sustancia" (p. 139), un sistema de valores puros donde el sentido se produce exclusivamente a través de la diferencia y la oposición dentro de una estructura cerrada. En el ecosistema de la inteligencia artificial y los Modelos de Lenguaje de

Gran Escala (LLM), esta "forma" saussureana es capturada, fragmentada y congelada por vectores matemáticos que reducen la riqueza polisémica y contextual de las lenguas originarias a meras probabilidades estadísticas dentro de un espacio latente de alta dimensionalidad. Como consecuencia de este proceso, el signo indígena pierde su anclaje vital en el territorio, la ritualidad y la memoria colectiva para ser re-semantizado por una lógica corporativa global.

Esta violencia estructural se profundiza al contrastarse con la conceptualización de la "gramática universal" de Chomsky (1965/1999), quien sostiene que la facultad del lenguaje es una propiedad cognitiva e innata dotada de reglas profundas que subyacen a la diversidad de todas las lenguas humanas. Sin embargo, la arquitectura actual de las redes neuronales de aprendizaje profundo no emula esta facultad generativa profunda; por el contrario, impone una gramática de máquina superficial que opera por fuerza bruta estadística. Al enfrentarse a la complejidad morfosintáctica de lenguas aglutinantes, polisintéticas o tonales no occidentales, el algoritmo es incapaz de decodificar su estructura y, por ende, las clasifica bajo la rúbrica tecnocrática de "lenguas de bajos recursos" o, en el peor de los casos, como "ruido" estadístico. Esto genera una exclusión jurídica e institucional inmediata, ampliamente documentada en los censos de vulnerabilidad sociolingüística (Instituto Nacional de Estadística y Geografía [INEGI], 2020), pues si los sistemas automatizados de traducción y atención del Estado no pueden procesar la estructura profunda de una lengua originaria, el sujeto que la habla queda despojado de su competencia lingüística activa y de su capacidad de interlocución ante las nuevas interfaces del poder digitalizado.

Esta exclusión no representa un error colateral o un fallo de programación corregible con el tiempo, sino que responde a la operatividad intrínseca de los Aparatos Ideológicos de Estado (AIE) descritos por Althusser (1970/2005). La IA funciona hoy como una de las infraestructuras ideológicas más potentes del capitalismo tardío, garantizando la reproducción de las relaciones de dominio coloniales sin necesidad de la intervención física directa, interpelando al individuo indígena exclusivamente como un "usuario", "perfil" o "consumidor" que debe amoldarse a las directrices del código occidental para ser validado como sujeto existente.

Aquí, el concepto de biopoder de Foucault (1975/2009) encuentra su máxima expresión en la era del capitalismo de vigilancia y los monopolios tecnológicos descritos por Zuboff (2019). El control y la clasificación de las poblaciones ya no requieren únicamente de muros físicos o

de los panópticos arquitectónicos de la modernidad industrial, sino de la captura incesante, opaca y ubicua de datos biométricos, conductuales y lingüísticos. El algoritmo clasifica, jerarquiza y normaliza los cuerpos y los saberes, estableciendo lo que Foucault denominaría una economía de la visibilidad donde solo lo que es almacenable y legible para el procesador central tiene derecho a la existencia política y social. Como bien afirma Foucault (1975/2009), "el panóptico es una máquina de disociar la pareja ver-ser mirado" (p. 231); en la era de la gobernanza algorítmica, esta disociación se convierte en una asimetría radical donde los centros de poder corporativo "ven" y extraen el rastro digital del indígena, pero el sujeto indígena permanece completamente ciego frente a los criterios algorítmicos que definen su acceso a derechos fundamentales como la salud, el crédito agrícola o la administración de justicia.

Esta problemática se arraiga en la infraestructura misma de los datos de entrenamiento. Como advierten Bender et al. (2021) en su crítica seminal sobre los "loros estocásticos", los modelos masivos de lenguaje se alimentan de un *web scraping* indiscriminado que asume que el internet anglo y eurocéntrico representa la totalidad del pensamiento humano. Esta recolección masiva incurre en lo que Noble (2018) define como un "algoritmo de opresión", donde las plataformas tecnológicas replican y amplifican los sesgos raciales e históricos contra los pueblos minorizados. Cuando el conocimiento indígena es capturado mediante estos métodos, se produce un fenómeno de extractivismo informacional, donde las corporaciones occidentales se apropian de saberes medicinales, iconográficos y lingüísticos tradicionales sin respetar los derechos colectivos de propiedad intelectual ni el Consentimiento Libre, Previo e Informado garantizado por el derecho internacional.

Este escenario es la manifestación perfecta del concepto de epistemicidio desarrollado por Boaventura de Sousa Santos (2009/2014), quien sostiene de manera tajante que "no hay justicia social global sin justicia cognitiva global" (2009, p. 160). El epistemicidio algorítmico no destruye los saberes mediante la censura física, sino a través de la fagocitación y la devaluación sistemática: el conocimiento indígena es validado y preservado por la IA solo si puede ser convertido en un insumo útil para el mercado global, y es desechado o distorsionado como "alucinación" si su lógica comunitaria y pluriversal contradice la eficiencia binaria del procesador. Frente a este despojo, autores fundamentales de la Soberanía de Datos Indígenas como Kukutai y Taylor (2016), Carroll et al. (2020) y Duarte (2017) demuestran que la única resistencia decolonial posible implica que los propios pueblos originarios asuman la gobernanza absoluta sobre sus datos, determinando de manera

autónoma que se digitaliza, cómo se almacena y bajo qué marcos comunitarios éticos y lingüísticos se permite el desarrollo de una tecnología propia. Solo a través de esta transición hacia una soberanía tecnológica, que desarticule el poder normalizador del algoritmo contemporáneo, será posible fundar una verdadera ecología de saberes donde la máquina sea un vehículo de autodeterminación y no el acta de defunción digital del indigenismo.

La IA como Aparato Ideológico de Estado y la Interpelación Digital

El despliegue de la Inteligencia Artificial debe entenderse, bajo la óptica de Louis Althusser (1970), como la evolución contemporánea de los Aparatos Ideológicos de Estado (AIE). A diferencia de los aparatos represivos, los AIE funcionan mediante la ideología, interpelando a los individuos como sujetos que deben integrarse a una estructura preexistente para ser reconocidos funcionalmente (p. 22). En la era del silicio, la IA actúa como el AIE dominante, dictando que la participación en la vida política, económica y educativa depende de la adopción de lenguas que el algoritmo sea capaz de procesar con eficiencia técnica, convirtiendo el código en la nueva gramática del poder estatal (Althusser, 1970, p. 25).

Esta interpelación genera un entorno donde el hablante de lenguas indígenas se ve forzado a la auto traducción o al abandono definitivo de su lengua materna. El aparato tecnológico no prohíbe explícitamente la lengua indígena, pero la hace "inservible" en la interfaz de la modernidad, logrando una asimilación más profunda que la represión directa de siglos anteriores. Aquí, la ideología de la "conectividad universal" oculta la realidad de una exclusión selectiva que prioriza las lenguas de mercado sobre el patrimonio cultural inmaterial de los pueblos (Althusser, 1970, p. 28).

Podemos entender por desplazamiento lingüístico inducido un proceso sociocultural complejo donde la presión sistémica de los aparatos tecnológicos altera la ecología del habla en favor de lenguas hegemónicas. Este desplazamiento no es una evolución natural de la comunicación, sino una consecuencia provocada por el diseño de herramientas digitales que omiten las gramáticas indígenas, forzando al sujeto a una migración semántica para evitar la muerte civil. Así, lo "inducido" refiere a la creación deliberada de condiciones económicas y educativas que vuelven inviable la transmisión intergeneracional de la lengua originaria ante la omnipotencia del algoritmo (Althusser, 1970, p. 32).

La eficacia de este aparato reside en su capacidad para naturalizar la jerarquía lingüística dentro del ecosistema digital, donde el usuario acepta el sesgo como una limitación técnica y no como una decisión política. Al considerar la IA como un AIE, desvelamos que la

extinción del indigenismo es una pieza clave en la reproducción de las condiciones de producción del capitalismo cognitivo globalizado. El sujeto interpelado por la máquina termina por reconocerse únicamente en la lengua del dominador, internalizando la obsolescencia de su propio patrimonio inmaterial frente a la potencia del código binario (Althusser, 1970, p. 30).

Asimismo, el desplazamiento lingüístico se ve reforzado por el carácter privado de los nuevos AIE, donde las empresas tecnológicas asumen funciones históricas del Estado en la educación y la comunicación social. Al ser corporaciones quienes dictan las reglas del procesamiento del lenguaje, el sesgo algorítmico se convierte en una ley de mercado que castiga la diversidad lingüística bajo el criterio de la falta de rentabilidad. El hablante queda atrapado en un ciclo de desvalorización donde su lengua materna es vista como un obstáculo para el acceso a servicios básicos que ahora son mediados por sistemas inteligentes (Althusser, 1970, p. 35).

Este fenómeno se manifiesta con mayor agresividad en el ámbito educativo, donde la introducción de herramientas de IA sin adaptaciones lingüísticas actúa como un catalizador del abandono de la lengua indígena. Los jóvenes hablantes, al no encontrar sus referentes culturales en el entorno digital que consumen a diario, sufren una crisis de identidad que facilita la transición hacia el español o el inglés. El AIE digital cumple así su función de homogeneizar a la población para que sea más dócil ante las lógicas de consumo y vigilancia que estructuran la sociedad de la información contemporánea (Althusser, 1970, p. 38).

Finalmente, el desplazamiento inducido debe ser visibilizado como una táctica de alienación que separa al individuo de su historia y su territorio a través de la interfaz de usuario. La IA no solo procesa datos, sino que procesa realidades, y al dejar fuera a las 68 lenguas indígenas de México, está diseñando un futuro donde el indigenismo es solo un recuerdo museificado. La lucha por la lengua es, en esencia, una lucha contra la interpelación ideológica de un aparato que busca reducir la inmensa riqueza del pensamiento humano a una serie de algoritmos predecibles y mercantilizables (Althusser, 1970, p. 40).

Gubernamentalidad Algorítmica y el Panóptico Lingüístico (Extensión)

La gubernamentalidad algorítmica se define como un modo de ejercicio del poder que no se dirige directamente a los individuos, sino a la gestión de sus datos y comportamientos probables a través de sistemas computacionales. En el contexto de las lenguas indígenas, esta forma de poder actúa mediante una "economía de la visibilidad" donde solo lo que es

cuantificable es considerado real para el aparato administrativo del Estado y las plataformas globales. Así, la gubernamentalidad algorítmica no prohíbe el habla indígena, sino que la despoja de su operatividad en la esfera pública digital, convirtiéndola en una lengua "muerta" para el sistema de interacción humana-computadora (Foucault, 1975, p. 195).

Este fenómeno se manifiesta a través del sesgo algorítmico, entendido como la inclinación sistémica y prejuiciosa incorporada en el diseño de los modelos de lenguaje que privilegia datos de lenguas hegemónicas. El sesgo no es una falla técnica menor, sino una decisión arquitectónica que refleja las jerarquías de poder del mundo analógico proyectadas al silicio. Para el algoritmo, la ausencia de una masa crítica de datos de lenguas indígenas, no es una pérdida cultural, sino una "ineficiencia" que debe ser filtrada; de este modo, el sesgo opera como un censor automatizado que valida la existencia de unas lenguas mientras borra activamente la ontología de las otras (Foucault, 1975, p. 198).

El panóptico digital, potenciado por este sesgo, ejerce una disciplina que Foucault describiría como "la mirada que vigila sin ser vista". El hablante indígena, al enfrentarse a una interfaz que no reconoce su sintaxis ni su léxico, internaliza la mirada del dominador y asume que su lengua es obsoleta para el futuro tecnológico. Esta presión invisible es la que induce el desplazamiento, pues el sujeto, buscando evitar la marginación digital, se autodisciplina adoptando el español o el inglés como única vía de reconocimiento. El sesgo algorítmico se convierte así en un instrumento de biopolítica que decide qué identidades lingüísticas son aptas para la supervivencia en la modernidad algorítmica (Foucault, 1975, p. 202).

Más allá de la exclusión estadística, la gubernamentalidad algorítmica fragmenta el Patrimonio Cultural Inmaterial al reducirlo a meras unidades de información despojadas de territorio y comunidad. Al procesar las lenguas indígenas bajo una lógica de estandarización, el poder algorítmico borra las variantes dialectales y las sutilezas culturales que no alcanzan el umbral de representatividad. El resultado es una "lengua en extinción" en el código, que pierde su vitalidad social al ser capturada por un sistema que solo la valora como insumo arqueológico y no como herramienta viva de comunicación. La soberanía lingüística es, por tanto, el principal enemigo de este panóptico que busca la transparencia total del sujeto bajo un solo código semántico (Foucault, 1975, p. 205).

El sesgo algorítmico funciona como un mecanismo de normalización que sanciona la diferencia. Aquellas lenguas que resisten la codificación binaria son etiquetadas como "ruido" o "error", justificando su exclusión de los servicios esenciales que hoy dependen de la IA,

desde la educación hasta la justicia. Esta discriminación automatizada es la forma más sofisticada de desplazamiento inducido, ya que se presenta como una "neutralidad técnica" que oculta su carácter extractivista y excluyente.

Epistemicidio y la Línea Abismal de la Inteligencia Artificial (Extensión)

El desplazamiento lingüístico inducido por la Inteligencia Artificial no debe interpretarse como una transición técnica, sino como la ejecución de un epistemicidio que anula al indigenismo como forma de existencia. Boaventura de Sousa Santos (2014) sostiene que la modernidad traza una "línea abismal" que divide el conocimiento humano: de un lado se encuentra el saber científico y tecnológico, y del otro, los saberes producidos como inexistentes. Cuando una lengua indígena es excluida del código binario, se decreta que la cosmovisión que esta porta carece de valor epistémico, condenando a los pueblos originarios a una invisibilidad radical donde su pensamiento ya no es consultable ni reproducible en la esfera digital (p. 45).

La extinción de la lengua implica la muerte de la cosmovisión, pues el lenguaje no es un mero vehículo de información, sino el tejido mismo que constituye la realidad. En las lenguas indígenas de México, la relación con la tierra, el tiempo y la comunidad se estructura a través de categorías lingüísticas que no tienen equivalencia en las lenguas hegemónicas. Al ser desplazadas por el algoritmo, estas formas de entender el cosmos desaparecen, dejando al sujeto indígena en una orfandad semántica donde debe interpretar su propia existencia a través de los conceptos del colonizador digital. La IA, al no "entender" estas estructuras, produce activamente la inexistencia de esos mundos posibles (Santos, 2014, p. 48).

Este proceso de extinción inducida por el sesgo tecnológico configura lo que Santos define como una "sociología de las ausencias", donde la riqueza lingüística es transformada en vacío. La IA opera bajo una razón metonímica que toma la parte (el español o el inglés) por el todo, borrando las alternativas de vida que el indigenismo ofrece frente a la crisis de la modernidad. Sin la lengua, el indigenismo se reduce a un folclorismo vacío de poder político y espiritual, pues la capacidad de nombrar el mundo de manera distinta es el primer acto de soberanía que la línea abismal del silicio intenta castrar (Santos, 2014, p. 52).

La pérdida de estas cosmovisiones representa un empobrecimiento de la biodiversidad cognitiva de la humanidad. Cada lengua que se extingue bajo la presión del aparato tecnológico es una ventana que se cierra a formas de conocimiento que podrían ser la solución a dilemas contemporáneos. El desplazamiento inducido no es, por tanto, una

consecuencia del progreso, sino una técnica de exclusión abismal que asegura que solo sobrevivan aquellas epistemologías que son funcionales al capital transnacional. El indigenismo, despojado de su lengua, queda recluido en el pasado, imposibilitado de habitar el futuro que la IA está construyendo para una sola humanidad uniforme (Santos, 2014, p. 55).

Finalmente, la lucha contra la extinción lingüística debe ser entendida como una exigencia de justicia cognitiva global. Romper la línea abismal en la IA significa reconocer que el Patrimonio Cultural Inmaterial es una fuente de saber vivo que debe permear la arquitectura misma de los algoritmos. Si el diseño tecnológico persiste en ignorar la pluralidad del pensamiento indígena, la IA no será más que el monumento final al epistemicidio colonial. La verdadera innovación no reside en la velocidad del procesamiento, sino en la capacidad de crear una "ecología de saberes" donde todas las lenguas y cosmovisiones tengan derecho a existir y coexistir en el horizonte de la técnica (Santos, 2014, p. 60).

CONCLUSIÓN

La investigación permite concluir que la extinción de las lenguas indígenas en México no es un fenómeno inercial, sino un proceso sociocultural complejo de desplazamiento inducido por una arquitectura tecnológica que opera como el nuevo gran Aparato Ideológico de Estado. Al contrastar la riqueza lingüística documentada por el INEGI (2020) con la lógica del silencio, se hace evidente que la Inteligencia Artificial actúa como un mecanismo de normalización biopolítica que invisibiliza cualquier ontología ajena al mercado global. Bajo la lente de Althusser (1970), la IA no es una herramienta neutra, sino un dispositivo de interpelación que fuerza al sujeto indígena a la auto-aniquilación semántica como requisito para su legibilidad y supervivencia en la modernidad digital (p. 28).

Este desplazamiento se agrava por un sesgo algorítmico que funciona como una "línea abismal" contemporánea, donde el conocimiento de los pueblos originarios es producido activamente como inexistente. Siguiendo a Sousa Santos (2014), hemos demostrado que la exclusión de estas lenguas del código binario constituye un epistemicidio, pues al mutilar la lengua se cercena la cosmovisión y la capacidad política de los pueblos para nombrar su propia realidad. La IA, al ignorar la densidad semántica del P'urhépecha, el Náhuatl o el Mixteco, no solo comete un error técnico, sino que ejecuta una forma de violencia epistémica que reduce el Patrimonio Cultural Inmaterial a un residuo arqueológico sin funcionalidad en el futuro (p. 52).

Desde la perspectiva de la gubernamentalidad de Foucault (1975), el algoritmo se erige como un panóptico que disciplina mediante la omisión. El silencio de las lenguas indígenas en las interfaces tecnológicas es la prueba fehaciente de un régimen de verdad que sanciona la diferencia y premia la estandarización. No se trata simplemente de una brecha digital de acceso, sino de una brecha ontológica de reconocimiento, donde el derecho a la lengua — protegido por marcos jurídicos internacionales— es vulnerado sistemáticamente por un determinismo técnico que prioriza la rentabilidad del dato sobre la dignidad de la cultura (p. 205).

Por consiguiente, la "evidente extinción" que muestran las estadísticas debe ser leída como un llamado a la acción jurídica y ética. Es imperativo transitar de una visión de "preservación" (que a menudo implica museificar la lengua en archivos inertes) hacia una estrategia de soberanía tecnológica indígena. La justicia lingüística en el siglo XXI exige que el diseño de los modelos de lenguaje rompa con el extractivismo de datos y permita una verdadera "ecología de saberes", donde las lenguas originarias no sean insumos para el entrenamiento de máquinas ajenas, sino plataformas vivas para la autonomía del pensamiento propio (Santos, 2014, p. 58).

Asimismo, el papel del Estado y de los organismos internacionales debe ser el de reguladores críticos de este Aparato Ideológico Digital. La protección del Patrimonio Cultural Inmaterial debe trascender el discurso romántico para insertarse en las políticas de gobernanza de la IA, obligando a una transparencia algorítmica que rinda cuentas sobre los sesgos de exclusión lingüística. Si el indigenismo ha de sobrevivir a la era de la IA, debe hacerlo como una fuerza que de-colonice el código, exigiendo que la tecnología sea un espejo de la pluralidad humana y no un martillo que aplaste la diversidad bajo el peso de la uniformidad binaria.

En última instancia, esta crítica jurídico-filosófica sostiene que la defensa de las lenguas indígenas es la defensa de la humanidad misma frente al empobrecimiento cognitivo de la técnica. Cada lengua desplazada es una posibilidad de futuro que se clausura para todos. La verdadera innovación de-colonial no reside en algoritmos más veloces, sino en sistemas lo suficientemente humanos para reconocer que el silencio de un pueblo es la derrota de la civilización.

REFERENCIAS

- Althusser, L. (2005). *Ideología y aparatos ideológicos de Estado*. Nueva Visión. (Obra original publicada en 1970).
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for indigenous data governance. *Data Science Journal*, 19(1), 1–12. <https://doi.org/10.5334/dsj-2020-043>
- Chomsky, N. (1999). *Aspectos de la teoría de la sintaxis*. Gedisa. (Obra original publicada en 1965).
- De Sousa Santos, B. (2009). *Una epistemología del sur: la reinención del saber y la emancipación social*. Siglo XXI.
- De Sousa Santos, B. (2014). *Epistemologías del Sur*. Akal.
- Duarte, M. E. (2017). *Network sovereignty: Building damages in Indian country*. University of Washington Press.
- Foucault, M. (2009). *Vigilar y castigar: Nacimiento de la prisión*. Siglo XXI. (Obra original publicada en 1975).
- Foucault, M. (2010). *Las palabras y las cosas: una arqueología de las ciencias humanas*. Siglo XXI. (Obra original publicada en 1966).
- Instituto Nacional de Estadística y Geografía [INEGI]. (2020). *Censo de Población y Vivienda 2020: Cuestionario ampliado y tabulados definitivos*. <https://www.inegi.org.mx/programas/ccpv/2020/>
- Kukutai, T., & Taylor, J. (Eds.). (2016). *Indigenous data sovereignty: Toward an agenda*. ANU Press. <https://doi.org/10.22459/CAEPR38.11.2016>

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing.

Saussure, F. (1945). *Curso de lingüística general*. Losada. (Obra original publicada en 1916).

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

© Los autores. Este artículo se publica en Prisma ODS bajo la Licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Esto permite el uso, distribución y reproducción en cualquier medio, incluidos fines comerciales, siempre que se otorgue la atribución adecuada a los autores y a la fuente original.



doi: <https://doi.org/10.65011/prismaods.v5.i2.247>

Cómo citar este artículo (APA 7ª edición):

Rodríguez González, O. . (2026). El Sesgo del Algoritmo ante el Desplazamiento Lingüístico: Una Crítica Jurídico-Filosófica sobre la Extinción del Indigenismo en la Era de la IA. *Prisma ODS: Revista Multidisciplinaria Sobre Desarrollo Sostenible*, 5(2), 724-741. <https://doi.org/10.65011/prismaods.v5.i2.247>