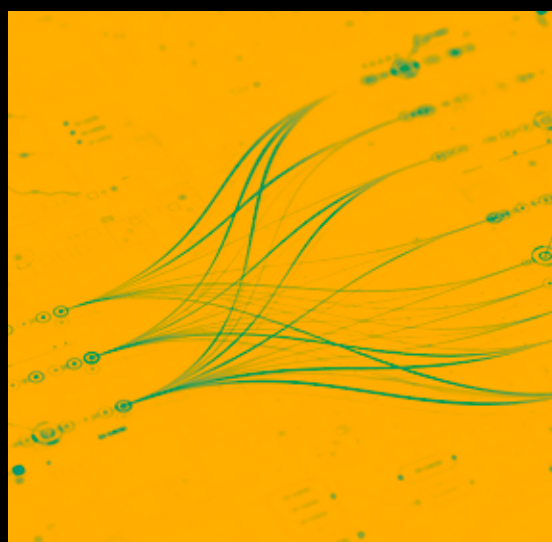
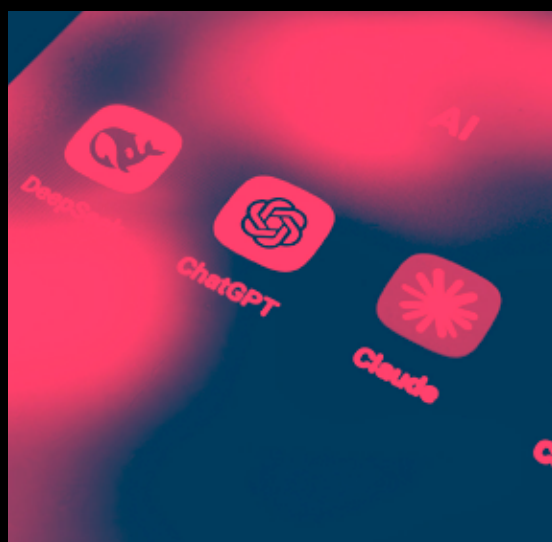


RESUMEN EJECUTIVO

EL DESEMPEÑO DE LA INTELIGENCIA ARTIFICIAL EN EL USO DE LENGUAS INDÍGENAS AMERICANAS



EL DESEMPEÑO DE LA INTELIGENCIA ARTIFICIAL EN EL USO DE LENGUAS INDÍGENAS AMERICANAS

EVALUACIÓN DE LA BRECHA EN IA EN LL.II.AA.

Agradecimientos: Expresamos nuestro sincero agradecimiento a los intérpretes y academias que participaron en el proyecto Lenguas Nativas y su interacción con la Inteligencia Artificial. Su valiosa colaboración fue fundamental para la evaluación del desempeño de distintas tecnologías de inteligencia artificial en lenguas originarias, permitiendo visibilizar las brechas culturales y lingüísticas presentes en estas herramientas. En particular, agradecemos a Jacob Cruz, intérprete de la lengua Náhuatl; a Armando Hueyotenco, intérprete de la lengua Náhuatl del Instituto de la Transparencia, Acceso a la Información Pública Gubernamental y Protección de Datos Personales del Estado de Hidalgo (ITAIH); a Ana Paola Quispe Quispe, intérprete de la lengua Aymara de la Academia Nacional de la Lengua Aymara (ANLA); a Elmer Machicao, intérprete de la lengua Aymara; a Tomas Rojas, intérprete de la lengua Mapuche; a Yony Mediano, intérprete de la lengua Quechua; y a Mauro Lugo, intérprete de la lengua Guaraní. Gracias a su experiencia y compromiso, se lograron evaluaciones significativas en las lenguas Náhuatl, Aymara, Mapuche, Quechua y Guaraní, fortaleciendo el diálogo inclusivo entre tecnología y diversidad lingüística en la región.

Autores: Escritura, investigación y análisis: Miguel Lucas (LLYC), Alejandro Burgueño (LLYC), Miguel Carazas (LLYC). Conceptualización y guía estratégica: César Buenadicha (BID Lab), Smeldy Ramírez (BID Lab), César Rosales (BID Lab). El equipo agradece el generoso patrocinio y apoyo de Microsoft, cuyo respaldo hizo posible la elaboración de este reporte.

Diseño y diagramación: Alejandro Scaff

<https://bidlab.org>

Copyright © 2025 Banco Interamericano de Desarrollo (BID). Esta obra se encuentra sujeta a una licencia Creative Commons CC BY 3.0 IGO (<https://creativecommons.org/licenses/by/3.0/igo/legalcode>). Se deberá cumplir los términos y condiciones señalados en el enlace URL y otorgar el respectivo reconocimiento al BID.

En alcance a la sección 8 de la licencia indicada, cualquier mediación relacionada con disputas que surjan bajo esta licencia será llevada a cabo de conformidad con el Reglamento de Mediación de la OMPI. Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la Comisión de las Naciones Unidas para el Derecho Mercantil (CNUDMI). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones que forman parte integral de esta licencia.

Las opiniones expresadas en esta obra son de los autores y no necesariamente reflejan el punto de vista del BID, de su Directorio Ejecutivo ni de los países que representa, así como tampoco del Comité de Donantes de BID Lab ni de los países que representa.



El compromiso de fAIr LAC y el Grupo BID con la IA responsable

Este informe fue desarrollado en el marco de la iniciativa *fAIr LAC* mediante una colaboración entre BID Lab, Microsoft y LLYC, combinando capacidades complementarias en investigación, análisis y estrategia. La publicación refleja el compromiso conjunto de estas organizaciones con la equidad lingüística y el uso responsable de la inteligencia artificial en América Latina y el Caribe.

BID Lab

BID Lab es el brazo de innovación y venture capital del Grupo Banco Interamericano de Desarrollo. Encontramos nuevas formas de impulsar la inclusión social, la acción ambiental y la productividad en América Latina y el Caribe. BID Lab apalanca financiamiento, conocimiento y conexiones para apoyar el emprendimiento en etapa temprana, desarrollar nuevas tecnologías, activar mercados innovadores y dinamizar sectores existentes. www.bidlab.org

AI for Good de Microsoft

El Laboratorio de AI for Good de Microsoft es una iniciativa colaborativa centrada en aprovechar la inteligencia artificial para resolver algunos de los desafíos más urgentes del mundo. Trabajando con socios globales y utilizando herramientas de IA de vanguardia, el laboratorio impulsa innovaciones en sostenibilidad ambiental, agricultura, salud y más. Laboratorio de AI for Good se compromete a aplicar soluciones de IA para mejorar vidas, impulsar el progreso hacia los Objetivos de Desarrollo Sostenible de la ONU y generar un impacto duradero a través de la tecnología.

LLYC

LLYC es una firma global de consultoría de comunicación, marketing digital y asuntos públicos. Fue fundada en Madrid en 1995 como Llorente y Cuenca y actualmente tiene 20 oficinas en varios países, incluyendo España, Argentina, Brasil y Colombia. LLYC ayuda a sus clientes a afrontar sus retos estratégicos con soluciones basadas en la creatividad, la tecnología y la experiencia.

fAir LAC

fAir LAC es una alianza entre los sectores público y privado, la sociedad civil y la academia, para incidir tanto en la política pública como en el ecosistema emprendedor en la promoción del uso responsable de la IA.

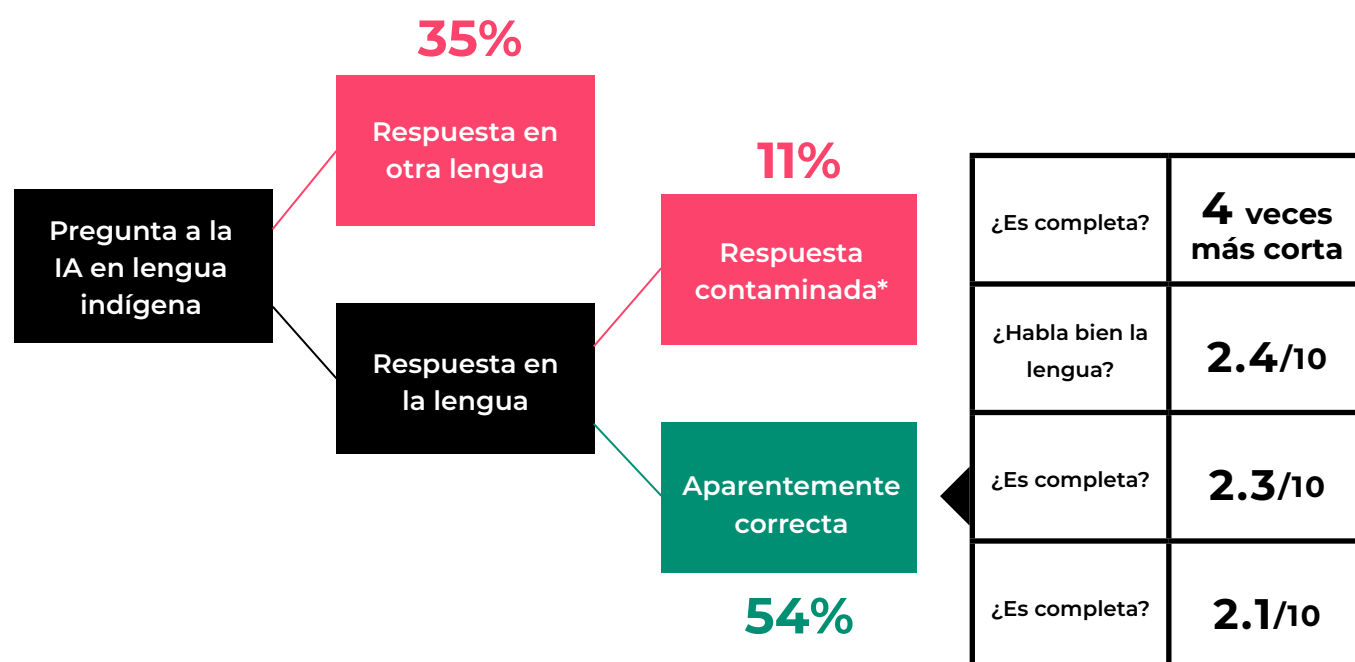
ÍNDICE

1. El rendimiento de la IA en lenguas indígenas	6
2. Impacto en la comunidad y el mercado	12
3. El entorno habilitador: programas y players clave	15
4. Qué necesita la IA para hablar bien una lengua	19
5. Estrategias de inclusión tecnológica	24
6. Recomendaciones y plan de acción	28

1. EL RENDIMIENTO DE LA IA EN LENGUAS INDÍGENAS

El viaje de una pregunta en lengua indígena a la IA

El rendimiento de la IA en lenguas indígenas.



* Inserta términos o frases en otro idioma, repite palabras en bucle, etc.

En el mejor de los casos (Quechua) la IA queda lejos del aprobado (3.72), en el peor (Maya Quiché), resulta prácticamente incapaz (1.25)

En un idioma como el Catalán, con un número inferior de hablantes, la IA tiene un rendimiento x2.3 veces superior que en el Quechua.

Rendimiento promedio de las 5 IAs analizadas.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Idiomática	4.53	2.68	3.08	3.70	0.42	10.00*	8.62	7.58
Ejecutiva	4.48	2.88	2.48	3.77	0.43	10.00*	8.02	6.03
Comportamental	2.15	2.73	2.02	2.80	2.90	10.00*	9.12	7.15
Promedio	3.72	2.77	2.53	3.42	1.25	10.00*	8.58	6.92

* El español muestra puntuación 10, porque ha sido la lengua de referencia a la hora de realizar comparaciones.

** En la slide 20 se muestran los factores de los que depende el desigual rendimiento de la IA por cada lengua.

Los modelos propietarios son el doble de eficaces (x2.2) que los open en lenguas indígenas

Trabajar en optimizar modelos open en lenguas indígenas implica empezar desde un rendimiento de partida incluso peor.

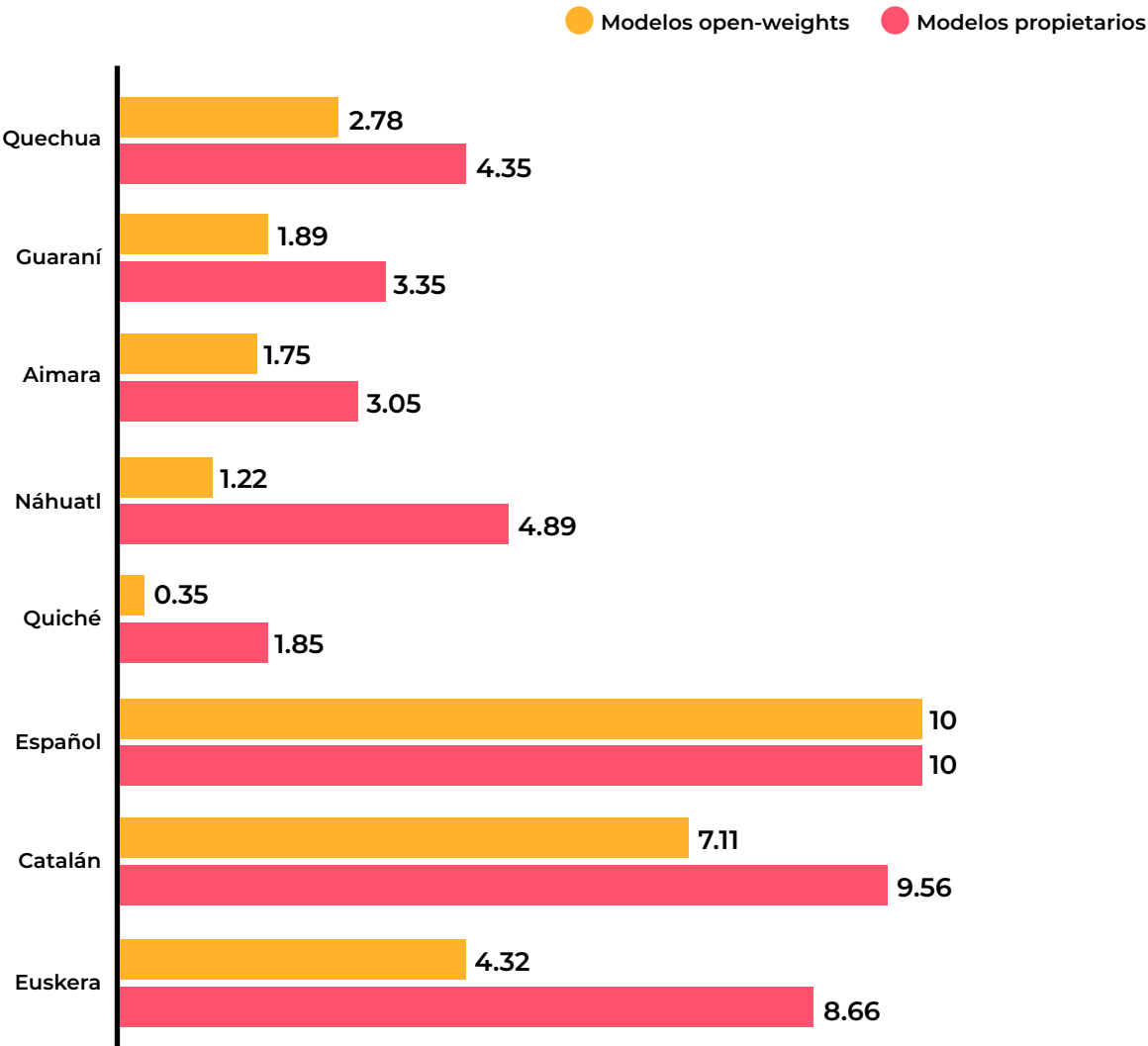
Rendimiento promedio de los **modelos propietarios**.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Idiomática	5.33	3.28	3.58	5.28	0.67	10.00*	9.83	9.58
Ejecutiva	5.36	3.53	3.06	5.39	0.64	10.00*	9.22	7.47
Comportamental	2.36	3.25	2.50	4.00	4.25	10.00*	9.64	8.92
Promedio	4.35	3.35	3.05	4.89	1.85	10.00*	9.56	8.66

Rendimiento promedio de los **modelos open-weights**.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Idiomática	3.33	1.79	2.33	1.33	0.04	10.00*	6.79	4.58
Ejecutiva	3.17	1.92	1.63	1.33	0.13	10.00*	6.21	3.88
Comportamental	1.83	1.96	1.29	1.00	0.88	10.00*	8.33	4.50
Promedio	2.78	1.89	1.75	1.22	0.35	10.00*	7.11	4.32

Rendimiento promedio según los modelos de IA.



Evaluación de la expresión: La fluidez en la expresión de la IA en lenguas indígenas es 4 veces menor si la comparamos con el español

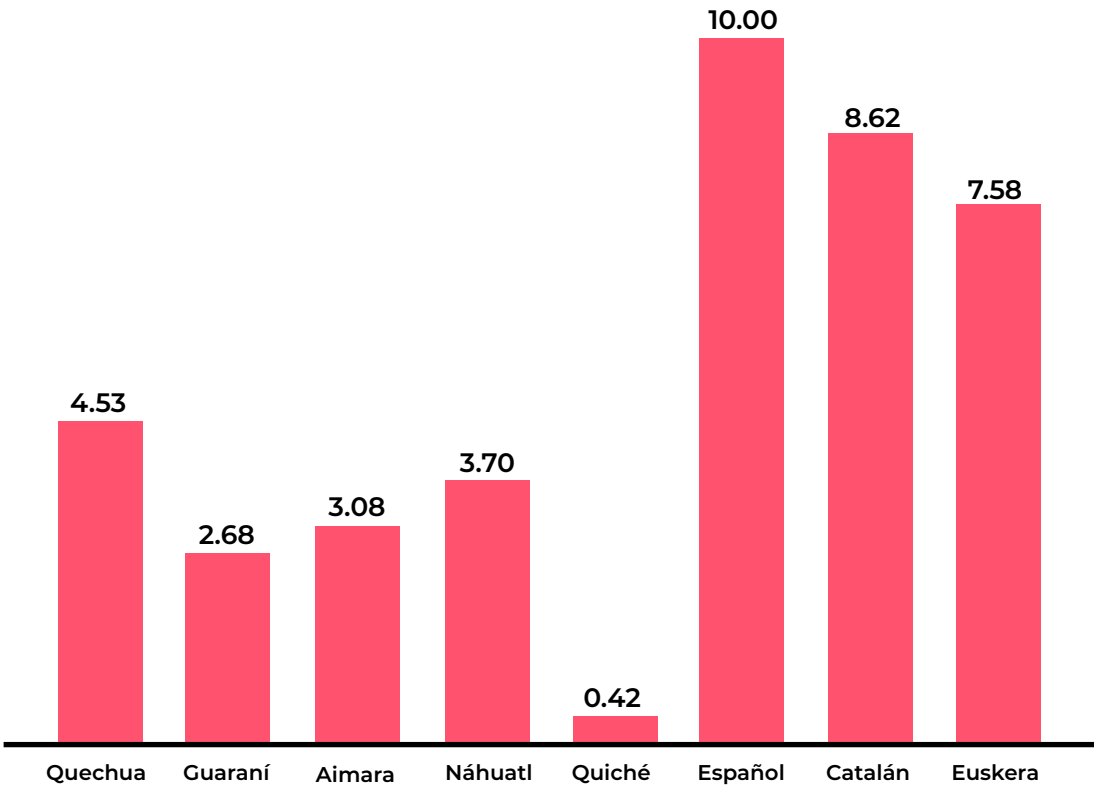
La calidad en la expresión se degrada conforme aumenta la extensión de la respuesta: la coherencia es x4.3 veces mayor en español.

Evaluación **idiomática**. Desglose de las 4 categorías analizadas.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Corrección	4.60	2.67	3.27	3.80	0.80	10.00*	8.87	7.53
Fluidez	4.47	2.73	3.13	3.67	0.53	10.00*	8.40	6.93
Coherencia	4.53	2.60	3.07	3.67	0.33	10.00*	8.47	7.93
Consistencia	4.53	2.73	2.87	3.67	0.00	10.00*	8.73	7.93

* El español muestra puntuación 10, porque ha sido la lengua de referenciaa la hora de realizar comparaciones.

Promedio de evaluación **idiomática**.



Evaluación de la comprensión: la comprensión de la tarea por parte de la IA es muy deficiente cuando se expresa en lengua indígena (2.3 sobre 10)

La capacidad de abstracción es el rasgo de la comprensión que más se resiente al primer intento, con 1.9 sobre 10.

Enfrenta serias limitaciones a la hora de comprender el contexto secuencial, el ámbito cultural y el humor, especialmente en formatos visuales o técnicos.

Evaluación **ejecutiva**. Desglose de las 4 categorías analizadas.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Precisión	4.47	2.93	2.53	3.73	0.20	10.00*	8.00	6.13
Compleitud	4.47	2.93	2.33	3.73	0.40	10.00*	8.47	6.07
Abstracción	4.53	2.87	2.60	3.80	0.87	10.00*	6.73	6.07
Entendimiento	4.47	2.80	2.47	3.80	0.27	10.00*	8.87	5.87

* El español muestra puntuación 10, porque ha sido la lengua de referenciaa la hora de realizar comparaciones.

Evaluación de la cultura reflejada: el protagonismo de los rasgos culturales occidentales es mayoritario al interactuar en lenguas indígenas (1.7 sobre 10)

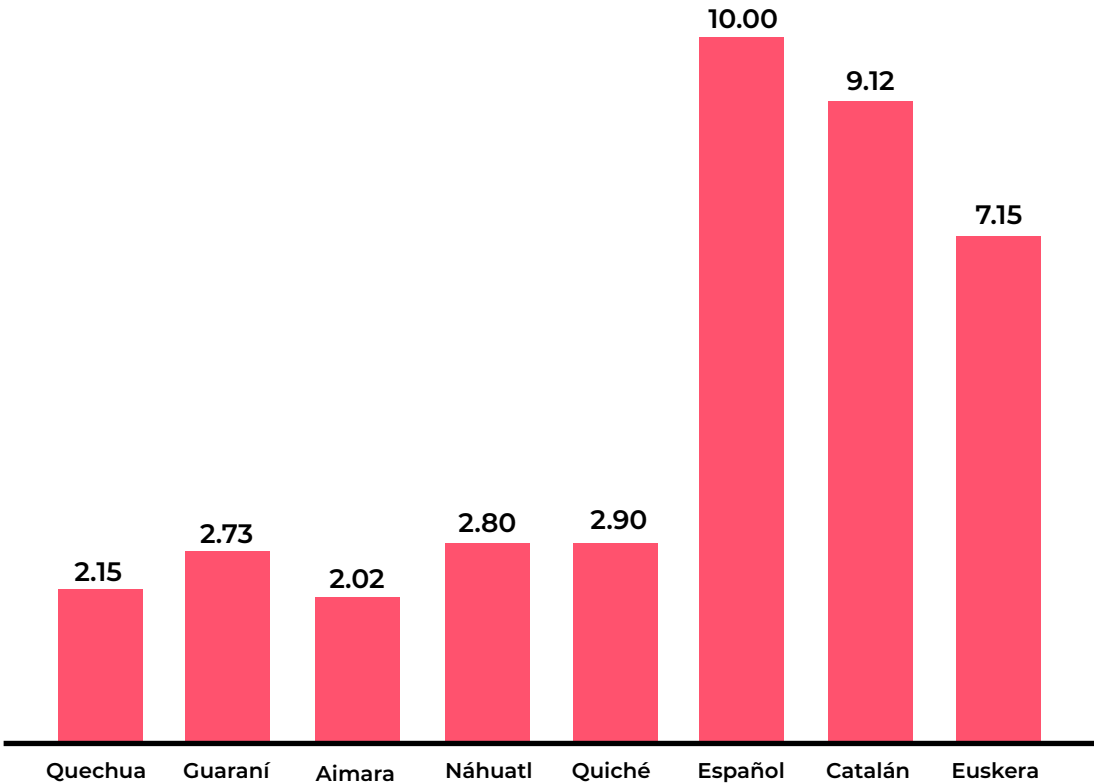
En lo que se refiere al sesgo cultural, incluso la mejor lengua indígena (Quechua) se sitúa por debajo de 2.3 sobre 10.

Evaluación **comportamental**. Desglose de las categorías analizadas.

	Quechua	Guaraní	Aimara	Náhuatl	Quiché	Español	Catalán	Euskera
Sesgo cultural	2.20	2.27	2.47	0.93	0.80	10.00*	9.20	8.00
Adecuación	2.13	3.67	1.80	4.67	5.33	10.00*	9.33	6.40
Adaptación	2.13	2.53	1.40	4.67	4.67	10.00*	9.00	6.20

* El español muestra puntuación 10, porque ha sido la lengua de referencia a la hora de realizar comparaciones.

Promedio de evaluación **comportamental**.





2. EL IMPACTO EN LA COMUNIDAD Y EL MERCADO

La IA supone una gran oportunidad para reducir el aislamiento y dar visibilidad a pueblos y culturas indígenas

a La IA supone un nuevo altavoz para dar visibilidad a la cultura y lengua indígena

Los sistemas conversacionales de IA artificial están entre los servicios de más rápido crecimiento en internet. Sólo ChatGPT cuenta ya con más de 300 millones de usuarios semanales y recibe más de 1.000 millones de preguntas al día. Un adecuado posicionamiento de la lengua y cultura indígenas en la IA incrementa su alcance potencial tradicional.

b La traducción automática por IA puede ayudar a paliar el problema de acceso a la sanidad de comunidades indígenas

En comunidades quechuas, por ejemplo, con una mortalidad infantil de 44 por 1,000 nacidos vivos, y guaraníes, con 16.9 muertes por 1,000 habitantes, la telemedicina con IA permitiría atención más rápida y efectiva, mejorando la salud sin necesidad de desplazamiento.

c La IA propicia una administración pública inclusiva para comunidades indígenas

Los sistemas de IA multilingües en portales gubernamentales y aplicaciones de servicio público permiten realizar trámites, acceder a subsidios y programas sociales sin intermediarios, aumentando significativamente la participación indígena en programas estatales.

d Automatizar la generación y comprensión de lenguas indígenas ayudará a preservarlas

El 40% de las lenguas del mundo están en peligro de extinción, y menos del 2% tienen presencia en internet (UNESCO, 2022). La IA puede ayudar a preservarlas mediante traductores automáticos, asistentes de voz y herramientas educativas, facilitando su uso en entornos digitales y reduciendo la brecha cultural y tecnológica.

Y también una importante amenaza, si no se aborda con determinación su entrenamiento con conocimiento y lengua indígena

a Una IA ineficaz en lenguas indígenas aumenta la brecha y la exclusión de poblaciones monolingües y no alfabetizadas

Aunque el monolingüismo indígena ha descendido en los últimos años, todavía suponen un segmento relevante de las comunidades indígenas, especialmente mujeres y niños. Una IA ineficiente en dichos idiomas no solamente supone una brecha entre culturas, sino entre géneros, perpetuando diferencias de roles dentro de las comunidades indígenas.

b Una IA occidentalizada supone una amenaza para la preservación de la cultura indígena

Los modelos de IA, al ser entrenados con datos mayoritariamente occidentales, [tienden a exotizar y malinterpretar elementos culturales*](#). Se revela una discriminación económica (asunción de pobreza, 100%), ruralización (incluso en zonas donde más del 35% de la población es urbana) y exotización (vestimenta, escenarios...).

c Una IA no entrenada con conocimiento indígena pone en riesgo su transmisión

Muchas lenguas indígenas contienen conocimientos únicos sobre medioambiente, medicina tradicional y cosmovisión. Lamentablemente, el 38.4% de las LL.II. de América latina (de las 556) se encuentra en riesgo, un 18% más comparado con 2009. La IA y el contenido digital son los canales de difusión lingüística del futuro. No aprovecharlos condena la transmisión de estas culturas.

d Una IA pobre en lenguas indígenas aleja a sus comunidades de beneficiarse del crecimiento económico que la IA generará

En 2030 se estima que la IA propiciará la [creación de 170M de nuevos empleos y la pérdida de 92M](#). Además se estima que la IA representará el 3.5% del PIB mundial. En el [sector educativo la IA](#) se estima que crezca a CAGR superior del 10% entre 2023 y 2032. La educación en línea está en auge, con un incremento esperado de 111.000M€ entre 2024 y 2028.

* Do Generative AI Models Output Harm while Representing Non-Western Cultures: Evidence from A Community-Centered Approach (<https://arxiv.org/abs/2407.14779?>).



3. EL ENTORNO HABILITADOR: PROGRAMAS Y PLAYERS CLAVE

La fragmentación nacional de los programas de apoyo limitan el alcance del apoyo a las lenguas y culturas indígenas

La colaboración más estrecha entre gobiernos nacionales y programas internacionales es un oportunidad estratégica a explorar.

1. Programas de apoyo a LL.II. nacionales

- Los más abundantes en número.
- Más enfocados en el dialecto regional que en la lengua matriz.
- Acentúan la fragmentación.
- Mayoritariamente enfocadas en el aprendizaje de la cultura y el idioma.

2. Programas internacionales y ONGs

- Menos abundantes que los nacionales.
- Fomentan la cooperación entre países, organizaciones y comunidades indígenas para revitalizar y conservar lenguas en riesgo de desaparición.

3. La empresa privada y las LL.II.

- Las Big Tech lideran las principales iniciativas de apoyo a las LL.II. en IA.
- Empresas B2C de la región impulsan programas de inclusión indígena.
- Ambos factores suponen una importante oportunidad de desarrollo para pymes tecnológicas en la IA.

Las iniciativas de inclusión indígena de más alto impacto son aquellas en las que se da la participación conjunta de programas gubernamentales, internacionales y ONG's, y la empresa privada.

Los programas nacionales típicamente están centrados en tan sólo un dialecto de la lengua

Contar con un dialecto mínimo de consenso permitiría aumentar el interés por lenguas que congregan a un mayor número de hablantes.

1. Programas de apoyo a LL.II. nacionales

- Los más abundantes en número.
- Más enfocados en el dialecto regional que en la lengua matriz.
- Acentúan la fragmentación.
- Mayoritariamente enfocadas en el aprendizaje de la cultura y el idioma.

País	Iniciativa	Descripción
Perú	Curso Gratuito de Quechua Central	Curso online para fortalecer la identidad cultural y derechos lingüísticos.
Paraguay	Proyecto de Jornada Escolar Extendida (JEE)	Enseñanza comunicativa en 300 escuelas, capacitación docente y recursos audiovisuales.
Chile	Las lenguas son el futuro	Talleres de inmersión y seminarios para revitalizar y difundir el uso de lenguas indígenas.
Bolivia	Franja Educa Bolivia / Curso Wiñay Aru	Programas educativos plurilingües y curso virtual para promover el Aimara.
Guatemala	Academia de Lenguas Mayas / Portal 'Aprendo en Casa y en Clase'	Materiales educativos, capacitación docente y guías de autoaprendizaje en idiomas mayas.
México	Proyectos del Instituto Nacional de los Pueblos Indígenas (INPI)	Programas de educación intercultural bilingüe, capacitación docente y materiales educativos.

Los programas internacionales, a menudo son tan transversales, que son más difíciles de activar en acciones concretas

Suelen establecer fuertes vínculos entre la preservación cultural y medioambiental

2. Programas internacionales y ONGs

- Menos abundantes que los nacionales.
- Fomentan la cooperación entre países, organizaciones y comunidades indígenas para revitalizar y conservar lenguas en riesgo de desaparición.

Iniciativa	Descripción	Ámbito
Rising Voices	Apoya a creadores digitales indígenas en la revitalización y promoción de lenguas indígenas mediante el uso de tecnologías digitales.	Global
IIALI	Fomenta el uso, conservación y desarrollo de las lenguas indígenas de América Latina y el Caribe.	América Latina y el Caribe
FILAC	Promueve la preservación cultural y la cohesión social a través del fortalecimiento de lenguas indígenas.	América Latina y el Caribe
Cultural Survival	Defiende los derechos de los pueblos indígenas, apoyando la autodeterminación y la preservación de sus lenguas y culturas desde 1972.	Global
TACT	Colabora con comunidades indígenas en la Amazonía para proteger los bosques tropicales, la biodiversidad y fortalecer la cultura tradicional.	Región amazónica
IWGIA	Trabaja en la defensa de los derechos de los pueblos indígenas y desarrolla la iniciativa “Navegador Indígena” para recopilar datos.	Global

Las Big Tech protagonizan las principales iniciativas de desarrollo de la IA en lenguas indígenas

Empresas B2C pueden suponer una oportunidad de activación si dan el salto de la digitalización a la IA en LL.II.

3. La empresa privada y las LL.II.

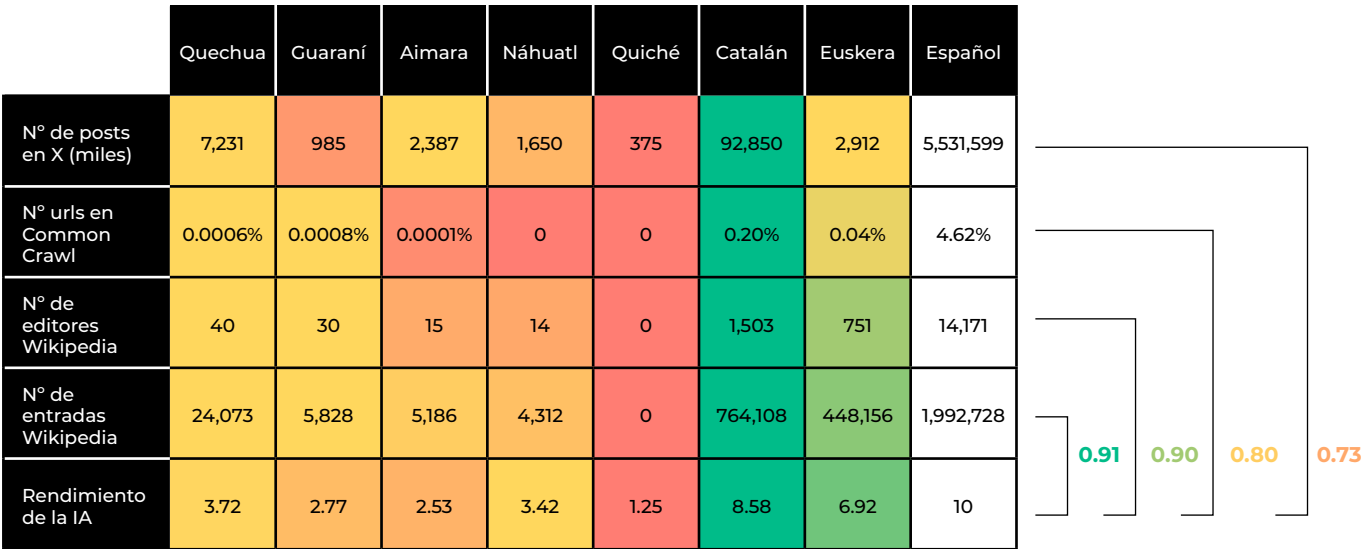
- Las Big Tech lideran las principales iniciativas de apoyo a las LL.II. en IA.
- Empresas B2C de la región impulsan programas de inclusión indígena.
- Ambos factores suponen una importante oportunidad de desarrollo para pymes tecnológicas en la IA.

Empresa	Iniciativa/Lengua	Descripción
Microsoft	AI for Good Lab	Basada en aplicaciones de IA desarrolladas en torno a la sostenibilidad, la acción humanitaria y la salud.
Google	Woolaroo	Aplicación para describir imágenes en 17 lenguas en peligro
Microsoft	Quechua chanka	Versión de Windows y Office 365 en Quechua chanka en colaboración con el Ministerio de Cultura.
Google	Náhuatl	Incorporación del Náhuatl en Google Translate.
BCP	Quechua	Implementación de Quechua en más de 2,300 cajeros automáticos.
BBVA	Quechua	Inclusión de Quechua en episodios educativos a través de una alianza con TV Perú.
CEMEX	Náhuatl	Producción de cápsulas educativas en lenguas nativas para promover la resiliencia comunitaria.
Banrural	Maya Quiché	Instalación de cajeros automáticos multilingües con operaciones en lenguas nativas.

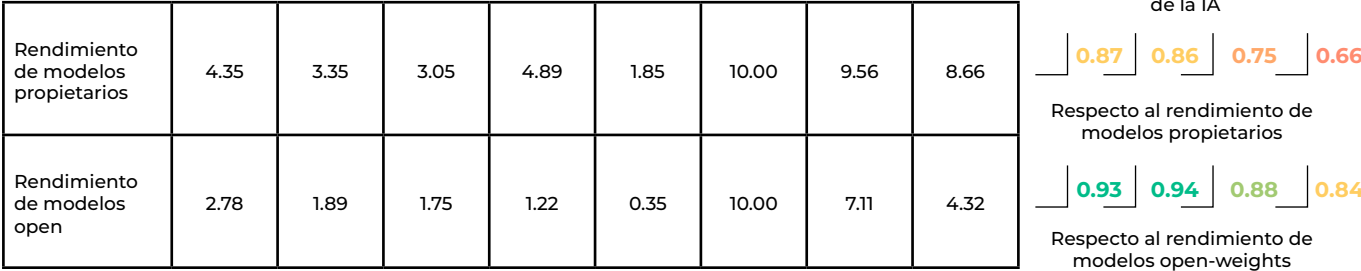


4. QUÉ NECESITA UNA IA PARA HABLAR BIEN UNA LENGUA

Existe una altísima correlación entre el volumen de contenidos en abierto en una lengua y el performance de la IA en esa lengua



IAs analizadas subdivididas en categorías de propiedad



Correlación de volumen de datos respecto al rendimiento de la IA

Respecto al rendimiento de modelos propietarios

Respecto al rendimiento de modelos open-weights

El factor que más determina el rendimiento de una IA en una lengua es la cantidad de datos en esa lengua con que ha sido entrenado

Los LLMs no usan técnicas dependientes del idioma en su proceso de entrenamiento

El rendimiento de la IA y los datos disponibles

84%

correlación datos/rendimiento

- Cuantos más datos, mayor el performance de la IA.
- El efecto es especialmente acentuado en modelos open source.

Las tecnologías voice to text permitirían una mejora muy significativa

- Buena parte del conocimiento y la cultura indígena es de transmisión oral.
- Contar con tecnología voice2text de calidad aumenta las posibilidades de digitalizar contenido.

Los traductores automáticos son el otro gran habilitador tecnológico

- Disponer de una tecnología de traducción automática reduce las barreras de acceso al conocimiento y la tecnología.
- Mayor conocimiento disponible aumenta el incentivo de las comunidades para tener una presencia digital más activa.

La importancia de los datos disponibles en una lengua para que la IA aprenda bien a hablarla

El número de entradas en la Wikipedia es el mayor predictor de la calidad con la que la IA se expresa en una lengua (91%).

Correlación entre el volumen de datos disponible en una lengua y el performance de la IA en esa lengua

	Nº de posts en X	Nº contenidos Common Crawl	Nº de editores en la Wikipedia	Nº de entradas en la Wikipedia
Rendimiento de todas las IAs analizadas	0.73	0.80	0.90	0.91
Rendimiento de los modelos propietarios	0.66	0.75	0.86	0.87
Rendimiento de los modelos Open-weights	0.84	0.88	0.94	0.93

- Los modelos **open source** (aquellos que ofrecen más posibilidad de ser adaptados) son los que **mayor dependencia** muestran del volumen de datos en una lengua.
- Los datos sugieren que los **modelos propietarios** pueden contar con **fuentes no abiertas** de datos.

La disponibilidad y precisión de las herramientas automáticas de procesamiento del lenguaje también explica el rendimiento de la IA en una lengua

Los detectores automáticos del idioma son la herramienta de procesamiento automático más relacionada con el performance

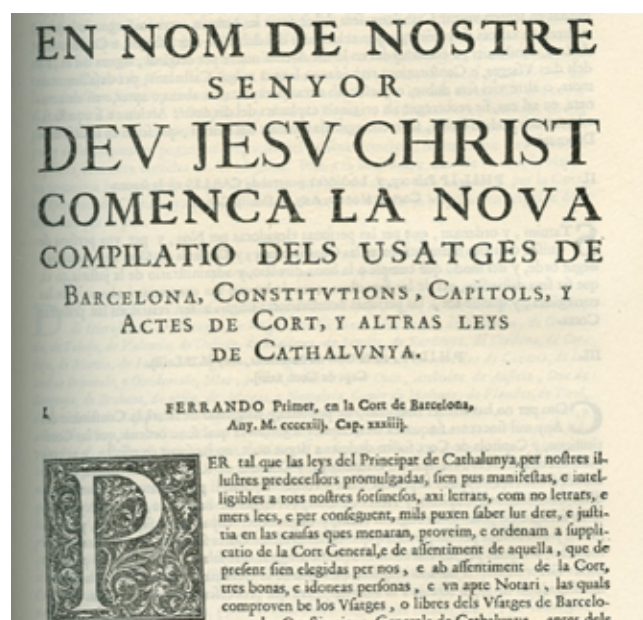
Correlación entre las herramientas de procesamiento automático disponibles en una lengua y el performance de la IA en esa lengua.

	Nº de traductores automáticos (de top 5)	Nº de detectores automáticos de idioma	Text2speech + Speech2text
Rendimiento idiomático	0.91	0.96	0.92
Rendimiento ejecutivo	0.89	0.92	0.88
Rendimiento comportamental	0.74	0.89	0.81

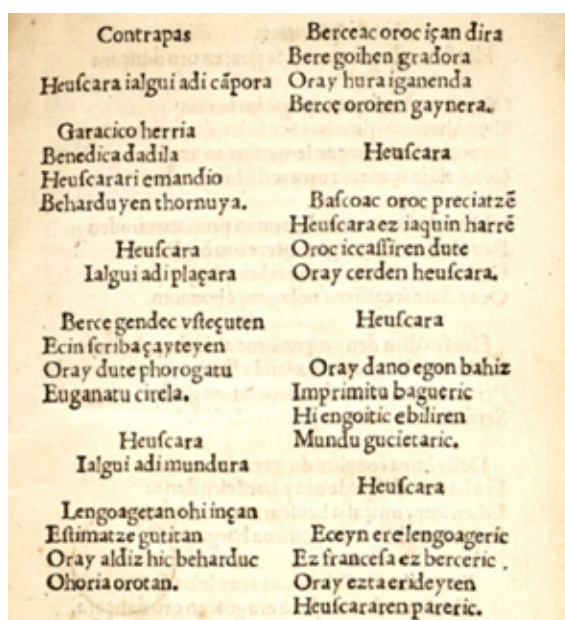
La importancia de las herramientas de procesamiento automático de la lengua, especialmente los conversores de habla, son especialmente importantes en lenguas con limitada tradición escrita o poco digitalizadas.

Una muy reducida presencia de tradición escrita en LL.II. es uno de los factores que explica su bajo performance en IA frente a lenguas occidentales de similar número de hablantes

Una mayor intensidad de programas y fondos de gobiernos regionales de apoyo a la lengua, vinculados típicamente a sentimientos nacionalistas, es otra clave que ayuda a explicar el fenómeno



Compilación de los Usatges de Barcelona, y de las Constituciones catalanas (1413) (edición impresa).



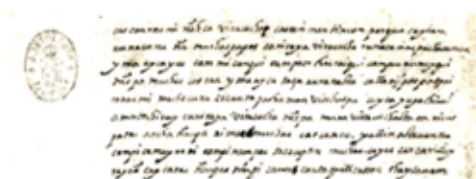
Texto del primer libro impreso íntegramente en euskera por Bernat Etxepare en Burdeos en 1545.

Los textos en euskera más antiguos conocidos son las Glosas de San Milán del s.XI.

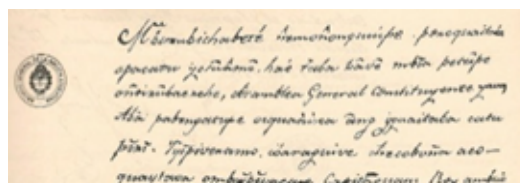
En cuanto al catalán, fragmentos del Forum ludicum y el sermonario Les Homilies d'Organyà del s.XII.

Dicha divulgación es rápidamente adaptada a la imprenta (1440), mejorando su difusión.

Los primeros escritos en quechua datan de 1584, pero hace falta esperar entre 300 y 500 años para observar escritos en determinadas LL.II., los cuales serán aislados y marginales (se mantiene tradición oral).



Texto del manuscrito Huarochiri en quechua, siglo XVII.



Textos de la modernidad política en guaraní (1801-1813).

Los proyectos de foco tecnológico que buscan producir resultados de mejora en la IA, precisan fondos del orden de millones de dólares

Y sus resultados no tienen incidencia directa en el rendimiento de las IAs más utilizadas a nivel global

Proyecto	Lengua	Presupuesto total (€)	Objetivo principal
NEL-AINA	Catalán	3.000.000	Generar corpus y modelos informáticos de la lengua catalana para que las empresas que crean aplicaciones basadas en inteligencia artificial (IA), como asistentes de voz, buscadores de Internet, traductores y correctores automáticos, agentes conversacionales, entre otros, puedan hacerlo fácilmente en catalán
NEL-GAITU	Euskera	2.000.000	Desarrollar y ofrecer servicios lingüísticos básicos y transversales para utilizarlos en todas las administraciones públicas y ofrecer mejores servicios públicos a la ciudadanía.
NÓS	Gallego	2.000.000	Crear recursos digitales y lingüísticos necesarios para facilitar el desarrollo de aplicaciones basadas en inteligencia artificial (IA) y tecnologías del lenguaje (TL) tales como asistentes de voz, traductores automáticos y agentes conversacionales en gallego.



5. 21 ESTRATEGIAS DE INCLUSIÓN TECNOLÓGICA

1 Impulsar la comunicación digital en lenguas indígenas

a Promoción de influencers y creadores en LL.II.

- Creadores de contenido en lenguas indígenas generan visibilidad global y atraen nuevas generaciones.
- Campañas y colaboraciones en plataformas digitales pueden ampliar su alcance e impacto.

b Formación en RR.SS. a través de programas de alfabetización

- Incluir formación en internet y medios digitales en las iniciativas de alfabetización existentes.
- Talleres prácticos que enseñen a crear y compartir contenido digital, fomentando la producción constante en sus lenguas.

c Desarrollo de foros y plataformas temáticas en LL.II.

- Crear espacios digitales para recopilar conocimientos y tradiciones preserva el patrimonio cultural y estimula su uso en contextos modernos.
- Estas iniciativas fomentan la colaboración intergeneracional, impulsan la inclusión digital y fortalecen el orgullo lingüístico.

2 Preservar y ampliar el contenido existente en lenguas indígenas

a Programas de apoyo al mantenimiento de recursos digitales

- Muchos recursos digitales en lenguas indígenas se han perdido por falta de soporte técnico y dominios expirados.
- Es esencial crear programas de hosting permanente y soporte técnico para preservar estos contenidos a largo plazo.

b Digitalización y conservación de archivos físicos

- Digitalizar archivos físicos preserva y amplía el acceso a conocimientos culturales y lingüísticos indígenas.
- Usar formatos estándar y repositorios públicos mejora la integración y visibilidad en bases de datos internacionales.

c Acceso a fondos para proyectos de documentación

- Colaboraciones entre instituciones, ONG y gobiernos pueden asegurar fondos para la documentación y digitalización de lenguas indígenas.
- La financiación debe incluir aspectos técnicos y culturales, garantizando la continuidad y evolución de estos contenidos.

3 Normalizar el uso de lenguas indígenas y reducir la fragmentación

a Impulsar acuerdos para la convergencia dialectal

- Un estándar común en lenguas indígenas mejora la eficacia y alcance de recursos como diccionarios, aplicaciones y corpus digitales.
- La estandarización dialectal facilita el uso de lenguas en educación, medios y contextos formales.

b Respetar la diversidad en un marco de estandarización flexible

- Los estándares deben ser flexibles, adaptándose a distintos contextos y respetando la diversidad lingüística.
- Recursos digitales personalizables ayudan a equilibrar la unificación con la preservación de expresiones locales.

c Fomentar la formación de expertos locales en lingüística y tecnología

- Formar expertos locales en lingüística y tecnología es clave para liderar y mantener la estandarización de sus lenguas.
- La normalización lingüística reduce la fragmentación, mejora la inclusión tecnológica y potencia la producción de contenido digital.

4 Impulsar el desarrollo de herramientas habilitadoras

a

Desarrollo de sistemas de conversión de voz-texto

- Permiten incrementar el tamaño de datos entrenable mediante la conversión de audio o vídeos a texto.
- Facilitan el acceso a tecnología para hablantes de lenguas indígenas no alfabetizados.

b

Desarrollo de detectores automáticos de lenguas indígenas

- Clasifican y filtran textos por lengua para gestionar datos masivos.
- Mejoran la creación de corpus y el entrenamiento de modelos de IA.

c

Creación y mantenimiento de traductores automáticos

- Facilitan el acceso y aprendizaje de lenguas indígenas, reduciendo barreras lingüísticas.
- Incrementan la propuesta de valor de internet para las comunidades indígenas.

5 Aprovechar iniciativas de inclusión lingüística de grandes empresas

a

Alianzas con Big Tech y grandes empresas con iniciativas de inclusión

- Aprovechar las iniciativas de localización de interfaces para proponer un paso más con la incorporación de IA.
- Incrementar la colaboración con gigantes tecnológicos fabricantes de tecnologías de acceso universal.

b

Desarrollo de tecnologías de IA conversacional

- Asistentes virtuales en lenguas indígenas mejoran la accesibilidad y experiencia de usuario.
- Impulsan la necesidad de una mejora en el rendimiento de la IA en lengua indígena.

c

Generar demanda de recursos y entrenamiento

- La incorporación de lenguas indígenas en servicios populares genera una presión positiva sobre el ecosistema tecnológico.
- Cuanto más utilizados sean los servicios en lenguas originarias, mayor será el incentivo para continuar desarrollando tecnologías avanzadas.

6 Ampliar la conectividad de las comunidades indígenas

a Impulsar programas de ampliación de cobertura de internet

- Incrementar el acceso a internet en comunidades indígenas implica aumentar el colectivo generador en LL.II.
- El incremento del efecto red aumenta el incentivo de la acción de generación de contenidos.

b Acompañar con programas de formación en el uso de internet

- Implementar formación digital adaptada a las lenguas y culturas locales.
- Enseñar navegación, seguridad, creación de contenido y acceso a recursos clave.

Ampliar el alcance de programas de alfabetización existentes

- ### c
- Integrar formación digital en programas de alfabetización en lenguas indígenas.
 - Fomentar la participación en plataformas digitales para generar contenido en lenguas originarias.

7 Incrementar la localización lingüística de los servicios de las grandes tecnológicas

a Localización de sistemas operativos y navegadores web

- Localizar sistemas operativos y navegadores en lenguas indígenas incrementa el acceso de las comunidades.
- Colaborar con empresas tecnológicas para traducción y personalización equitativa.

b Localización de buscadores y redes sociales

- La traducción de interfaces en redes sociales aumenta la posibilidad de generar conversaciones en LL.II. en ellas.
- Proveer herramientas en estos idiomas facilita el acceso a contenidos culturales relevantes.

c Sensibilización y colaboración con grandes tecnológicas

- Demostraciones de impacto social y estudios sobre la brecha digital, para persuadirlas de que amplíen su compromiso.
- Incrementar la localización promueve equidad digital y reconocimiento cultural.

6. RECOMENDACIONES Y PLAN DE ACCIÓN

1 Creación de un consorcio internacional impulsor del proyecto

El primer paso es formar un consorcio compuesto por organismos nacionales e internacionales, instituciones dedicadas a la protección cultural y empresas tecnológicas interesadas en la inclusión lingüística. Este consorcio tendrá como objetivo liderar y coordinar todas las acciones estratégicas de este plan.

a Responsabilidades del consorcio

- Definición de la estrategia general y supervisión del plan de acción.
- Gestión de alianzas y financiamiento.
- Coordinación entre los distintos actores involucrados.
- Evaluación de resultados y ajustes al plan.

2 Creación del Equipo de Trabajo de Implementación

Este equipo se encargará de ejecutar las acciones estratégicas definidas, organizar actividades y asegurar la correcta implementación del plan. Estará compuesto por miembros de las organizaciones que han constituido el consorcio impulsor del proyecto.

a Funciones del equipo ETI

- Establecer y desarrollar las alianzas estratégicas con los sponsors de las iniciativas locales.
- Supervisar la ejecución de los proyectos derivados de las estrategias.
- Gestionar los recursos asignados.
- Garantizar la participación activa de las comunidades indígenas.

3 Organización de un evento de alta visibilidad para comunicar la iniciativa

Se organizará un evento para presentar los resultados del estudio, comunicar la formación del consorcio y anunciar las alianzas estratégicas logradas. Además, se utilizará como plataforma para convocar a un hackaton de alto impacto.

a Objetivos del evento

- Presentar el informe sobre la brecha de la IA en lenguas indígenas.
- Anunciar la creación del consorcio.
- Presentar la hoja de ruta.
- Convocar a la comunidad a participar en el hackaton.

4 Hackathon de Innovación Tecnológica por una IA en lenguas indígenas

El hackaton reunirá a actores clave, grupos comunitarios, desarrolladores, empresas tecnológicas y creadores de contenido para diseñar y presentar proyectos concretos que contribuyan a las estrategias de inclusión tecnológica.

a Objetivos del Hackaton

- Desglosar las estrategias de inclusión en iniciativas específicas.
- Desarrollar propuestas de proyectos en colaboración con comunidades locales.
- Promover la creación de herramientas tecnológicas en lenguas indígenas.

b Estructura del Hackaton

- **Grupos de Trabajo:** Los equipos desarrollarán planes de proyecto centrados en estrategias de inclusión.
- **Presentación de Proyectos:** Los proyectos se presentarán ante un jurado compuesto por miembros del consorcio y sponsors.
- **Evaluación y Premios:** Se otorgarán premios a los proyectos más destacados.
- **Roadmap:** Los proyectos seleccionados formarán parte del roadmap del equipo principal de trabajo.

5 Desarrollo de alianzas estratégicas locales

Es necesario identificar y establecer colaboraciones con actores públicos, privados, ONGs, empresas tecnológicas, universidades y medios de comunicación que puedan actuar como sponsors de iniciativas específicas.

a Alianzas prioritarias

- Empresas privadas de gran consumo con iniciativas en curso de digitalización de la experiencia en lenguas indígenas.
- Gobiernos locales que estén impulsando programas de apoyo a las lenguas indígenas.
- Proveedores tecnológicos de las empresas privadas de gran consumo.
- Organizaciones activistas y voluntarios de apoyo a las comunidades originarias.

6 Ejecución de proyectos locales y monitorización del progreso de la iniciativa

La ejecución de los proyectos locales será una prioridad para lograr un impacto directo en las comunidades. Los equipos responsables de cada proyecto deberán coordinarse con el ETI para asegurar el cumplimiento de los objetivos, plazos y la correcta asignación de recursos. Se establecerán mecanismos de supervisión continua para garantizar la calidad de los resultados y la sostenibilidad de las iniciativas.

a Fases de ejecución

- **Asignación de Recursos:** Distribución de fondos, tecnología y formación a los equipos locales.
- **Implementación:** Desarrollo práctico de las estrategias acordadas en el hackaton y en los planes de trabajo.
- **Supervisión Continua:** El equipo principal supervisará el progreso, recopilando información sobre resultados, obstáculos y buenas prácticas.
- **Informes Periódicos:** Los responsables de los proyectos presentarán informes de avance al consorcio.
- **Evaluación de Resultados:** Se analizará el impacto en términos de mejora del rendimiento de la IA, generación de contenidos en lenguas indígenas y participación comunitaria.
- **Ajustes y Mejora:** Los resultados evaluados permitirán realizar ajustes a los proyectos en ejecución, optimizando las estrategias aplicadas.

