

ILENIA

IMPULSO DE LAS LENGUAS
EN LA INTELIGENCIA ARTIFICIAL

Estudio de casos de uso de alto impacto de las Tecnologías del Lenguaje

©2024 Hitz Zentroa

Atribución/Reconocimiento-CompartirIgual 4.0 Internacional (CC-BY-SA 4.0)

<https://creativecommons.org/licenses/by-sa/4.0/legalcode.es>

HiTZ Zentroa

Facultad de Informática de la UPV/EHU

Paseo Manuel Lardizabal, 1

20018 Donostia

hitz@ehu.eus

Autoría:

Hitz Zentroa, Langune y Emun

Promotores:

Hitz Zentroa

1 Origen y objetivos del proyecto	1
2 Conceptualización	5
2.1 <i>Inteligencia Artificial</i>	5
2.2 <i>Tecnologías del Lenguaje</i>	5
3 Diagnóstico	7
3.1 <i>Fases del proyecto</i>	7
3.2 <i>Metodología</i>	8
3.3 <i>Resultados de la fase de escucha</i>	11
4 Propuesta de criterios	19
4.1 <i>Clasificación y ponderación de los criterios</i>	19
4.2 <i>Tablas de ponderación</i>	20
5 Demostradores de alto impacto	23
5.1 <i>Ficha técnica de demostradores</i>	23
6 Conclusiones	43
Anexos	47

Abreviaciones

IA	Inteligencia Artificial
LLM	Modelo de lenguaje de gran tamaño (del inglés <i>Large Language Model</i>)
PLN	Procesamiento del Lenguaje Natural
STT	Reconocimiento del habla (del inglés <i>speech-to-text</i>)
TA	Traducción Automática
TL	Tecnologías del lenguaje
TTS	Síntesis de voz (del inglés <i>text-to-speech</i>)
TRL	Nivel de madurez de la tecnología (del inglés <i>Technology Readyness Level</i>)

Niveles de madurez tecnológica

- 1 Principios básicos observados y reportados
- 2 Concepto y/o aplicación tecnológica formulada
- 3 Función crítica analítica y experimental y/o prueba de concepto característica
- 4 Validación de componente y/o disposición de los mismos en entorno de laboratorio
- 5 Validación de componente y/o disposición de los mismos en un entorno relevante
- 6 Modelo de sistema o subsistema o demostración de prototipo en un entorno relevante
- 7 Demostración de sistema o prototipo en un entorno real
- 8 Sistema completo y certificado a través de pruebas y demostraciones
- 9 Sistema probado con éxito en entorno real

ORIGEN Y OBJETIVOS DEL PROYECTO

Desde hace unos pocos años el aprendizaje profundo está revolucionando la Tecnología del Lenguaje (TL) y la Inteligencia Artificial (IA). Esta tecnología, que puede aplicarse entre otras al procesamiento del lenguaje natural, la traducción automática y la síntesis y reconocimiento de la voz, está transformando radicalmente la manera de comunicarnos, acceder a la información e interactuar con nuestro entorno social y cultural.

El proyecto coordinado **ILENIA**¹ se enmarca dentro del Proyecto Estratégico para la Recuperación y Transformación Económica (PERTE) de la Nueva Economía de la Lengua. Su objetivo principal es desarrollar los recursos lingüísticos necesarios, como corpus y modelos computacionales, para facilitar la incorporación de las lenguas oficiales de España en las más avanzadas aplicaciones de IA.



¹<https://www.proyectoilenia.es>

En un contexto donde una gestión inadecuada de las lenguas podría invisibilizar la diversidad lingüística de nuestra sociedad, **ILENIA** persigue el desarrollo de recursos y aplicaciones multilingües, en especial modelos multilingües de texto, voz y traducción automática que den respuesta a las necesidades de la sociedad y estén en línea con la tecnología actual, en la que el multilingüismo y las lenguas juegan un papel primordial.

El proyecto **ILENIA** se desarrolla de forma coordinada en cuatro centros de investigación punteros en el desarrollo de actividades del ámbito de la tecnología lingüística: el Barcelona Supercomputer Center - Centro Nacional de Supercomputación (**BSC-CNS**²; proyecto AINA³), el Centro Vasco de Tecnología de la Lengua (**HiTZ**⁴; proyecto IKER-GAITU⁵), la Universidad de Santiago de Compostela (**USC**⁶; proyecto NÓS⁷) y el Centro de Inteligencia Digital de Alicante (**CENID**⁸; proyecto VIVES⁹).

Así, el proyecto coordinado **ILENIA** se presenta como una oportunidad única y singular para dotar a nuestra sociedad de la infraestructura lingüística necesaria que garantice la inclusión de nuestras lenguas oficiales en las aplicaciones más avanzadas de IA actuales y futuras. Podemos resumir los resultados de **ILENIA** en los siguientes puntos:

1. Impulsar el sector de la IA centrada en el lenguaje, mejorando su competitividad a nivel europeo y global.
2. Hacer rentable y atractiva la incorporación de nuestras lenguas oficiales en cualquier aplicación de IA, tanto para las grandes multinacionales como para las pequeñas y medianas empresas.
3. Conseguir que nuestra sociedad pueda participar en el mundo digital utilizando cualquiera de nuestras lenguas oficiales, al mismo nivel que las lenguas globales.
4. Fomentar que la investigación en IA centrada en el lenguaje esté suficientemente preparada para responder al reto de la sociedad digital basada en el conocimiento.



²<https://www.bsc.es>

³<https://www.projecteaina.cat>

⁴<https://www.hitzeus>

⁵<https://www.hitzeus/iker-gaitu>

⁶<https://www.usc.gal>

⁷<https://www.nos.gal>

⁸<https://www.cenid.es>

⁹<https://www.vives.gplsi.es>

El presente informe constituye el **Entregable 5.1** de HiTZ para el proyecto **ILENIA**.

El objetivo de este informe es **identificar los casos de uso de alto impacto** con el asesoramiento de los principales agentes de Euskadi en TL/IA, junto con el sector industrial directamente interesado en los objetivos del proyecto **ILENIA**. Concretamente, se propone la identificación, selección y desarrollo de casos de uso y aplicaciones innovadoras de alto impacto en la industria de la lengua que se basen en las tecnologías que se pretenden desarrollar o mejorar dentro de **ILENIA**.

A partir de la propuesta del centro HiTZ en colaboración con **LANGUNE**¹⁰ y **EMUN**¹¹ se ha podido identificar una lista inicial de posibles demostradores de alto impacto y definir una batería de indicadores para medir su impacto tanto económico como social.

¹⁰<https://www.langune.eus>

¹¹<https://www.emun.eus>

CONCEPTUA- LIZACIÓN

Inteligencia Artificial

La Inteligencia Artificial (IA) es una rama de la informática que busca crear sistemas capaces de realizar tareas que, tradicionalmente, requieren inteligencia humana, como el reconocimiento de patrones, la toma de decisiones, la comprensión del lenguaje y el aprendizaje a partir de la experiencia. Estos sistemas se basan en algoritmos que procesan datos para la toma de decisiones. Hoy en día, gran parte de los datos disponibles está compuesta por información no estructurada en forma de documentos escritos o hablados en distintos idiomas.

Así, el lenguaje y su procesamiento automático son fundamentales para el desarrollo de la IA. A su vez, la IA constituye la base tecnológica que permite el desarrollo y mejora continua de las [Tecnologías del Lenguaje](#). Juntos, sustentan numerosas aplicaciones modernas que facilitan la interacción entre humanos y máquinas, desde la traducción automática y los motores de búsqueda hasta los asistentes virtuales.

La IA centrada en el lenguaje se ha consolidado como una tecnología clave y es, posiblemente, uno de los campos más innovadores de la IA, con un impacto económico en rápido crecimiento. Son varias las empresas tecnológicas que ya han implantado en producción grandes modelos de lenguaje pre-entrenados, con el fin de mejorar tanto las funcionalidades como la experiencia de usuario de sus productos. Estos modelos tienen el potencial de transformar múltiples sectores de la sociedad, pero no se han de olvidar los importantes [desafíos técnicos, sociales y éticos](#) que plantean.



Tecnologías del Lenguaje

Las Tecnologías del Lenguaje (TL) constituyen un campo interdisciplinario que investiga y desarrolla sistemas y aplicaciones que permiten a las máquinas interpretar y generar lenguaje humano en todas sus formas: escrito, hablado, o de signos, entre otras. Las TL pueden clasificarse de diferentes maneras y abarcan distintos aspectos del lenguaje, tanto escrito como hablado.

El presente estudio se centra en analizar las TL multilingües en tres áreas interrelacionadas de un alto potencial de impacto:

- **La traducción automática (TA):** Sistemas que convierten automáticamente texto de un idioma a otro, facilitando la comunicación entre diferentes lenguas.

- **Las tecnologías de la voz:** Sistemas para la transcripción o reconocimiento del habla (STT, del inglés *speech-to-text*), que convierten el habla humana en texto escrito; y sistemas de síntesis del habla (TTS, del inglés *text-to-speech*), que generan voces artificiales a partir de texto.
- **Los modelos de lenguaje de gran tamaño** (LLM, del inglés *large language model*): Modelos de IA de gran escala entrenados con enormes cantidades de datos textuales, capaces de comprender y generar lenguaje natural de manera avanzada.

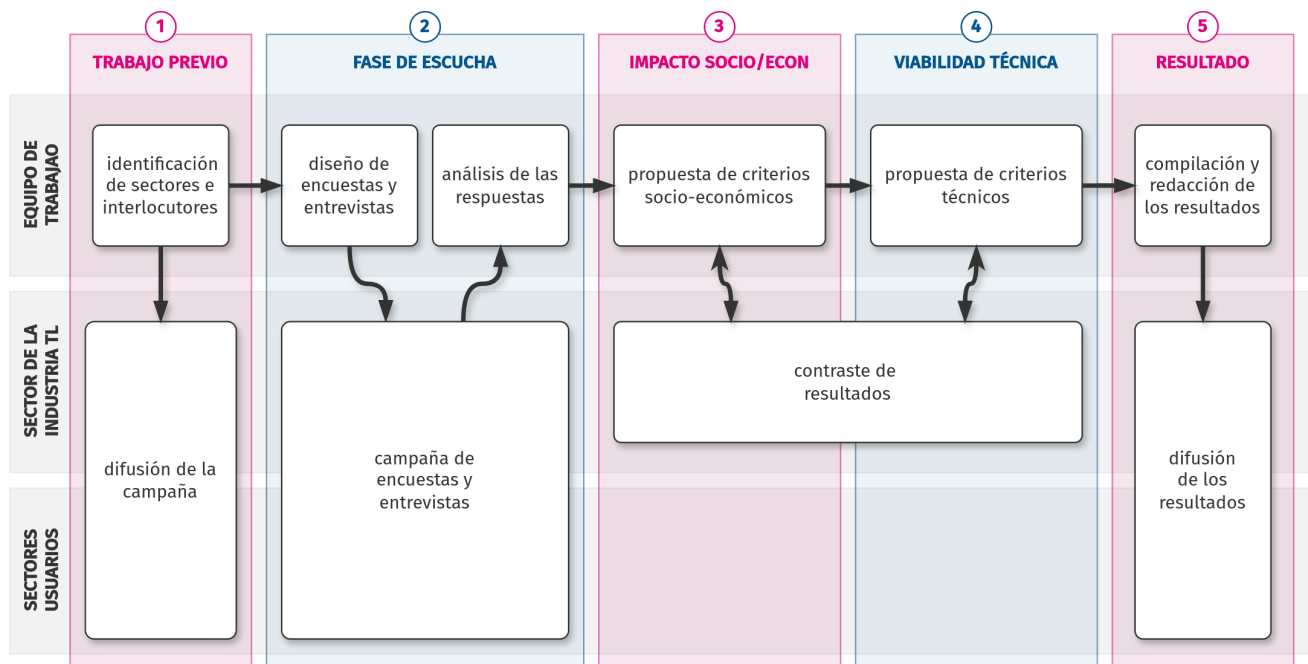
Cada una de estas áreas desempeña un papel importante en el avance de la interacción entre humanos y máquinas, impulsando el progreso del estado del arte de la tecnología más avanzada de aprendizaje profundo hacia la comprensión y la generación del lenguaje natural.

Del mismo modo, cada vez es más necesaria la creación de recursos multilingües, en especial modelos multilingües de texto, voz y traducción automática, que den respuesta a las necesidades de la sociedad y estén acorde con la tecnología actual, en cuyo desarrollo el multilingüismo y la transferencia entre lenguas juegan un papel primordial.

DIAGNÓSTICO

Fases del proyecto

En este trabajo, la **fase de escucha** ha tenido un papel primordial, donde entrevistas y cuestionarios han servido de base metodológica. Hemos querido obtener un resultado cualitativo y más profundo en las entrevistas, completado con un resultado más cuantitativo mediante los cuestionarios.



Metodología

En primer lugar, se han identificado dos sectores principales de interés: el sector dedicado a la investigación y desarrollo de las Tecnologías del Lenguaje –es decir, la industria de las Tecnologías del Lenguaje– y los ámbitos que están empleando o aplicando estas tecnologías: por ejemplo, empresas de servicios e industria, medios de comunicación, el sector educativo, de la administración pública, o judicial.

En segundo lugar, se han analizado y seleccionado los perfiles concretos de esos sectores, para posteriormente realizar las entrevistas y cuestionarios. Así, se ha obtenido una visión más amplia de la situación actual en torno a las Tecnologías del Lenguaje, identificando una serie de retos, limitaciones y objetivos, así como los criterios a tener en cuenta a la hora de desarrollar o seleccionar un tipo u otro de tecnología.

Entrevistas

La herramienta principal utilizada en este trabajo han sido las entrevistas. Se ha tenido en cuenta que la técnica cualitativa sería la más adecuada para obtener información más variada, global y útil para esta elaboración.

El guion utilizado en las entrevistas (ver Anexo A) se ha estructurado en tres grandes bloques: primero, el conocimiento de los casos de uso actuales y potenciales; segundo, el análisis de los retos, objetivos y limitaciones de la sociedad respecto a las tecnologías y; tercero, el análisis de los impactos o criterios a tener en cuenta en el desarrollo o aplicación de estas tecnologías.

Todas las entrevistas han comenzado con una introducción. Dicha introducción establece los objetivos de la investigación y el uso de los datos, en este caso confidenciales. Eso ha ayudado a generar confianza con las personas entrevistadas y crear las condiciones para hacer una lectura más completa de sus vivencias y necesidades. A continuación, se han abordado los temas previstos en el guión, con flexibilidad para adaptarse a los intereses de las personas entrevistadas.

En total se han realizado **17** entrevistas individuales o grupales. Se han tenido en cuenta diferentes perfiles y sectores, obteniendo el testimonio de **30** personas referentes.



- 7 al sector de la industria de las TL.
- 2 al sector industrial y de servicios.
- 3 al sector educativo.
- 1 al sector de los medios de comunicación.
- 4 a organismos del sector público (administración pública, salud y justicia).

La técnica empleada para recoger la información de las entrevistas ha sido la grabación oral, junto a la toma de notas. En todas las entrevistas se ha solicitado la firma de autorización para la grabación (ver Anexo B).

Encuestas

En la elaboración de los cuestionarios se han tenido distintos criterios. Entre ellos, su flexibilidad y atractivo.

Al inicio de los cuestionarios se han explicado de forma breve y concisa sus objetivos o finalidad, dejando claro que son totalmente anónimos. El cuidado de la redacción también se ha considerado imprescindible.

3 cuestionarios, según los perfiles:

- **Cuestionario C.1:** a miembros de la industria de las tecnologías lingüísticas para profundizar en la comprensión de las tendencias y necesidades del sector.
- **Cuestionario C.2:** a miembros de empresas de servicios e industria para recoger su visión sobre el uso de las nuevas tecnologías y sus impactos.
- **Cuestionario C.3:** distribuido entre el alumnado universitario para analizar el conocimiento y la actitud de los futuros profesionales con respecto a estas tecnologías.

Las preguntas de las tres encuestas realizadas no han sido iguales entre sí, ya que cada una ha adaptado al sector correspondiente, aunque las tres han seguido la misma secuencia de preguntas (ver Anexo C).

En primer lugar, se han formulado preguntas de carácter general o de identificación. Así, se han obtenido algunos datos de segmentación sobre las personas que van a cumplimentar el cuestionario, para tenerlos en cuenta en el análisis posterior de datos.

Posteriormente, una vez desarrollado el trabajo previo en profundidad, se ha recurrido a preguntas concretas sobre el tema a investigar, que son la base de la investigación. Se mantiene el equilibrio entre preguntas cerradas y abiertas. Las opciones de respuesta a preguntas cerradas son diferentes, con respuestas posibles variadas, y mediciones del grado de intensidad de las tendencias o comportamientos.

Las preguntas concretas se dividen en dos bloques: por un lado, las que tienen por objeto po-

sicionarse y sintonizar con el tema, que se entienden y responden sin problemas; y, por otro, las referidas al tema llevado al ámbito de trabajo, principales inputs del estudio. Al igual que en las entrevistas, algunas preguntas de los cuestionarios están basadas en el estudio sobre la industria de las lenguas en Europa (ELIS survey, 2023) realizado por diversas organizaciones de la Industria de la Tecnología Lingüística en el año 2023.

Por último, y al final de cada cuestionario, hay una pregunta abierta en la que las personas encuestadas han podido dejar comentarios, observaciones o opiniones si lo desean

**Los cuestionarios
están basados en el
estudio sobre la
industria de las
lenguas en Europa
(ELIS survey, 2023)
realizado por diversas
organizaciones de la
Industria de la
Tecnología Lingüística
en el año 2023.**

Los cuestionarios se han elaborado en una plataforma online y han sido recibidos por correo electrónico en bilingüe (euskera y castellano) para que las personas usuarias puedan elegir el idioma.

La difusión se ha realizado de diversas formas. Por un lado, se ha utilizado la red de Langune para acercarse a la industria de las tecnologías de la lengua. Además, como parte de la asociación Gaia, se ha podido extenderla a las empresas industriales y de servicios. Por otro lado, se ha utilizado la lista de correos electrónicos de CLARIAH para continuar con dicha difusión ma-

siva y, finalmente, también se ha publicado en redes sociales para llegar a un público más amplio. A continuación, se ha realizado un despliegue más directo en función de perfiles específicos, sobre todo a empresas industriales y de servicios, bancos, prestadoras de servicios informáticos... Finalmente, se ha difundido en el campo de la educación, se ha solicitado al profesorado que cumplimente el cuestionario para las organizaciones industriales y de servicios, y se ha enviado otro cuestionario dirigido al alumnado.

La información de los cuestionarios se ha organizado en una hoja de cálculo para tener ordenados y estructurados los datos de cada pregunta de la encuesta y poder analizarlos posteriormente.

En los siguientes apartados de este documento se presenta la información de las entrevistas y cuestionarios obtenidos por medio de esta metodología, así como los resultados de los datos obtenidos.

Traducido a números, se han enviado cerca de **1.000** cuestionarios a través de 3 canales principales (canal de socios de GAIA y Langune, canal de CLARIAH y canal de Mondragon Unibertsitatea), y han respondido **281** personas, el **28 %**.

C.1, Industria de las Tecnologías del Lenguaje:

20 personas

C.2, Empresas de servicios e industriales:

96 personas

C.3, Estudiantes universitarios:

165 personas

Resultados de la fase de escucha

Resultados de las entrevistas

Participación

En total, se han realizado **17** entrevistas individuales o grupales. Se han tenido en cuenta la opinión y experiencia de diversos perfiles, obteniendo el testimonio de un total de **30** personas referentes en alguno de los siguientes sectores:

- Tecnologías del Lenguaje
- Industria y servicios
- Educación
- Medios de comunicación
- Sector público: AAPP, sanidad y justicia

Resumen de las respuestas

Los resultados obtenidos de las entrevistas se han dividido a su vez en las tres áreas de estudio del presente informe: traducción automática (TA), tecnologías de la voz (STT o TTS), y grandes modelos de lenguaje (LLM). Concretamente, se han registrado las opiniones recabadas en relación a los indicadores actuales, desafíos e indicadores y desafíos futuros correspondientes a cada una de las áreas.

En cuanto a la **traducción automática**, los entrevistados consideran que se ha alcanzado un buen nivel de desarrollo que satisface la necesidad básica, también en euskera. Este tipo de sistemas son de uso común en la administración pública, en los medios de comunicación, y en el sector educativo. Por lo tanto, actualmente se busca mejorar los sistemas existentes, incluyendo más pares de idiomas y minimizando los errores. Uno de los mayores desafíos detectados es la creación de traductores especializados en áreas como la medicina o el derecho.

El reconocimiento del habla y los sistemas que integran el **uso del lenguaje hablado** han sido consideradas las áreas más emergentes, indispensables y con mayor proyección dentro de las

TLs analizadas. Tanto el reconocimiento como la síntesis del habla han sido mencionados en todas las entrevistas. También se ha hecho hincapié en el hecho de que este tipo de tecnología está más avanzada de lo que se piensa o de lo que se utiliza, incluso en euskera. Se han detectado indicadores de uso de tecnologías del habla para interactuar con robots y altavoces inteligentes, transcribir actas y videos, crear voces, así como proyectos para el uso de este tipo de sistemas en distintos sectores como la AAPP, justicia, salud y educación.

Los **LLMs** se han mencionado como parte de productos comerciales de los cuales los entrevistados hacen un uso personal. Las entrevistas han revelado que tanto los desarrolladores como los usuarios de TLs conocen las limitaciones y desafíos de los LLMs actuales, tales como la capacidad de cómputo, la seguridad en el manejo de datos y la adaptación a distintos dominios. Por ahora, no se ha detectado una actividad destacable en el desarrollo de LLMs propios, en parte debido a las reticencias de las empresas a compartir sus datos privados con los desarrolladores de estos modelos, lo que dificulta su adaptación a entornos específicos.

Finalmente, se han identificado varios aspectos comunes en todas las áreas analizadas. En primer lugar, se ha destacado la importancia de que la interacción con las máquinas sea lo más natural posible, evitando que el usuario deba adaptar su registro, dialecto o acento para obtener un buen rendimiento. Además, los entrevistados han subrayado la importancia de identificar las emociones del interlocutor, sea en lenguaje hablado o escrito, y de desarrollar tecnologías capaces de interpretarlas. Por último, se considera esencial ofrecer soluciones que no sólo estén adaptadas a las necesidades de las empresas y usuarios finales, sino que también se integren eficazmente en los sistemas propios.

Resultados de las encuestas

Figura 3.1: Edad

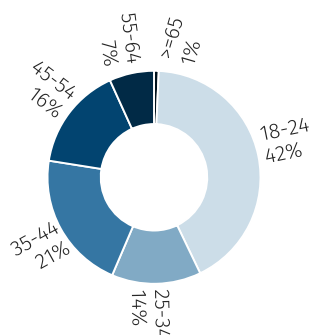


Figura 3.2: Genero

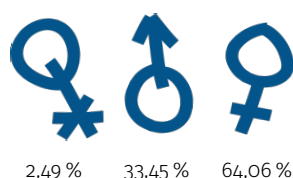
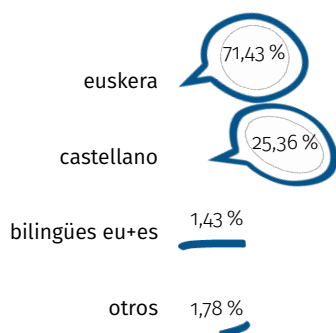


Figura 3.3: Lengua materna



Cuadro 3.1: Sector (C.2)

Academia	35.46 %
Servicios	29.17 %
Industria	13.54 %
Medios de comunicación	5.21 %
Investigación	5.21 %
Otro	5.21 %
Administración pública	4.17 %
No contesta	1.04 %

Participación

En total, se han enviado alrededor de **1.000** encuestas utilizando tres canales principales: el canal de los socios de GAIA y Langune, el canal de la red CLARIAH y el canal de Mondragon Unibertsitatea. Además, se ha difundido la encuesta a través de las redes sociales LinkedIn y X. Finalmente, **281** personas han respondido a uno de los tres cuestionarios: el 7,1 % se ha identificado como experto en Tecnologías del Lenguaje (cuestionario C.1); el 33,3 % ha respondido al cuestionario C.2, dirigido a miembros de empresas de servicios e industria, es decir, usuarios profesionales de las tecnologías; y el 59,6 % corresponde a alumnos universitarios (cuestionario C.3). La distribución por sectores para los participantes de perfil C.2 puede consultarse en el Cuadro 3.1. En cuanto a la edad de los participantes, el 42.0 % tiene entre 18 y 24 años, el 50.2 % está en el rango de 25 a 55 años, y el 7,8 % restante es mayor de 56. Respecto a la identidad de género, el 2,5 % se identifica como no binario, el 33,4 % como hombre y el 64,1 % como mujer. En cuanto a la lengua materna, el 71,4 % de los participantes se identifica con el euskera, el 25,4 % con el castellano, el 1,4 % es bilingüe (euskera y castellano) y el 1,8 % tiene otra lengua materna.

Resumen de las respuestas

¿Qué aplicaciones basadas en Tecnologías del Lenguaje (TL) se utilizan hoy en día? Todos los grupos de encuestados utilizan predominantemente aplicaciones para la corrección de textos y traducción, además de herramientas para la generación de contenido y el subtítulo automático (ver Figura 3.4). No se han encontrado diferencias significativas en el uso de estas tecnologías entre los expertos en TL (C.1) y resto de profesionales (C.2), ni entre expertos y estudiantes (C.3). Sí se observa que los profesionales de servicios e industria priorizan las aplicaciones de corrección más que los estudiantes [Mann-Whitney U, $p < .001$].

¿Cómo y por qué se utilizan? En cuanto a los beneficios que justifican el uso de estas tecnologías, los aspectos más valorados por todos los grupos son el aumento de la productividad y la mejora en la eficacia de la comunicación, seguidos por la automatización de tareas repetitivas. Sin embargo, existen algunas diferencias significativas entre grupos. Los estudiantes universitarios priorizan la mejora en la eficacia de la comunicación y la corrección textual significativamente más que los expertos en TL [χ^2 , $p < .05$], mientras que la capacidad de realizar acciones mediante voz es significativamente más valorada por la industria que por los usuarios profesionales

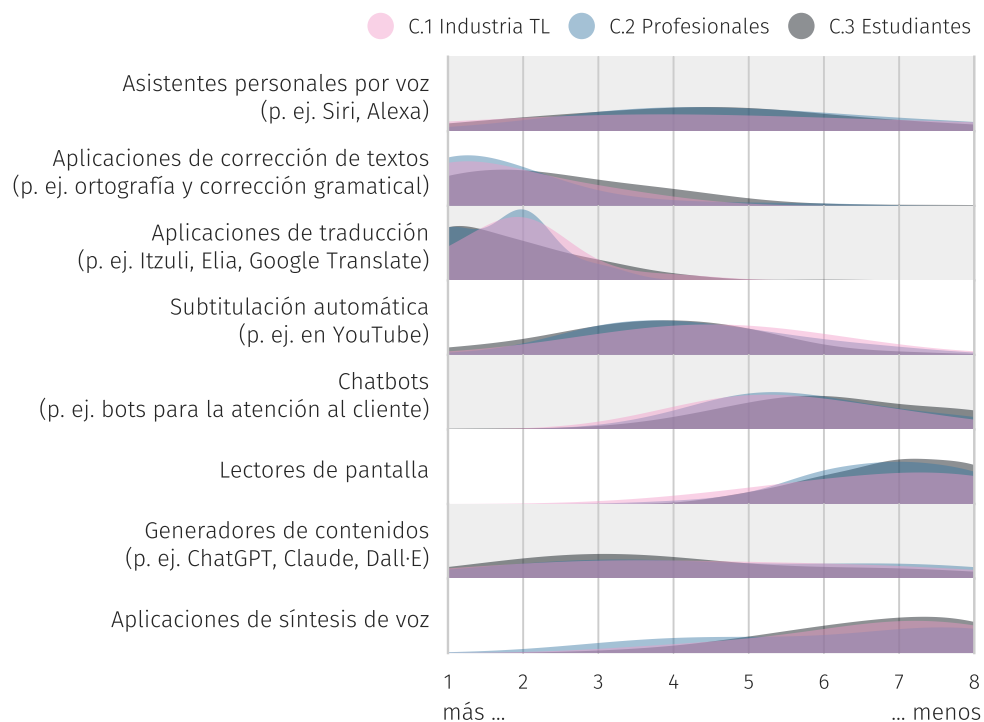


Figura 3.4: ¿Qué tecnologías del lenguaje utilizas en tu día a día? Ordena las opciones de mayor a menor.

	C.1 Industria TL	C.2 Profesionales	C.3 Estudiantes
Aumentar mi productividad	78%	66%	68%
Mejorar la comunicación y calidad del texto	50%	75%	78%
Automatizar tareas repetitivas	39%	24%	21%
Poder realizar acciones por voz	17%	1%	6%
Ampliar los servicios lingüísticos	6%	17%	24%
Poder personalizar las interacciones	0%	3%	4%

Figura 3.5: ¿Cuál de los siguientes beneficios priorizas para justificar su uso? Elige todos los que apliquen.

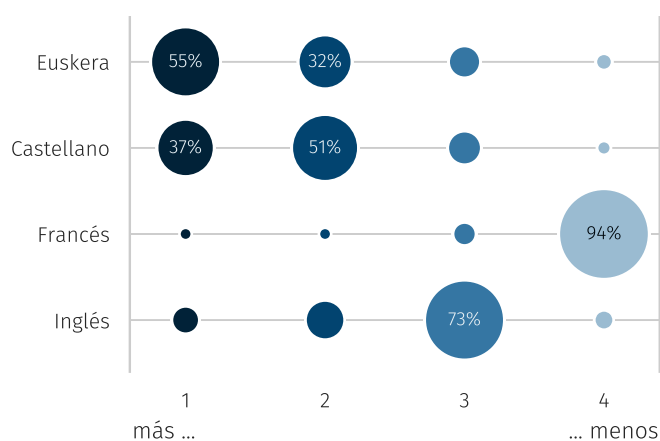


Figura 3.6: En su uso diario, ¿en qué idioma las utilizas principalmente? Ordena las opciones de mayor a menor.

Figura 3.7:
¿Por qué?

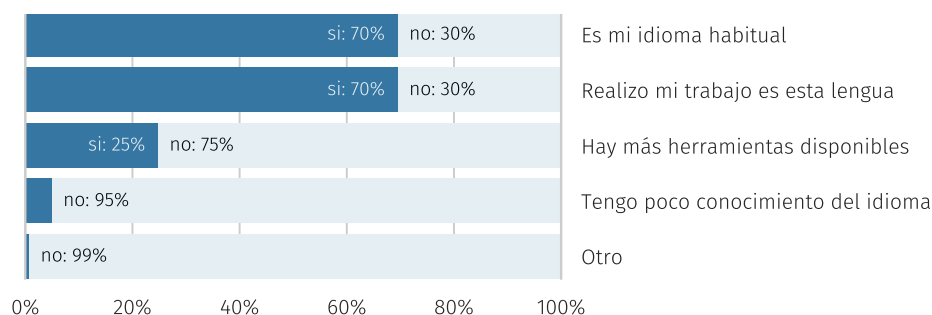
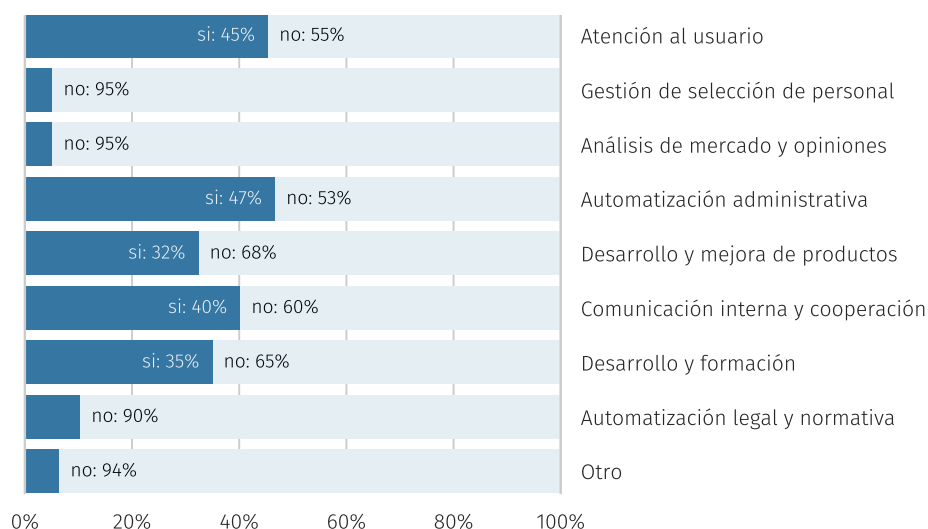


Figura 3.8:
¿Para qué utilizas las TL en el ámbito profesional?



[χ^2 , $p < .05$]. La personalización de las interacciones resulta ser el beneficio menos priorizado por todos los grupos. Estos resultados se ilustran en la Figura 3.5. En relación al uso de idiomas en estas tecnologías, los datos muestran un uso predominante del castellano y el euskera, seguidos por el inglés y, en menor medida, el francés (ver Figura 3.6). Es destacable que los profesionales utilizan significativamente más el castellano que los estudiantes [Mann-Whitney U, $p < .001$]. En cuanto a las razones de uso de cada idioma (ver Figura 3.7), las dos principales son que se trata del idioma habitual del usuario y que es el idioma en el que realiza sus actividades profesionales. En menor medida, también influye la disponibilidad de herramientas y, de manera marginal, el uso de idiomas en los que se tiene poco conocimiento.

¿Para qué se utilizan las TLs en el ámbito laboral? Como se muestra en la Figura 3.8, la adopción de las TLs presenta un patrón variado en diferentes áreas profesionales (C.2). Las aplica-

ciones con mayor penetración se concentran en dos ámbitos principales: la automatización administrativa y documental (47 %) y la atención al cliente y asistencia técnica (45 %). Les sigue de cerca la mejora de la comunicación interna y cooperación. El desarrollo y formación, así como el desarrollo y mejora de productos, muestran niveles de adopción más moderados. En contraste, existe una adopción notablemente baja en áreas como la gestión y selección de recursos humanos, el análisis de mercado y opiniones, y la automatización del cumplimiento legal y normativo, donde menos del 10 % de los encuestados reportan su uso.

¿Qué dificultades se encuentran los usuarios a la hora de adoptar TLs en el ámbito laboral? El análisis de las respuestas a la encuesta sobre la adopción de TLs en el entorno laboral revela un panorama complejo donde confluyen desafíos técnicos, organizativos y de capacitación profesional. La interrelación entre estos aspectos sugiere la necesidad de un enfoque integral

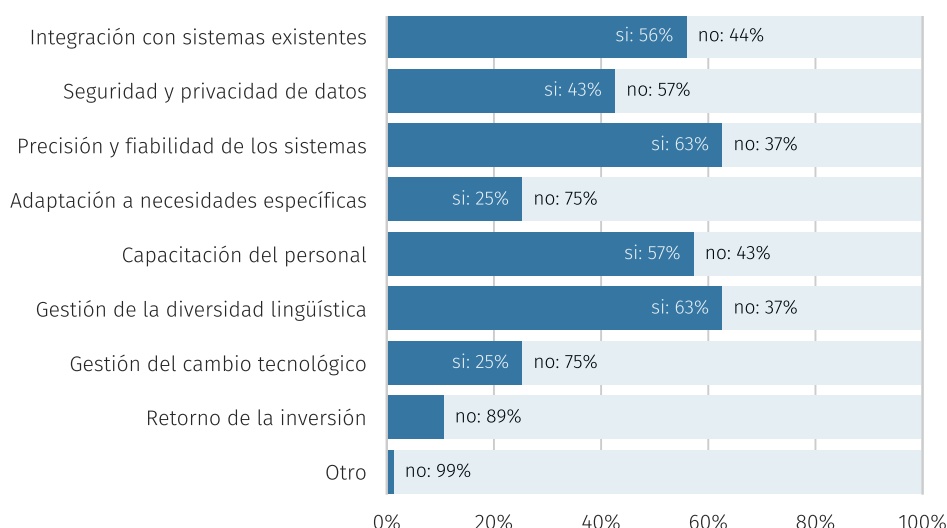


Figura 3.9:
¿Cuáles crees que son los principales retos a la hora de implantar las TLs en una entidad como la tuya?

en la implementación de estas tecnologías.

- **Desafíos organizativos y de integración:** La integración de las TL con los sistemas y plataformas tecnológicas existentes emerge como uno de los desafíos más significativos, siendo identificado por el 56 % de los encuestados (ver Figura 3.9). Este dato cuantitativo se refleja en la realidad organizativa actual, que muestra un panorama de adopción fragmentado y en fase inicial. Los encuestados indican que la mayoría de las organizaciones se encuentran en etapas muy preliminares de integración, con implementaciones principalmente informales a nivel individual, sin una estrategia organizativa coherente. No obstante, algunas entidades están comenzando a abordar la integración de manera más sistemática, incorporando estas tecnologías en sus flujos de trabajo, especialmente en procesos de traducción y documentación.

- **Desafíos técnicos:** La fiabilidad y precisión de los sistemas de procesamiento del lenguaje constituye una preocupación fundamental, señalada por el 63 % de los participantes. Este desafío técnico se manifiesta particularmente en el procesamiento del euskera, donde se observan limitaciones significativas tanto en la calidad de las traducciones como en el manejo de dialectos. La necesidad de desarrollo tecnológico en este ámbito es patente, especialmente en lo referente a la adaptación a las particularidades lingüísticas locales. Los problemas de fiabilidad se extienden también a la necesidad de verificación constante de los resultados, especial-

mente en contextos complejos o con condiciones adversas. Esto se complica aún más cuando se trata de tareas específicas como la transcripción de conversaciones con múltiples participantes o la adaptación a contextos organizativos específicos, aspecto este último identificado como reto por el 25 % de los encuestados.

- **Desafíos competenciales:** La dimensión humana de la adopción de TL se refleja en que el 57 % de los participantes identifica la capacitación del personal como un reto significativo. Las carencias formativas constituyen un obstáculo fundamental, que se manifiesta tanto en la falta de conocimientos técnicos como en la necesidad de una mayor alfabetización tecnológica. Este aspecto se ve agravado por barreras motivacionales, incluyendo la falta de tiempo para el aprendizaje y la experimentación con nuevas herramientas. La adaptación a nuevas formas de trabajo representa otro desafío importante, aunque curiosamente solo el 25 % identifica la gestión del cambio como un reto, lo que podría indicar una subestimación de este factor. La necesidad de desarrollar un criterio crítico en el uso de estas tecnologías emerge también como un aspecto crucial, reconociendo que las TL son herramientas de apoyo que requieren supervisión y criterio humano en su aplicación.

¿Qué dificultades se encuentran los expertos en TLs a la hora de desarrollar la tecnología? Los expertos en TLs (C.1) señalan que, si bien las tecnologías son potencialmente extensibles a cual-

Figura 3.10: Actualmente, ¿cuánto influyen las TLs en tu ámbito de trabajo?

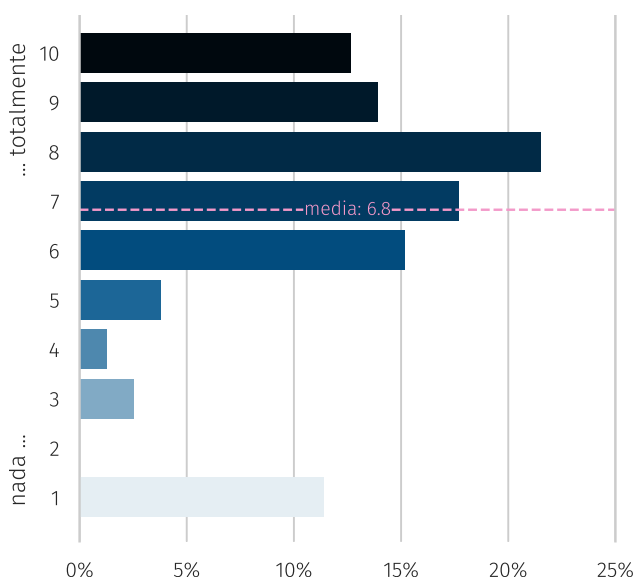
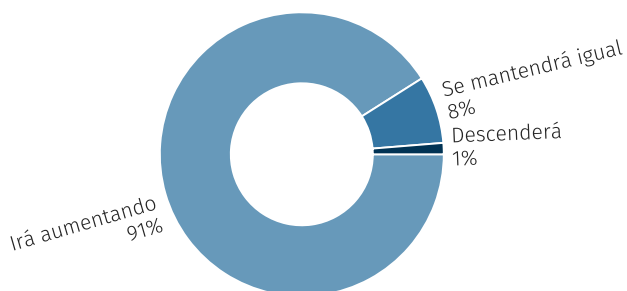


Figura 3.11: Y en el futuro...



quier ámbito profesional, su implementación efectiva se ve condicionada por dos grandes desafíos que requieren superar importantes barreras tanto técnicas como económicas. Por un lado, **La disponibilidad y calidad de los datos emerge como el obstáculo técnico fundamental.** Los expertos señalan la dificultad de obtener datos suficientes y adecuados para cada caso de uso específico, lo que explica en gran medida los problemas de precisión y adaptación que reportan los usuarios finales. Este desafío es especialmente relevante para lenguas con menos recursos como el euskera, donde la escasez de datos de calidad limita el desarrollo de soluciones robustas. Por otro lado, la financiación representa un reto significativo en el desarrollo de proyectos de TL. Los criterios económicos son determinantes para el 100% de los expertos a la hora de decidir si ejecutar un proyecto, por encima

de consideraciones tecnológico-técnicas (67%) o sociales (22%). En este sentido, los proyectos dependen principalmente de subvenciones públicas y fondos propios, mientras que la inversión privada juega un papel más limitado.

¿Qué impacto tienen las TLs actualmente en el ámbito profesional? ¿Que impacto tendrán el futuro? El análisis de las respuestas de los profesionales (C.2) revela percepciones diversas sobre el impacto de las TLs en su ámbito laboral, con una valoración media de 6,8 en una escala de 1 a 10. La distribución de las respuestas es bastante amplia, con un grupo destacable (11.4%) que percibe un impacto mínimo o nulo, aunque predominan las valoraciones moderadas y altas, incluyendo una presencia notable de puntuaciones 9 y 10. A pesar de esta variabilidad en la percepción del impacto actual, **existe un consenso claro sobre su evolución futura: el 91% de los profesionales cree que el impacto de estas tecnologías irá aumentando**, mientras que solo un 8% cree que se mantendrá igual y apenas un 1% prevé que disminuirá (Figura 3.11).

¿Cuáles son los principales retos de futuro en relación a las TLs? Facilitar la comunicación y colaboración es un reto clave en opinión de los usuarios profesionales (C.2), quienes le otorgan mayor importancia que los expertos en Tecnologías del Lenguaje (C.1) [χ^2 , $p < .05$]. Asimismo, los usuarios identifican como reto principal mejorar la accesibilidad e inclusión, aunque en este caso la diferencia no resulta significativa. Ambos grupos coinciden en destacar como áreas relevantes a abordar la mejora de la seguridad mediante la tecnología de detección de fraudes y análisis de riesgos basados en el lenguaje, y la optimización de los procesos de prestación de servicio al cliente. En contraste, la automatización de grandes volúmenes de datos para la toma de decisiones, el fomento de la innovación en el mercado y la creación de interfaces más intuitivas se consideran retos de menor prioridad para ambos grupos. Estos resultados se ilustran en la Figura 3.12. Además, los encuestados han expresado una vez más una preocupación específica por los retos relacionados con las lenguas minoritarias, especialmente el euskera. Los usu-

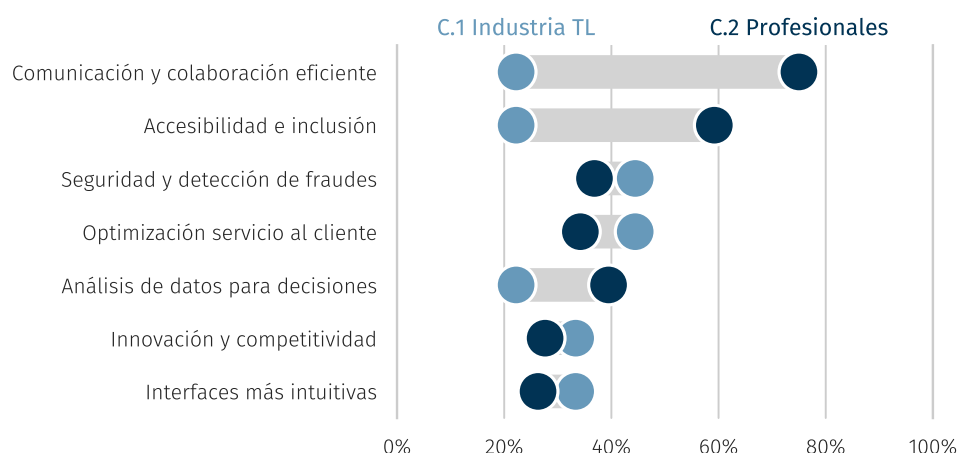


Figura 3.12: ¿Cuáles crees que son los principales retos de futuro en relación con las TLs? Elige todas las que apliquen.

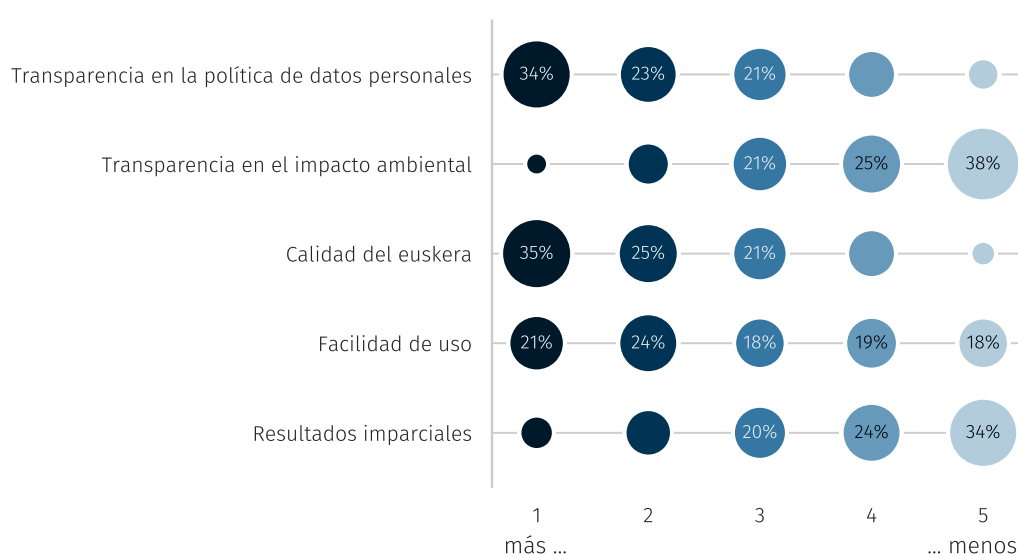


Figura 3.13: ¿A cuáles de los siguientes criterios le darías prioridad?

arios enfatizan la importancia de lograr una paridad en la calidad y cantidad de servicios entre lenguas como el euskera y lenguas dominan-es como el inglés, prestando especial atención a la naturalidad lingüística y la preservación de variantes dialectales más allá del euskera batua.

¿Qué criterios deberían guiar el desarrollo futuro de las TLs? Según la perspectiva de los estudiantes universitarios, futuros profesionales del sector (cuestionario C.3), **la calidad del euskera emerge como el criterio más valorado**, con un 35% de los encuestados situándolo en las posiciones de máxima prioridad (Figura 3.13). Esta preocupación por la calidad lingüística se alinea con los retos previamente identificados respecto a las lenguas minoritarias. De manera similar, la transparencia en la política de datos personales recibe una alta consideración (34% en niveles superiores de prioridad), reflejando una creciente concienciación sobre la seguridad y privacidad. La facilidad de uso muestra una distribución más equilibrada a lo largo de la escala, aunque con una ligera tendencia hacia la alta prioridad (21% en nivel máximo). En contraste, la transparencia en el impacto ambiental y los resultados imparciales tienden a recibir una prioridad relativamente menor, con un 38% y 34% respectivamente situándolos en los niveles más bajos de la escala.

¿Qué aplicaciones basadas en las TLs quedan por explorar y desarrollar? Las respuestas revelan varios ámbitos de interés. Un aspecto fundamental señalado es la necesidad de desarrollar **sistemas que integren diferentes modalidades de interacción** y que puedan incorporarse de manera natural en aplicaciones cotidianas, tanto en el hogar como en el trabajo o en la calle. En este contexto, se destaca el potencial de desarrollar sistemas de interacción inteligente y traducción directa de voz. En el entorno empresarial y administrativo, se identifica la necesidad de desarrollar **sistemas más eficientes para la gestión documental multilingüe**, incluyendo la clasificación automática de documentos, la extracción de metadatos y la implementación de búsquedas en lenguaje natural. También se demanda la integración de herramientas de comunicación interna más sofisticadas y el desarrollo de **asistentes virtuales especializados** para diferentes sectores, como educación y sanidad, que puedan facilitar tareas como el análisis de datos, la generación de informes y la atención al cliente. Por otro lado, destaca especialmente el interés por **aplicaciones con impacto social**, con un énfasis particular en la accesibilidad e inclusión. En este ámbito, se propone el desarrollo de herramientas para personas con diversidad funcional, como sistemas de control por voz y conversión texto-voz bidireccional, así como tecnologías específicamente adaptadas a las necesidades de las personas mayores, considerando las particularidades de su habla. También se señala la importancia de crear sistemas para simplificar y hacer más accesible el lenguaje administrativo, especialmente para personas no familiarizadas con la terminología legal. Finalmente, se identifican dos áreas de desarrollo adicionales: herramientas para la detección automática de desinformación y aplicaciones específicas para el aprendizaje de idiomas, con especial atención a facilitar la integración lingüística de personas migrantes en comunidades bilingües.

PROPUESTA DE CRITERIOS

Clasificación y ponderación de los criterios

Cuando se selecciona una tecnología o proyecto, es importante tener claros los criterios que se deben priorizar. Durante las entrevistas realizadas en la fase de escucha se han analizado estos criterios y clasificado las respuestas en 3 grandes grupos.

Los resultados muestran que se da mayor importancia al **criterio económico**; sin una base económica sólida, incluso las tecnologías más prometedoras pueden fracasar. Los entrevistados expertos en Tecnología del Lenguaje han destacado la inversión económica necesaria tanto para el desarrollo como para la implementación por parte de los usuarios. Las empresas, por ejemplo, deben evaluar si pueden mantener a largo plazo una tecnología, considerando las actualizaciones y el soporte técnico necesario. Además, los costes incluyen no sólo la adquisición de software o hardware, sino también la formación del personal y su integración con los sistemas existentes.

Respecto al **criterio social**, se han identificado diversas consideraciones de gran relevancia. Dado que estas tecnologías pueden tener un impacto tanto positivo como negativo en la sociedad, resulta crucial considerar aspectos relacionados especialmente con la orientación a la inclusión, equidad y ética. Esto implica, entre otras cosas, garantizar que las tecnologías sean accesibles para todas las comunidades lingüísticas y no potencien sesgos o exclusiones. Los criterios sociales han sido mencionados de manera consistente por todos los perfiles entrevistados.

En cuanto a los **criterios técnicos**, destaca la palabra “viable” entre los encuestados de la industria de las Tecnologías del Lenguaje. Los dos criterios anteriores son importantes, pero la solución debe ser técnicamente posible: considerando las posibilidades tecnológicas actuales, la disponibilidad de datos, y el nivel computacional requerido. La solución debe adaptarse a nuestras capacidades y, sobre todo a la disponibilidad de datos. Los entrevistados con perfil de usuarios han subrayado la frecuencia de uso y la integración de los sistemas con sus propios entornos de trabajo.

En las siguientes páginas se detalla cada uno de estos criterios –económico, social y técnico. Para cada uno se incluye una descripción y se indica su nivel de importancia en una escala del 1 al 3, siendo 3 el nivel más alto según lo identificado durante la fase de escucha.

Tablas de ponderación

Criterios BÁSICOS

PARÁMETROS	DESCRIPCIÓN
Que la herramienta esté disponible en euskera	Garantizar el uso de la herramienta en euskera, ser una herramienta que permita la localización, disponible para su uso en distintos idiomas.
Viables con presupuesto aprobado	Planteamiento realizable mediante subvención o con la cantidad ofertada en el "CALL".

Criterios ECONÓMICOS

PARÁMETROS	DESCRIPCIÓN
Mercado potencial	- Verificar si se basa en un diagnóstico o análisis que responda a una necesidad del mercado. - Evaluar el potencial de la tecnología y las oportunidades de crecimiento en el mercado. - Determinar si el proyecto permite estandarización o es altamente personalizado. - Comprobar si se ha definido el modelo de negocio.
Valor económico	- Evaluar el nivel de inversión para valorar la factibilidad de la idea como filtro previo al desarrollo. - Determinar el nivel de inversión necesario para el lanzamiento de la solución. - Analizar la inversión o gasto destinado al mantenimiento de la solución. - Evaluar el riesgo de obsolescencia y su plazo estimado. - Comprobar la viabilidad económica en cada fase de la solución.
Retorno de la inversión	Evaluar la capacidad de la tecnología para generar beneficios medibles, como aumento de productividad, de mercado o reducción de costes.
RESULTADO	3

Criterios SOCIALES

PARÁMETROS	DESCRIPCIÓN
Accesibilidad y usabilidad	• Verificar la accesibilidad de la tecnología para todas las personas, incluidas aquellas con discapacidad, mediante soluciones de diseño inclusivo. • Determinar la disponibilidad de la tecnología en función del número de personas usuarias y su frecuencia de uso en sectores específicos. • Comprobar que la tecnología sea fácil de utilizar, adaptándose a diferentes niveles de habilidad y experiencia mediante interfaces intuitivas. • Asegurar que el precio y acceso a la tecnología sean asequibles y justos para el mayor número de personas. • Revisar la localización e idioma de la tecnología, adecuando contenidos e interfaces a la lengua y cultura locales.
La ética en la inteligencia artificial	• Garantizar el cumplimiento de la legislación actual. • Evaluar la mirada para evitar sesgos en los sistemas de IA. • Verificar la transparencia en el funcionamiento y decisiones automatizadas del sistema. • Comprobar si existen medidas implementadas para prevenir la discriminación. • Evaluar las características que aumentan la confianza en la tecnología. • Verificar si la tecnología se ha desarrollado para multiplicar las capacidades humanas.
Impacto en la comunidad	• Determinar si la propuesta favorece a la comunidad local y su desarrollo, evitando modelos hegemónicos. • Comprobar si la solución responde al contexto comunitario. • Evaluar el impacto tanto positivo como negativo de la solución en la comunidad.
Colaboración	• Verificar si se plantea cooperación entre agentes públicos, privados y la ciudadanía. • Asegurar que el proyecto contribuya al desarrollo del ecosistema.
Privacidad y seguridad de los datos	• Garantizar el uso seguro y respetuoso de los datos de las personas usuarias, incluyendo encriptación y políticas de privacidad robustas.
Medio ambiente	• Evaluar los efectos positivos o negativos que la solución produce en el medio ambiente. • Identificar las posibilidades de multiplicación o reducción de efectos.
RESULTADO	1

Criterios TÉCNICOS

PARÁMETROS	DESCRIPCIÓN
Precisión y eficacia de la tecnología	- Identificar el tipo de datos que se van a utilizar y la precisión y eficacia de la tecnología. - Evaluar el uso de tecnologías existentes y su nivel, con el objetivo de adaptarlas o emplear una tecnología externa.
Escalabilidad - posibilidades de integración	- Comprobar la capacidad de la tecnología para crecer y adaptarse al aumento de la demanda sin perder rendimiento. - Evaluar la capacidad de expansión e integración de la tecnología con otras organizaciones o sectores. - Asegurar la escalabilidad de la tecnología, adaptándola para funcionar en dispositivos como móviles, tablets y ordenadores mediante diseño optimizado.
Innovación y originalidad	Incorporar propuestas innovadoras que promuevan exploraciones no habituales y abran nuevas líneas de investigación.
Apertura tecnológica	Reconocer el grado de apertura de la tecnología, tanto en programación, como en modelos y datos, y su alineación con estándares reconocidos como CC o copyright.
RESULTADO	2

DEMOSTRADORES DE ALTO IMPACTO

En la siguiente tabla se presenta la relación de los demostradores que más se han repetido en la fase de escucha, en respuesta a las principales necesidades, carencias y retos actuales.

En la ficha técnica de los demostradores se presentan:

- el uso o necesidad a cubrir por cada demostrador
- usuario de dicho demostrador
- aplicaciones
- nivel de madurez tecnológica (TRL)
- recursos necesarios
- principales retos técnico-científicos

Ficha técnica de demostradores

Demostrador 1: Subtitulación automática

Demostrador 2: Generador de actas

Demostrador 3: Traducción instantánea del habla

Demostrador 4: Chatbot empático

Demostrador 5: Clonación de la voz

Demostrador 6: Traductor centrado en el usuario

Demostrador 7: Texto adaptativo

Demostrador 8: Buscador propio

Demostrador 9: Automatización de procesos ofimáticos

Subtitulación automática

Tecnología principal: STT; opcionalmente, TA

Aplicación de alto impacto: Subtitulación con traducción en tiempo real para reuniones virtuales

- | | |
|---|---|
| ¿Qué necesidades cubre? | <p>1. Responde a la necesidad de hacer el contenido audiovisual más accesible y comprensible para una amplia gama de audiencias. Esta tecnología posibilita el consumo de contenido sin audio, permitiendo superar barreras lingüísticas y ayudando a personas con discapacidad auditiva, entre otros. Además, la subtitulación automática aumenta el alcance global del contenido (por ejemplo, mejorando la indexación para motores de búsqueda), y ofrece una alternativa rápida y económica a la subtitulación manual, permitiendo que una mayor cantidad de material audiovisual sea accesible para todos. Es importante destacar que esta tecnología se desarrolla en dos ejes principales: por un lado, la subtitulación puede realizarse en tiempo real (online) o de forma diferida (offline), adaptándose a diferentes necesidades y contextos; por otro lado, puede implementarse con o sin traducción automática, lo que amplía aún más su versatilidad, permitiendo tanto la accesibilidad dentro de un mismo idioma como la comprensión entre distintas lenguas.</p> |
| Usuarios | <p>2. Esta tecnología puede aplicarse en cualquier sector. Está dirigida a un amplio grupo de usuarios, incluidas empresas u organizaciones que producen contenido visual a gran escala y necesitan hacer su contenido accesible, cumpliendo con las regulaciones correspondientes, así como aquellas que organizan eventos en vivo, conferencias, o webinars y desean proporcionar subtítulos en tiempo real.</p> |
| Aplicaciones | <p>3.</p> <ul style="list-style-type: none"> • Proporcionar subtítulos traducidos en tiempo real para reuniones virtuales, mejorando la comprensión en entornos multilingües. • Proporcionar subtítulos en tiempo real para transmisiones en directo, como noticias o deportes. • Proporcionar subtítulos para la educación online, como cursos y conferencias en línea, mejorando la accesibilidad para estudiantes con discapacidad auditiva o que aprenden en un idioma no nativo. • Proporcionar subtítulos y traducir vídeos informativos en museos, monumentos o puntos de interés turístico. |
| Nivel de madurez tecnológica (TRL) | <p>4. La madurez de esta tecnología puede variar significativamente al considerar sus distintas variantes, especialmente en el contexto de lenguas con pocos recursos:</p> <ul style="list-style-type: none"> • Subtitulación offline sin traducción: TRL 7 a 9 • Subtitulación online sin traducción: TRL 6 a 7 • Subtitulación offline con traducción: TRL 6 a 7 • Subtitulación online con traducción: TRL 4 a 6 <p>Estos TRLs son estimaciones y pueden variar dependiendo de factores específicos como la calidad del audio y la complejidad del contenido (ver debajo Retos científico-técnicos).</p> |

BÁSICOS

- **Corpus de audio y texto alineado:** Grandes cantidades de audio y sus transcripciones, en el idioma y dominio de aplicación, para la creación o adaptación de un modelo acústico o un modelo STT “extremo a extremo”.
- **Software de segmentación:** Para dividir la transcripción automáticamente en subtítulos de características apropiadas, como la longitud, la composición gramatical, y el tiempo en pantalla.
- **Software de capitalización y puntuación:** Para enriquecer los subtítulos de forma que resulten más legibles y naturales.

OTROS

- **Corpus de texto libre:** Grandes cantidades de texto libre en el idioma y dominio de aplicación, para la creación o adaptación de un modelo de lenguaje.
- **Diccionario de pronunciación:** Mapa de unidades fonéticas a palabras en el idioma de aplicación.
- **Corpus paralelo:** Grandes cantidades de texto alineado en varios idiomas, para el entrenamiento o adaptación de un módulo de TA.

5.

¿Qué recursos o herramientas se necesitan?

1. **Captación de audios de calidad:** En grabaciones captadas con dispositivos básicos (por ejemplo, el micrófono de un ordenador portátil) y en entornos con condiciones adversas (como ecos y ruidos de fondo), la calidad del audio tiende a deteriorarse. Esto afecta negativamente a la precisión del STT. Una solución técnicamente eficaz es el uso de micrófonos individuales de alta calidad; sin embargo, esta solución limita gravemente la usabilidad del sistema.
2. **Habla espontánea de múltiples hablantes:** Las disfluencias como repeticiones, muletillas y pausas en el habla son comunes en la conversación cotidiana. Además, cuando hay múltiples hablantes, éstos pueden interrumpirse entre sí y hablar al mismo tiempo. Estas disfluencias son comunes y pueden confundir a los sistemas de STT si no están diseñados para manejarlas adecuadamente.
3. **Vocabulario técnico y especializado:** En contextos específicos, como discusiones médicas o legales, se utilizan terminología que pueden no aparecer en los datos de entrenamiento estándar de los sistemas de STT. Es crucial que estos sean capaces de reconocer vocabularios especializados con precisión para ser útiles en sus dominios.
4. **Traducción del habla:** El habla tiende a ser más informal y contextualmente dependiente que el texto escrito. La interpretación del significado puede requerir un entendimiento profundo del contexto, lo cual puede ser más difícil de capturar automáticamente en comparación con el texto narrativo más formal y estructurado, medio de aplicación más común para la traducción.
5. **Sincronización en subtítulos traducidos:** Las traducciones pueden variar en longitud respecto al texto original, pero deben ajustarse igualmente para caber en pantalla sin sacrificar la legibilidad ni la sincronización con el audio.
6. **Tiempo real y latencia:** En aplicaciones en tiempo real, los subtítulos deben generarse rápidamente con una mínima latencia. Esto implica optimizar la integración de la cadena de procesamiento y los distintos módulos que la componen para mejorar la velocidad de procesamiento sin comprometer la calidad del resultado.

6.

Retos científico-técnicos

Generador de actas

Tecnología principal: STT y PLN; opcionalmente, TA

Aplicación de alto impacto: Generación de actas paralelas para plenos en territorios con lenguas cooficiales

- | | |
|---|--|
| ¿Qué necesidades cubre? | <p>1. Esta tecnología permite una generación eficiente de registros formales o actas que resumen las ideas principales, las decisiones clave y los puntos de acción derivados de distintos tipos de reuniones o encuentros, facilitando el seguimiento y la referencia futura de acuerdos y tareas. Específicamente, permite la captura estructurada de las intervenciones de los diferentes participantes mediante la atribución automática de sus contribuciones, y resume el intercambio de ideas, destacando elementos importantes y eliminando lo redundante o secundario. Esta tecnología permite además que todos los participantes se involucren plenamente en el objeto de la reunión, sin la distracción derivada de la toma de notas. Finalmente, en combinación con la traducción automática, permite que las actas sean comprendidas por hablantes de diferentes idiomas en contextos internacionales o multilingües.</p> |
| Usuarios | <p>2. Esta tecnología puede aplicarse en cualquier sector. Está dirigida a un amplio grupo de usuarios, siendo idóneo para organizaciones que gestionan reuniones, conferencias y entrevistas. Es de especial interés para aquellos que necesitan documentar y sintetizar de forma rápida y eficaz juntas, reuniones estratégicas o de seguimiento, entre otros. Específicamente en el ámbito del sector público y gubernamental, la generación automática de actas paralelas en varios idiomas puede ayudar a mejorar la transparencia y el acceso a la información.</p> |
| Aplicaciones | <p>3.</p> <ul style="list-style-type: none"> • Generar actas multilingües para plenos municipales, donde los interlocutores utilizan indistintamente cualquiera de las lenguas oficiales del territorio. • Generar actas de sesiones lectivas, identificando los temas clave presentados, las preguntas de los estudiantes, y los conceptos que habrían requerido más explicación para ser asimilados. • Generar actas detalladas de reuniones de seguimiento de proyectos técnicos, capturando decisiones clave, planes de contingencia, asignaciones de tareas y plazos de ejecución. • Generar actas de sesiones de terapia grupal o grupos de apoyo anonimizados, permitiendo a los terapeutas o mediadores analizar patrones y progresos sin comprometer la privacidad de los participantes. |
| Nivel de madurez tecnológica (TRL) | <p>4. El nivel de madurez tecnológica de los generadores automáticos de actas varía según los componentes específicos que integren el sistema. Es importante entender que esta tecnología combina varias disciplinas de I+D, cada una en diferentes etapas de desarrollo. A continuación, se presenta un desglose de los principales componentes y sus TRLs para contextos de pocos recursos como el euskera:</p> <ul style="list-style-type: none"> • Transcripción automática: TRL 8 a 9 • Identificación de hablantes: TRL 6 a 8 • Comprensión de transcripciones y generación de actas: TRL 2 • Traducción automática: TRL 8 a 9 |

BÁSICOS

- **Corpus de audio y texto alineado, y las actas correspondientes:** Grandes cantidades de audio con sus correspondientes transcripciones y actas resultantes, en el idioma y dominio de aplicación, para la creación o adaptación de a) un modelo acústico o un modelo STT extremo-a-extremo, y b) modelos de comprensión de transcripciones y generación de actas mediante PLN.
- **Muestras de los hablantes a identificar:** Grabaciones breves pero variadas (un total de 30 a 60 segundos suele ser suficiente) de cada participante a registrar en el sistema de identificación de hablantes, para que el sistema sea capaz de reconocer las distintas voces y atribuir cada turno de la conversación al hablante correspondiente.

OTROS

- **Software de identificación de idiomas:** Para distinguir el contenido verbalizado en un idioma u en otro, y aplicar las herramientas de procesamiento correspondientes.
- **Corpus de texto libre:** Grandes cantidades de texto no anotado en el idioma y dominio de aplicación, para la creación o adaptación de un modelo de lenguaje del componente STT.
- **Diccionario de pronunciación:** Mapa de unidades fonéticas a palabras en el idioma de aplicación.
- **Software de capitalización y puntuación:** Para enriquecer las transcripciones de forma que resulten más legibles y naturales.
- **Corpus paralelo:** Grandes cantidades de texto alineado en varios idiomas, para el entrenamiento o adaptación de un módulo de TA.

5.

¿Qué recursos o herramientas se necesitan?

1. **Captación de audios de calidad:** Las grabaciones con dispositivos básicos y en entornos adversos deterioran la calidad del audio, afectando negativamente la precisión del reconocimiento. Si bien los micrófonos individuales de alta calidad son efectivos, limitan la usabilidad y adopción del sistema.
2. **Habla espontánea de múltiples hablantes:** Las disfluencias como repeticiones, muletillas y pausas son habituales en conversaciones, especialmente con múltiples hablantes que pueden interrumpirse. Estos elementos pueden confundir a los sistemas de STT si no están preparados para gestionarlos.
3. **Vocabulario técnico y especializado:** Los términos técnicos usados en contextos específicos (científicos, médicos o legales) pueden no aparecer en los datos de entrenamiento estándar. Es esencial que los sistemas reconozcan este vocabulario para ser útiles en sus dominios.
4. **Subjetividad de la tarea:** La generación de actas implica la interpretación y selección de información relevante, lo cual puede ser subjetivo según la perspectiva de diferentes usuarios y la intención de las discusiones. Esta subjetividad afecta tanto a la precisión del modelado como a su evaluación.
5. **Falta de contexto visual:** En reuniones presenciales, los participantes utilizan gestos, expresiones faciales y referencias como “tú” o “esto”. La ausencia de este contexto visual dificulta que los sistemas automáticos interpreten estos referentes, afectando a la precisión y coherencia de las actas.
6. **Explicabilidad del resultado:** La subjetividad inherente en la generación de actas subraya la importancia de la explicabilidad. Los sistemas deben clarificar cómo se determinaron los puntos clave, proporcionando transparencia sobre qué partes de la transcripción influyeron en el resumen final.

6.

Retos científico-técnicos

Traducción instantánea del habla

Tecnología principal: STT y TA; opcionalmente, TTS

Aplicación de alto impacto: Traducción instantánea voz a voz entre idiomas oficiales del estado e idiomas de principales comunidades migrantes

- | | |
|---|--|
| ¿Qué necesidades cubre? | <p>1. Esta tecnología aborda la necesidad de comunicación fluida entre personas que hablan diferentes idiomas, eliminando barreras lingüísticas en tiempo real. Facilita el entendimiento mutuo en contextos multilingües o internacionales, permitiendo interacciones más naturales y eficientes en situaciones donde la comunicación inmediata es esencial, desde negociaciones empresariales hasta asistencia médica de emergencia en países extranjeros. Además, cuando se incorpora tecnología de síntesis del habla, se posibilita una comunicación completamente oral. Esta opción es particularmente valiosa en situaciones donde la lectura de texto traducido puede ser un inconveniente o imposible, como en interacciones donde mantener el contacto visual es importante, o para personas con discapacidad visual.</p> |
| Usuarios | <p>2. Esta tecnología puede aplicarse en una amplia gama de sectores, incluyendo educación, salud, administración pública, turismo, negocios internacionales y servicios de emergencia. Los principales beneficiarios son personas con limitaciones en el habla local, como inmigrantes, refugiados y visitantes temporales (turistas, personas que viajan por trabajo, etc.).</p> |
| Aplicaciones | <p>3.</p> <ul style="list-style-type: none"> • En centros de salud y hospitales, facilitando la comunicación entre personal médico y pacientes extranjeros. • En instituciones educativas, mejorando la comunicación entre docentes y padres de alumnos inmigrantes. • En el núcleo familiar, facilitando que padres extranjeros se involucren activamente en el proceso educativo de sus hijos. • En servicios gubernamentales y administrativos, ayudando a inmigrantes y refugiados en trámites esenciales. • En entornos laborales con equipos internacionales. <p>En general, se trata de tecnología que puede ayudar en la promoción de la inclusión de personas con discapacidad visual en entornos multilingües, facilitando su acceso a información y servicios en diferentes idiomas.</p> |
| Nivel de madurez tecnológica (TRL) | <p>4. Una de las variables que más directamente impactan en el nivel de madurez tecnológica es el par de idiomas que el sistema debe ser capaz de entender y traducir. En este sentido, los idiomas de las principales comunidades migrantes en España varían en dificultad, por ejemplo, desde el chino estándar (un idioma rico en recursos), pasando por el urdu, rumano y el ucraniano (idiomas con bastantes recursos) al dariya (o árabe marroquí) y sus distintas variedades (con pocos recursos). A esto se le añade la escasez de recursos paralelos de calidad para la traducción entre estos idiomas y los idiomas oficiales del estado distintos al español. En consecuencia, esta tecnología se sitúa en TRLs bajos y medios, entre 2 y 4. En el caso de los idiomas hegemónicos (inglés, francés), el TRL puede ascender hasta 8 o 9.</p> |

BÁSICOS

- **Corpus de audio y texto alineado:** Grandes cantidades de audio con sus correspondientes transcripciones, en el idioma y dominio de aplicación, para la creación o adaptación de un modelo acústico o un modelo STT “extremo a extremo”. Si el corpus fuera paralelo en los idiomas de interés, permitiría el desarrollo de un modelo de transcripción y traducción “extremo a extremo”.
- **Corpus paralelo:** Grandes cantidades de texto alineado en varios idiomas, para el entrenamiento o adaptación de un módulo de TA.
- **Software de capitalización y puntuación:** Para enriquecer las transcripciones de forma que resulten más legibles y naturales.

OTROS

- **Corpus de texto libre:** Grandes cantidades de texto libre en el idioma y dominio de aplicación, para la creación o adaptación de un modelo de lenguaje.
- **Diccionario de pronunciación:** Mapa de unidades fonéticas a palabras en el idioma de aplicación.

5.

¿Qué recursos o herramientas se necesitan?

1. **Captación de audios de calidad:** Las grabaciones con dispositivos básicos y en entornos adversos deterioran la calidad del audio, afectando negativamente la precisión del reconocimiento. Si bien los micrófonos individuales de alta calidad son efectivos, limitan la usabilidad y adopción del sistema.
2. **Habla espontánea de múltiples hablantes:** Las disfluencias como repeticiones, muletillas y pausas son habituales en conversaciones, especialmente con múltiples hablantes que pueden interrumpirse. Estos elementos pueden confundir a los sistemas de STT si no están preparados para gestionarlos.
3. **Vocabulario técnico y especializado:** Los términos técnicos usados en contextos específicos (científicos, médicos o legales) pueden no aparecer en los datos de entrenamiento estándar. Es esencial que los sistemas reconozcan este vocabulario para ser útiles en sus dominios.
4. **Manejo de variedades lingüísticas:** Los escenarios de aplicación se caracterizan por la diversidad de hablantes, que pueden utilizar diferentes dialectos, acentos y jergas, suponiendo un reto para el reconocimiento y la traducción. El entrenamiento requiere datos representativos de estas variedades, aunque su recopilación y etiquetado es costoso y laborioso.
5. **Traducción del habla:** El habla tiende a ser más informal y contextualmente dependiente que el texto escrito. La interpretación del significado requiere un entendimiento profundo del contexto, más difícil de automatizar que en el texto formal y estructurado.
6. **Traducción bidireccional:** Para ser útil en escenarios reales, la tecnología debe tener la capacidad de traducir en ambas direcciones con una precisión aceptable, lo cual multiplica el esfuerzo y la dificultad de su desarrollo.
7. **Tiempo real y latencia:** En aplicaciones en tiempo real, las transcripciones traducidas y la síntesis de voz deben generarse rápidamente con una mínima latencia. Esto implica optimizar la integración de la cadena de procesamiento para mejorar la velocidad sin comprometer la calidad.
8. **Expectativa de infalibilidad:** Debido a la criticidad de muchos de los escenarios de aplicación, se espera un rendimiento casi perfecto. Errores en la traducción pueden tener consecuencias graves en contextos como la atención médica o trámites legales y, al no ser posible la post-edición, esto crea una presión significativa para lograr una precisión extremadamente alta.

6.

Retos científico-técnicos

Chatbot empático

Tecnología principal: STT y PLN; opcionalmente, TTS

Aplicación de alto impacto: Monitorización del bienestar psicológico y recopilación de información fidedigna de individuos que padecen afecciones mentales

¿Qué necesidades cubre?	1. Los chatbots empáticos abordan la necesidad de una interacción más humana y comprensiva en las comunicaciones hombre-máquina. Al identificar y analizar automáticamente las emociones de las personas a través de la voz, estos chatbots mejoran la experiencia del usuario, proporcionando apoyo emocional, reduciendo la frustración y aumentando la eficacia de la comunicación.
Usuarios	2. Los chatbots empáticos son particularmente útiles en contextos donde el apoyo emocional y la empatía son cruciales, como en el soporte al cliente, la atención médica, y la educación.
Aplicaciones	3. • Chatbots empáticos en contextos de aislamiento social, proporcionando compañía y apoyo emocional constante. • Chatbots empáticos para el seguimiento de la salud mental, facilitando la captura de datos reales en pacientes con trastornos mentales. • Chatbots empáticos para situaciones de alto estrés, como en la atención al cliente, en procesos educativos o en entornos industriales, ofreciendo respuestas que se ajustan mejor al estado emocional del usuario y reducen así su frustración. • Chatbots empáticos para videojuegos, permitiendo mejorar la experiencia del jugador mediante interacciones más naturales y emocionantes, enriqueciendo así la inmersión y realismo del juego.
Nivel de madurez tecnológica (TRL)	4. El nivel de madurez tecnológica de los chatbots empáticos varía según los componentes específicos que integren el sistema. Esta tecnología abarca múltiples disciplinas de I+D, cada una en diferentes etapas de desarrollo. Además, algunos casos de uso son más sencillos de implementar, como los chatbots dirigidos para atención al cliente, mientras que el desarrollo de chatbots capaces de mantener conversaciones abiertas sobre cualquier tema presenta una complejidad significativamente mayor. A continuación, se presenta un desglose de los principales componentes y sus TRLs para contextos de pocos recursos como el euskera: • Transcripción automática: TRL 8 a 9 • Reconocimiento de emociones en el habla: TRL 5 a 7 • Comprensión del lenguaje y generación de respuestas empáticas: TRL 3 a 5 • Síntesis de voz emocional: TRL 4 a 6

BÁSICOS

- **Corpus de audio y texto alineado:** Grandes cantidades de audio y sus transcripciones, en el idioma y dominio de aplicación, para la creación o adaptación de un modelo acústico o un modelo STT “extremo a extremo”.
- **Corpus de datos de voz emocional:** Grabaciones de voces etiquetadas con diferentes emociones para entrenar o adaptar modelos de reconocimiento emocional sobre el habla. Si el corpus está transcrito, también podrá explotarse la información textual con este fin.

Dependiendo del tipo de chatbot, será necesario uno u otro de los siguientes:

- **Sistema de diálogo basado en reglas:** Son más adecuados para flujos de conversaciones limitados, como es el caso de la atención al cliente, donde el chatbot es capaz de realizar un número determinado de tareas.
- **Corpus de conversaciones y LLM pre-entrenado:** Conjuntos de interacciones etiquetadas con información emocional para ajustar o adaptar el comportamiento de un LLM pre-entrenado.

OTROS

- **Corpus de texto libre:** Grandes cantidades de texto libre en el idioma y dominio de aplicación, para la creación o adaptación de un modelo de lenguaje.
- **Diccionario de pronunciación:** Mapa de unidades fonéticas a palabras en el idioma de aplicación.
- **Software de capitalización y puntuación:** Para enriquecer las transcripciones de forma que resulten más legibles y naturales.

5.

¿Qué recursos o herramientas se necesitan?

1. **Privacidad y seguridad:** El procesamiento de las emociones es un tema altamente sensible, por tratarse de información inherentemente personal que puede utilizarse indebidamente para abusar, manipular o explotar a los usuarios. Por tanto, es imprescindible asegurar la protección de los datos y cumplir las regulaciones (Ley de IA de la EU y la GDPR, por ejemplo), aunque esto pueda perjudicar el desarrollo de la tecnología.
2. **Captación de audios de calidad:** Las grabaciones con dispositivos básicos y en entornos adversos deterioran la calidad del audio, afectando negativamente la precisión del reconocimiento. Si bien los micrófonos individuales de alta calidad son efectivos, limitan la usabilidad y adopción del sistema.
3. **Habla espontánea:** Las disfluencias como repeticiones, muletillas y pausas en el habla son comunes en la conversación cotidiana. Estas disfluencias son comunes y pueden confundir a los sistemas de STT si no están diseñados para manejarlas adecuadamente.
4. **Manejo de variedades lingüísticas:** Los escenarios de aplicación se caracterizan por la diversidad de hablantes, que pueden utilizar diferentes dialectos, acentos y jergas, suponiendo un reto para el reconocimiento y la traducción. El entrenamiento requiere datos representativos de estas variedades, aunque su recopilación y etiquetado es costoso y laborioso.
5. **Subjetividad de la tarea:** La interpretación de las emociones y la generación de respuestas empáticas son inherentemente subjetivas, a menudo difíciles de interpretar incluso para los humanos. Esta característica de la tarea no solamente afecta a la precisión del modelado del problema, si no que dificulta también su evaluación.
6. **Tiempo real y latencia:** En aplicaciones en tiempo real, las transcripciones traducidas y la síntesis de voz deben generarse rápidamente con una mínima latencia. Esto implica optimizar la integración de la cadena de procesamiento para mejorar la velocidad sin comprometer la calidad.

6.

Retos científico-técnicos

Clonación de la voz

Tecnología principal: TTS

Aplicación de alto impacto: Proporcionar voces sintéticas personalizadas a personas con discapacidades orales

1. ¿Qué necesidades cubre? La clonación de voz consiste en replicar las características de una cierta voz en una voz sintética obtenida mediante técnicas de conversión de texto a voz (TTS). Esta tecnología permite **ampliar la gama de voces sintéticas disponibles, superando las limitaciones de tener que utilizar un conjunto reducido de voces**. Esta necesidad surge con especial relevancia para los usuarios de dispositivos de comunicación alternativa y aumentativa, que habitualmente ofrecen un catálogo de voces sintéticas muy reducido.

2. Usuarios El uso de voces sintéticas personalizadas tiene multitud de aplicaciones en diferentes ámbitos. Un caso de uso de alto impacto social es el **uso de voces personalizadas por parte de personas que son diagnosticadas con ELA**, una enfermedad en la que se produce una pérdida progresiva de la capacidad de hablar. Estas personas van a utilizar un dispositivo de comunicación aumentativa y alternativa en su vida diaria. El empleo de una voz de características adecuadas es fundamental para favorecer el mantenimiento de la comunicación con el entorno social de estas personas. Estas personas reciben su diagnóstico cuando todavía tienen el habla intacta, por lo que pueden beneficiarse de esta tecnología si realizan un conjunto limitado de grabaciones a tiempo.

3. Aplicaciones

- En conjunto con el demostrador 3 (traducción instantánea del habla), permitir que el habla de una persona sea traducida y reproducida en otro idioma mientras conserva las características únicas de su voz original.
- Restaurar o preservar voces de figuras históricas para exposiciones educativas y museísticas.

4. Nivel de madurez tecnológica (TRL) El nivel de madurez tecnológica de un sistema de clonación de voz varía según los componentes específicos que integren el sistema. En el campo de la síntesis personalizada de habla neutra (sin emociones) y en habla no dialectal, podemos hablar de un TRL 8 a TRL 9. Para sistemas con personalización más avanzada podemos hablar de TRL 6.

BÁSICOS

- **Equipamiento de grabación:** Dispositivos de grabación y herramientas de captura para la obtención de las muestras de la voz a clonar.
- **Modelos TTS base:** Modelos abiertos de conversión de texto a voz que permitan su reentrenamiento y adaptación con las grabaciones disponibles.
- **Copus de entrenamiento:** Grabaciones de audio de alta calidad con diversidad de contenidos y hablantes para la adaptación de los modelos.
- **Herramientas de despliegue:** Software para la obtención de modelos capaces de ejecutarse en dispositivos portátiles, y para su integración en los sistemas operativos habituales como Windows o Android.

5.

¿Qué recursos o herramientas se necesitan?

1. **Despliegue “on-device”:** La mayoría de las tecnologías no funcionan en dispositivos locales y, aunque están en la nube, no son compatibles con las herramientas habituales de las personas con discapacidad..
2. **Obtención de muestras:** La pérdida previa de la capacidad de hablar dificulta la obtención de muestras de voz para la clonación. Existen diferentes alternativas según el grado y tipo de discapacidad oral:
 - En casos de pérdida total: recuperación de grabaciones previas (vídeos, mensajes de WhatsApp, conversaciones), aunque la calidad final puede verse limitada por la calidad de estas muestras.
 - En casos de pérdida total sin muestras disponibles: identificación de voces similares y aplicación de técnicas de conversión de voz.
 - En casos de pérdida parcial con disartria: aplicación de técnicas de conversión para mejorar la inteligibilidad.

La clonación de voz típicamente adapta un sistema TTS base usando muestras de la voz objetivo. Este sistema requiere grandes cantidades de grabaciones previas que pueden no estar disponibles en todos los idiomas. Por ello, es necesario desarrollar técnicas que minimicen los requisitos tanto para el sistema base como para la clonación específica.

3. **Lenguas minoritarias:** Los sistemas TTS se concentran principalmente en lenguas mayoritarias como inglés, español y chino. La disponibilidad es menor en lenguas minoritarias como euskera o gallego, debido a la menor demanda y la dificultad para obtener recursos. Las variedades dialectales, importantes para la identidad y personalidad de la voz, suponen un desafío adicional.
4. **Expresividad de la voz:** Los sistemas TTS actuales tienen limitaciones para controlar la expresividad y emoción. La naturalidad del habla incluye elementos como risas y carraspeos que aún no se integran adecuadamente.
5. **Privacidad y seguridad:** Al ser la voz un elemento biométrico, su clonación puede emplearse fraudulentamente en sistemas de identificación vocal o comunicaciones telefónicas. El desarrollo de técnicas “anti-spoofing” para distinguir entre voces sintéticas y humanas es un área de investigación en pleno desarrollo.

6.

Retos científico-técnicos

Traductor centrado en el usuario

Tecnología principal: TA

Aplicación de alto impacto: Traducción automática de documentación especializada en el ámbito judicial y sanitario con adaptación dinámica de terminología y estilo

- | | |
|---|---|
| ¿Qué necesidades cubre? | <p>1. La traducción automática (TA) ha experimentado un avance significativo en los últimos años, convirtiéndose en una tecnología ampliamente accesible. Sin embargo, los sistemas actuales son mayoritariamente genéricos y presentan limitaciones en entornos profesionales especializados. El traductor automático centrado en el usuario aborda uno o más de estos desafíos pendientes, enfocándose en la adaptación a dominios específicos, la mejora continua mediante el aprendizaje de datos reales de usuarios, y/o la estimación precisa de la calidad de traducción. Esta aproximación busca elevar la utilidad y fiabilidad de la TA en contextos profesionales donde la precisión y la adecuación terminológica y de estilo son cruciales.</p> |
| Usuarios | <p>2. Los profesionales y organizaciones que requieren traducciones de alta calidad y especificidad se benefician significativamente de esta tecnología. En ámbitos como la salud y la justicia, donde la precisión es crucial, su impacto sería especialmente notable. La versatilidad de esta solución se manifiesta en su capacidad para abordar la traducción de una amplia gama de documentos especializados, incluyendo historiales clínicos, sentencias judiciales, artículos científicos, contratos mercantiles, patentes, informes financieros y contenido periodístico de actualidad.</p> |
| Aplicaciones | <p>3.</p> <ul style="list-style-type: none"> • Adaptación de modelos genéricos a modelos específicos mediante la inyección de terminología especializada • Aprendizaje online a partir del feedback de usuarios para la mejora ágil y continua de los modelos • Personalización de modelos de traducción basada en el historial y preferencias del usuario • Estimación de calidad automática para identificar segmentos que requieren revisión humana |
| Nivel de madurez tecnológica (TRL) | <p>4. El sistema integra varias tecnologías con diferentes niveles de madurez, que pueden variar significativamente entre lenguas con abundantes recursos y aquellas con recursos limitados:</p> <ul style="list-style-type: none"> • Núcleo de traducción automática: TRL 8 a 9 • Adaptación a dominios específicos e inyección de terminología: TRL 5 a 7 • Aprendizaje en tiempo real con feedback de usuarios: TRL 4 a 5 • Estimación de calidad automática: TRL 4 a 6 |

BÁSICOS

- **Modelo base de traducción automática:** un modelo de traducción automática pre-entrenado que sirva como base para el resto de adaptaciones a realizar (en su defecto, grandes cantidades de texto alineado en varios idiomas, para el entrenamiento de un modelo).
- **Corpus paralelo especializado:** Grandes cantidades de texto del dominio de aplicación relevante y alineado en varios idiomas, para la adaptación del modelo base y evaluación de los distintos desarrollos.
- **Glosarios o bases terminológicas especializadas:** Colecciones de términos de interés para el usuario y sus traducciones en los pares de idiomas relevantes, en forma de listados de términos o repositorios estructurados de terminología.
- **Corpus de traducciones post-editadas:** Colección de traducciones automáticas que han sido manualmente corregidas por humanos. Este recurso sirve para desarrollar tanto sistemas de estimación de la calidad como flujos de aprendizaje online.

5.**¿Qué recursos o herramientas se necesitan?**

1. **Adaptación sin olvido catastrófico:** El olvido catastrófico ocurre cuando el modelo, al adaptarse a nuevos datos o dominios, pierde repentinamente su capacidad para realizar correctamente traducciones que previamente dominaba. Por lo tanto, el reto radica en desarrollar técnicas de aprendizaje online que permitan al modelo incorporar nuevo conocimiento y adaptarse a nuevos escenarios sin degradar su rendimiento en contextos previamente superados.
2. **Modelado de la coherencia en contextos largos:** Los modelos actuales de traducción automática tienden a funcionar bien a nivel de oración o párrafo corto, pero a menudo fallan en capturar y mantener coherentemente elementos narrativos a lo largo de textos más extensos, un problema particularmente crítico en los entornos mencionados anteriormente, donde los documentos a traducir (p. e.j, informes clínicos, patentes) pueden ser largos y complejos.
3. **Expectativa de infalibilidad:** Debido a la criticidad de muchos de los escenarios de aplicación, se espera un rendimiento casi perfecto. Errores en la traducción pueden tener consecuencias graves en contextos como la atención médica o trámites legales y, al no ser posible la post-edición, esto crea una presión significativa para lograr una precisión extremadamente alta.

6.**Retos científico-técnicos**

Texto adaptativo

Tecnología principal: PLN (LLMs)

Aplicación de alto impacto: Adaptación automática de documentos administrativos y textos legales a formatos de lectura fácil para personas con diversidad funcional cognitiva

¿Qué necesidades cubre? 1. La adaptación de documentos de texto aborda la necesidad de **hacer la información más accesible y comprensible para diversos públicos**. Esta tecnología permite transformar contenidos complejos en versiones más sencillas, resumidas o adaptadas a lenguaje claro, facilitando la comprensión y el acceso a la información para personas con diferentes niveles de alfabetización, especialización, capacidades cognitivas o necesidades específicas.

Usuarios 2. Esta tecnología beneficia a un amplio espectro de la sociedad, incluyendo estudiantes, personas mayores, personas con discapacidades cognitivas, y público general que se enfrenta a textos complejos.

Aplicaciones 3.

- Simplificación y/o resumen de **materiales educativos** para mejorar la comprensión de temas complejos.
- Adaptación de **información médica** para facilitar la comunicación entre profesionales de salud y pacientes.
- Simplificación de **documentos legales** para mejorar el acceso a la justicia.
- Creación de versiones adaptadas de noticias y contenidos de actualidad **para personas con capacidades cognitivas diversas**.
- Generación de resúmenes de documentos extensos **para mejorar la eficiencia en entornos profesionales**.
- Adaptación de textos **para facilitar la comunicación intergeneracional**.

Nivel de madurez tecnológica (TRL) 4. El nivel de madurez tecnológica varía según los módulos específicos. En entornos más favorables, se han llegado a alcanzar los siguientes niveles:

- Sumarización: TRL 7 a 8
- Simplificación: TRL 6 a 7
- Adaptación a lenguaje claro: TRL 5 a 6

Para idiomas con pocos recursos, el TRL general puede descender a 2 o 3 debido a la menor disponibilidad de datos y herramientas específicas.

BÁSICOS

- **Modelo de lenguaje de gran tamaño (LLM):** Constituyen la base fundamental para las tareas de adaptación de texto. Estos modelos, pre-entrenados en el idioma objetivo, comprenden y generan texto de manera avanzada. Se pueden ajustar para tareas específicas como simplificación o adaptación a lenguaje claro, aprovechando su conocimiento general del lenguaje y su capacidad para entender textos complejos.
- **Corpus de referencia:** Se trata de colecciones de textos paralelos (original y adaptado) en el idioma objetivo. Además, estos corpus deben ser representativos del público objetivo, incluyendo variaciones en complejidad, temática y estilo que reflejen las necesidades de diferentes grupos de usuarios.

OTROS

- **Diccionarios y tesauros:** Proporcionan recursos léxicos que pueden facilitar la simplificación y adaptación del vocabulario. Aunque no son estrictamente necesarios para tareas como la sumarización, pueden ser útiles para encontrar sinónimos más sencillos o explicar términos complejos.

5.

¿Qué recursos o herramientas se necesitan?

1. **Preservación del significado:** Los modelos de lenguaje pueden generar información no presente en el texto original, lo que compromete la preservación del significado. Es un reto crucial evitar estas “alucinaciones” para mantener la fidelidad del mensaje original durante la simplificación, adaptación o resumen de un texto. Además, los modelos actuales a menudo pierden coherencia y cohesión al procesar textos extensos, debido a sus limitaciones por diseño. Esto plantea un reto para mantener la estructura lógica y las relaciones entre ideas en documentos largos.
2. **Dominios especializados:** La adaptación de textos técnicos requiere un conocimiento profundo del dominio de aplicación para mantener la precisión mientras se simplifica o resume. El reto surge de la dificultad de los modelos para adquirir y aplicar conocimiento experto en campos específicos como la medicina, el derecho o la tecnología.
3. **Adaptación a públicos objetivo diversos:** La efectividad de la adaptación del texto varía significativamente según las características del lector (edad, nivel educativo, capacidades cognitivas, etc.). Los sistemas actuales tienden a simplificar de manera uniforme, sin considerar las necesidades específicas de los diferentes públicos objetivo, resultando en adaptaciones que son demasiado simples para algunos usuarios y aún demasiado complejas para otros y limitando la utilidad y accesibilidad real de los textos adaptados.
4. **Evaluación constructiva:** La naturaleza subjetiva de la “buena” adaptación de un texto hace que sea difícil crear métricas puramente automáticas que sean informativas y constructivas. El desafío radica en implementar medidas que capturen aspectos cualitativos como la claridad, fidelidad y la adecuación al público objetivo.

6.

Retos científico-técnicos

Buscador propio

Tecnología principal: PLN (LLMs); opcionalmente, visión por computadora

Aplicación de alto impacto: Buscador multilingüe de jurisprudencia y casos precedentes

- | | |
|--|--|
| 1. ¿Qué necesidades cubre? | Un buscador personalizado basado en modelos de lenguaje de gran tamaño (LLMs), opcionalmente multimodal, aborda la creciente necesidad de acceder y procesar eficientemente grandes volúmenes heterogéneos de información interna en organizaciones, incluyendo tanto texto como contenido multimedia . Aunque los motores de búsqueda tradicionales han mejorado significativamente, a menudo carecen de la comprensión contextual y la capacidad de manejar consultas complejas, ambiguas o formuladas en idiomas distintos al contenido de la base de conocimiento sobre la que realizar la búsqueda. Esta tecnología aprovecha los avances en LLMs para proporcionar búsquedas más precisas y relevantes, comprendiendo el contexto y la intención detrás de la consultas, posiblemente en diversos idiomas. |
| 2. Usuarios | Empresas, instituciones de investigación, agencias gubernamentales, el sector jurídico y el de la atención médica pueden ser los principales usuarios, especialmente para empleados, gerentes de proyectos y analistas que necesitan una búsqueda eficiente de documentación y datos. |
| 3. Aplicaciones | <ul style="list-style-type: none"> • Asistentes legales que facilitan el acceso rápido y preciso a jurisprudencia y casos precedentes, permitiendo a los profesionales del derecho analizar eficientemente grandes volúmenes de información legal para fundamentar sus argumentos. • Como apoyo al diagnóstico médico, permitiendo a los profesionales de la salud encontrar y comparar casos clínicos similares, acceder a historiales médicos relevantes y analizar antecedentes. • Para la gestión del conocimiento en medios y periodismo, unificando el acceso a archivos multimedia y textuales, facilitando la investigación periodística, la contextualización de las noticias. • Para la recuperación de información corporativa, optimizando la búsqueda y acceso a trabajos, proyectos e informes internos acumulados a lo largo del tiempo, mejorando la colaboración entre equipos y evitando la duplicación de esfuerzos. • Como soporte en procesos comerciales, agilizando la elaboración de presupuestos y propuestas comerciales mediante el acceso rápido a ofertas anteriores y datos históricos de productos o servicios, facilitando así la toma de decisiones estratégicas en ventas y marketing. |
| 4. Nivel de madurez tecnológica (TRL) | El nivel de madurez de un buscador personalizado basado en modelos de lenguaje de gran tamaño (LLMs) varía significativamente dependiendo del contexto lingüístico y de aplicación. La tecnología base de LLMs está bastante establecida para idiomas con muchos recursos, alcanzando un TRL de 7 a 8 en escenarios más favorables. Sin embargo, cuando se trata de adaptar estos modelos a idiomas con pocos recursos como el euskera o el gallego, o a dominios altamente especializados como el jurídico o el médico, así como a tareas específicas de análisis visual en estos campos, el TRL puede descender hasta 2 o 3. |

BÁSICOS

- **Modelo de lenguaje de gran tamaño (LLM):** Es el componente central que genera representaciones vectoriales de textos y consultas. Si es multimodal, procesa también contenido visual; si es multilingüe, permite búsquedas entre diferentes idiomas.
- **Corpus de datos internos:** Documentos y datos organizacionales que conforman la base de conocimiento para búsquedas. También sirve para adaptar el modelo de lenguaje al vocabulario técnico específico del sector.
- **Herramientas de procesamiento de documentos con formato:** Convierten documentos en formatos como PDF, DOCX o PPT a texto plano, permitiendo su vectorización e indexación por modelos de lenguaje.
- **Motor de indexación y recuperación de vectores:** Sistema que indexa y recupera representaciones vectoriales de documentos, facilitando búsquedas semánticas eficientes en grandes volúmenes de datos.

OTROS

- **Modelo de visión por computadora:** Complementa a LLMs no-multimodales procesando contenido visual. Extrae representaciones vectoriales de imágenes y vídeos para su indexación y búsqueda.
- **Herramientas de preprocesamiento de imágenes y vídeos:** Software para extraer fotogramas clave, realizar OCR en imágenes con texto y preparar contenido visual para su análisis. Especialmente relevante con LLMs no-multimodales.

5.

¿Qué recursos o herramientas se necesitan?

1. **Eficiencia y escalabilidad:** Mantener tiempos de respuesta rápidos en búsquedas en tiempo real es un reto, dado el alto consumo de recursos de los LLMs. Es necesario optimizar el rendimiento del sistema para manejar consultas concurrentes y datos en crecimiento continuo.
2. **Actualización continua:** El sistema debe actualizarse sin reconfiguraciones completas, ajustando el LLM de forma incremental para aprender nuevos datos y contextos, y mantener la relevancia ante cambios en la distribución de datos ("data drift") o conceptuales ("concept drift").
3. **Manejo de documentos complejos:** Indexar semánticamente documentos largos y variados es clave para la eficacia del buscador. Esto requiere métodos que representen estructuras, fragmentos multilingües, tablas, listas y notas, devolviendo resultados relevantes y contextualizados.
4. **Documentos de baja calidad:** El sistema debe manejar documentos en diversos estados de conservación y calidad, incluyendo fotos o escaneados de documentos mecanizados o manuscritos. Es fundamental contar con procesamiento de imágenes robusto para condiciones adversas.
5. **Integración multilingüe efectiva:** Para un sistema de búsqueda multilingüe, debe garantizarse la comprensión y procesamiento homogéneos en todos los idiomas, proporcionando respuestas precisas y coherentes sin importar el idioma de las consultas o documentos.
6. **Alineación semántica entre modalidades:** Es crucial crear una representación semántica unificada para relacionar texto e imágenes, permitiendo búsquedas cruzadas que encuentren imágenes a partir de texto o viceversa, mediante comprensión multimodal y espacios vectoriales compartidos.

6.

Retos científico-técnicos

Automatización de procesos ofimáticos

Tecnología principal: PLN y visión por computadora

Aplicación de alto impacto: Automatización multilingüe de la gestión documental en licitaciones públicas en comunidades con lenguas cooficiales

- | | |
|---|---|
| ¿Qué necesidades cubre? | 1. La automatización de procesos ofimáticos aborda la demanda de eficiencia y productividad en entornos empresariales y organizativos. Esta tecnología busca agilizar tareas administrativas repetitivas, reduciendo errores humanos y liberando tiempo para actividades de mayor valor añadido . Su implementación permite una gestión más efectiva de la información, mejoras en la toma de decisiones y una respuesta más rápida a las necesidades de la organización. |
| Usuarios | 2. Esta tecnología está dirigida a un amplio grupo de usuarios, incluidas empresas de todos los tamaños, organizaciones gubernamentales, instituciones educativas y profesionales independientes que manejan grandes volúmenes de datos y documentación o que se dedican a la gestión logística. La automatización de procesos ofimáticos abarca aplicaciones muy diversas, que pueden agruparse en dos grandes bloques tecnológicos: |
| Aplicaciones | 3. <ul style="list-style-type: none">• Clasificación automática de tipos de documentos relevantes al usuario o sector de aplicación (p. ej., facturas, albaranes y recibos, pero también currículums, títulos educativos, formularios de inscripción, etc.)• Extracción automática de datos específicos y metadatos contenidos en dichos documentos, para su volcado automático en las bases de datos correspondientes y el desencadenamiento automático de los procesos pertinentes. |
| Nivel de madurez tecnológica (TRL) | 4. El nivel de madurez para este tipo de aplicaciones basado en modelos de lenguaje de gran tamaño (LLMs) varía significativamente dependiendo del contexto lingüístico y de aplicación. La tecnología base de LLMs está bastante establecida para idiomas con muchos recursos, alcanzando un TRL de 7 a 8 en escenarios más favorables. Sin embargo, cuando se trata de adaptar estos modelos a idiomas con pocos recursos como el euskera o el gallego, o a dominios altamente especializados como el jurídico o el médico, así como a tareas específicas de análisis visual en estos campos, el TRL puede descender hasta 2 o 3. |

BÁSICOS

- **Modelo de lenguaje de gran tamaño (LLM):** Se trata del componente que dota al sistema con la capacidad para procesar y generar representaciones matemáticas de los documentos. Si es multimodal, además puede procesar y generar representaciones para contenido visual (imágenes y vídeos). Si es multilingüe, permitirá automatizar el procesamiento de documentos en varios idiomas.
- **Corpus de documentos ofimáticos:** Colecciones de documentos y datos específicos de la organización que representen la variedad de formatos y tipos de documentos que se manejan en los procesos ofimáticos a automatizar (facturas, formularios, informes, etc.). Este corpus sirve tanto para adaptar el modelo de lenguaje al dominio específico como para probar y evaluar el sistema.
- **Herramientas de procesamiento de documentos con formato:** A menudo, las organizaciones almacenan su conocimiento en documentos legibles para humanos como PDFs, DOCX o PPTs, formatos que no son directamente procesables por modelos de lenguaje basados en texto, por lo que es necesario procesarlos con herramientas que extraen texto plano para su posterior procesamiento.

OTROS

- **Modelo de visión por computadora:** Complementa al LLM no-multimodal para el análisis y procesamiento de contenido visual. Este modelo se encarga de extraer representaciones vectoriales de imágenes y vídeos, permitiendo su posterior tratamiento automático por los algoritmos de clasificación o extracción de la información.
- **Herramientas de preprocesamiento de imágenes y vídeos:** Software especializado para extraer fotogramas clave de vídeos, realizar reconocimiento óptico de caracteres (OCR) en imágenes con texto, y preprocesar imágenes para su análisis por el modelo de visión. Estas herramientas cobran especial relevancia si el LLM no es multimodal.

5.

¿Qué recursos o herramientas se necesitan?

1. **Eficiencia y escalabilidad:** Es necesario diseñar arquitecturas y estrategias de implementación que permitan el uso eficiente de LLMs multimodales en entornos empresariales. Esto implica la reducción de los tiempos de respuesta y la capacidad de manejar grandes volúmenes de documentos simultáneamente, todo ello manteniendo la calidad y precisión del procesamiento.
2. **Privacidad y seguridad:** Debido a la presencia de información sensible como nombres, números de identificación, datos médicos o información confidencial de empresas, el acceso a corpus de datos reales necesarios para el entrenamiento puede ser limitado o inexistente. Este reto implica desarrollar estrategias para obtener y utilizar datos representativos sin comprometer la privacidad, posiblemente mediante la creación de datos sintéticos o mediante la anonimización de datos reales.
3. **Variabilidad en formatos y contenidos:** Los sistemas deben de ser capaces de procesar y extraer información relevante de documentos con estructuras y contenidos muy diversos dentro de una misma categoría (p. ej., diferentes formatos de facturas por composición, colores, tamaños de letra, etc.).
4. **Documentos de baja calidad:** La eficacia del sistema depende de su capacidad para manejar documentos en diversos estados de conservación y calidad, incluyendo fotografías o escaneados de calidad variable de documentos tanto mecanizados como manuscritos. Para ello, se requieren sistemas de procesamiento de imágenes y algoritmos de OCR robustos que puedan funcionar con fiabilidad en condiciones subóptimas.

6.

Retos científico-técnicos

CONCLUSIONES

En este apartado se presentan las conclusiones derivadas del estudio, que sintetiza los datos y opiniones recabadas a través de entrevistas y cuestionarios. El análisis revela cómo se utilizan y valoran actualmente las herramientas basadas en Tecnologías del Lenguaje. Nuestro estudio se ha centrado tanto en analizar los usos actuales e ideales de estas herramientas como en identificar las necesidades y expectativas manifestadas por los usuarios.

A continuación, se presentan las principales conclusiones para cada una de las tres bloques tecnológicos analizados:

Traducción automática

- La traducción automática en euskera ha alcanzado un nivel de desarrollo notable, hasta el punto de que las necesidades básicas pueden considerarse satisfechas. En consecuencia, las entrevistas han puesto el foco principalmente en nuevos desarrollos y aspectos clave para áreas especializadas. Los participantes señalan que el principal reto actual reside en mejorar la precisión y calidad del servicio, minimizando al máximo los errores.

Procesamiento de voz

- Esta tecnología ha adquirido un protagonismo destacado en la mayoría de las entrevistas, donde se reconoce el habla como capacidad intrínseca humana y elemento fundamental para los desarrollos futuros. Los resultados indican que la tecnología está más avanzada de lo que se percibe o utiliza, incluso en euskera. En este ámbito, resulta significativo que tanto usuarios como profesionales del sector de las Tecnologías del Lenguaje han centrado su atención en los demostradores actuales y en las necesidades ya cubiertas. La existencia de diversos proyectos de colaboración entre usuarios y desarrolladores ha difuminado las diferencias de perspectiva entre ambos grupos.

Modelos de Lenguaje de Gran Tamaño (LLM)

- En el ámbito de los usuarios, se observa que la implementación propia es aún limitada, tanto en organizaciones industriales y de servicios como en el sector público y académico. El uso actual se concentra principalmente en aplicaciones personales, destacando especialmente las soluciones Copilot y ChatGPT. Sin embargo, existe un consenso claro sobre su potencial como tecnología clave para el futuro. Los participantes identifican múltiples aplicaciones prometedoras: desde la automatización y la integración de hardware en distintas herramientas hasta la implementación de soluciones para usos diversos. También se destaca su potencial para incrementar el valor organizacional, ya sea mejorando la comunicación o en el propio diseño de productos, así como su aplicación en diversos objetivos estratégicos. Se percibe como un campo con gran potencial de evolución, aunque se reconoce la existencia de importantes desafíos por abordar.

En las siguientes líneas se recogen algunas conclusiones generales:

1. Las tecnologías de traducción automática y los correctores lideran el uso actual.

Estas herramientas se utilizan principalmente para mejorar la productividad y la comunicación. La percepción general de las personas encuestadas es que la importancia e influencia de las Tecnologías del Lenguaje seguirán creciendo, hasta convertirse en herramientas imprescindibles en numerosos ámbitos de la vida cotidiana.

La percepción general de las personas encuestadas es que la importancia e influencia de las Tecnologías del Lenguaje seguirán creciendo

2. Diferentes sectores presentan necesidades y desafíos específicos.

El estudio ha combinado cuestionarios generales con entrevistas presenciales sectorizadas, comenzando por las industrias lingüísticas como referentes en el campo, y extendiéndose a los sectores industrial, judicial, sanitario, público, medios de comunicación y educativo. Cada sector presenta necesidades particulares y realidades diversas.

- Los sectores **judicial y sanitario** enfatizan la necesidad de datos, aunque la sensibilidad de su información requiere un desarrollo tecnológico especialmente cuidadoso. El sector público también prioriza la protección de datos.
- En el sector **industrial**, aunque la protección de datos es relevante, son los requisitos de seguridad laboral los que condicionan principalmente la implementación tecnológica, incluyendo aspectos de ciberseguridad y medidas preventivas.
- El sector **educativo** presenta necesidades variables. Si bien valora positivamente las tecnologías emergentes como apoyo al aprendizaje, enfrenta el desafío de la fiabilidad de las fuentes y la gestión de la incertidumbre.

3. Es necesario aumentar el nivel de conocimiento y concienciación de la sociedad, y para ello se requiere un proceso de cambio de hábitos.

Se ha detectado un bajo nivel de conocimiento sobre digitalización en las organizaciones, siendo crucial mejorar la comprensión de su utilidad y aplicación. Esto implica replantearse hábitos establecidos en diversos ámbitos. En el caso específico del euskera, aunque las tecnologías lingüísticas son fundamentales para su normalización, se requiere el compromiso de trabajar en este idioma y una formación específica para su uso eficaz en las organizaciones.

4. El euskera cobra gran relevancia respecto al uso y despliegue de estas tecnologías, tanto en la cuestión de los desafíos como en los casos de uso futuros.

El estudio ha prestado especial atención al euskera, dado que constituye la primera lengua de la mayoría de los participantes y su principal herramienta en el uso de Tecnologías del Lenguaje. Los resultados indican un nivel de uso comparable al castellano, con un desarrollo que satisface las necesidades básicas. Sin embargo, se identifica la necesidad de reforzar capacidades y asegurar un uso adecuado en ámbitos especializados y nuevos desarrollos.

5. De cara al futuro, los objetivos, preocupaciones y desafíos de estas tecnologías cobran un gran protagonismo.

Los principales retos identificados incluyen:

- La necesidad garantizar precisión y fiabilidad en los sistemas.
- El imperativo de gestionar adecuadamente la diversidad lingüística, un desafío que se intensificará por la creciente incorporación de nuevas lenguas y niveles de uso.
- El desarrollo de sistemas que sean adaptables a entornos computacionales reducidos y que faciliten la integración con otros sistemas, especialmente con los ya existentes en las organizaciones.
- La importancia de optimizar los procesos para mejorar la experiencia del usuario final.
- El reto fundamental de obtener corpus de datos suficientes, especialmente relevante ante la creciente tendencia hacia la personalización, que requerirá conjuntos de datos específicos para diferentes dominios.

6. El criterio económico es el que ha recibido una mayor importancia en las valoraciones, mientras que el criterio social es el que menor peso tiene.

Se pueden ver esos resultados en la tabla de ponderación. Hay que tener en cuenta que el principal perfil que ha participado en las entrevistas y cuestionarios son empresas u organizaciones que compiten en el mercado. Además, hemos entrado en su espacio laboral, pues tanto las entrevistas como los cuestionarios se han realizado mediante el email de trabajo o en el propio lugar de trabajo y, por tanto, es comprensible que el componente económico sobresalga. Hemos hecho las entrevistas en un momento en el que está pensando en la productividad y, en consecuencia, se han buscado soluciones para ser más eficientes, y en muchos casos nuestras entrevistas también se han dirigido en ese sentido.

7. Aun así, la perspectiva social es fundamental al considerar las Tecnologías del Lenguaje, dado su significativo potencial de impacto en la sociedad.

En efecto, las Tecnologías del Lenguaje se han convertido en elementos clave de nuestra vida cotidiana y continuarán su evolución, por lo que resulta importante tener presente la perspectiva social. A partir de las entrevistas a profesionales del sector y del análisis de distintos informes, se evidencia que los criterios sociales son fundamentales en la investigación y desarrollo de estas tecnologías. En el estudio de los demostradores actuales, se han identificado tecnologías orientadas a la inclusión. Como ejemplo, un organismo del sector ha desarrollado una solución que permite al alumnado con dificultades motoras realizar redacciones mediante dictado. Siguiendo esta línea, se están desarrollando soluciones que mejoran la calidad de vida de diversos colectivos: sistemas basados en voz para asistir a personas mayores, tecnologías de clonación de la voz para personas que han perdido el habla tras un accidente o enfermedad, y sistemas que facilitan la comunicación.

Las Tecnologías del Lenguaje se han convertido en elementos clave de nuestra vida cotidiana y continuarán su evolución, por lo que resulta importante tener presente la perspectiva social

Si bien los criterios sociales no han tenido una presencia destacada en las entrevistas y cuestionarios, estos adquieren gran importancia en los informes analizados, constituyendo un criterio que no se debe subestimar. De hecho, al hablar de inteligencia artificial en contextos cotidianos, uno de los aspectos más relevantes es precisamente su potencial para beneficiar a la sociedad.

En consecuencia, este informe, además de presentar demostradores de casos de uso de alto impacto, busca enfatizar la importancia de la relación entre la digitalización y la perspectiva social.

Anexos

Anexo A: Guión de entrevistas

Anexo B: Hoja de autorización para la grabación entrevistas

Anexo C: Cuestionarios

Anexo A: Guión de entrevistas

Objetivo

- Completar un menú de posibles aplicaciones - demostradores – mapa actual / futuros
- Menú de posibles criterios para una adecuada medición de impactos
- Identificar posible casos avanzados de uso - nuevos sectores o ámbitos de aplicación

Actualmente

1. ¿Qué tecnologías estáis desarrollando/trabajando?
2. ¿Qué hay actualmente disponible? ¿En el mercado?
3. ¿Cuál ha sido el nivel de implementación? ¿Se ha transferido a empresas o entidades?
4. ¿A qué sectores o ámbitos son extensibles estas aplicaciones?

Retos y limitaciones actuales

5. ¿Cuáles son los principales retos actuales en el campo del procesamiento del lenguaje natural?
 - Mejorar la eficacia de la comunicación, cuidar la corrección del texto
 - Para aumentar mi nivel de productividad
 - Enfocar acciones de diferentes tipos porque me resulta más cómodo por voz
 - Personalizar las interacciones con los usuarios
 - Automatización de tareas repetitivas
 - Ampliar el alcance de los servicios lingüísticos
6. ¿Cuáles son los principales obstáculos o retos que encontráis en vuestros proyectos de desarrollo?

Promoción de lenguas minoritarias

7. ¿Cómo están gestionando las tecnologías la diversidad lingüística?
8. En cuanto a las lenguas minoritarias (euskera, p. ej.), ¿cuáles son las principales ventajas de las Tecnologías del Lenguaje?
 - Aumentar la exposición de las personas a estas lenguas
 - Mejorar las capacidades comunicativas de las personas con discapacidad
 - Mejorar el comercio en los países donde se hablan esas lenguas
 - Incrementar el compromiso de las personas con las actividades sociales, de ocio y profesionales de sus lenguas
 - Mejorar la alfabetización de las lenguas minoritarias o regionales
 - Mejorar la comunicación entre hablantes de diferentes lenguas
9. ¿Qué aplicación podría crearse que tuviera un impacto en la comunidad vasca?
10. ¿Cómo pueden ayudar las Tecnologías del Lenguaje en la supervivencia y promoción de las lenguas minoritarias?
11. ¿Cuáles son los retos que se encuentran a la hora de digitalizar estas lenguas?

Interdisciplinariedad

12. ¿Con qué otra disciplina crees que debería haber una mayor colaboración en el desarrollo de las Tecnologías del Lenguaje?

Tendencias de futuro

13. ¿Qué dirección tomará la investigación en Tecnologías del Lenguaje en las próximas décadas?
14. ¿Cuáles son las principales tendencias de las Tecnologías del Lenguaje para los próximos años?
15. ¿Qué nuevas aplicaciones o mejoras se pueden esperar?

Casos de Uso y Aplicaciones Prácticas

16. ¿Cuáles son, en tu opinión, los casos de uso más afectados por las Tecnologías del Lenguaje en la actualidad? ¿Por qué?
17. ¿Cómo te imaginas el papel de las Tecnologías del Lenguaje en la sociedad y en la tecnología dentro de 10 años? ¿Qué cambios o aplicaciones esperas?
18. ¿Qué soluciones/aplicaciones desarrollarías utilizando las Tecnologías del Lenguaje existentes si tuvieras todos los recursos necesarios?
19. ¿Qué recursos priorizarías en este bloque si tuvieras todos los recursos necesarios?

Criterios y Medición de Impacto

20. ¿Qué criterios tenéis en cuenta a la hora de elegir las Tecnologías del Lenguaje?
 - Inclusión y accesibilidad
 - Equidad lingüística
 - Protección de datos y privacidad
 - Responsabilidad social y medioambiental
21. ¿Cómo elegirías un proyecto u otro?
 - I. ¿Qué priorizarías entre lo económico, lo social y lo técnico? ¿Por qué?
 - II. ¿Cuáles son los principales criterios para decidir la puesta en marcha o desarrollo de un proyecto en el sector o en tu organización en la actualidad?

Anexo B: Hoja de autorización para la grabación entrevistas

	Emun, Koop.E. Datu Pertsonalen Babesa Irudia/ahotsa tratatzeko baimena	Edizioa: 2024/04/01 Erref.: IRU-003 Orrialdea: 1 / 1
---	--	---

PROTECCIÓN DE DATOS PERSONALES - TRATAMIENTO DE IMAGEN Y VOZ

En cumplimiento de la legislación de Protección de Datos de Carácter Personal, le informamos que:

1. Toda la información recogida en esta entrevista se utilizará exclusivamente para el estudio de casos de aplicación de las Tecnologías Lingüísticas por parte de Langune y Emun y para los fines del proyecto GBL que coincide con este tema. La información será confidencial, independientemente de su soporte. La información no será pública ni se difundirá sin el consentimiento previo de las personas entrevistadas.
2. De acuerdo con la normativa de protección de datos de carácter personal, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril ("Reglamento General de Protección de Datos" o "RGPD") y la Ley Orgánica de Protección de Datos de Carácter Personal. Toda la información que contenga datos personales será anonimizada tan pronto como sea posible. Asimismo, todo soporte que contenga datos personales e información confidencial será destruido una vez redactado el informe final del proyecto.
3. En la utilización y elaboración de la información se anonimizarán cualquier organismo o empresa citando expresiones del tipo "la empresa suministradora" los "órganos competentes o superiores". Tampoco se citarán localizaciones de empresas u organismos al objeto de que no sean identificables.
4. Los datos personales recogidos o tratados en el proyecto no podrán ser cedidos a terceros.
5. Este compromiso de confidencialidad se mantendrá también una vez finalizada la relación entre Langune-Emun y las personas entrevistadas.
6. La aceptación de este documento supone la autorización expresa a EMUN para la realización de los tratamientos y comunicaciones descritos. La base legal para el tratamiento de sus datos es su consentimiento.

En, a de..... de 2024

Interesado/a
Nombre y apellidos:
DNI:

Proveedor/a de servicios
Nombre y apellidos:
DNI:

Anexo C: Cuestionarios

Introducción

Las Tecnologías del Lenguaje basadas en la inteligencia artificial y en el procesamiento del lenguaje natural han evolucionado notablemente en las últimas décadas permitiendo aplicaciones como la traducción automática, el conocimiento y síntesis de la voz, el análisis y la creación de textos propios de grandes modelos lingüísticos como ChatGPT, etc. Estas tecnologías las utilizamos en nuestro día a día, tanto en el ámbito personal como profesional. Una gestión inadecuada de las lenguas puede invisibilizar la diversidad lingüística de nuestra sociedad, es por ello que el **proyecto ILENIA** pretende desarrollar aplicaciones multilingües en las lenguas del estado español. Desde **LANGUNE**, Asociación de Industrias del Lenguaje, queremos abrir un espacio de reflexión a través del siguiente cuestionario.

- **Para qué:** Identificar a través de las Tecnologías del Lenguaje los casos de uso de alto impacto socio-económico que puedan tener en la sociedad, en el sector industrial o en otros sectores. Determinar lo que se hace actualmente para, a partir de ahí, detectar nuevas zonas o usos extensibles.
- **Cómo:** Preguntas cerradas y abiertas. Es totalmente anónimo y se utilizará exclusivamente para este estudio. Se trata de visualizar nuevos ámbitos de uso futuro a partir del uso actual.
- **Qué:** Se ha puesto en marcha un estudio y este cuestionario será un canal para escuchar a agentes de diferentes sectores. El resultado del estudio se recogerá en un informe público.
- **Plazo para cumplimentar el cuestionario:** 27 de mayo de 2024

Información general

[C-1] Edad: ☐ 18-24 ☐ 25-34 ☐ 35-44 ☐ 45-54 ☐ 55-64 ☐ >=65	[C-2] Género: ☐ Mujer ☐ Hombre ☐ No binario ☐ Otro (<i>específica</i>)	[C-3] Lengua materna: ☐ Euskera ☐ Castellano ☐ Francés ☐ Otro (<i>específica</i>)	[C-4] Sector: *Sólo C.2 ☐ Industrial ☐ Servicios ☐ Medios de comunicación ☐ Académico ☐ Administración pública ☐ Otro (<i>específica</i>)
---	--	---	--

Sintonizando con el tema...

- [C-5] ¿Qué Tecnologías del Lenguaje utilizas en tu día a día?
- Ordena las opciones de mayor a menor.*
- ≡ Aplicaciones de asistente personal, órdenes por voz (p. ej. Siri, Alexa)
 - ≡ Aplicaciones de corrección de textos (p. ej. ortografía y corrección gramatical)
 - ≡ Aplicaciones de traducción (p. ej. Itzuli, Elia, Google Translate, DeepL)
 - ≡ Subtitulación automática (p. ej. reportaje, YouTube)
 - ≡ Chatbots (p. ej. atención al cliente)
 - ≡ Lectores de pantalla
 - ≡ Generador de contenidos (p. ej. ChatGPT, Claude, Dall·E, Gemini)

[C-6] ¿Cuál de los siguientes beneficios es el que priorizas para justificar su uso?

Elige 2.

- ☐ Mejorar la eficacia de la comunicación, cuidar la corrección del texto
- ☐ Para aumentar mi nivel de productividad
- ☐ Porque a través de la voz me resulta más cómodo dirigir acciones de diferente tipo
- ☐ Personalizar las interacciones con los usuarios
- ☐ Automatización de tareas repetitivas
- ☐ Aumentar el alcance de los servicios lingüísticos
- ☐ Otro (*especifica*)

[C-7] En su uso diario, ¿en qué idioma los utilizas principalmente?

Ordena las opciones de mayor a menor.

- ≡ Euskera
- ≡ Castellano
- ≡ Francés
- ≡ Inglés

[C-8] ¿Por qué?

Elige todas las opciones que apliquen.

- ☐ Por ser el idioma habitual que uso en mi día a día
- ☐ Por realizar mis actividades profesionales en dicha lengua
- ☐ Porque hay más herramientas tecnológicas en dicha lengua
- ☐ Porque es una lengua en la que tengo poco conocimiento
- ☐ Otro (*especifica*)

En el trabajo...

C.1 Industria de las tecnologías de la lengua

[C.1-9] ¿Qué tecnología lingüística habéis desarrollado o aplicado y qué uso le habéis dado?

Responde brevemente.

[C.1-10] ¿Cuáles son los principales obstáculos o retos que encontráis en vuestros proyectos de desarrollo? *Responde brevemente.*

[C.1-11] ¿A qué sectores o ámbitos son extensibles estas aplicaciones?

Responde brevemente.

[C.1-12] ¿Cuáles crees que son los principales retos de futuro en relación con las Tecnologías del Lenguaje? *Elige 3.*

- ☐ Mejorar la accesibilidad e inclusión, facilitar la vida
- ☐ Facilitar la comunicación y colaboración
- ☐ Automatizar la clasificación y análisis de grandes volúmenes de datos textuales y de voz para la toma de decisiones
- ☐ Desarrollo de interfaces usuales más intuitivas y naturales
- ☐ Optimizar los procesos de prestación del servicio al cliente, por ejemplo, para la comprensión y resolución de consultas complejas
- ☐ Fomentar la innovación y la competitividad en el mercado
- ☐ Mejorar la seguridad de los datos utilizando tecnologías de detección de fraudes y análisis de riesgo basados en el lenguaje (biometría, identificación vocal...)

☐ Otro (*específica*)

[C.1-13] ¿Cuáles crees que son las principales opciones de aplicación no exploradas en el campo de las Tecnologías del Lenguaje? *Responde brevemente.*

[C.1-14] ¿Qué solución/aplicación inventarías si tuvieras todos los recursos necesarios? *Responde brevemente.*

[C.1-15] ¿Cuáles son actualmente los principales criterios para decidir si poner en marcha el proyecto de desarrollo en el sector o en su organización? *Elige tantas opciones como apliquen.*

- ☐ Criterios económicos
- ☐ Criterios sociales
- ☐ Criterios tecnológico-técnicos

[C.1-16] ¿Con qué instituciones colabora su organización en el desarrollo de las Tecnologías del Lenguaje? *Elige tantas opciones como apliquen.*

- ☐ Academia
- ☐ Organizaciones industriales y de servicios
- ☐ Administración
- ☐ Otro (*específica*)

[C.1-17] ¿Cuáles son las principales fuentes de financiación de su organización para el desarrollo de nuevas Tecnologías del Lenguaje? *Ordena las opciones de mayor a menor.*

- ≡ Subvenciones públicas (Europa, España, País Vasco...)
- ≡ Fondos propios (ventas, autofinanciación)
- ≡ Aportación de los inversores

[C.1-18] Comentarios / observaciones / opiniones

C.2 Empresas industriales y de servicios

[C.2-9] ¿Cómo/cuánto influyen actualmente las Tecnologías del Lenguaje en su ámbito de trabajo?
1 = *nada*, 10 = *totalmente*.

[C.2-10] Y el grado de influencia futura...
Elige 1.

- ☐ Irá aumentando
- ☐ Se mantendrá igual
- ☐ Descenderá

[C.2-11] ¿En qué áreas concretas utilizáis las Tecnologías del Lenguaje en la organización?
Elige tantas opciones como apliquen.

- ☐ Atención al cliente y asistencia técnica
- ☐ Para la gestión y selección de recursos humanos
- ☐ Análisis de mercado y visión de las opiniones
- ☐ Para procesos de automatización administrativa y documental
- ☐ Desarrollo y mejora de productos
- ☐ Mejorar la comunicación interna y la cooperación
- ☐ Facilitar el desarrollo y la formación
- ☐ Automatización del cumplimiento legal y normativo

☐ Otro (*específica*)

[C.2-12] ¿Cómo habéis integrado o aplicado las Tecnologías del Lenguaje en vuestra organización con otras tecnologías o sistemas? *Responde brevemente.*

[C.2-13] ¿Cuáles son los principales obstáculos o retos que encontráis a la hora de dicha aplicación? *Responde brevemente.*

[C.2-14] ¿Cuáles crees que son los principales retos de futuro en relación con las Tecnologías del Lenguaje? *Elige 3.*

- ☐ Mejorar la accesibilidad e inclusión, facilitar la vida
- ☐ Facilitar la comunicación y colaboración
- ☐ Automatizar la clasificación y análisis de grandes volúmenes de datos textuales y de voz para la toma de decisiones
- ☐ Desarrollo de interfaces usuales más intuitivas y naturales
- ☐ Optimizar los procesos de prestación del servicio al cliente, por ejemplo, para la comprensión y resolución de consultas complejas
- ☐ Fomentar la innovación y la competitividad en el mercado
- ☐ Mejorar la seguridad de los datos utilizando tecnologías de detección de fraudes y análisis de riesgo basados en el lenguaje (biometría, identificación vocal...)
- ☐ Otro (*específica*)

[C.2-15] Teniendo en cuenta los 3 principales objetivos que has elegido en la pregunta anterior, ¿qué uso/aplicación se te ocurre? *Responde brevemente.*

[C.2-16] ¿Cuáles crees que son los principales retos a la hora de implantar las Tecnologías del Lenguaje en una empresa como la tuya? *Elige máximo 4.*

- ☐ Integración con los sistemas y plataformas tecnológicas existentes
- ☐ Garantizar la seguridad y privacidad de los datos
- ☐ Garantizar la precisión y fiabilidad de los sistemas de procesamiento del lenguaje
- ☐ Adaptación a las necesidades específicas de la empresa y de los clientes
- ☐ Capacitar al personal para utilizar eficazmente las nuevas herramientas y tecnologías
- ☐ Gestión de la diversidad lingüística
- ☐ Aprobación de nuevas tecnologías y gestión del cambio
- ☐ Medición y justificación del tipo de reembolso de la inversión
- ☐ Otro (*específica*)

[C.2-17] ¿Cuáles crees que son las principales opciones de aplicación no exploradas en el campo de las Tecnologías del Lenguaje y en vuestro sector/sociedad? *Responde brevemente.*

[C.2-18] ¿Qué solución/aplicación inventarías si tuvieras todos los recursos necesarios? *Responde brevemente.*

[C.2-19] Comentarios / observaciones / opiniones

C.3 Sector académico

[C.3-9] ¿Cómo crees que las Tecnologías del Lenguaje pueden favorecer el proceso educativo?

Elige tantas opciones como apliquen.

- ☐ Mejorando la disponibilidad de materiales educativos para el alumnado con diferentes habilidades y necesidades
- ☐ Facilitando la traducción automática de los recursos educativos para el alumnado internacional
- ☐ Automatizando la retroalimentación en las evaluaciones escritas, como ensayos y proyectos
- ☐ Personalizando el aprendizaje mediante sistemas de tutoría virtual basados en TIs
- ☐ Proporcionando enseñanzas más interactivas utilizando tecnologías de reconocimiento y conversión de voz
- ☐ Impulsando la inclusión del alumnado con diferentes perfiles
- ☐ Facilitando la colaboración y comunicación en tele-educación
- ☐ Desarrollando el aspecto lúdico de las competencias lingüísticas
- ☐ Otro (*especifica*)

[C.3-10] ¿Cuál crees que es el mayor reto de las Tecnologías del Lenguaje en la educación o en la sociedad? *Elige tantas opciones como apliquen.*

- ☐ Falta de infraestructuras tecnológicas adecuadas en las instituciones educativas.
- ☐ La resistencia al cambio del profesorado y de los y las administradoras educativas.
- ☐ Necesidad de formación y capacitación específica para el uso eficiente de estas tecnologías.
- ☐ Preocupación por la integridad académica y el plagio en el uso de sistemas automatizados.
- ☐ Reconocimiento de la diversidad lingüística y de la interculturalidad
- ☐ Privacidad y seguridad de datos
- ☐ Coste y accesibilidad de estas tecnologías
- ☐ Otro (*especifica*)

[C.3-11] ¿A cuáles de los siguientes criterios le darías prioridad?

Ordena las opciones de mayor a menor.

- ≡ Explicar qué hace la aplicación con los datos personales
- ≡ Descripción clara del impacto ambiental de la aplicación
- ≡ Calidad y nivel de euskera de la aplicación
- ≡ La “sencillez” de la aplicación para llegar al mayor número de personas
- ≡ No dar sesgos en los resultados de la aplicación (p. ej. sesgo de género)

[C.3-12] ¿Qué tecnología lingüística o aplicación inventarías para afrontar un reto social?

Responde brevemente.

[C.3-13] ¿Qué otras aplicaciones innovadoras se pueden crear con las Tecnologías del Lenguaje?

Responde brevemente.

[C.3-14] Comentarios / observaciones / opiniones



IMPULSO DE LAS LENGUAS
EN LA INTELIGENCIA ARTIFICIAL

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

HiTZ

Hizkuntza Teknologiako Zentroa
Basque Center for Language Technology