



ODESIA

Proyecto Espacio de Observación de Inteligencia Artificial en Español

Ámbito 3. Nivel de adopción. Informe Año 1.

1. INTRODUCCIÓN

En este documento se presenta el informe sobre el nivel de adopción de las tecnologías de la lengua por parte de empresas y ciudadanos. Este estudio se ha desarrollado en el marco del proyecto del Espacio de Observación de Inteligencia Artificial en Español, en concreto "Ámbito 3 Nivel de Adopción", y recoge las conclusiones del primer año.

En los últimos tiempos, la Inteligencia Artificial (IA) basada en lenguaje se ha convertido en un tema cada vez más relevante en el mundo, y se espera que su impacto en la forma en que las empresas operan y compiten continúe creciendo en el futuro. En este trabajo nos centramos en dar respuesta a dos preguntas fundamentales: ¿Se están adoptando estas herramientas de manera dispar en las empresas de habla hispana e inglesa? ¿Existen diferencias de adopción entre los ciudadanos de ambos idiomas?

Para contestar estas preguntas, se realiza un análisis de la adopción de herramientas de tecnologías del lenguaje en empresas de habla hispana e inglesa, utilizando una variedad de fuentes y métodos. En primer lugar, se identifican una serie de empresas representativas y se analizan menciones de herramientas de IA en sus presentaciones e informes de resultados corporativos. A continuación, se realiza un análisis similar de las menciones en medios de comunicación y se calcula el impacto de la adopción de tecnologías del lenguaje. Además, se llevan a cabo encuestas en España y Estados Unidos para conocer el grado de adopción de herramientas por parte de los ciudadanos. Finalmente, se presentan las conclusiones de este estudio, y se calcula la brecha existente en el nivel de adopción.

En este documento se presentan los resultados obtenidos para cada indicador. La metodología seguida para desarrollar y aplicar los indicadores, así como los detalles de los cálculos realizados para obtener los resultados se describen en el documento "Ámbito 0.2 Diseño y cálculo de la métrica agregada para medir la brecha español/inglés en tecnologías de la lengua. Informe Año 1".

1. LISTADO DE EMPRESAS

Para medir el nivel de adopción de las soluciones y tecnologías por parte de las empresas se han tomado las empresas de la [Tabla 1](#) como referencia. El criterio de selección ha sido el de mayor capitalización bursátil, de esta manera se capta la realidad de las grandes empresas en cada uno de los países. Se seleccionan 20 compañías del IBEX-35 [1] y 20 compañías de S&P 500 [2].

	IBEX 35	S&P 500
1	Inditex	Apple
2	Iberdrola	Microsoft
3	Santander	Alphabet
4	BBVA	Amazon
5	Caixabank	Berkshire Hathaway
6	Naturgy	UnitedHealth Group
7	Amadeus	Johnson & Johnson
8	Cellnex	Exxon Mobil
9	ArcelorMittal	Tesla
10	Telefonica	Visa
11	Repsol	NVIDIA
12	Endesa	Walmart
13	Ferrovial	JPMorgan Chase
14	Aena	Lilly
15	Siemens Gamesa	Procter & Gamble
16	Acciona	Chevron
17	Red Eléctrica	MasterCard
18	IAG	Home Depot
19	Grifols	Meta
20	ACS	Pfizer

Tabla 1. Empresas seleccionadas para el análisis del nivel de adopción.

2. ANÁLISIS DE MENCIONES EN PRESENTACIONES E INFORMES DE RESULTADOS CORPORATIVOS

En esta sección se analizan las menciones de las soluciones de mercado y de tecnologías del lenguaje en las presentaciones e informes de resultados corporativos de las empresas seleccionadas en el apartado anterior y publicados en los tres últimos años (2020, 2021 y 2022).



Seleccionar los tres últimos años garantiza un volumen suficiente para que las brechas calculadas sean estadísticamente significativas.

Menciones de soluciones en informes corporativos

En el lado de las soluciones de mercado, las áreas de aplicación y las herramientas analizadas son aquellas seleccionadas en el “Ámbito 2 Soluciones de mercado”.

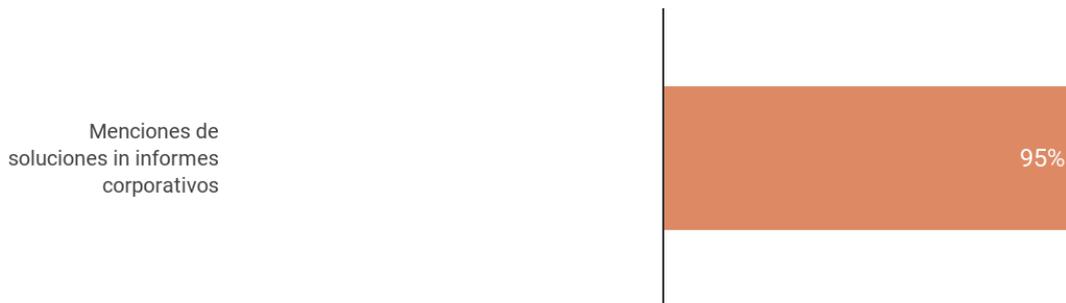


Figura 1. Brecha en menciones de soluciones en informes corporativos.

En la [Figura 1](#) se muestra la brecha en menciones de soluciones en informes corporativos. 19 de las 48 soluciones seleccionadas han sido desarrolladas por cuatro de las empresas que aparecen en la [Tabla 1](#): Apple, Microsoft, Alphabet y Amazon. 30 de las 38 menciones identificadas en S&P se corresponden con soluciones desarrolladas por estas cuatro empresas. Solo se han encontrado 2 menciones a soluciones en las presentaciones e informes de resultados corporativos de las empresas españolas y, también, son menciones a una de las soluciones desarrolladas por una de estas cuatro empresas. El hecho de tener empresas que son desarrolladoras y adoptantes de las soluciones puede tener un impacto en la brecha, pero no hay volumen de coincidencias suficientes para determinarlo, ya que al eliminar dichas soluciones de la ecuación nos quedamos con un número de menciones muy reducido: 8 para empresas de S&P y 0 para empresas del IBEX 35.

Es posible que la brecha medida en el indicador I.A.1 tenga dos motivos principales:

1. Que las empresas erradicadas en Estados Unidos están adoptando más herramientas relevantes desarrolladas por terceros.
2. Que las empresas erradicadas en Estados Unidos están adoptando más herramientas relevantes desarrolladas por ellos mismos.



No obstante, debido al reducido número de coincidencias, no podemos calcular la brecha en los dos componentes.

Menciones de tecnologías del lenguaje en informes corporativos

Las tecnologías del lenguaje se han identificado mediante una bolsa de palabras que contiene 181 expresiones y que permite identificar 20 grupos de tecnologías diferentes:

- Análisis de sentimientos
- Análisis sintáctico
- Análisis de texto
- Análisis morfológico
- Comprensión del lenguaje natural
- Generación de resúmenes
- Lingüística de corpus
- Lingüística estadística
- Menciones de PLN no específicas
- Modelado de lenguaje
- Procesamiento de voz
- Reconocimiento de entidades
- Reconocimiento óptico de caracteres
- Respuesta a preguntas
- Semántica del discurso
- Semántica léxica
- Semántica relacional
- Sistemas de diálogo
- Sistemas generativos
- Traducción

La bolsa de palabras que se ha utilizado para identificar cada una de las tecnologías puede encontrarse en el documento "Ámbito 0.2 Diseño y cálculo de la métrica agregada para medir la brecha español/inglés en tecnologías de la lengua. Informe Año 1".

Solo se han identificado 9 menciones en los informes de las empresas de S&P y 4 menciones en los informes de las empresas del IBEX 35.

En la [Figura 2](#) se muestra la brecha en menciones de tecnologías del lenguaje en informes corporativos.

INDICADOR I.A.2: BRECHA EN MENCIONES DE TECNOLOGÍAS DEL LENGUAJE EN INFORMES CORPORATIVOS

56%



Figura 2. Brecha de menciones de tecnologías del lenguaje en informes corporativos.

3. ANÁLISIS DE MENCIONES EN MEDIOS DE COMUNICACIÓN

Adicionalmente, se han analizado las menciones en los medios de comunicación de las soluciones de mercado y de tecnologías del lenguaje en los medios de comunicación de las empresas seleccionadas en los tres últimos años (2020, 2021 y 2022).

Menciones de soluciones en medios de comunicación

En total, las empresas de S&P seleccionadas han sido mencionadas un total de 87.826.500 veces en los medios, y las empresas del IBEX 35 seleccionadas han sido mencionadas un total de 5.486.395 veces en los medios. Para evitar sesgar las medidas por la popularidad de las empresas o la capacidad de generar noticias de los medios en un país u otro, las menciones serán normalizadas dividiendo por el número total de noticias de las empresas de cada índice.

La tabla con las menciones de cada una de las soluciones en medios de comunicación se encuentra en el documento "Ámbito 0.2 Diseño y cálculo de la métrica agregada para medir la brecha español/inglés en tecnologías de la lengua. Informe Año 1".





Figura 3. Brecha en menciones de soluciones en medios de comunicación por área de aplicación.

En la [Figura 3](#) se muestra la brecha en menciones de soluciones en medios de comunicación para cada una de las áreas de aplicación, cuya media da una brecha del 76%.

Al igual que en el análisis de menciones en informes corporativos, parte de la brecha corresponde a la adopción de herramientas de terceros y parte de la brecha corresponde a la adopción de herramientas propias. En este caso sí podemos detallar la brecha en en cada una de las dos situaciones.

En la parte superior de la [Figura 4](#), se muestra la brecha en menciones de herramientas de terceros, aquellas que no han sido desarrolladas por ninguna empresa de la [Tabla 1](#); y, en la parte inferior, la brecha en menciones de las 19 herramientas que han sido desarrolladas por Apple, Microsoft, Alphabet o Amazon. Como se puede observar, la brecha es un 15% mayor en el segundo caso.



Figura 4. Brecha en menciones de soluciones en medios de comunicación por tipo de solución.

Menciones de tecnologías del lenguaje en medios de comunicación

La brecha de menciones de tecnologías del lenguaje también es muy significativa, aunque algo más reducida que la de soluciones:



Figura 5. Brecha en menciones de tecnologías del lenguaje en medios de comunicación.

En la [Figura 5](#) se muestra la brecha en menciones de tecnologías del lenguaje en medios de comunicación.

Tal y como se puede ver la [Figura 6](#), los grupos de tecnologías que tienen una brecha a favor del español son: lingüística estadística, asociado a los términos del tratamiento cuantitativo y

estadístico de textos; y respuesta a preguntas y sistemas de diálogo, que son dos de las tecnologías necesarias para crear soluciones como chatbots. En los otros 17 grupos de tecnologías existe una brecha a favor del inglés y es especialmente acentuada para: análisis sintáctico, semántica de discurso, semántica léxica, modelado de lenguaje, generación de resúmenes, procesamiento de voz y sistemas generativos.

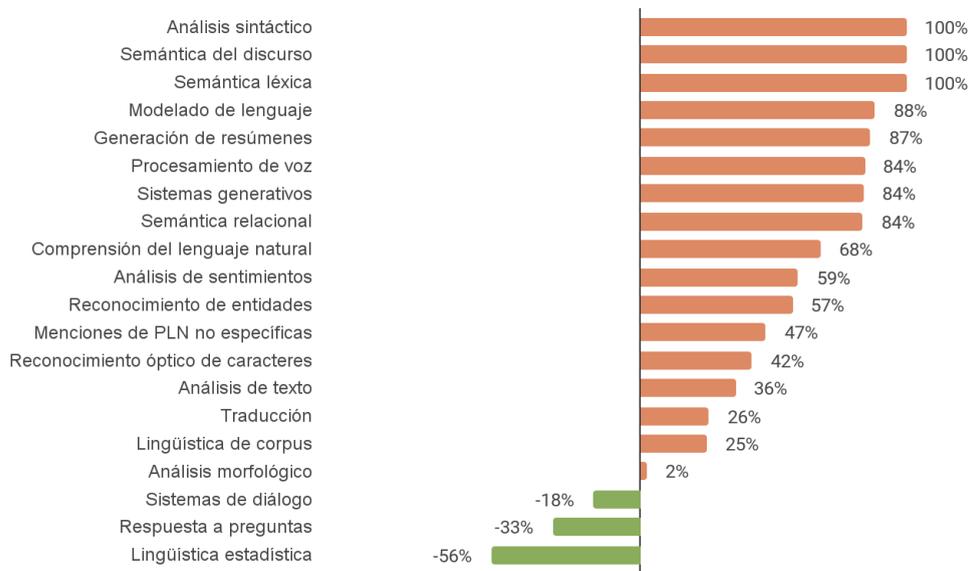


Figura 6. Brecha en menciones en medios de comunicación de cada una de las tecnologías del lenguaje identificadas.

Impacto de adopción de tecnologías del lenguaje

El impacto que tiene la adopción de cada uno de los grupos de tecnologías se ha calculado aplicando el criterio experto de científicos de datos y atendiendo a las siguientes características:

1. **Proximidad al mercado:** se refiere a lo cerca que se encuentra una tecnología de ser adoptada en el mercado. Esta proximidad puede ser alta si ya existen soluciones similares en el mercado o si no requiere de otras tecnologías para llegar a construir una solución; y será baja si no existen soluciones comercialmente viables en el mercado o si requiere la combinación de muchas otras tecnologías para llegar a construir una solución.
2. **Viabilidad técnica:** se refiere a si la tecnología puede ser desarrollada y utilizada en la práctica, teniendo en cuenta factores como la disponibilidad de recursos técnicos y financieros, la complejidad del desarrollo y los posibles obstáculos técnicos que puedan surgir durante el proceso de implementación.
3. **Demanda de mercado:** se refiere a la cantidad de clientes potenciales que están interesados en la tecnología y a la cantidad de ventas o ingresos que se pueden esperar de la adopción de la tecnología.

4. Potencial disruptivo: se refiere a la capacidad de la tecnología para alterar significativamente un mercado o industria existente, ya sea creando nuevas oportunidades o amenazando la existencia de soluciones y productos existentes.

Siguiendo este criterio, se ha otorgado la categoría mostrada en la [Tabla 2](#) a cada una de las tecnologías identificadas.

Grupo de tecnologías	Impacto
Análisis de sentimientos	alto
Análisis sintáctico	bajo
Análisis de texto	bajo
Análisis morfológico	bajo
Comprensión del lenguaje natural	muy alto
Generación de resúmenes	medio
Lingüística de corpus	muy bajo
Lingüística estadística	bajo
Menciones de PLN no específicas	medio
Modelado de lenguaje	bajo
Procesamiento de voz	alto
Reconocimiento de entidades	medio
Reconocimiento óptico de caracteres	medio
Respuesta a preguntas	alto
Semántica del discurso	medio
Semántica léxica	bajo
Semántica relacional	bajo
Sistemas de diálogo	alto
Sistemas generativos	alto
Traducción	alto

Tabla 2. Impacto asignado a cada una de las tecnologías.

Para otorgar un valor numérico a cada una de las categorías se ha seguido el criterio de medición de impacto del modelo de puntuación RICE [3]. El modelo de puntuación RICE es un marco diseñado para determinar qué productos, características y otras iniciativas priorizar según cuatro factores: alcance, impacto, confianza y esfuerzo. Siguiendo el criterio establecido por el modelo para la medición del impacto, se dará el siguiente valor numérico a cada una de las categorías:

- Muy bajo: 0.25
- Bajo: 0.5
- Medio: 1

- Alto: 2
- Muy alto: 3



Figura 7. Brecha en impacto de las tecnologías en la empresa.

En la [Figura 7](#) se muestra la brecha en impacto de las tecnologías en la empresa. La brecha medida del impacto de las tecnologías del lenguaje ([Figura 7](#)) es un 17% más grande que en el caso de la adopción ([Figura 5](#)).

4. ENCUESTAS

Se han realizado 900 encuestas en Estados Unidos sobre el uso de soluciones en inglés y 900 encuestas en España sobre el uso de soluciones en español, preguntando por el nivel de adopción, cuyos resultados se recogen en este apartado, y la experiencia de usuario, cuyos resultados se recogen en el “Ámbito 4 Experiencia de usuario”.

A cada una de las 1800 personas encuestadas se le ha preguntado por todas las soluciones de las 5 áreas de aplicación. Se ha seleccionado la cifra de 900 para cada idioma porque era el mínimo necesario para tener resultados estadísticamente significativos.

Las preguntas sobre el nivel de adopción de cada una de las áreas de aplicación se han dividido en dos bloques:

1. Si el encuestado ha utilizado alguna herramienta del área.
2. Si el encuestado ha utilizado alguna de las herramientas específicas seleccionadas en el “Ámbito 2 Soluciones de mercado”.

Los resultados de las preguntas del primer bloque se recogen en la [Figura 8](#), y podemos observar que existen diferencias significativas a favor del uso de en inglés en el área de

buscadores web y, sobre todo, en el análisis de opiniones. Por el contrario, la brecha en el área de traducción automática es muy significativa a favor del español.

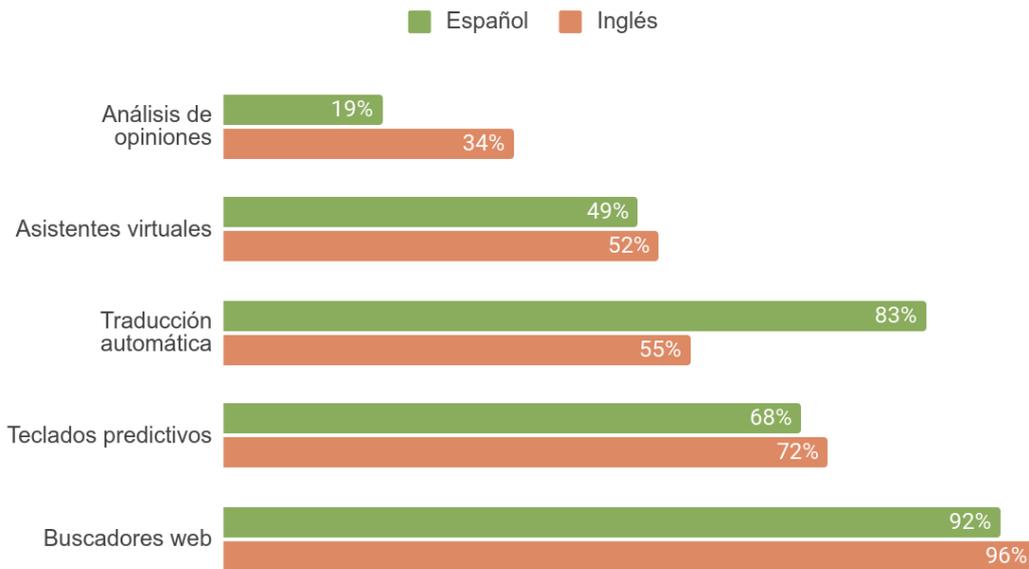


Figura 8. Primer bloque de la encuesta: adopción de cada una de las áreas de aplicación.

Respecto a la adopción de herramientas por género, en base a los datos que se recogen en la [Tabla 3](#), se observa poca diferencia entre hombres y mujeres. Entre los usuarios de herramientas en español, sólo se observa mayor adopción de teclados predictivos en las mujeres. Entre los usuarios de herramientas en inglés, hay mayor proporción en los hombres en herramientas de traducción automática y análisis de opinión.

	español			inglés		
	Total	Masculino	Femenino	Total	Masculino	Femenino
Análisis de opiniones	18.8%	19.1%	18.4%	34.4%	39.4%	29.4%
Asistentes virtuales	49.1%	46.6%	51.3%	51.6%	55.6%	47.7%
Traducción automática	83.2%	81.8%	84.4%	55.3%	60.8%	50.1%
Teclados predictivos	68.4%	63.6%	72.9%	71.6%	71.8%	71.3%



Buscadores web	92.0%	91.4%	92.7%	96.3%	95.5%	97.1%
Promedio	62.3%	60.5%	63.9%	61.8%	64.6%	59.1%

Tabla 3. Primer bloque de la encuesta: adopción de cada una de las áreas de aplicación por género.

Respecto a la edad, como se puede observar en la [Tabla 4](#), entre los usuarios de herramientas en español la adopción de herramientas es similar entre los grupos de edad, excepto para los asistentes virtuales, que son menos usados por las personas de 46 años o más. Entre los usuarios de herramientas en inglés, la situación se muestra diferente con una adopción claramente mayor entre los más jóvenes, que se reduce a medida que aumenta la edad, en todas las herramientas excepto en el uso de buscadores web.

	español				inglés			
	Total	18 a 29 años	30 a 45 años	46 años o más	Total	18 a 29 años	30 a 45 años	46 años o más
Análisis de opiniones	18.8%	20.8%	18.9%	16.8%	34.4%	53.3%	35.5%	14.3%
Asistentes virtuales	49.1%	53.5%	53.2%	40.8%	51.6%	65.5%	56.5%	33.2%
Traducción automática	83.2%	85.8%	84.2%	79.6%	55.3%	72.0%	58.1%	35.9%
Teclados predictivos	68.4%	71.3%	68.7%	65.1%	71.6%	86.3%	73.4%	55.1%
Buscadores web	92.0%	92.1%	91.2%	92.8%	96.3%	95.0%	97.0%	97.0%
Promedio	62.3%	64.7%	63.2%	59.0%	61.8%	74.4%	64.1%	47.1%

Tabla 4. Primer bloque de la encuesta: adopción de cada una de las áreas de aplicación por rango de edad.

En las siguientes subsecciones se detalla el segundo bloque que consiste en el análisis de la adopción de cada una de las soluciones específicas y el cálculo de los indicadores. El nivel de adopción de las soluciones seleccionadas en cada uno de los idiomas, diferenciando si ha sido para uso personal o profesional, se encuentra detallado en el documento "Ámbito 0.2 Diseño y cálculo de la métrica agregada para medir la brecha español/inglés en tecnologías de la lengua. Informe Año 1".

Análisis de opiniones

Dentro de su baja proporción general de adopción, sí observamos que Talkwalker destaca en ambos países/idiomas, aunque en Estados Unidos Linkfluence lo supera ligeramente. La solución de menor adopción en ambos casos es NetBase Quid.



En español, el uso personal o adopción ciudadana supera en todas las herramientas al uso profesional. En Estados Unidos, Linkfluence y Talkwalker muestran mucho más uso personal o adopción ciudadana que profesional.

Asistentes virtuales

Google Assistant y Alexa destacan en ambos países/idiomas con 7 de cada 10 personas como usuarios, con lo que muestra la existencia de personas con contacto con ambos asistentes. Siri presenta en Estados Unidos porcentajes similares a los de las otras dos herramientas. La adopción en Estados Unidos es similar a la adopción en España de Google Assistant, Alexa y Cortana. Para el resto la adopción es mucho mayor en Estados Unidos que en España.

Tanto en España como en Estados Unidos, la utilización personal o adopción ciudadana supera de manera muy importante y en todas las herramientas al uso profesional. En Estados Unidos, Google Assistant y Siri presentan la mayor cantidad de personas que les dan ambos usos (personal y profesional).

Traducción automática

Google Translate destaca en ambos países/idiomas. En España, la frecuencia con la que se mencionan otras herramientas es baja, mientras que en Estados Unidos las menciones a otras herramientas son mucho mayores.

En España, Google Translate destaca en el uso personal o adopción ciudadana superando al uso combinado (personal y profesional), mientras que en otras herramientas no se observan tendencias claras. Para el caso de Estados Unidos, también el uso personal o adopción ciudadana supera de manera importante al uso profesional, especialmente en Google Translate, Bing Translator y Amazon Translate. En Google Translate es importante la proporción de personas que la adoptan para ambos usos (personal y profesional) en Estados Unidos.

Teclados predictivos

Gmail destaca del resto, ocupando la primera posición en ambos países/idiomas, con proporciones similares. Para el resto de teclados predictivos, Estados Unidos tiene una adopción superior a España, excepto en Gboard.

Tanto en España como en Estados Unidos y para todos los teclados predictivos analizados, destaca el uso personal o adopción ciudadana frente al uso profesional y al combinado (personal y profesional). Las herramientas de Microsoft (Outlook y Office) son las de mayor uso profesional y combinado (personal y profesional).

Buscadores web

Google ocupa la primera posición en ambos países/idiomas, en proporciones similares. Para el resto de buscadores web, Estados Unidos tiene una adopción superior a España.

En español, Google destaca en el uso personal o adopción ciudadana y en el uso combinado (personal y profesional). Bing y Yahoo Search presentan porcentajes similares de adopción y muestran un importante nivel de uso personal. En Estados Unidos, para todos los buscadores, el uso personal o adopción ciudadana destaca frente al uso profesional y al combinado



(personal y profesional). Analizando a Google en inglés, observamos que es el buscador con mayor proporción de uso combinado (personal y profesional).

Brecha en la adopción de soluciones por los ciudadanos

Tal y como se puede observar en la [Figura 9](#), existe una brecha notable entre el inglés y el español del 80% en el uso de herramientas de análisis de opiniones y del 63% en el uso de asistentes virtuales para el uso profesional. También presentan un gap significativo a favor del inglés el resto de las áreas: 18% en traducción automática, 44% en teclados predictivos y 26 % en buscadores web. La media da una brecha en adopción de soluciones para uso profesional del 46%.

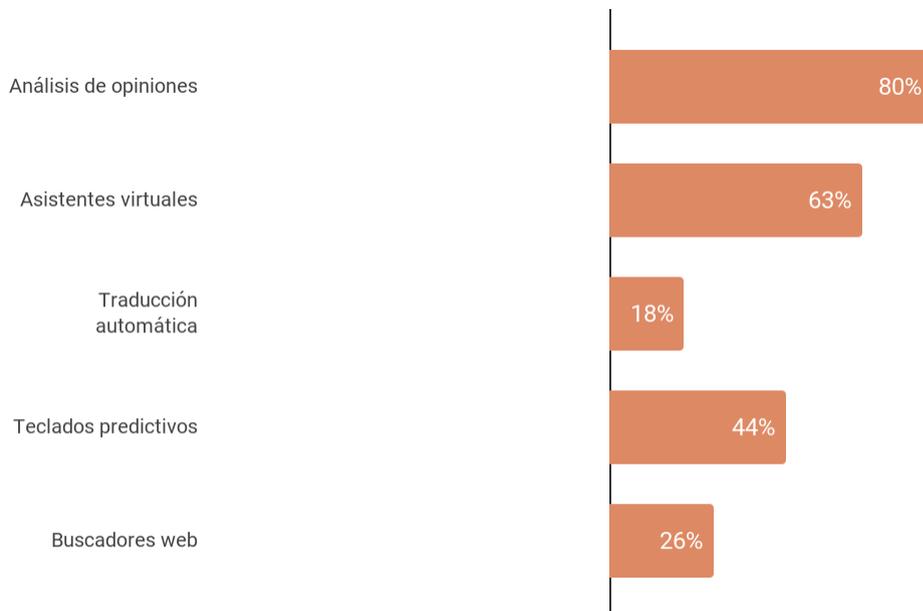


Figura 9. Segundo bloque de la encuesta: brecha en adopción de las soluciones de cada una de las áreas de aplicación para uso profesional.

Como se puede observar en la [Figura 10](#), las diferencias en el caso del uso personal se reducen en todos los casos salvo en el de buscadores web, que aumenta de un 26% a un 32%. La media



de cada una de las áreas de aplicación da una brecha en adopción de soluciones para uso personal del 33%.

INDICADOR I.A.7: BRECHA EN ADOPCIÓN DE SOLUCIONES PARA USO PERSONAL

33%

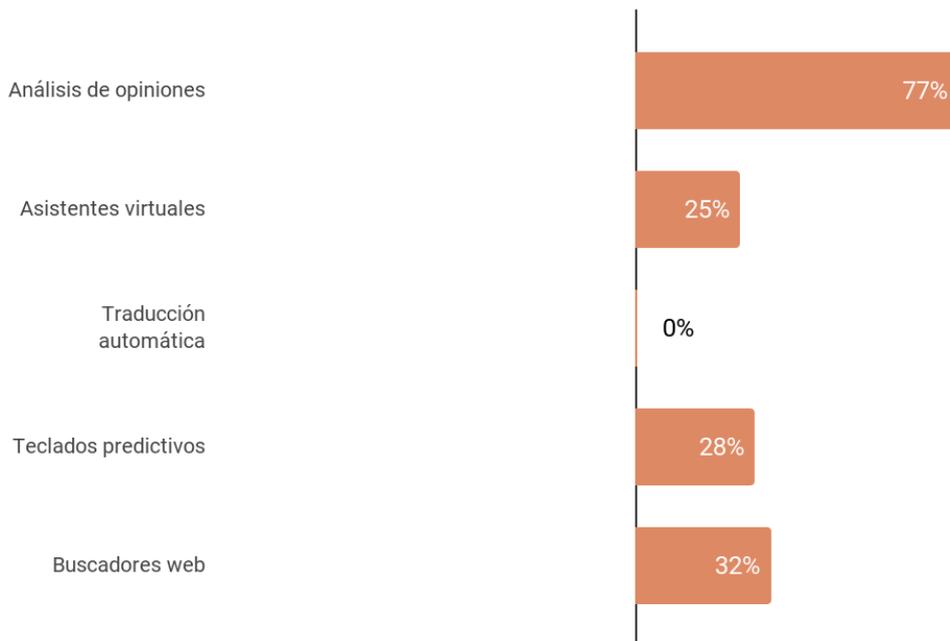


Figura 10. Segundo bloque de la encuesta: brecha en adopción de las soluciones de cada una de las áreas de aplicación para uso personal.

Cuando se pregunta por el uso de traductores de manera general (primer bloque), los usuarios de herramientas en español manifiestan un mayor uso que los usuarios de herramientas en inglés. Sucede lo contrario al preguntar por la utilización de herramientas específicas (segundo bloque) para uso profesional, donde los usuarios en inglés utilizan más los traductores que los usuarios en español. Por último, al preguntar por la utilización de herramientas para uso personal la brecha es nula entre ambos idiomas. Dada la gran cuota de mercado que abarcan las diez herramientas utilizadas, las diferencias a la hora de preguntar sobre el uso general o sobre herramientas particulares puede deberse a que los usuarios estadounidenses tienen un mayor recuerdo de las herramientas utilizadas que los usuarios españoles.



5. CONCLUSIONES

En base a los resultados obtenidos con la metodología aplicada para este estudio podemos establecer varios tipos de conclusiones.

Existe una brecha alta en la adopción de soluciones y tecnología de IA en español respecto al inglés. La brecha a la hora de medir la adopción mediante menciones de soluciones es superior al 75% en ambos indicadores. Al medir la adopción de tecnologías también tenemos brechas considerables de entorno al 50% con un impacto del 66%.

Los ciudadanos de habla inglesa también manifiestan hacer un mayor uso de soluciones de IA basadas en tecnologías del lenguaje que los ciudadanos de habla hispana. Esta brecha es mayor en el uso profesional de las herramientas (46%) que en el uso personal (33%).

En resumen, la brecha existente, sobre todo en el ámbito profesional y empresarial, puede generar una desventaja competitiva para las empresas y economías de habla hispana en el futuro. Por otro lado, existe una gran oportunidad de actuación en la adopción de tecnologías de lenguaje en empresas y organizaciones de habla hispana.

La media de los indicadores I.A.1, I.A.2, I.A.3, I.A.4, I.A.5, I.A.6 y I.A.7, nos da una brecha global de nivel de adopción del 60%. La metodología detallada con la que se han calculado los indicadores particulares y el agregado se encuentra en el documento "Ámbito 0.2 Diseño y cálculo de la métrica agregada para medir la brecha español/inglés en tecnologías de la lengua. Informe Año 1".

BRECHA EN NIVEL DE ADOPCIÓN

60%

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Unión Europea - NextGenerationEU a través del "Plan de Recuperación, Transformación y Resiliencia", por el Ministerio de Asuntos Económicos y Transformación Digital y por la UNED. Sin embargo, los puntos de vista y las opiniones expresadas son únicamente los del autor o autores y no reflejan necesariamente los de la Unión Europea o la Comisión Europea. Ni la Unión Europea ni la Comisión Europea pueden ser consideradas responsables de los mismos.

REFERENCIAS

- [1] «IBEX 35 Market capitalization | Markets Insider», *markets.businessinsider.com*. https://markets.businessinsider.com/index/market-capitalization/ibex_35 (accedido 23 de diciembre de 2022).
- [2] «S&P 500 Market capitalization | Markets Insider», *markets.businessinsider.com*. https://markets.businessinsider.com/index/market-capitalization/s&p_500 (accedido 23 de diciembre de 2022).
- [3] «RICE Scoring Model». <https://www.productplan.com/glossary/rice-scoring-model/>



(accedido 3 de marzo de 2023).

AUTOR



Beñat San Sebastián
Gerente de Deep Learning -
Innovación en LLYC