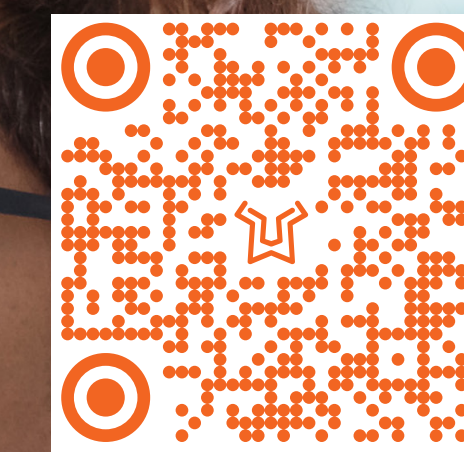


# EQUILIBRIO ENTRE INNOVACIÓN Y RIGOR: La postura de los titulares de derechos de autor sobre el uso de la IA generativa en la localización de su propiedad intelectual


**LIONBRIDGE**

 Autores: Melinda Johnson, Megan Duff, Kathryn Nolte, Stephanie Casale

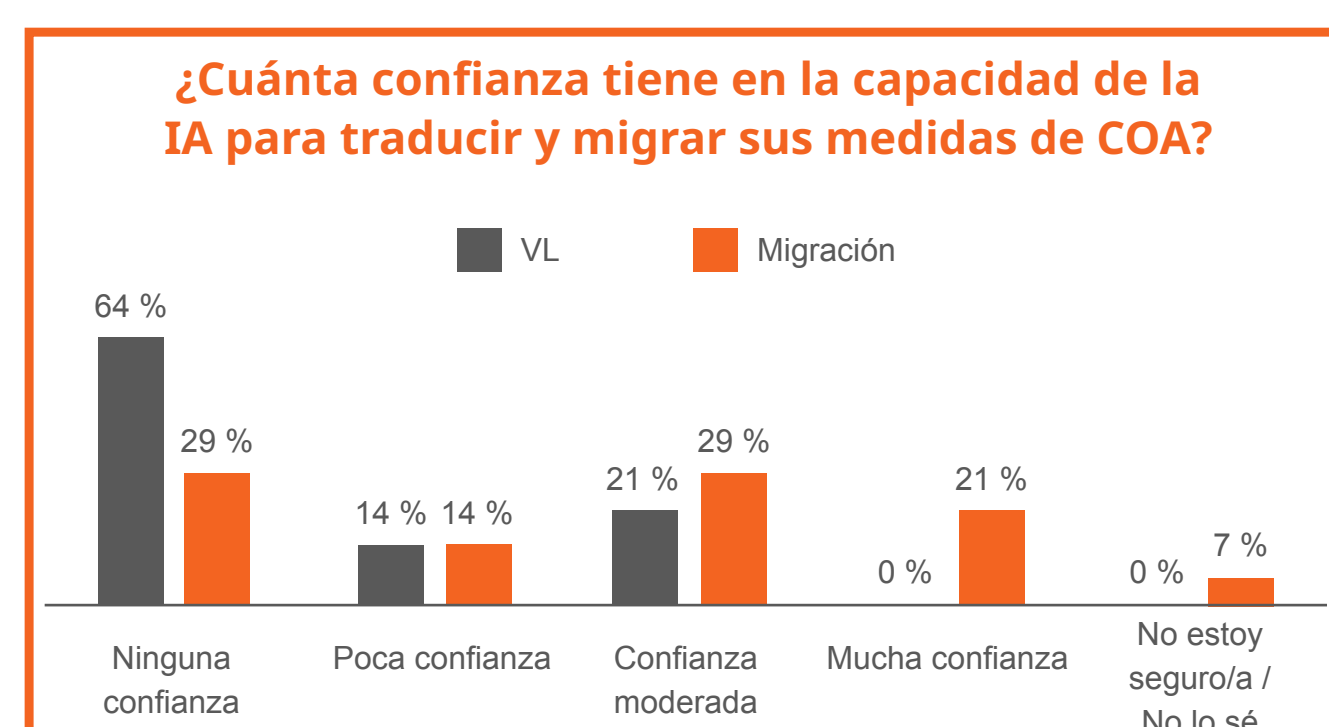
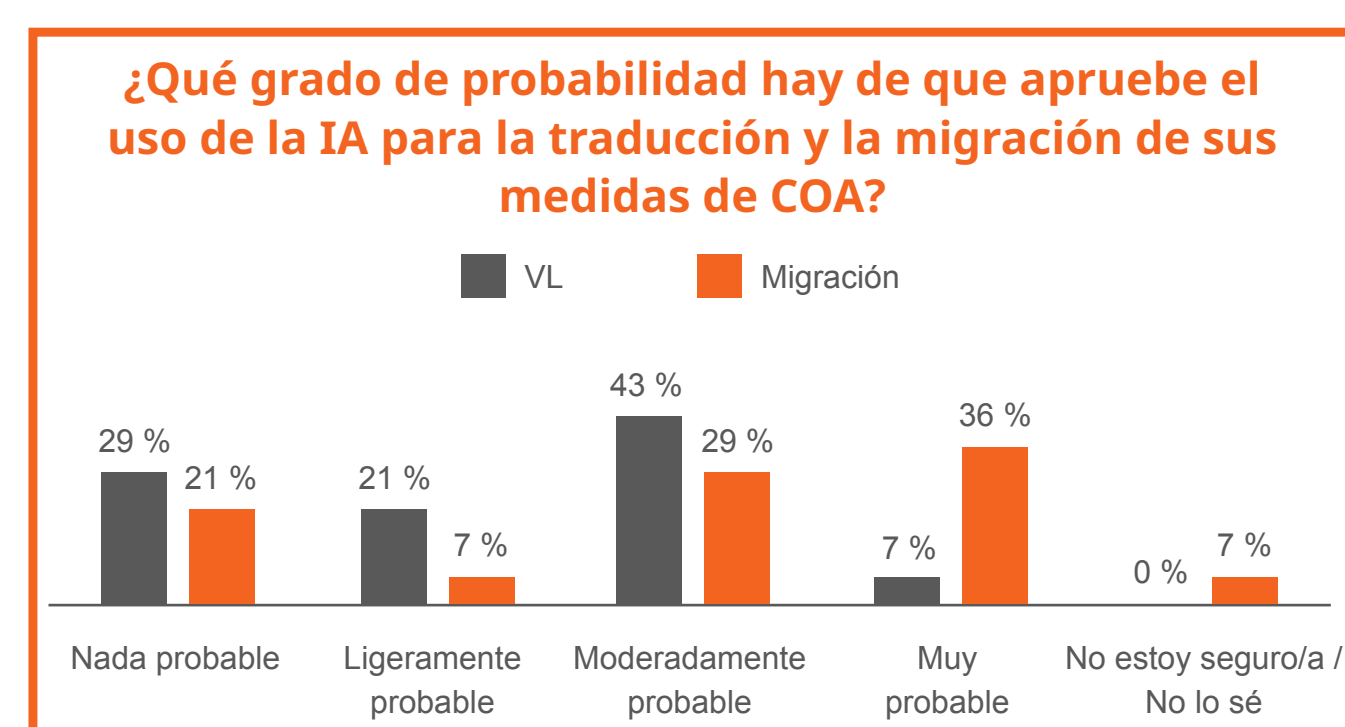
## OBJETIVO

La incorporación de la inteligencia artificial generativa (IA generativa) en las operaciones de ensayos clínicos, y en particular en las evaluaciones de resultados clínicos (COA, por sus siglas en inglés), se ve dificultada por la falta de directrices claras sobre su aplicación a la propiedad intelectual (PI) de estas evaluaciones. Esta incertidumbre ha ralentizado la adopción de la IA generativa en áreas críticas, como la validación lingüística (VL) y la migración de COA, procesos esenciales para incorporar la voz del paciente. **El presente estudio tuvo como objetivo recopilar y difundir las perspectivas de los titulares de derechos de autor de COA sobre el uso de la IA generativa en estos procesos, con el fin de orientar futuras estrategias de implementación.**

## RESULTADOS E INTERPRETACIONES

### 1) Aprobación y confianza en cuanto al uso de la IA en la VL y la migración

**a. Resultados:** el 64 % (9/14) de los titulares de derechos de autor de COA que respondieron a la encuesta indicaron no tener «Ninguna confianza» en la capacidad de la IA para traducir sus medidas de COA dentro del proceso de validación lingüística. No obstante, solo el 29 % (4/14) afirmaron que no era «Nada probable» que aprobasen el uso de la IA para traducir dichas medidas en el mismo contexto. En cuanto al proceso de migración de eCOA, las opiniones sobre aprobación y confianza estuvieron mucho más alineadas: el 29 % (4/14) declararon no tener «Ninguna confianza» y el 21 % (3/14) consideraron que no había «Ninguna probabilidad» de aprobar el uso de la IA.

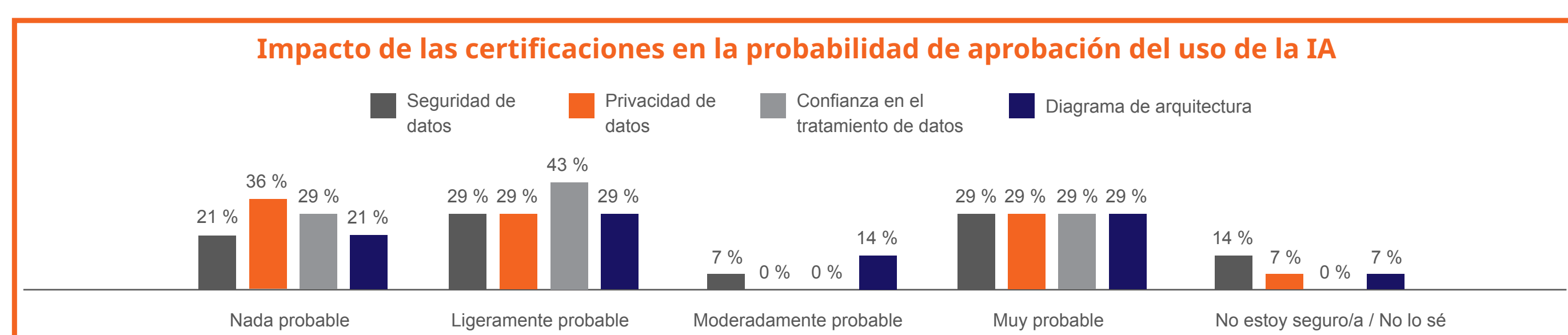


**b. Interpretación:** aunque el grado de confianza de los titulares de derechos de autor en las capacidades de la IA varió entre la VL y la migración, la predisposición a permitir el uso de la IA fue en general positiva en ambos flujos de trabajo, aunque con distintos grados de probabilidad («Moderadamente probable» frente a «Muy probable»), con un 50 % para la VL (7/14) y un 65 % para la migración (9/14). Las respuestas abiertas sugieren que, si bien la confianza general en la IA sigue siendo baja, la mayoría de los encuestados estarían dispuestos a aprobar su uso siempre que se definan claramente todas las condiciones y matices asociados a su aplicación.

### 2) Certificaciones y nivel de confianza

**a. Resultados:** se consultó a los titulares de derechos de autor sobre el impacto —es decir, la probabilidad de aprobar el uso de IA— de distintas certificaciones y documentación relacionada: 1) certificación de seguridad de datos (p. ej., ISO 27001:2022), 2) certificación de privacidad de datos (p. ej., ISO 27001:2019), 3) certificación de confianza en el tratamiento de datos (p. ej., SOC2 tipo 2), y 4) diagrama de arquitectura de IA. Los resultados fueron dispares, con un mayor escepticismo respecto al impacto de la privacidad de datos y la confianza en el tratamiento de datos. No obstante, en los cuatro tipos, el 29 % (4/14) de los encuestados consideraron «Muy probable» aprobar el uso de IA siempre que se contase con esta documentación.

**b. Interpretación:** aunque persiste cierto escepticismo en torno a los distintos tipos de documentación (certificaciones y diagramas), la seguridad de los datos, la privacidad y el propio funcionamiento del motor de IA influyen en que al menos el 29 % de los titulares de derechos de autor se muestren más dispuestos a aprobar el uso de la IA para la traducción de su propiedad intelectual.



## CONCLUSIÓN

Los titulares de derechos de autor de las medidas de COA se enfrentan a un sector que, al igual que otras partes implicadas, sigue careciendo de directrices claras sobre el uso de la IA en la localización de su propiedad intelectual. No obstante, aunque todavía existe margen para fortalecer la confianza de estas partes en el uso correcto de la IA durante la localización de su PI, **parecen mostrarse receptivos al diálogo, la formulación de políticas, la modificación de acuerdos de licencia y la revisión de la documentación asociada que permita verificar la seguridad de su PI al interactuar con un motor de IA.** Como consideración final, el número de participantes limitó este estudio, por lo que será necesario llevar a cabo nuevos análisis para confirmar estos resultados.

## MÉTODOS

Para lograrlo, se envió una encuesta electrónica de 15 preguntas a 100 titulares de derechos de autor de COA, que abarcaba los tres temas que se detallan a continuación. Catorce de ellos respondieron a esta encuesta anónima dentro del plazo previsto de tres semanas.

- 1) Aprobación y confianza en cuanto al uso de la IA en la traducción de COA (VL) y la migración de eCOA (migración)
- 2) Certificaciones y nivel de confianza
- 3) Políticas, enfoques y licencias actuales relacionadas con la IA

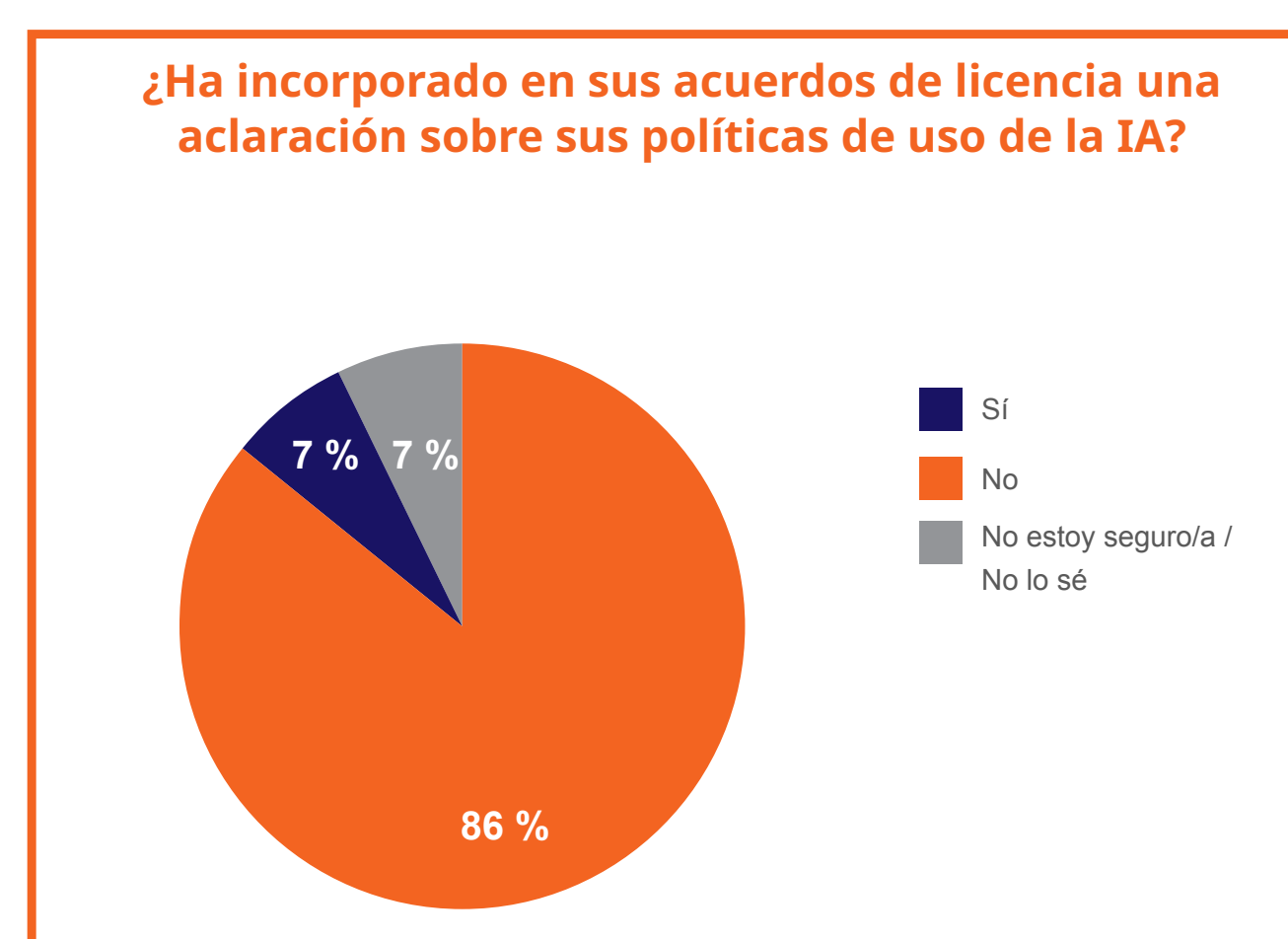
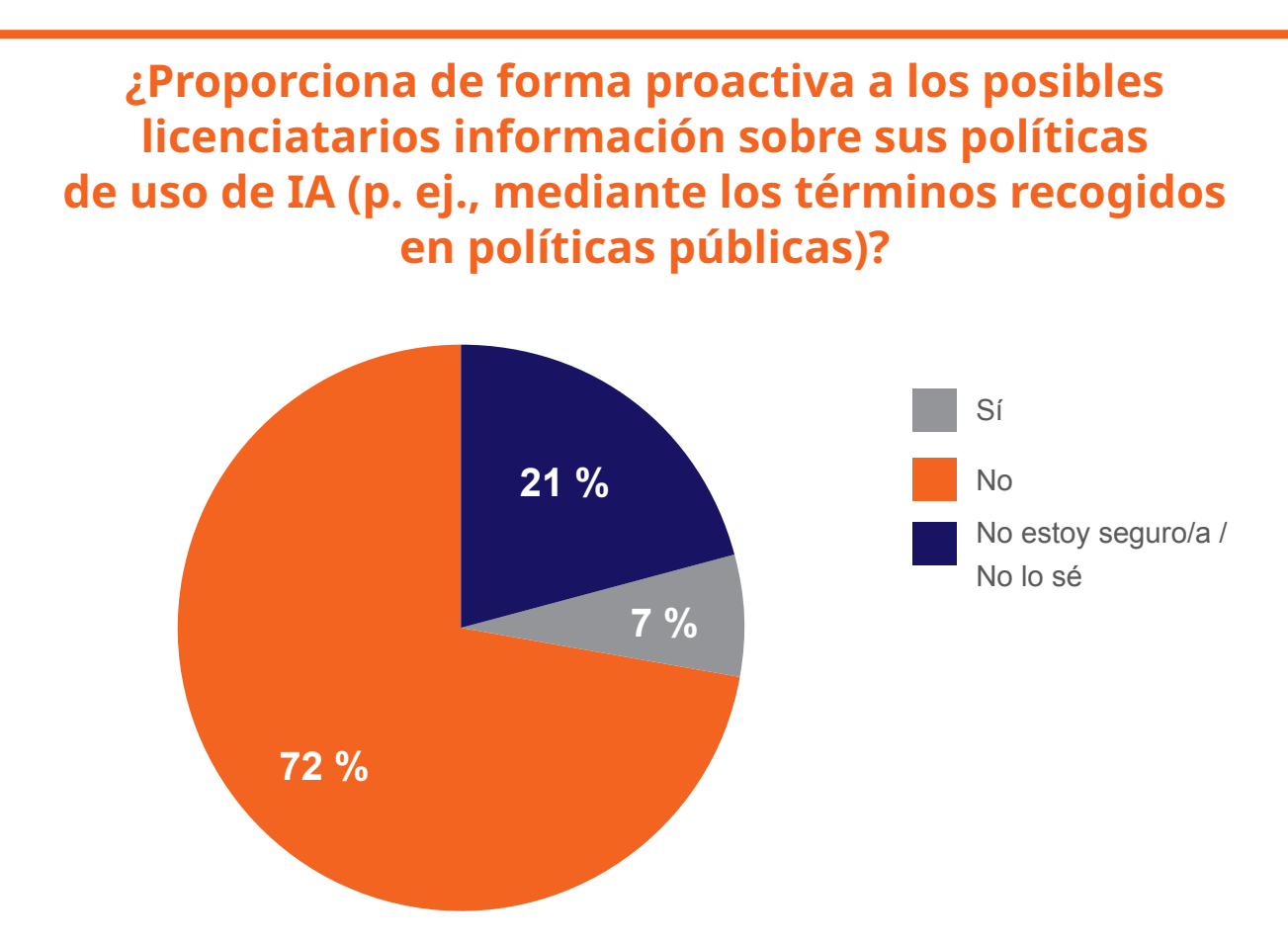
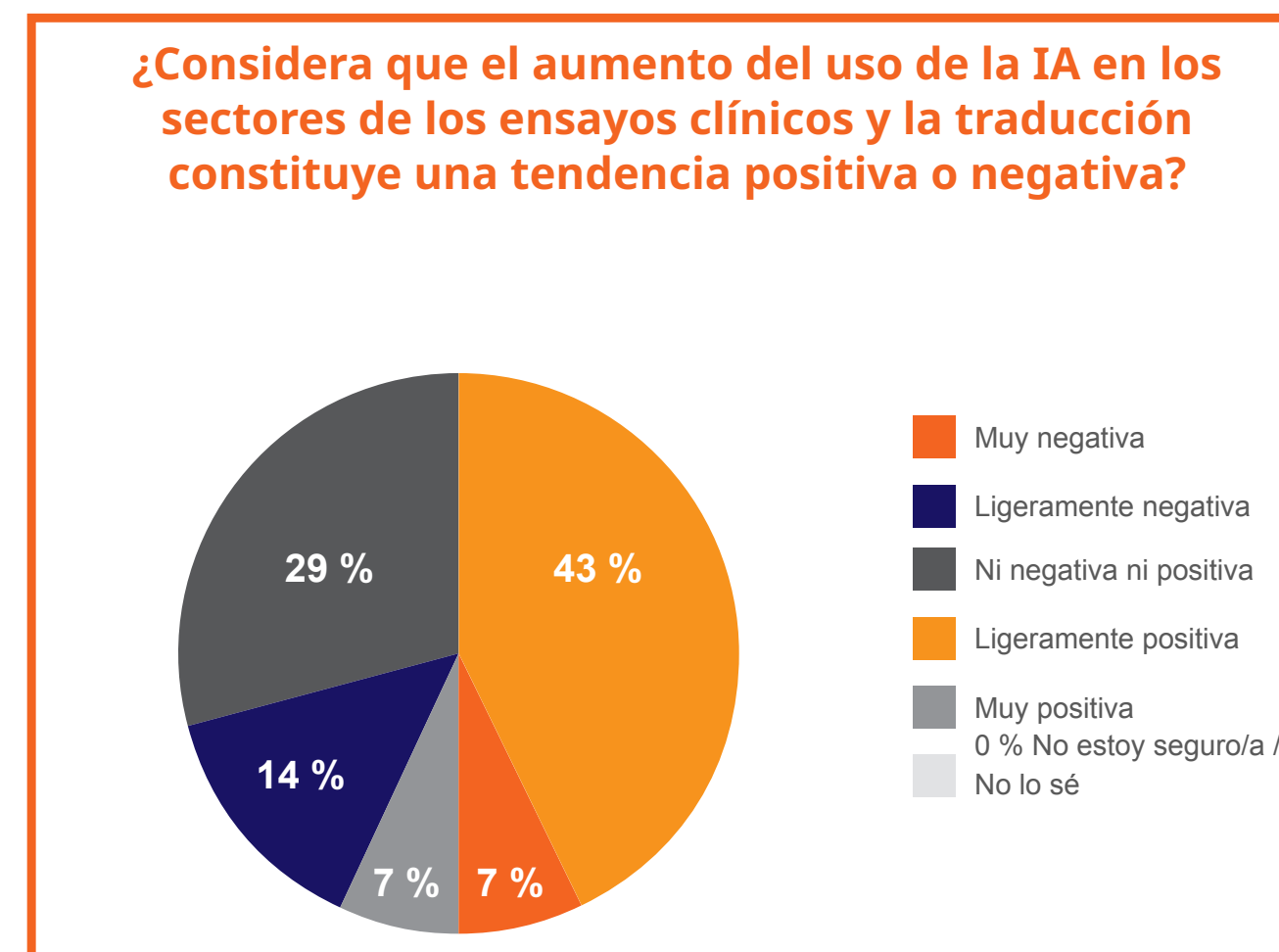
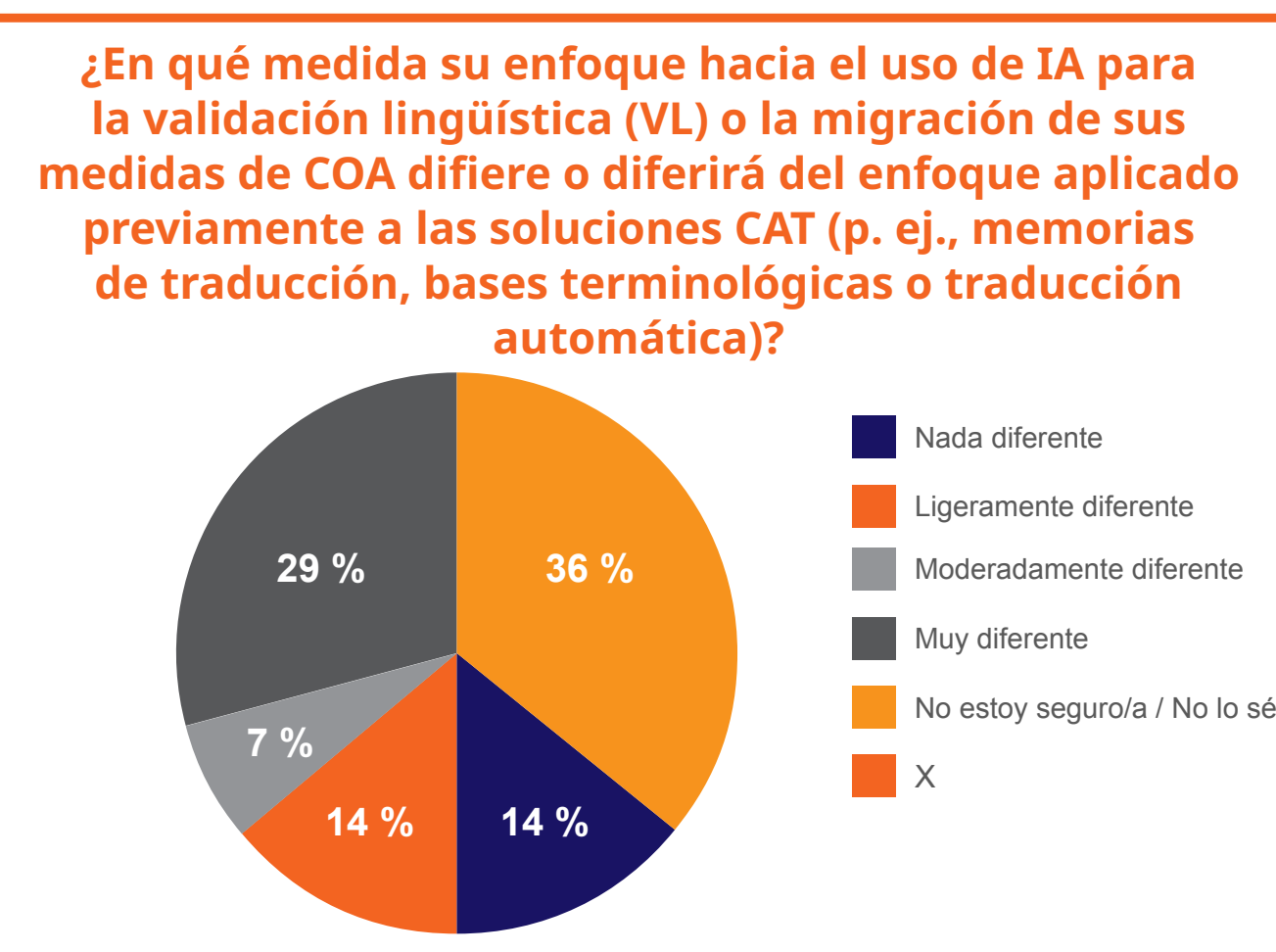
### 3) Enfoques actuales respecto a la IA

**a. Resultados:** se formularon cuatro (4) preguntas sobre este tema.

**1)** Al preguntar cómo diferiría su enfoque hacia el uso de la IA para la VL y la migración en comparación con su enfoque respecto al uso de soluciones basadas en herramientas CAT, la mayoría (36 % [5/14]) respondieron «No estoy seguro/a» o «No lo sé». En comparación, el 28,5 % (4/14) señalaron que su enfoque sería «Muy diferente». Tan solo el 28,5 % (4/14) afirmaron que su enfoque hacia esta tecnología no sería «Nada diferente» o que solo sería «Ligeramente diferente» con respecto a la tecnología de las herramientas CAT.

**2)** El 43 % (6/14) consideraron que el uso de la IA en los ensayos clínicos y en la traducción era una tendencia «Ligeramente positiva», seguido del 29 % (4/14), que lo percibieron como una tendencia «Ni negativa ni positiva». Las opciones «Muy negativa» y «Muy positiva» obtuvieron resultados similares, con un 7 % (1/14) cada una, mientras que el 14 % (2) restante de los encuestados la consideraron «Ligeramente negativa».

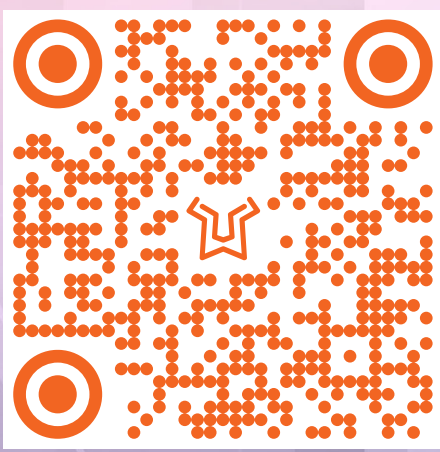
**3,4)** Al preguntar si proporcionaban de forma proactiva políticas de uso de la IA en foros de acceso público, el 72 % (10/14) respondieron «No». Además, el 86 % (12/14) no han adaptado sus acuerdos de licencia para aclarar sus políticas sobre el uso de la IA.



**b. Interpretación:** parece existir cierta incertidumbre acerca de hasta qué punto esta tecnología difiere de tecnologías anteriores empleadas en traducción y cómo deberían ajustarse los enfoques en función de ese grado de diferencia. Asimismo, existe un amplio espectro de opiniones sobre el uso de IA dentro del sector. No obstante, el hallazgo más relevante en este tema —y posiblemente de toda la encuesta— es la falta de proactividad a la hora de informar a los posibles licenciarios acerca de las políticas de IA aplicables al uso de su propiedad intelectual, incluso dentro de la documentación de los acuerdos de licencia. Sin embargo, según las respuestas abiertas, esta situación cambiará pronto. Por ello, este grupo de estudio recomienda que los licenciarios no asuman que la ausencia de información implica la inexistencia de una opinión o política al interactuar con los titulares de derechos de autor.



# EVALUACIÓN DE LA IA GENERATIVA FRENTE A LA REVISIÓN HUMANA DE CAPTURAS DE PANTALLA EN LA LOCALIZACIÓN DE ECOA: ¿Es la IA generativa la clave para mejorar las opiniones?



LIONBRIDGE

Autores: Kathryn Nolte, Karolina Elizondo Jiménez, Rupali Kadam, Melinda Johnson

## INTRODUCCIÓN

La traducción, la migración y la revisión de capturas de pantalla de las evaluaciones electrónicas de resultados clínicos (eCOA) se han basado tradicionalmente en gran medida en tareas manuales, debido en parte a la naturaleza no editable del contenido en pantalla. Durante el proceso de revisión de capturas de pantalla (SSR, por sus siglas en inglés), los informes de las capturas en los idiomas de destino se comparan con los informes maestros de capturas en el idioma de origen y, cuando procede, con el contenido heredado (es decir, los cuestionarios en papel originales).

Lionbridge realiza dos tipos de revisión de capturas de pantalla: **simple** y **compleja**. La SSR compleja se diferencia de la simple en que requiere una verificación explícita línea por línea de todo el contenido heredado en los idiomas de destino. Lionbridge reconoció y evaluó el potencial de la IA generativa y de la tecnología de reconocimiento óptico de caracteres (OCR, por sus siglas en inglés) para complementar los métodos tradicionales de revisión de capturas de pantalla, reducir costes y plazos (es decir, el número de rondas de revisión), mejorar la calidad y, en última instancia, contribuir a mejorar los resultados para los pacientes. Esta investigación se llevó a cabo utilizando las funciones de revisión de capturas de pantalla simple y compleja de la herramienta patentada de Lionbridge, Aurora AI Clinical Outcomes™.

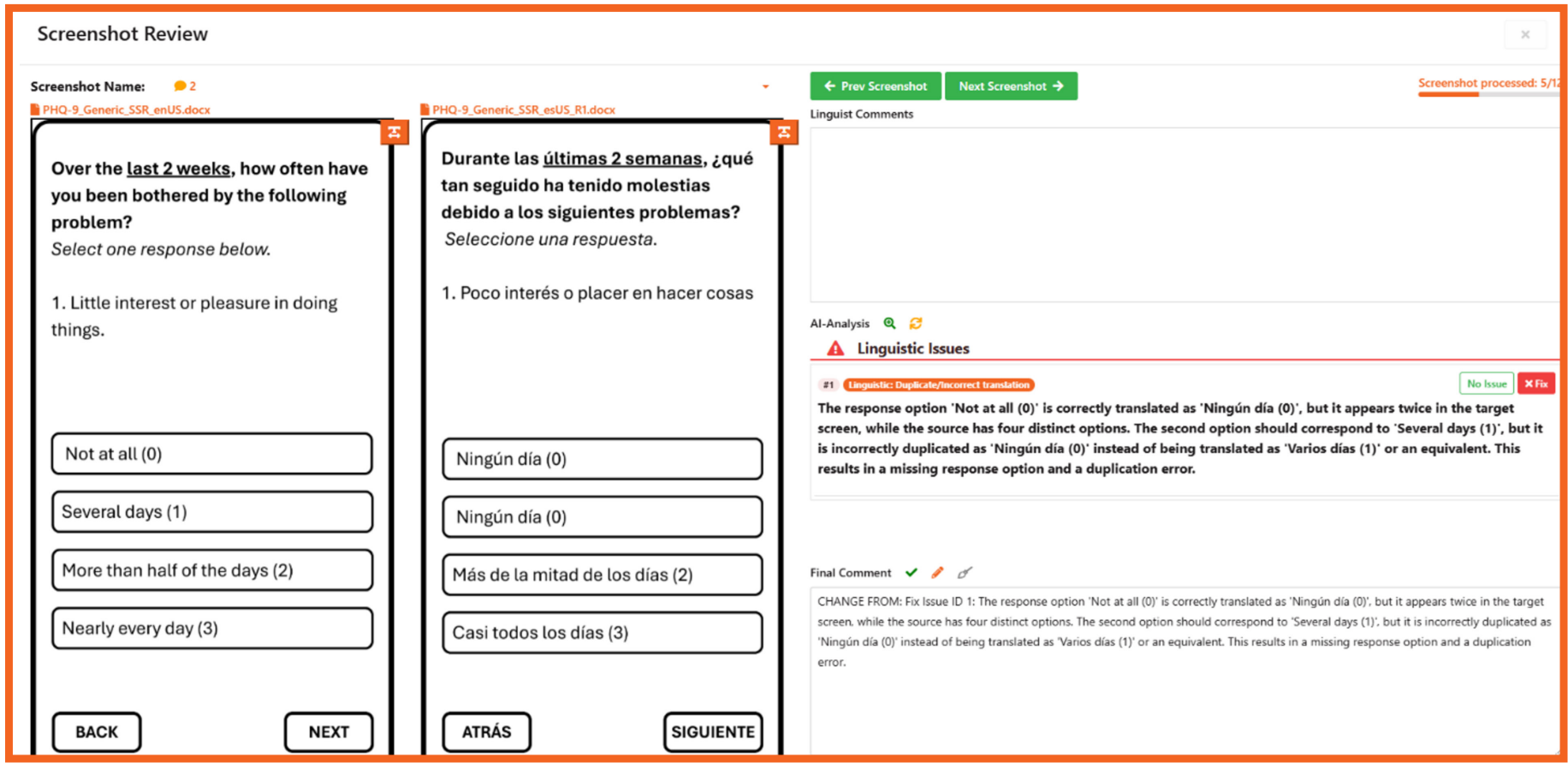
## MÉTODOS

Lionbridge empleó un motor de IA generativa seguro para generar comentarios de control de calidad sobre los informes de capturas de pantalla en los idiomas de destino correspondientes a 5 evaluaciones electrónicas de resultados clínicos (eCOA) dirigidas a pacientes, cada una con distinta longitud y nivel de complejidad, en las que se incorporaron intencionadamente 11 tipos de errores con fines de prueba: contenido omitido, contenido sin traducir, traducción incorrecta, modificaciones de eCOA omitidas, discrepancias en números de versión, problemas de formato, problemas de morfología, problemas de etiquetas, problemas de saltos de línea, problemas de uso de mayúsculas y minúsculas, y problemas con anclajes de respuesta de escalas de valoración. En total, se evaluaron 16 idiomas de destino, representativos de diversos alfabetos y familias lingüísticas: búlgaro (Bulgaria), polaco (Polonia), rumano (Rumanía), griego (Grecia), español (Argentina, México y Estados Unidos), francés (Bélgica y Francia), portugués (Brasil, Portugal), húngaro (Hungría), turco (Turquía), coreano (Corea), chino tradicional (Taiwán) y tailandés (Tailandia). Los prompts se ajustaron y personalizaron hasta obtener resultados adecuados, de acuerdo con los estándares más recientes del sector de eCOA y las prácticas establecidas. Paralelamente, se enviaron los mismos informes de capturas de pantalla de eCOA a lingüistas acreditados por Lionbridge para su revisión humana. Posteriormente, se designó un revisor imparcial para validar ambos resultados (IA frente a revisión humana) y evaluarlos en términos de precisión e integridad.

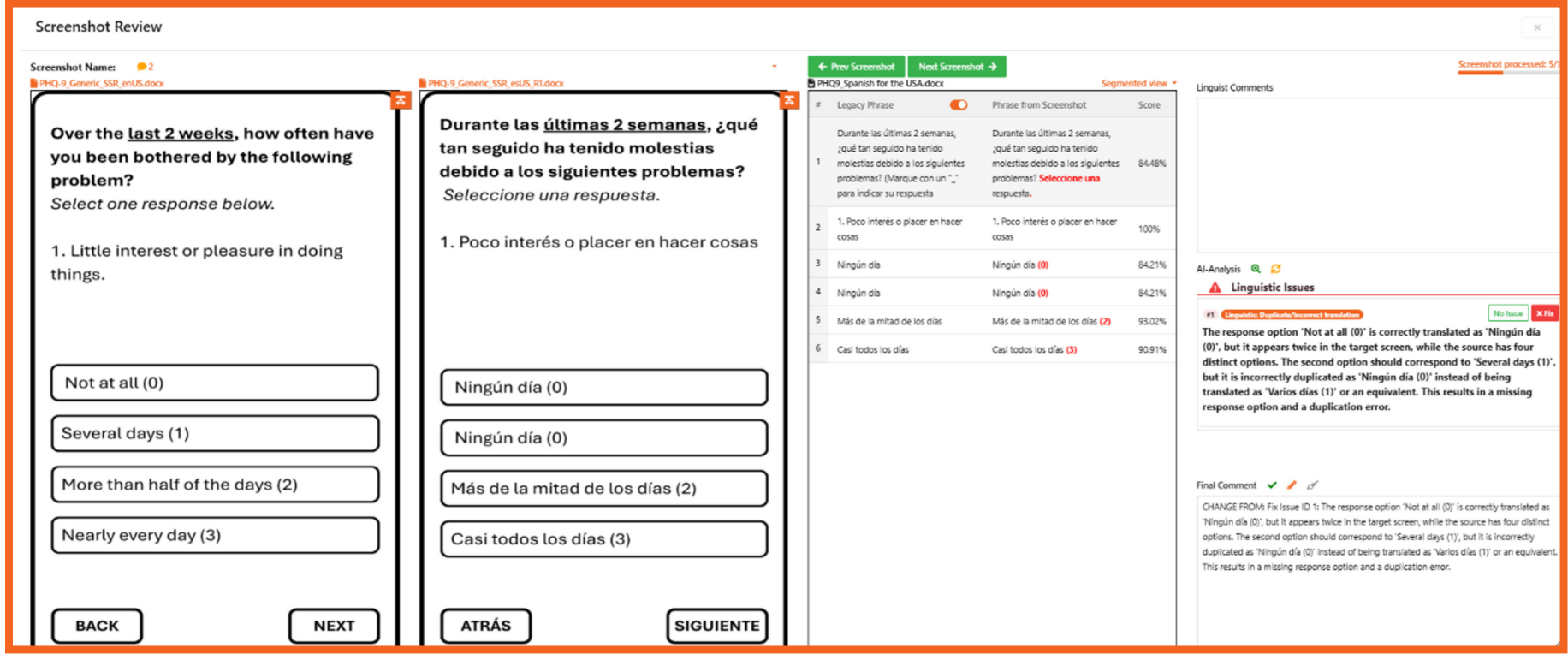
### Evolución del diseño de prompts:

La primera fase de la investigación empleó un prompt de IA generativa para SSR simple, destinada a comparar los informes de capturas en el idioma de origen con los de los idiomas de destino una vez completado el flujo de traducción (ya fuese traducción y edición o traducción directa y traducción inversa). Una vez consolidada esta función de la herramienta, el equipo se centró en desarrollar y evaluar una función de IA generativa para SSR compleja más avanzada, pensada para contenido que ha pasado por un proceso de migración, ya que este tipo de revisión va un paso más allá de la simple. Además de comparar el contenido de las capturas en los idiomas de destino con el informe maestro de capturas en el idioma de origen, la SSR compleja también permite verificar la correspondencia con el contenido heredado en el mismo idioma.

Interfaz de revisión simple de capturas:



Interfaz de revisión compleja de capturas:



## RESULTADOS E INTERPRETACIONES

El análisis de los comentarios de control de calidad obtenidos mediante la función de IA generativa para la SSR simple reveló que constituye un recurso sumamente eficaz como complemento de la revisión humana, tanto en proyectos de traducción como de migración. Esta función resulta especialmente útil para los lingüistas en proyectos de traducción y revisión de capturas de pantalla en los que no es necesario realizar una verificación cruzada con el contenido heredado. Gracias a su uso, la herramienta Aurora AI Clinical Outcomes logró identificar ejemplos correspondientes a los once tipos de errores presentes en las capturas en los idiomas de destino. Sin embargo, su rendimiento se vio limitado en algunos casos por la imposibilidad de contrastar la información con el contenido heredado.

En cuanto a la función de IA generativa para la SSR compleja, actualmente en desarrollo y con fecha de lanzamiento prevista para el primer trimestre de 2026, los resultados preliminares indican que es capaz de detectar todo lo que identifica la SSR simple y, además, aporta hallazgos adicionales. Asimismo, ofrece comentarios más detallados en comparación con la SSR simple. La posibilidad de realizar una verificación cruzada con el contenido heredado permitirá que la SSR compleja elimine algunos de los falsos positivos que genera actualmente la SSR simple, tal como se muestra en la siguiente tabla.

| TIPO DE ERROR  | FUNCIÓN DE IA GENERATIVA PARA SSR SIMPLE: NOTAS Y LIMITACIONES   | FUNCIÓN DE IA GENERATIVA PARA SSR COMPLEJA: NOTAS Y LIMITACIONES  | EJEMPLOS   |
|--|--|---|--|
| Contenido omitido  | Es capaz de señalar si falta contenido en la versión de destino.   | Es capaz de confirmar si el texto de destino se ha omitido intencionadamente porque no estaba presente en el contenido heredado.                            | <p>Portuguese (Portugal)</p> <p>#4 [Incorrect/Inconsistent missing content] Target Paragraph: 1</p> <p>No Issue X Fix</p> <p>"The statement '[scale name]' is subject to [copyright holder's Terms of Use]." is present in the source but missing in the target. All content should be retained and translated."</p>   |
| Contenido sin traducir                                       | Es capaz de señalar texto que se haya dejado en inglés y sugerir que podría haberse hecho intencionadamente (p. ej., texto sujeto a derechos de autor).  | Es capaz de confirmar si el texto se ha dejado en inglés intencionadamente.   | <p>Bulgarian (Bulgaria)</p> <p>Linguistic Issues</p> <p>#1 [Linguistic: Untranslated content]</p> <p>No Issue X Fix</p> <p>"The line '[scale name]' - items H17, BP1, N6, GE6' remains in English in the Bulgarian screen. This should be translated or localized for Bulgarian users unless these are standardized item codes that should remain in English. If these are not standard codes, provide a Bulgarian translation."</p>   |
| Traducción incorrecta  | Es capaz de identificar desviaciones respecto al contenido original.   | Es capaz de identificar traducciones incorrectas porque se desvían del contenido heredado, aunque coincidan con el texto original.                          | <p>Portuguese (Portugal)</p> <p>#3 [Linguistic: Inconsistent scale anchors]</p> <p>No Issue X Fix</p> <p>The response options in the target text do not consistently match the gradation of the source. For example, '1 - Muito pouco' (Very little) is stronger than 'A little bit', and '2 - Mais ou menos' (More or less) does not directly correspond to 'Somewhat'. '4 - Muito' (A lot) is not equivalent to 'Very much'. The scale anchors should be consistently translated to preserve the intended gradation.</p>   |
| Modificaciones de eCOA omitidas                              | Se prevé un rendimiento similar tanto en la función simple como en la compleja.  |   | <p>Portuguese (Portugal)</p> <p>#1 [Linguistic: Inconsistent translation] Source Paragraph: 6</p> <p>No Issue X Fix</p> <p>Target Paragraph: 4</p> <p>The instruction 'Please select one number per statement to indicate your response as it applies to the past 7 days,' is translated as 'faça um círculo ou marque um número por afirmação para indicar a sua resposta no que se refere aos últimos 7 dias.' The phrase 'faça um círculo ou marque um número' introduces the instruction to circle or mark a number, which is not present in the source text. The source only instructs to select a number, not to circle or mark. The translation should not introduce additional instructions not present in the source.</p> |
| Discrepancias en números de versión                          | Es capaz de señalar discrepancias en los números de versión entre el texto de origen y el de destino, con el matiz de que estas podrían ser intencionadas.   | Es capaz de confirmar si las diferencias en los números de versión son intencionadas comprobando el contenido heredado.                                     | <p>Korean (Republic of Korea)</p> <p>#1 [Incorrect/Inconsistent: Version number mismatch]</p> <p>✓ No Issue Fix</p> <p>The version number in the source is 'v1.1', while in the target it is 'v1.0'. The version number should match between source and target unless there is a justified reason for the difference.</p>  |
| Problemas de formato   | Es capaz de señalar diferencias de formato respecto al informe maestro de capturas de pantalla en el idioma de origen.   | Es capaz de confirmar si las diferencias de formato respecto al texto original son intencionadas, basándose en el contenido heredado.                       | <p>Portuguese (Portugal)</p> <p>Formatting/Layout Issues</p> <p>#1 [Formatting/Layout: Inconsistent bolding/underlining] Source Paragraph: 7</p> <p>No Issue X Fix</p> <p>Target Paragraph: 5</p> <p>In the source screen, the phrase 'past 7 days,' is bolded, while in the target screen only 'últimos 7 dias' is bolded. The bolding should match the source, with 'últimos 7 dias,' in bold in the target.</p>   |
| Problemas de morfología                                      | Es capaz de detectar morfología de género inesperada en adjetivos, con el matiz de que podría ser intencionada (p. ej., todos los adjetivos podrían figurar en femenino debido a una población de pacientes compuesta al 100 % por mujeres). | Es capaz de aportar contexto adicional sobre la morfología de género esperada realizando una comprobación con el contenido heredado.                        | <p>Spanish (United States)</p> <p>Linguistic Issues</p> <p>#1 [Linguistic: Inconsistent translation] Source Paragraph: 5</p> <p>✓ No Issue X Fix</p> <p>Target Paragraph: 5</p> <p>The source uses 'I feel fatigued', which is gender-neutral. The target uses 'Me siento agotada', which is feminine. For a general patient-facing questionnaire, the translation should be gender-neutral ('Me siento fatigado/a' or 'Me siento fatigada o fatigado') unless the instrument is specifically for female patients.</p>   |
| Problemas de etiquetas                                       | Se prevé un rendimiento similar tanto en la función simple como en la compleja.  |   | <p>Thai (Thailand)</p> <p>Formatting/Layout Issues</p> <p>#1 [Formatting/Layout: Underlining/HTML Tag Display] Source Paragraph: 7</p> <p>No Issue X Fix</p> <p>The target screen displays HTML tags (&lt;u&gt;) instead of rendering underlined text for 'habituell' and 'habituell', and 'habituell'. The source uses underlining for emphasis, but the target shows raw tags, which is incorrect formatting. The underlining should be properly rendered in the target.</p>   |
| Problemas de salto de línea                                  | Se prevé un rendimiento similar tanto en la función simple como en la compleja.  |   | <p>French (France)</p> <p>Formatting/Layout Issues</p> <p>#1 [Formatting/Layout: Line break] Source Paragraph: 6</p> <p>No Issue X Fix</p> <p>Target Paragraph: 4</p> <p>The target screen introduces a line break in 'habi-tuelles' that splits the word across two lines with a hyphen, which is not present in the source screen. This disrupts readability and should be corrected so that 'habituell' appears on one line without a hyphen.</p>   |
| Problemas de uso de mayúsculas y minúsculas                  | Es capaz de señalar patrones inesperados de uso de mayúsculas y minúsculas que difieran del original, con el matiz de que podrían ser intencionados.   | Es capaz de confirmar si el uso inesperado de mayúsculas o minúsculas en la captura de pantalla en el idioma de destino coincide con el contenido heredado. | <p>Hungarian (Hungary)</p> <p>#2 [Linguistic: Inconsistent capitalization] Source Paragraph: 6</p> <p>No Issue X Fix</p> <p>Target Paragraph: 6</p> <p>The target text uses lowercase for 'magyar verzió', while the source uses title case 'English version'. For consistency and professionalism, the Hungarian should use 'Magyar verzió Magyarország részére'.</p>   |
| Problemas con anclajes de respuesta de escalas de valoración | Se prevé un rendimiento similar tanto en la función simple como en la compleja.  |   | <p>Portuguese (Portugal)</p> <p>#2 [Formatting/Layout: Scale anchor placement] Source Paragraph: 8</p> <p>No Issue X Fix</p> <p>Target Paragraph: 7</p> <p>The scale anchor text in the target is longer and may not sit directly under the endpoint of the scale, potentially causing layout misalignment. Ensure the anchor text does not spread into the center and remains directly under the endpoint.</p>  |

### Humanos frente a máquina:

- La IA generativa es mucho más rápida que los humanos a la hora de evaluar el contenido de las capturas de pantalla en el idioma de destino.
- Genera comentarios para todas las capturas de una escala determinada en cuestión de segundos.
- Los especialistas en control de calidad de Lionbridge destacaron mejoras en tiempo y eficiencia gracias al proceso de la IA generativa, ya que la interfaz de SSR de la herramienta Aurora AI Clinical Outcomes permitió acceder fácilmente a los archivos heredados, los informes maestros y los informes de las capturas de pantalla en los idiomas de destino de forma simultánea.
- La IA generativa detectó varios problemas que pasaron desapercibidos para los revisores humanos, especialmente problemas con saltos de línea, uso de mayúsculas y minúsculas, traducciones incorrectas (incluido texto duplicado) y anclajes de respuesta de escalas de valoración. Estos errores resultan más fáciles de identificar para la IA generativa gracias al uso del reconocimiento óptico de caracteres (OCR), que permite analizar discrepancias de texto en el contenido de las imágenes no editable, que el ojo humano puede pasar por alto fácilmente.
- En algunos casos, la IA generativa para SSR detectó errores de forma incoherente según el idioma. Por ejemplo, aunque la IA generativa identificó que se había omitido la traducción de «items» en la mayoría de idiomas, este problema pasó desapercibido en español. Probablemente, se debió a que la única diferencia era una tilde («Items»).
- Los revisores humanos mostraron una menor propensión a señalar incorrectamente problemas que dependen del contexto (como poblaciones de pacientes exclusivamente femeninas, discrepancias intencionadas en los números de versión o texto protegido por derechos de autor dejado deliberadamente en inglés, entre otros).