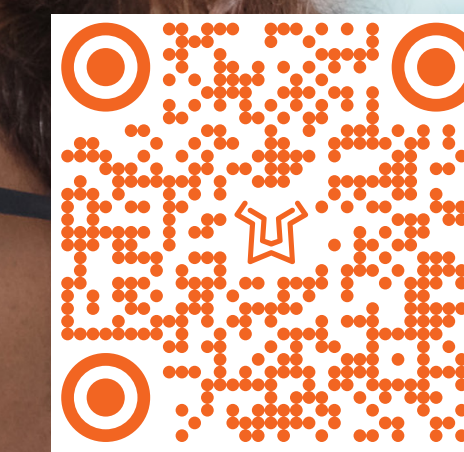


INNOVATION VERSUS GEISTIGES EIGENTUM: Was Urheberrechtsinhaber über den Einsatz generativer KI für die Lokalisierung ihres geistigen Eigentums denken


LIONBRIDGE

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

ZIEL

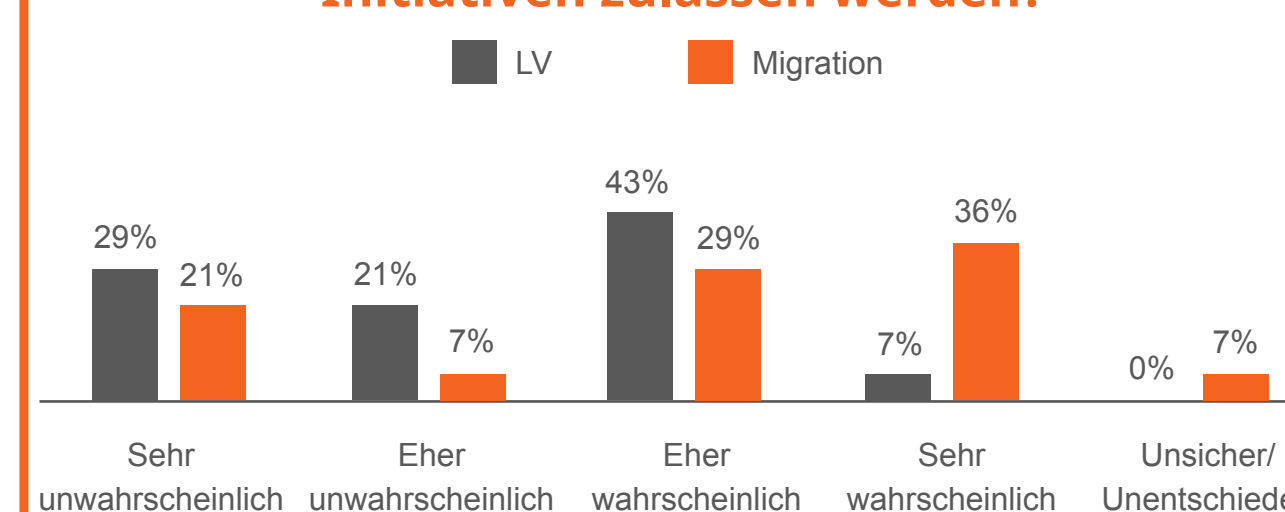
Der Einsatz generativer künstlicher Intelligenz (generative KI) für klinische Prüfungen und insbesondere für Bewertungen klinischer Ergebnisse (Clinical Outcome Assessment, COA) leidet unter dem Mangel an klaren Leitlinien zur Nutzung von KI für COA, die geistiges Eigentum darstellen. Diese Unsicherheit hat die Einführung von KI in kritischen Bereichen verlangsamt. Dazu zählen die linguistische Validierung und Migration von COA, die der Einbeziehung der Patientenurteile dienen. **Ziel dieser Studie war es, die Perspektiven der COA-Urheberrechtsinhaber zum Einsatz von KI in diesen Prozessen zusammenzustellen und zu veröffentlichen, um diese Informationen in künftigen Bereitstellungsstrategien berücksichtigen zu können.**

ERGEBNISSE UND AUSWERTUNG

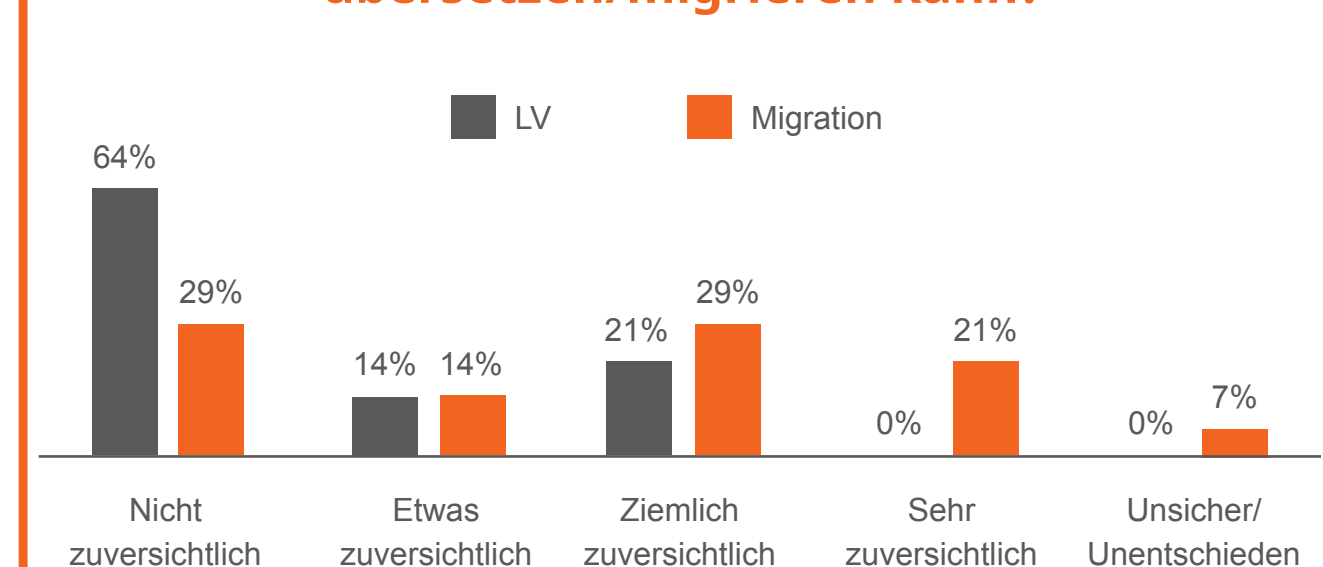
1) Zulässigkeit des Einsatzes von KI bei der COA-Übersetzung (LV) und der eCOA-Migration (Migration) sowie Vertrauen in die Ergebnisse

a. Ergebnisse: 64 % (9 von 14) der antwortenden COA-Urheberrechtsinhaber waren hinsichtlich des KI-Einsatzes zur Übersetzung von COA-Initiativen im Rahmen der linguistischen Validierung (LV) nicht zuversichtlich. Allerdings hielten es nur 29 % (4 von 14) für nicht wahrscheinlich, KI für COA-Übersetzungen im Rahmen der linguistischen Validierung zuzulassen. Eindeutiger waren die Antworten zum Zulassen des KI-Einsatzes für die eCOA-Migration und zum Vertrauen auf die Ergebnisse: 29 % (4 von 14) hielten eine solche Zulassung für sehr unwahrscheinlich, und 21 % (3 von 14) würden den Einsatz höchstwahrscheinlich nicht zulassen.

Wie wahrscheinlich ist es, dass Sie den Einsatz von KI für die Übersetzung/Migration von COA-Initiativen zulassen werden?



Wie zuversichtlich sind Sie, dass KI COA-Initiativen übersetzen/migrieren kann?



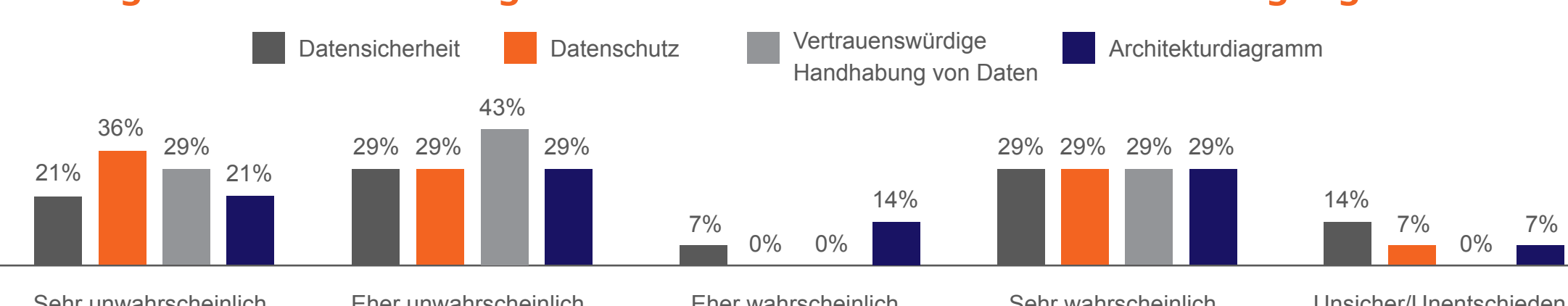
b. Auswertung: Das Vertrauen der Urheberrechtsinhaber in die Fähigkeiten der KI ist für LV und Migration deutlich unterschiedlich ausgeprägt. Sie standen dem Einsatz der KI in diesen beiden Workflows grundsätzlich, wenn auch mit unterschiedlicher Ausprägung, positiv gegenüber: mit 50 % (7 von 14) für die LV und 65 % (9 von 14) für die Migration. Den Anmerkungen in Textform zufolge ist das Vertrauen insgesamt gering. Die meisten Befragten waren aber bereit, den Einsatz von KI zuzulassen, sofern er durch einen nuancierten Dialog vorbereitet wird.

2) Zertifizierungen und Vertrauen

a. Ergebnisse: Wir haben die Urheberrechtsinhaber nach den Auswirkungen von Zertifizierungen und/oder Dokumentationen auf die Wahrscheinlichkeit einer Zulassung des KI-Einsatzes gefragt: 1) Datensicherheitszertifizierung (z. B. ISO 27001:2022), 2) Datenschutzzertifizierung (z. B. ISO 27001:2019), 3) Zertifizierung der vertrauenswürdigen Handhabung von Daten (z. B. SOC2 Typ 2) und 4) KI-Architekturdiagramm. Die verschiedenen Zertifizierungen/Dokumentationen wurden in Bezug auf die Ausgangsfrage unterschiedlich beurteilt. Am spektischsten wurden dabei die Auswirkungen von Zertifizierungen zu Datenschutz und vertrauenswürdiger Handhabung von Daten beurteilt. Für alle vier Zertifizierungs-/Dokumentationstypen waren aber immerhin 29 % (4 von 14) der Befragten mit hoher Wahrscheinlichkeit bereit, die KI-Nutzung zu genehmigen, wenn die jeweiligen Dokumentationen vorliegen.

b. Auswertung: Obwohl die verschiedenen Dokumentationstypen (Zertifizierung/Diagramme), Datensicherheit, Datenschutz und die Funktionsweise der KI-Engine skeptisch betrachtet werden, erhöht das Vorliegen dieser Dokumentationstypen die Wahrscheinlichkeit des Einsatzes von KI für die Übersetzung geistigen Eigentums bei mindestens 29 % der Urheberrechtsinhaber.

Auswirkungen der Zertifizierung auf die Wahrscheinlichkeit der Genehmigung des KI-Einsatzes



FAZIT

Die Urheberrechtsinhaber müssen COA-Initiativen in einer Branche abwickeln, in der es nach wie vor an klaren Leitlinien für den Einsatz von KI bei der Lokalisierung geistigen Eigentums mangelt. Dies gilt auch für andere Stakeholder in diesem Segment. Und obwohl das Vertrauen dieser Stakeholder in die angemessene Nutzung von KI bei der Lokalisierung ihres geistigen Eigentums noch vertieft werden muss, **scheinen sie Diskussionen, der Ausarbeitung von Richtlinien, Ergänzungen der Lizenzverträge und der Überprüfung von Dokumentationen offen gegenüberzustehen, um für die Sicherheit ihres geistigen Eigentums bei Nutzung einer KI-Engine zu sorgen.** Abschließend sei angemerkt, dass die geringe Anzahl an Teilnehmern den Erkenntniswert dieser Studie beschränkt und weitere Untersuchungen erforderlich sind.

METHODEN

Um dieses Ziel zu erreichen, wurde eine elektronische Umfrage mit fünfzehn Fragen zu den drei unten aufgeführten Themen an 100 COA-Urheberrechtsinhaber verschickt. Vierzehn dieser Urheberrechtsinhaber haben innerhalb der vorgegebenen drei Wochen an dieser anonymen Umfrage teilgenommen.

- 1) Zulässigkeit des Einsatzes von KI bei der COA-Übersetzung (LV) und der eCOA-Migration (Migration) sowie Vertrauen in die Ergebnisse
- 2) Zertifizierungen und Vertrauen
- 3) Aktuelle KI-Ansätze, -Richtlinien und -Lizenzierung

3) Aktuelle KI-Ansätze

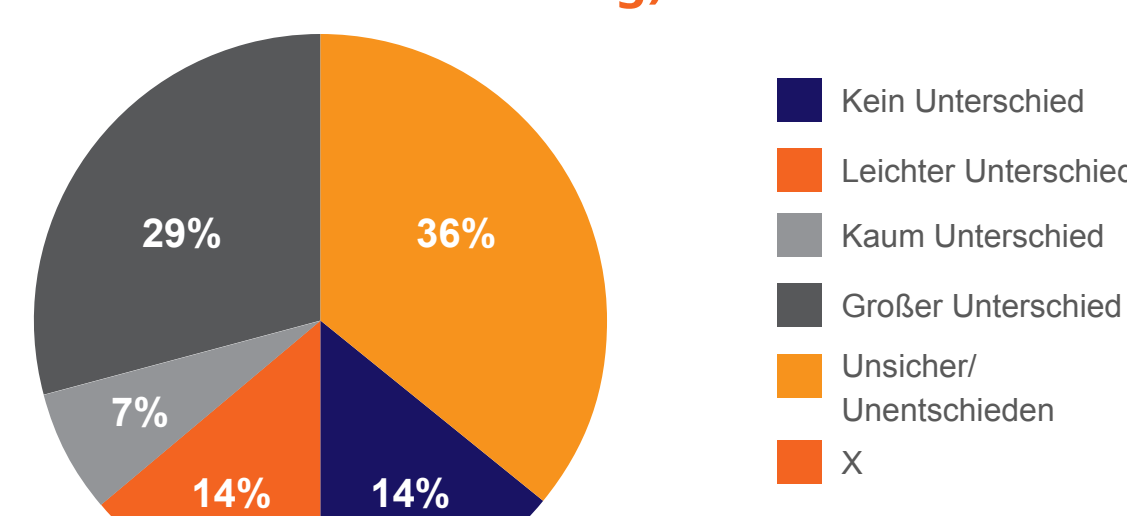
a. Ergebnisse: Zu diesem Thema wurden vier (4) Fragen gestellt.

1) Auf die Frage, inwieweit sich ihre Herangehensweise an die KI-Nutzung für LV und Migration vom Einsatz von CAT-Tool-Lösungen unterscheiden würde, war sich die Mehrheit der Umfrageteilnehmer (36 % (5 von 14)) unsicher. Nur 28,5 % (4 von 14) gehen von einer deutlich abweichenden Herangehensweise aus. Und ebenfalls 28,5 % (4 von 14) gingen davon aus, diese Technologie nicht oder nur geringfügig anders als die CAT-Tool-Technologie einsetzen zu wollen.

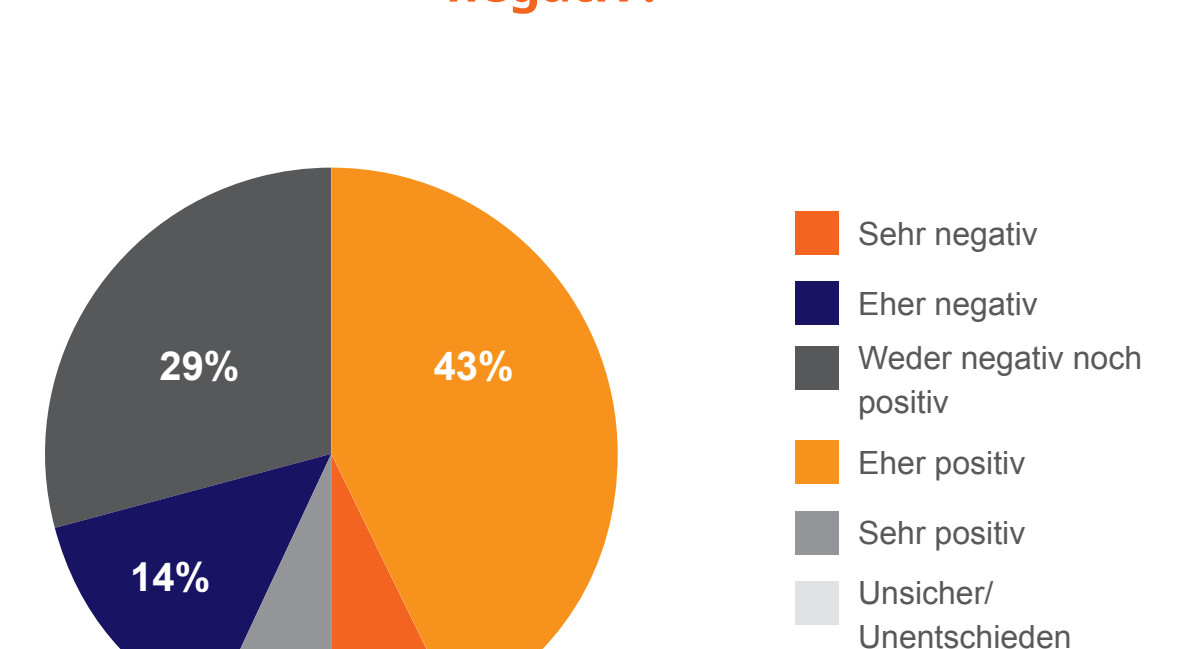
2) 43 % (6 von 14) waren der Meinung, dass der Einsatz von KI in klinischen Studien und Übersetzungen leicht positive Resultate haben könne, gefolgt von 29 % (4 von 14), die weder negative noch positive Resultate erwarten. Als sehr negativ bzw. sehr positiv schätzte jeweils ein Umfrageteilnehmer (entspricht 7 % (1 von 14)) ein, während die übrigen 14 % (2) der Befragten dem Einsatz eher negativ gegenüberstanden.

3, 4) Die Frage, ob sie proaktiv Richtlinien zur KI-Nutzung in öffentlich zugänglichen Foren bereitstellen würden, verneinten 72 % (10 von 14). Und 86 % (12 von 14) haben ihre Lizenzverträge nicht präzisiert, um die Richtlinien zur KI-Nutzung zu berücksichtigen.

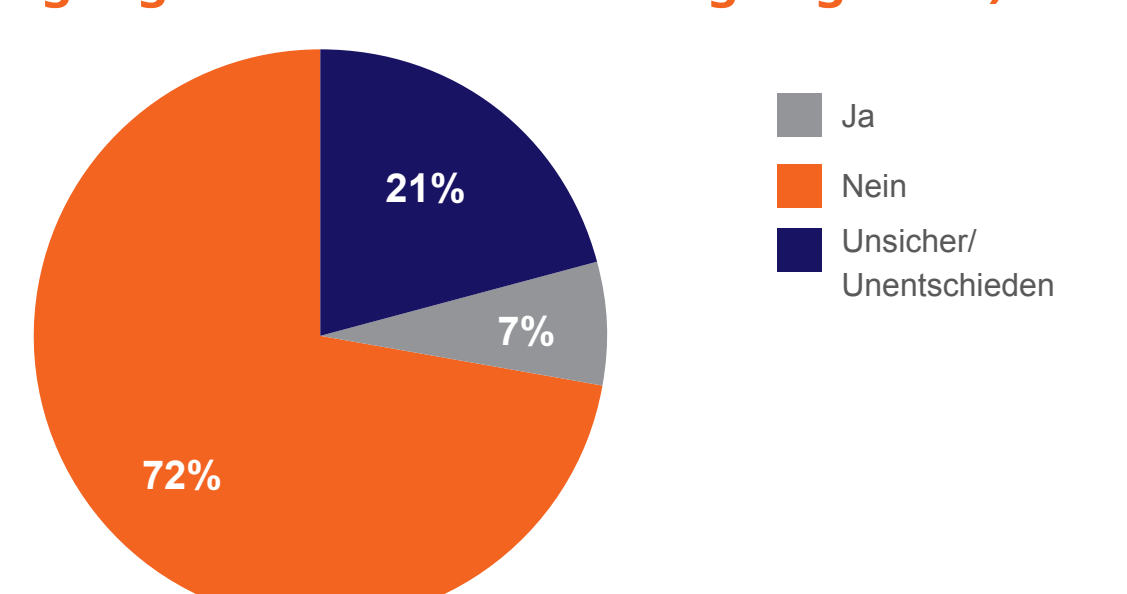
Unterscheidet sich Ihre Herangehensweise an die KI für LV oder Migration von COA-Initiativen von der bisherigen Nutzung von CAT-Tool-Lösungen (z. B. Translation Memory/Glossar/Maschinelle Übersetzung)?



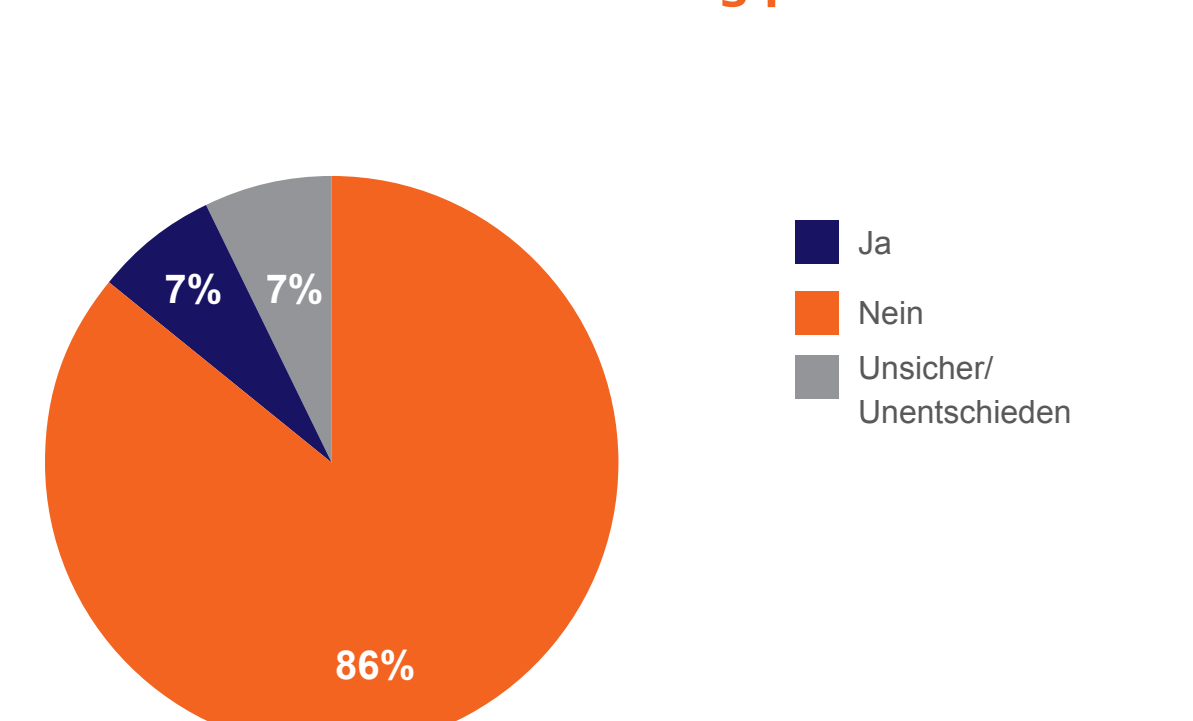
Bewerten Sie den zunehmenden Einsatz von KI für klinische Forschung und Übersetzung positiv oder negativ?



Informieren Sie potenzielle Lizenznehmer proaktiv über Ihre Richtlinien zur KI-Nutzung (z. B. im Rahmen der Bedingungen, die in öffentlich zugänglichen Richtlinien dargelegt sind)?



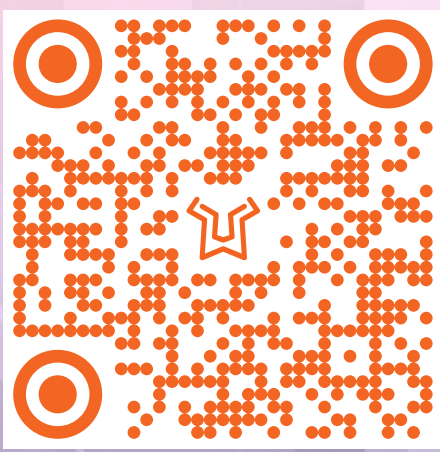
Haben Sie Ihre Lizenzverträge in Bezug auf die Richtlinien zur KI-Nutzung präzisiert?



b. Auswertung: Offensichtlich bestehen Unklarheiten über den Einsatz dieser Technologie im Vergleich zu älteren Übersetzungstechnologien und darüber, wie diese Herangehensweisen aufgrund dieser Unterschiede angepasst werden müssen. Die Branche begegnet dem Einsatz von KI mit einem breitgefächerten Meinungsspektrum. Das auffälligste Ergebnis zu diesem Themenkreis und möglicherweise der gesamten Umfrage war jedoch die Tatsache, dass potenzielle Lizenznehmer nicht proaktiv über KI-Richtlinien zur Nutzung von KI für geistigen Eigentums informiert werden, nicht einmal in den Dokumentation, die Teil der Lizenzverträge ist. Die Anmerkungen deuten jedoch darauf hin, dass sich dies bald ändern wird. Diese Studiengruppe empfiehlt Lizenznehmern deshalb, die fehlenden Informationen beim Kontakt mit Urheberrechtsinhabern als unzureichende Meinungsbildung und als Mangel an diesbezüglichen Richtlinien zu interpretieren.



AUSWERTUNG DER KORREKTUREN VON ECOA-LOKALISIERUNGEN DURCH GENERATIVE KI IM VERGLEICH ZUM MENSCHLICHEN REVIEW VON SCREENSHOTS: Ist KI der Schlüssel zu einem besseren Feedback?



LIONBRIDGE

Autoren: Kathryn Nolte, Karolina Elizondo Jimenez, Rupali Kadam, Melinda Johnson

EINLEITUNG

Übersetzung, Migration und Screenshot-Reviews von eCOA (Electronic Clinical Outcome Assessments) basierten bislang weitgehend auf menschlichen Tätigkeiten. Das liegt zum Teil daran, dass der Bildschirmcontent nicht bearbeitet werden kann. Im Screenshot-Review (SSR) werden die Berichte mit den Screenshots in Zielsprache mit den entsprechenden Berichten in der Ausgangssprache und gegebenenfalls mit dem Altmaterial (also den ursprünglichen Fragebögen in Papierform) abgeglichen.

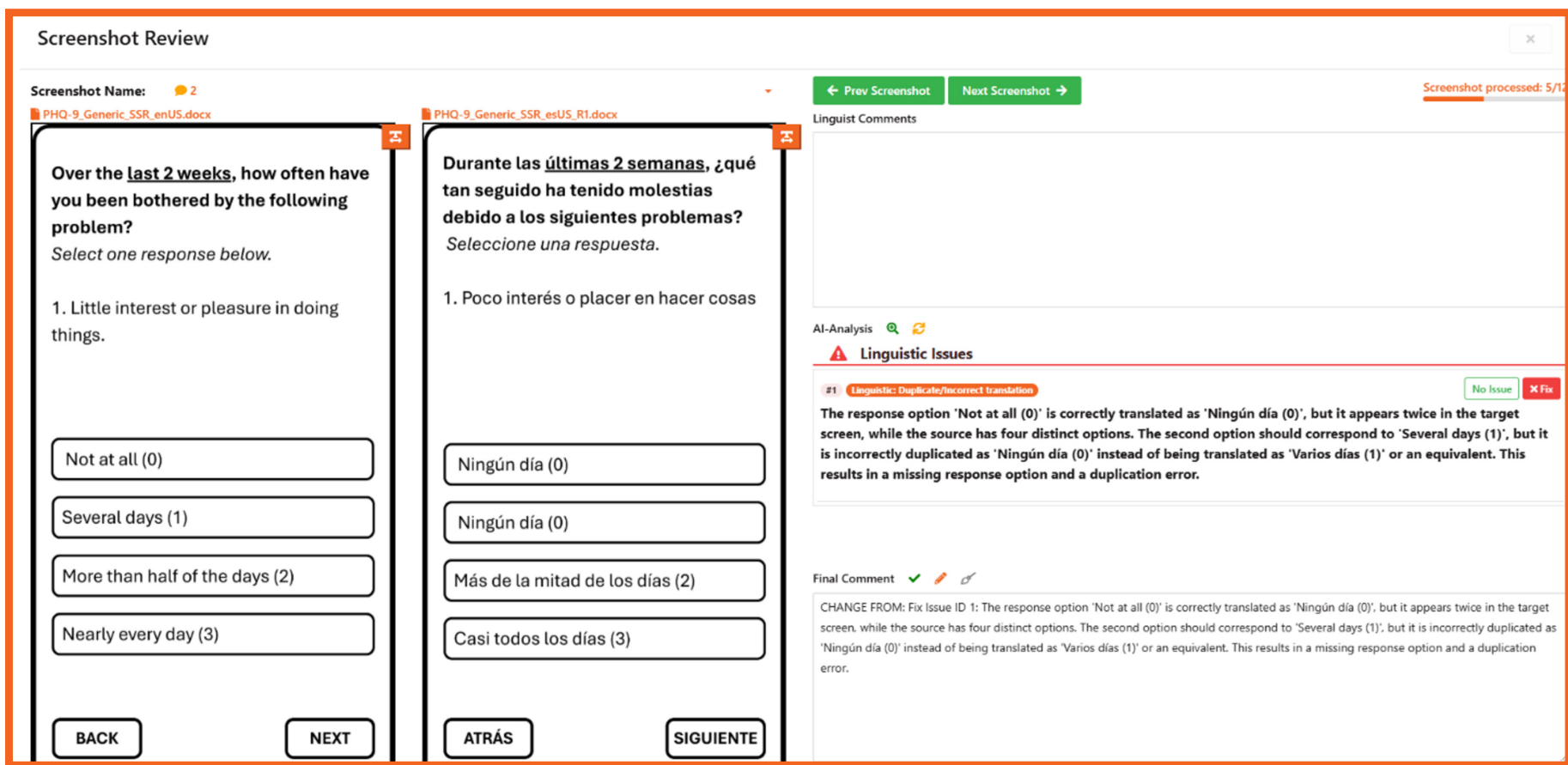
Lionbridge bietet zwei Arten von Screenshot-Review an: **einfach** und **komplex**. Komplexe Reviews unterscheiden sich von einfachen Reviews darin, dass ein zeilenweiser Abgleich mit dem gesamten Altmaterial erfolgt. Lionbridge hat das Potenzial der generativen KI und der optischen Zeichenerkennung (OCR) erkannt, herkömmliche Screenshot-Prüfverfahren zu ergänzen, Kosten und Zeitaufwand (d. h. die Anzahl der Reviews) zu reduzieren, die Qualität zu verbessern und letztendlich zu besseren Patientenergebnissen zu kommen. Die entsprechenden Tests und Studien wurden anhand einfacher und komplexer Screenshot-Reviews mit den Funktionen unseres Tools Aurora AI Clinical Outcomes™ durchgeführt.

METHODEN

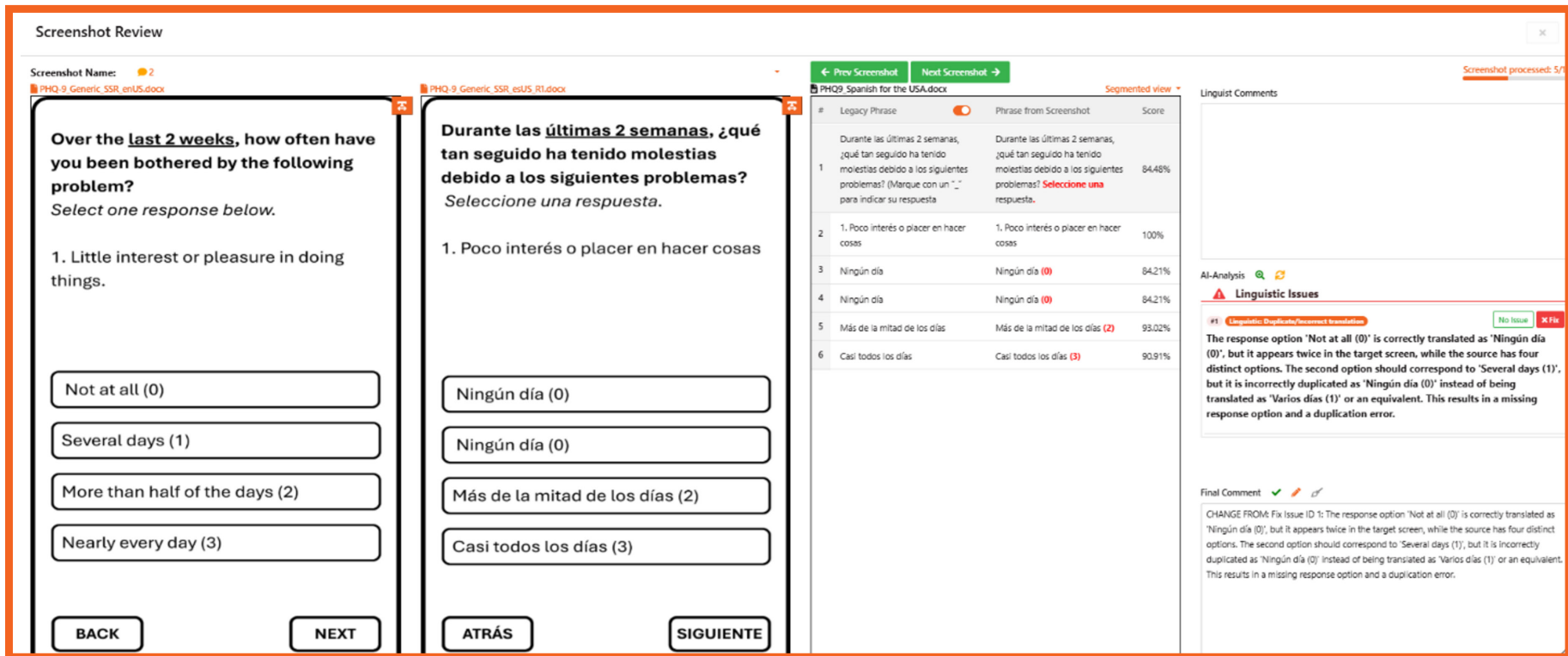
Lionbridge nutzte eine sichere Engine auf Basis generativer KI, um QS-Feedback für Screenshots in Zielsprache für fünf patientenbezogene eCOA (Electronic Clinical Outcome Assessments) unterschiedlicher Länge und Komplexität zu generieren. Zu Testzwecken wurden Fehler elf verschiedener Typen eingefügt: Fehlender Content, Nicht übersetzter Content, Falsche Übersetzung, Fehlende eCOA-Bearbeitungen, Versionsnummern stimmen nicht überein, Formatierungsprobleme, Morphologische Probleme, Tag-Probleme, Probleme mit Zeilenumbrüchen, Probleme bei Groß- und Kleinschreibung und Layoutprobleme. Insgesamt wurden sechzehn Zielsprachen mit verschiedenen Alphabeten und aus unterschiedlichen Sprachfamilien getestet: Bulgarisch (Bulgarien), Polnisch (Polen), Rumänisch (Rumänien), Griechisch (Griechenland), Spanisch (Argentinien, Mexiko, Vereinigte Staaten von Amerika), Französisch (Belgien, Frankreich), Portugiesisch (Brasilien, Portugal), Ungarisch (Ungarn), Türkisch (Türkei), Koreanisch (Korea), Traditionelles Chinesisch (Taiwan) und Thailändisch (Thailand). Die Prompts wurden angepasst, bis die Ausgaben den aktuellen eCOA-Branchenstandards und etablierten Verfahren entsprachen. Gleichzeitig haben wir dieselben Berichte mit eCOA-Screenshots für eine Überprüfung durch Menschen an von Lionbridge geprüfte Linguisten gesendet. Ein unparteiischer Prüfer wurde dann damit beauftragt, die beiden Ergebnisse (KI und Mensch) zu validieren und hinsichtlich ihrer Genauigkeit und Vollständigkeit zu bewerten.

Evolution des Prompt-Aufbaus:
In der ersten Phase unserer Studie haben wir einen Prompt für das einfache Screenshot-Review auf Basis generativer KI verwendet, der sich ideal für den Vergleich von Berichten mit Quelltext- und Zieltextscreenshots nach dem Übersetzungsworkflow (entweder Übersetzung/ Bearbeitung oder Vorwärtsübersetzung/Rückübersetzung) eignet. Nach Ausführung dieses Prompts konzentrierten wir uns auf die Entwicklung und Erprobung einer differenzierteren Funktion für das komplexe Screenshot-Review auf Basis generativer KI, die sich besser für migrierten Content eignet, da sie einen Schritt weiter geht als die Funktion für einfaches SSR. Ergänzend zum Vergleich der Zielscreenshots mit den Screenshots des Originals vergleicht sie die Zielscreenshots auch mit dem Zielcontent im Altmaterial.

Benutzeroberfläche für einfaches Screenshot-Review:



Benutzeroberfläche für komplexes Screenshot-Review:



ERGEBNISSE UND AUSWERTUNG

Die Analyse des QS-Feedbacks für das einfache SSR mittels generativer KI ergab, dass es sich um eine äußerst hilfreiche QS-Funktion handelt, die zusammen mit menschlichem Feedback sowohl für Übersetzungs- als auch für Migrationsprojekte eingesetzt werden kann. Einfaches SSR eignet sich besonders gut zur Unterstützung von Linguisten bei Übersetzungs-/Screenshot-Reviewprojekten, bei denen kein Abgleich mit Altmaterial erforderlich ist. Mit dieser Funktion konnte das Tool Aurora AI Clinical Outcomes etliche Instanzen aller elf Fehlertypen in den Zielsprache-Screenshots erfassen. Die Erkennung einiger Fehler erwies sich als problematisch, weil das Altmaterial nicht zum Abgleich herangezogen werden konnte.

Das komplexe SSR auf Basis generativer KI befindet sich noch in der Entwicklung. Die Einführung ist für das erste Quartal 2026 geplant. Vorläufige Ergebnisse zeigen allerdings bereits, dass diese Funktion noch mehr Fehler erfassen kann als das einfache SSR. Außerdem liefert es im Vergleich zum einfachen SSR differenzierteres Feedback. Durch die Möglichkeit des Abgleichs mit dem Altmaterial kann die Funktion für komplexes SSR einige der derzeit von der Funktion für einfaches SSR falsch positiv gemeldeten Ergebnisse ausschließen, siehe folgende Tabelle.

FEHLERTYP	EINFACHES SSR AUF BASIS GENERATIVER KI: HINWEISE UND EINSCHRÄNKUNGEN	KOMPLEXES SSR AUF BASIS GENERATIVER KI: HINWEISE UND EINSCHRÄNKUNGEN	BEISPIELE
Fehlender Content	Fehlender Zielcontent wird gemeldet	Kann anhand des Vergleichs mit dem Altmaterial bestätigen, ob Zieltext gewollt fehlt	Portuguese (Portugal) #4 [Redundancy: Missing content] Target Paragraph: 1 "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."
Nicht übersetzter Content	Kann nicht übersetzten Text melden und bewerten, ob dies beabsichtigt sein könnte (z. B. Urheberrechtsvermerk).	Kann bestätigen, ob Text absichtlich unübersetzt blieb	Bulgarian (Bulgaria) Linguistic Issues #1 [Linguistic: Untranslated content] "The line '[scale name]' - items B17, B11, 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."
Falsche Übersetzung	Kann Abweichungen von der Quelle erkennen	Kann falsche Übersetzungen aufgrund der Abweichungen vom Altmaterial identifizieren, auch wenn sie dem Ausgangstext entsprechen	Portuguese (Portugal) #3 [Linguistic: Inconsistent scale anchors] 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.
Fehlende eCOA-Bearbeitungen	Gleiche Funktionalität für einfaches und komplexes SSR erwartet		Portuguese (Portugal) #1 [Linguistic: Inconsistent translation] Source Paragraph: 6 Target Paragraph: 6 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.
Versionsnummern stimmen nicht überein	Kann Versionsnummernkonflikte zwischen Quelle und Ziel melden, diese können aber beabsichtigt sein	Kann durch Abgleich mit dem Altmaterial ermitteln, ob Unterschiede in den Versionsnummern beabsichtigt sind	Korean (Republic of Korea) #3 [Redundancy: Verbose number mismatch] 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.
Formatierungsprobleme	Kann Formatierungen melden, die vom Bericht mit den Screenshots in Originalsprache abweichen	Kann durch Abgleich mit dem Altmaterial verifizieren, ob Formatierung absichtlich vom Quelltext abweicht	Portuguese (Portugal) Formatting/Layout Issues #1 [Formatting/Layout: Inconsistent bolding/underlining] Source Paragraph: 7 Target Paragraph: 5 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.
Morphologische Probleme	Kann unerwartete Geschlechtsmorphologie bei Adjektiven erkennen, was möglicherweise beabsichtigt ist (wenn z. B. alle Adjektive aufgrund einer zu 100 % weiblichen Patientenpopulation in der weiblichen Form stehen)	Kann durch Abgleich mit dem Altmaterial zusätzlichen Kontext zur erwarteten Geschlechtsmorphologie bereitstellen	Spanish (United States) Linguistic Issues #1 [Linguistic: Inconsistent translation] Source Paragraph: 7 Target Paragraph: 5 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.
Tag-Probleme	Gleiche Funktionalität für einfaches und komplexes SSR erwartet		Thai (Thailand) Formatting/Layout Issues #1 [Formatting/Layout: Underlining/HTML Tag Display] Source Paragraph: 7 Target Paragraph: 5 The target screen displays HTML tags (<u>) instead of rendering underlined text for 'การรับรู้ความรู้สึก', 'ความรู้สึก', and 'การรับรู้ความรู้สึก'. 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.
Probleme mit Zeilenumbrüchen	Gleiche Funktionalität für einfaches und komplexes SSR erwartet		French (France) Formatting/Layout Issues #1 [Formatting/Layout: Line break] Source Paragraph: 6 Target Paragraph: 4 The target screen introduces a line break in 'habituels' 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 'habituels' appears on one line without a hyphen.
Probleme bei Groß- und Kleinschreibung	Kann unerwartete Groß-/ Kleinschreibungen melden, die vom Original abweichen, aber möglicherweise beabsichtigt sind	Kann verifizieren, ob unerwartete Groß-/ Kleinschreibung in Screenshots in der Zielsprache dem Altmaterial entspricht	Hungarian (Hungary) #2 [Linguistic: Inconsistent capitalization] Source Paragraph: 6 Target Paragraph: 6 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'.
Layoutprobleme	Gleiche Funktionalität für einfaches und komplexes SSR erwartet		Portuguese (Portugal) #2 [Formatting/Layout: Scale anchor placement] Source Paragraph: 8 Target Paragraph: 7 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.

Mensch versus Maschine:

- Generative KI kann den Content von Screenshots in der Zielsprache viel schneller bewerten als Menschen. Sie generiert innerhalb von Sekunden Feedback für alle übergebenen Screenshots.
- Die QS-Spezialisten von Lionbridge stellen Zeit- und Effizienzgewinne durch den Prozess auf Basis generativer KI fest, weil sie über die SSR-Schnittstelle des Tools Aurora Clinical Outcomes gleichzeitig auf Dateien mit Altmaterial sowie auf Berichte mit den Screenshots in der Ausgangssprache und in der Zielsprachen zurückgreifen konnten.
- Die generative KI hat mehrere Probleme erkannt, die menschlichen Prüfern entgangen sind – insbesondere Zeilenumbrüche, Groß-/ Kleinschreibung, falsche Übersetzungen (einschließlich duplizierten Textes) und Bewertungsskalen. Diese Probleme sind für generative KI mutmaßlich leichter zu erkennen, da das Tool die optische Zeichenerkennung (OCR) verwendet, um Diskrepanzen in nicht bearbeitbarem Bildcontent zu analysieren, die dem menschlichen Auge leicht entgehen können.
- Die Erkennung von Fehlern mittels SSR auf Basis generativer KI verlief nicht immer über alle Sprachen einheitlich. Während die generative KI beispielsweise die fehlende Übersetzung von „items“ in den meisten Sprachen erkannte, verpasste sie den Fehler im Spanischen. Wahrscheinlich lag dies daran, dass der Unterschied hier nur in einem Akzentzeichen („ítems“) bestand.
- Menschliche Prüfer bemerkten kontextabhängige Probleme seltener (z. B. weibliche Patientenpopulationen, absichtliche Abweichungen bei Versionsnummern, absichtlich auf Englisch belassene Urheberrechtsvermerke usw.).