

L'ÉQUILIBRE ENTRE INNOVATION ET PROPRIÉTÉ :

l'avis des détenteurs de droits d'auteur sur l'utilisation de l'IA générative lors de la localisation de leur propriété intellectuelle



LIONBRIDGE

Auteures : Melinda Johnson, Megan Duff, Kathryn Nolte et Stephanie Casale

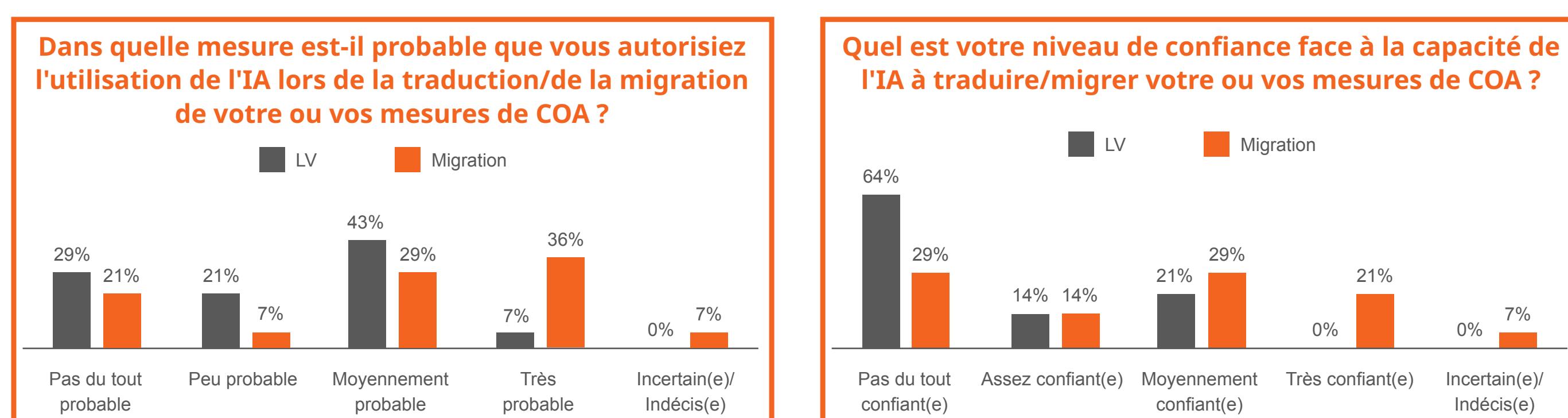
OBJECTIF

La question de l'intégration de l'intelligence artificielle générative (IA générative) dans les opérations d'essais cliniques, en particulier concernant les évaluations des résultats cliniques (COA), est ralenti par l'absence de directives claires concernant l'application de l'IA générative à la propriété intellectuelle (PI) des COA. Cette incertitude a ralenti l'adoption de l'IA générative dans des domaines essentiels tels que la migration et la validation linguistique de COA, qui sont essentielles pour intégrer la voix du patient. Le but de cette étude était de recueillir et de partager les points de vue de détenteurs de droits d'auteur de COA sur l'utilisation de l'IA générative dans ces processus, afin de soutenir l'élaboration de stratégies de déploiement futures.

RÉSULTATS ET INTERPRÉTATIONS

1) Autorisation et confiance envers l'utilisation de l'IA dans la LV et la migration

a. Résultats : 64 % (9/14) des détenteurs de droits d'auteur de COA qui ont répondu ont indiqué n'être Pas du tout confiants dans le fait que l'IA puisse être utilisée pour traduire leurs mesures de COA au sein du processus de validation linguistique. Cependant, seuls 29 % (4/14) n'étaient Pas du tout susceptibles d'autoriser l'utilisation de l'IA pour la traduction de leurs mesures de COA au sein du processus de validation linguistique. Les réponses pour la Confiance et l'Autorisation de l'utilisation de l'IA au sein du processus de migration d'eCOA étaient bien plus proches : 29 % (4/14) n'étaient Pas du tout confiants et 21 % (3/14) Pas du tout susceptibles de l'autoriser.

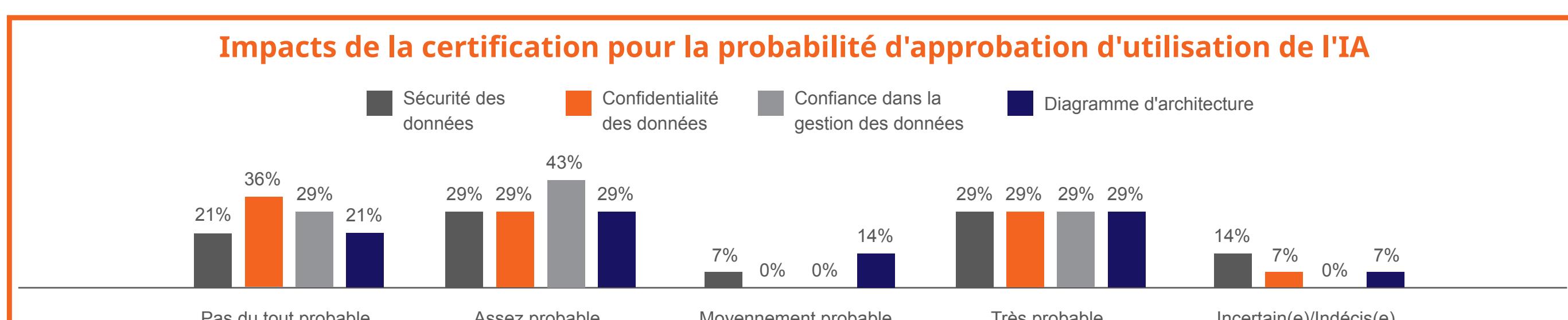


b. Interprétation : Bien que la confiance des détenteurs de droits d'auteur envers les fonctionnalités de l'IA ait varié selon qu'il s'agissait de la LV ou de la migration, la probabilité d'autoriser l'utilisation de l'IA dans ces deux processus était invariablement positive, même si les degrés variaient (Moyennement probable et Très probable), avec 50 % pour la LV (7/14) et 65 % pour la migration (9/14). Selon les réponses en texte libre, le niveau de confiance est faible de manière générale. Cependant, la plupart des participants étaient prêts à autoriser l'utilisation de l'IA après un dialogue nuancé.

2) Certifications et confiance

a. Résultats : Nous avons interrogé les détenteurs de droits d'auteur sur l'impact (au niveau de la probabilité qu'ils approuvent l'utilisation de l'IA) de plusieurs certifications et/ou documentations : 1) certification en sécurité des données (par ex., ISO 27001:2022), 2) certification en confidentialité des données (par ex., ISO 27001:2019), 3) certification de confiance pour la gestion des données (par ex., SOC2 Type 2), et 4) diagramme d'architecture de l'IA. Les résultats ont varié selon les différentes certifications/documentations, celles faisant l'objet du scepticisme le plus marqué étant la confidentialité des données et la confiance dans la gestion des données. Cependant, dans les quatre types de certifications, 29 % (4/14) des participants se sont dits Très susceptibles d'approuver l'utilisation de l'IA avec cette documentation en place.

b. Interprétation : Bien qu'il y ait un certain scepticisme autour des différents types de documentation (certification/diagrammes), de la sécurité des données, de la confidentialité et de la manière dont le moteur d'IA fonctionne, ces types de documentation augmentent la probabilité qu'au moins 29 % des détenteurs de droits d'auteurs deviennent plus favorables à l'utilisation de l'IA lors de la traduction de leur PI.



CONCLUSION

Les détenteurs de droits d'auteur sur des mesures de COA font face à un secteur qui subit toujours un manque de directives claires concernant l'utilisation de l'IA lors de la localisation de leur PI, tout comme les autres parties prenantes. Cependant, bien qu'il existe une marge d'amélioration de leur confiance vis-à-vis d'une utilisation appropriée de l'IA lors de la localisation de leur PI, ces parties prenantes semblent ouvertes aux discussions, à la création de politique, aux modifications des accords de licence et à la révision de la documentation pour vérifier la sécurité de leur PI lors de l'interaction avec un moteur d'IA. Enfin, notons que le nombre de participants a limité cette étude, et d'autres analyses seront nécessaires pour vérifier ces résultats.

MÉTHODES

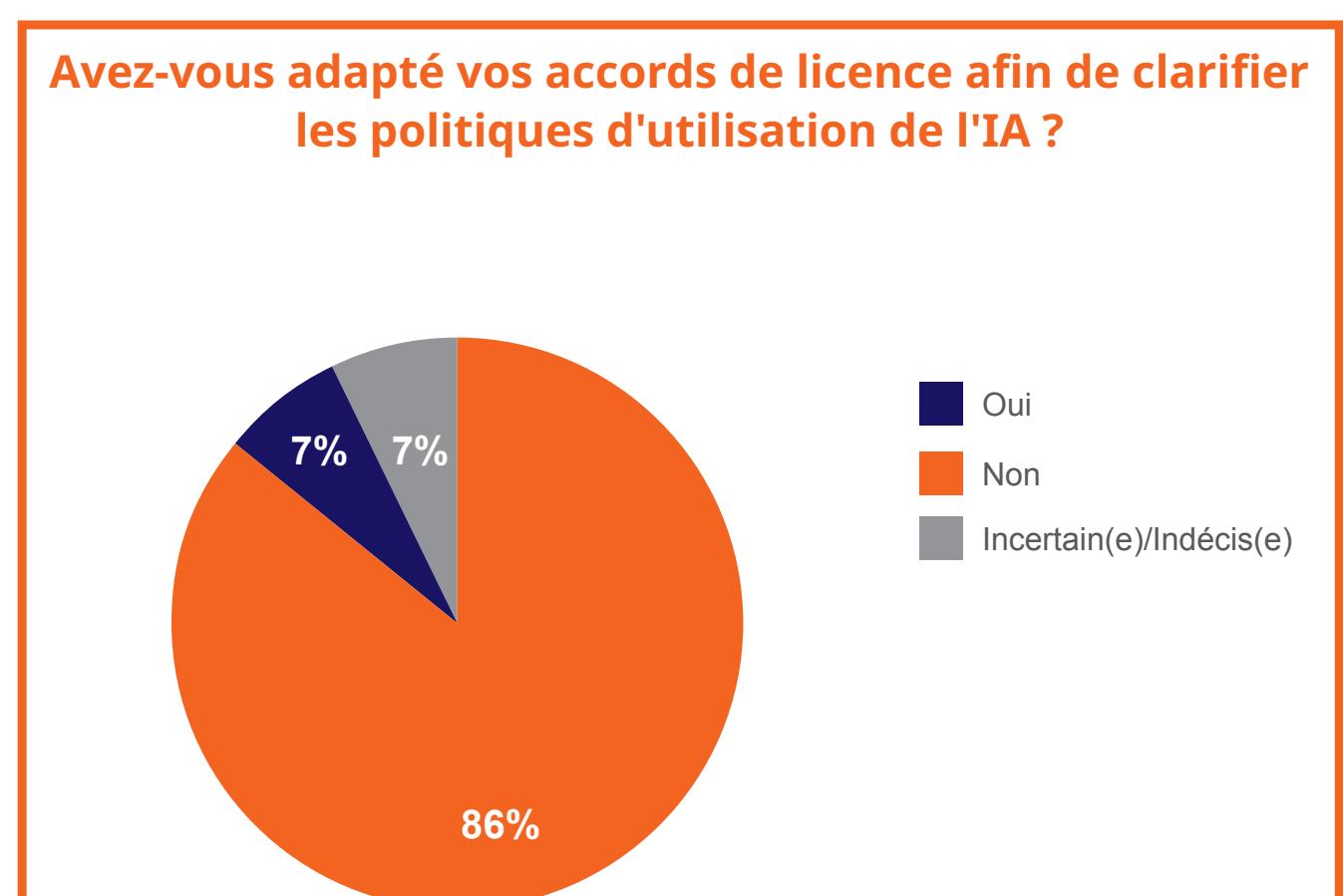
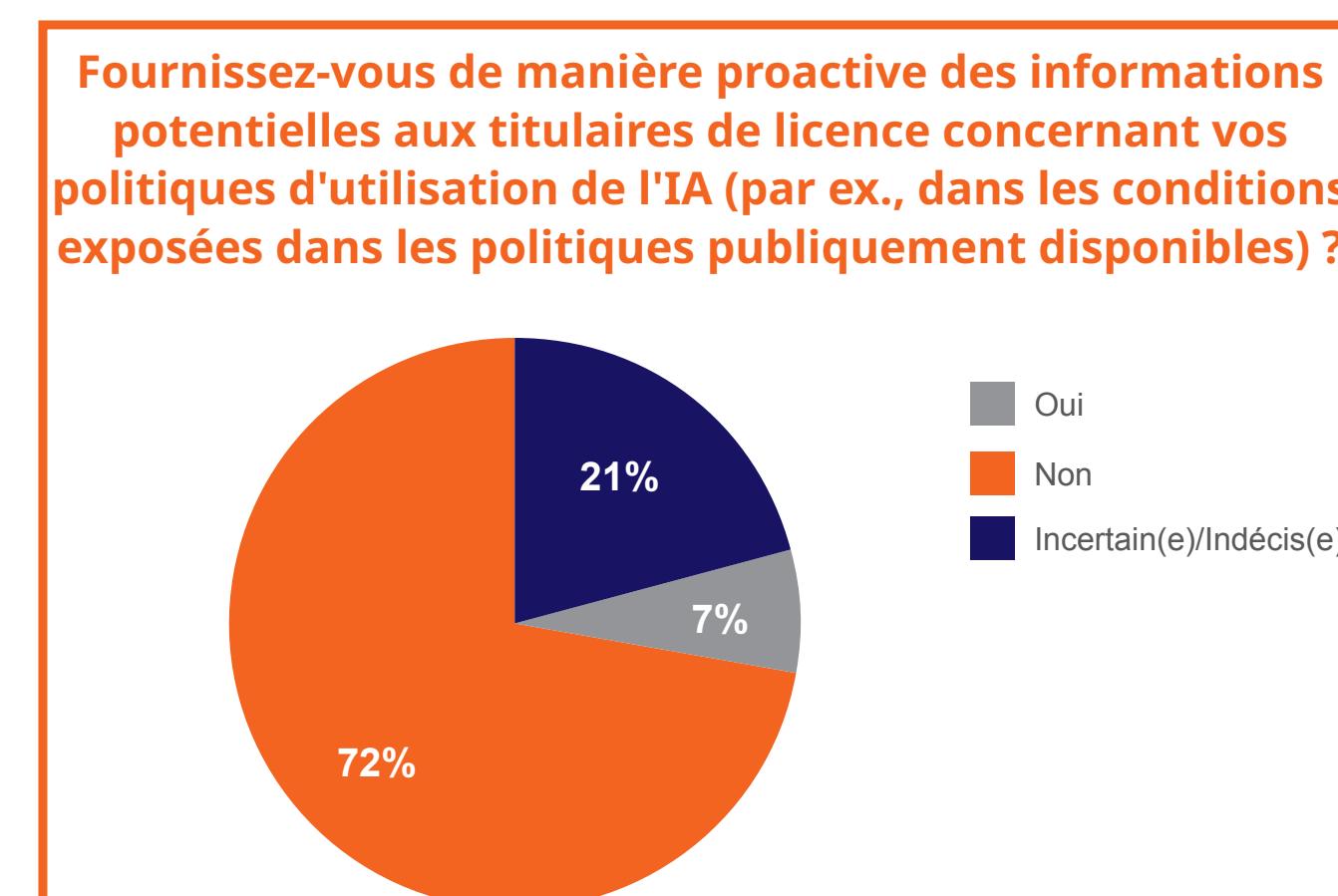
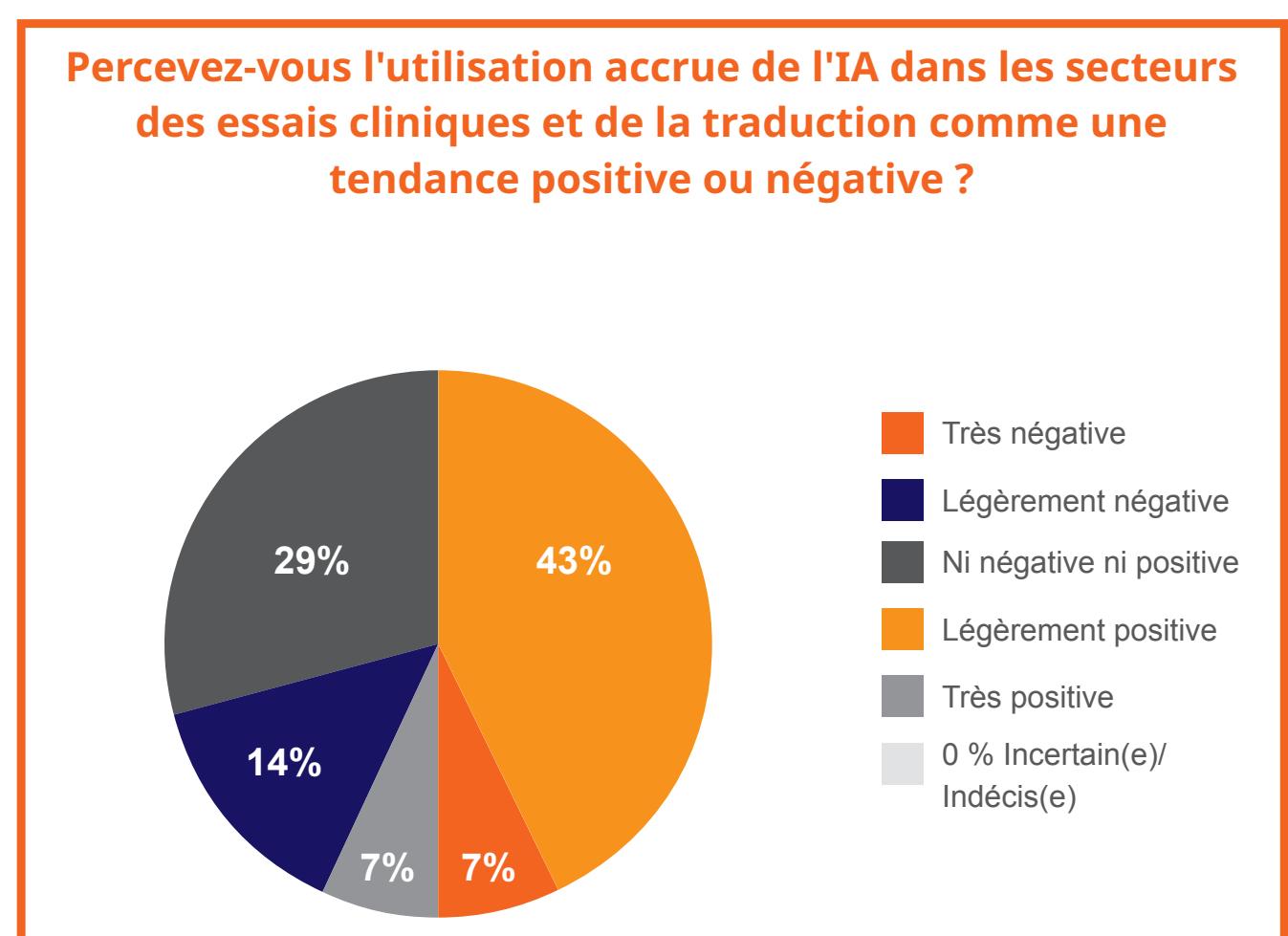
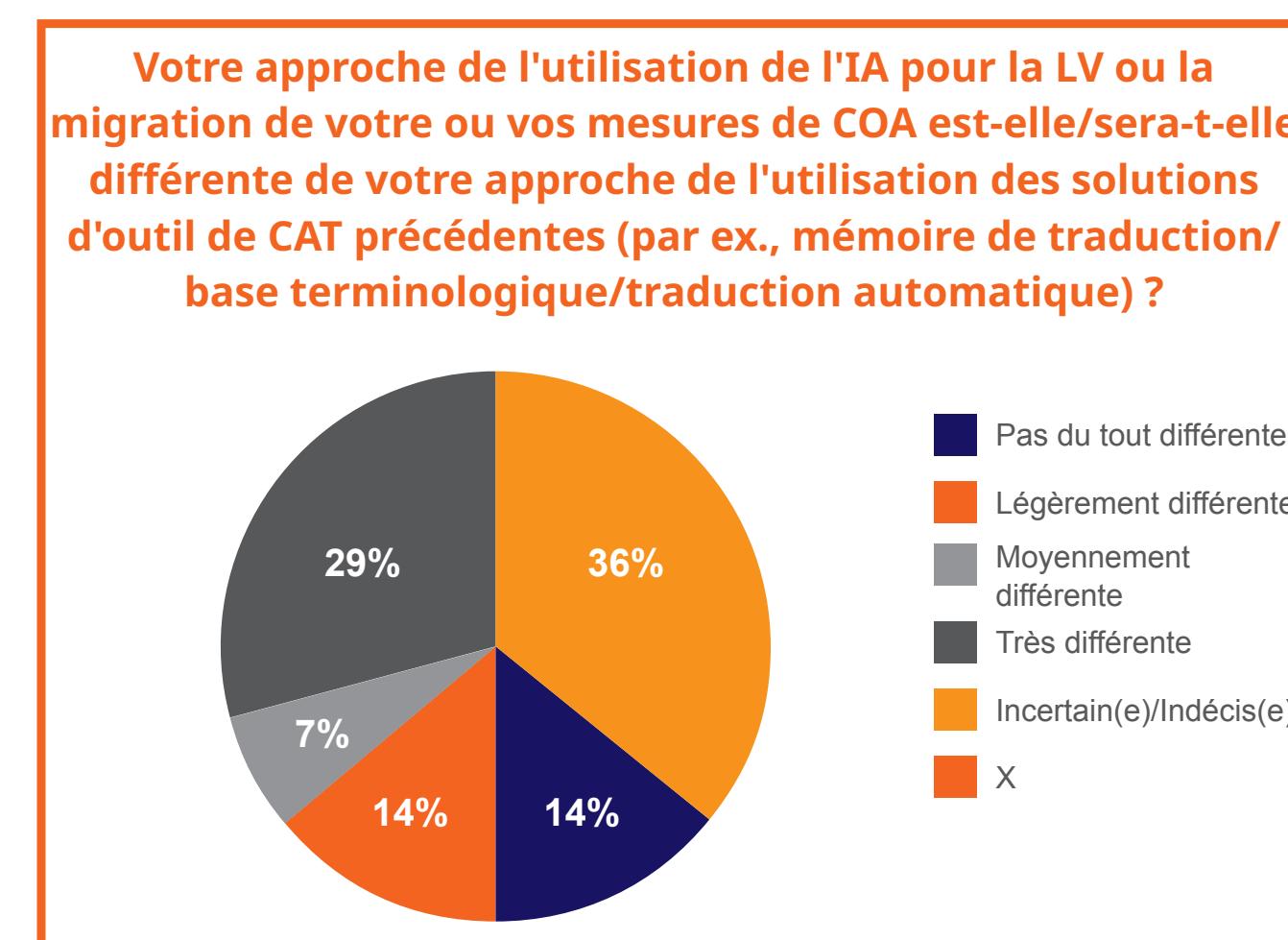
Pour atteindre cet objectif, une enquête électronique en 15 questions, abordant les trois thèmes ci-dessous, a été envoyée à 100 détenteurs de droits d'auteur de COA. Quatorze d'entre eux ont répondu à cette enquête anonyme dans le délai fixé de trois semaines.

- 1) Autorisation et confiance envers l'utilisation de l'IA dans la traduction de COA (LV) et la migration d'eCOA (migration)
- 2) Certifications et confiance
- 3) Approches actuelles de l'IA, politiques et gestion des licences

3) Approches de l'IA actuelles

a. Résultats :

- Quatre (4) questions ont été posées à ce sujet.
- 1) Lorsqu'on a demandé aux participants dans quelle mesure leur approche de l'utilisation de l'IA pour la LV et la migration serait différente par rapport à leur approche de l'utilisation d'une solution avec outil de CAT, une majorité (36 % (5/14)) a répondu Incertain(e)/Indécis(e). En comparaison, 28,5 % (4/14) ont indiqué que leur approche serait Très différente. Seuls 28,5 % (4/14) ont indiqué qu'ils n'aborderaient pas cette technologie d'une manière différente ou qu'ils l'aborderaient d'une manière légèrement différente par rapport à la technologie avec outil de CAT.
 - 2) 43 % (6/14) ont indiqué que l'utilisation de l'IA dans les essais cliniques et la traduction était Légèrement positive, et 29 % (4/14) ont répondu qu'elle n'était Pas négative ou qu'elle était Positive. Les réponses Très négative et Très positive ont chacune été choisies par 7 % des participants (1/14), tandis que les autres participants, soit 14 % (2), ont opté pour la réponse Légèrement négative.
 - 3,4) Lorsqu'on leur a demandé s'ils fournissaient activement des politiques d'utilisation de l'IA dans des forums publiquement disponibles, 72 % (10/14) ont répondu « Non ». En outre, 86 % (12/14) n'ont pas adapté leurs accords de licence pour clarifier leurs politiques sur l'utilisation de l'IA.

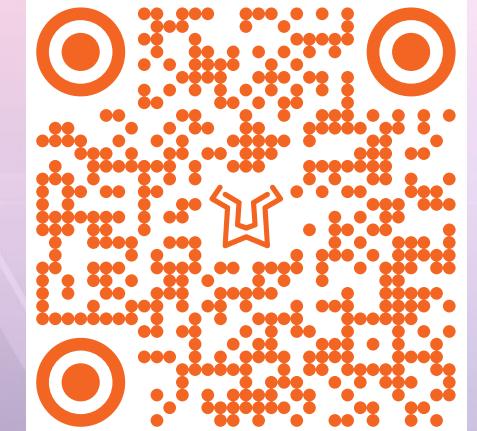


b. Interprétation : Il semblerait qu'il y ait une certaine incertitude dans la manière dont cette technologie diffère de celle, plus ancienne, utilisée dans la traduction, et dans la manière d'ajuster les approches avec ce niveau de différence à l'esprit. Il existe également un large spectre d'opinions autour de l'utilisation de l'IA dans le secteur. Cependant, le résultat le plus important sur ce thème, et potentiellement de toute l'enquête, a été le manque de proactivité pour informer les titulaires de licence potentiels des politiques sur l'IA portant sur l'utilisation de leur PI, y compris dans la documentation des accords de licence eux-mêmes. Néanmoins, cela va bientôt changer d'après les réponses en texte libre. Ce groupe d'étude recommande donc aux titulaires de licence d'admettre qu'un manque d'informations est synonyme d'un manque d'opinion ou de politique dans leurs interactions avec des détenteurs de droits d'auteur.



CAPTURE D'ÉCRAN ÉVALUATION DE L'IA GÉNÉRATIVE PAR RAPPORT À LA TRADUCTION HUMAINE RÉSULTATS DES RÉVISIONS DANS LA LOCALISATION D'ECOA :

L'IA générative est-elle la clé pour améliorer les commentaires ?



LIONBRIDGE

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

INTRODUCTION

Auparavant, la traduction, la migration et la révision des captures d'écran des Évaluations électroniques des résultats cliniques (eCOA) reposaient fortement sur les activités humaines manuelles, en partie en raison de la nature non modifiable du contenu à l'écran. Au cours du processus de révision de capture d'écran (SSR, Screenshot Review Process), les rapports d'écran dans la langue cible sont vérifiés par rapport aux rapports d'écran principaux source et, le cas échéant, par rapport au contenu existant (c.-à-d. les questionnaires papier d'origine).

Lionbridge exécute deux types de révisions de capture d'écran : **Simple** et **Complexe**. La SSR complexe diffère de la SSR simple en ce qu'elle requiert une vérification ligne par ligne explicite par rapport à tout le contenu existant cible. Lionbridge a reconnu et testé le potentiel de l'intelligence artificielle générative (IA générative) et de la technologie de reconnaissance optique de caractères (OCR, Optical Character Recognition) pour compléter les méthodes de révision de capture d'écran traditionnelles, réduire les coûts et les délais (c.-à-d. le nombre de tours de révision), améliorer la qualité, et mener enfin à de meilleurs résultats pour les patients. Cette recherche a été effectuée en utilisant les fonctionnalités de révision de capture d'écran simple et complexe de l'outil propriétaire de Lionbridge, Aurora AI Clinical Outcomes™.

MÉTHODES

Lionbridge a utilisé un moteur d'IA générative sécurisé afin de générer des commentaires d'assurance qualité concernant les rapports de capture d'écran cible pour 5 Évaluations électroniques des résultats cliniques (eCOA) orientées patient de différents niveaux de longueurs et complexité, avec 11 types d'erreurs intentionnellement intégrés à des fins de tests : Contenu manquant, Contenu non traduit, Traduction incorrecte, Modifications d'eCOA manquantes, Non-correspondances de numéro de version, Problèmes de mise en forme, Problèmes de genre, Problèmes de balises, Problème de saut de ligne, Problèmes d'utilisation des majuscules, et Problèmes d'ancre d'échelle. 16 langues cibles ont été testées au total, représentant différents alphabets et familles de langues : bulgare (Bulgarie), polonais (Pologne), roumain (Roumanie), grec (Grèce), espagnol (Argentine, Mexique, États-Unis), français (Belgique, France), portugais (Brésil, Portugal), hongrois (Hongrie), turc (Türkiye), coréen (Corée), chinois traditionnel (Taïwan) et thaï (Thaïlande). Les invités ont été personnalisées jusqu'à l'obtention d'un résultat convenable, dans le respect des dernières normes du secteur en matière d'eCOA et des pratiques établies. Simultanément, nous avons envoyé les mêmes rapports d'écran d'eCOA à des linguistes approuvés par Lionbridge pour une révision humaine. Un réviseur impartial a ensuite eu la mission de valider les deux résultats (IA et humains) et de les noter en matière d'exactitude et d'exhaustivité.

Évolution de la conception de l'invite :

La première phase de notre recherche a utilisé une invite d'IA générative pour la révision de capture d'écran simple, l'idéal pour comparer les rapports d'écran source à ceux cible après le processus de traduction (soit traduction/modification, soit traduction directe/rétro-traduction). Une fois cette fonctionnalité de l'outil finalisée, nous avons tourné notre attention vers le développement et les tests d'une fonctionnalité d'IA générative de révision de capture d'écran complexe plus nuancée, mieux adaptée au contenu migré, car elle rajoute une étape par rapport à la version simple : en plus de comparer le contenu d'écran cible au rapport d'écran principal source, la fonctionnalité complexe compare également le contenu d'écran cible au contenu existant cible.

Interface de révision de capture d'écran simple :

Interface de révision de capture d'écran complexe :

CONCLUSION

Notre recherche a démontré que l'intégration de l'IA générative à la révision humaine est la stratégie la plus efficace pour la révision de capture d'écran dans la traduction et la migration d'eCOA.

L'invite est affinée en continu afin de détecter un nombre croissant de problèmes au fil du temps. Cependant, il est actuellement toujours recommandé de conserver une intervention humaine dans le processus de révision de capture d'écran afin d'assurer une qualité optimale pour ce contenu hautement sensible. L'IA générative peut fortement améliorer la vitesse, l'efficacité et l'exactitude en identifiant les erreurs préliminaires, tandis que l'expertise humaine reste essentielle pour une prise de décision nuancée et pour identifier les problèmes que l'IA aurait pu manquer. Cette approche avec intervention humaine peut raccourcir les délais en diminuant le nombre de tours de révision requis. Cela réduit les coûts en faisant gagner du temps aux linguistes lors de leurs propres contrôles qualité, et surtout, en renforçant notre capacité à recueillir la voix du patient avec plus de précision et de qualité. Cela assure une expérience centrée sur le patient plus fiable.

RÉSULTATS ET INTERPRÉTATIONS

L'analyse des commentaires d'assurance qualité pour la fonction de SSR simple par l'IA générative a révélé qu'il s'agit d'une vérification d'AQ très utile, qui peut être utilisée en collaboration avec les commentaires humains pour les projets de traduction et de migration. La fonction simple est particulièrement adaptée pour aider les linguistes dans les projets de révision de traduction/capture d'écran qui ne nécessitent pas de vérification croisée par rapport au contenu existant. En utilisant cette fonction, l'outil Aurora AI Clinical Outcomes a capturé plusieurs instances des 11 types d'erreurs reflétés dans les écrans cible. Cependant, il était limité dans certains cas par son incapacité à vérifier le contenu existant.

Bien que la fonction d'IA générative pour la SSR complexe soit toujours en cours de création (déploiement prévu au 1er trimestre 2026), les résultats provisoires montrent qu'elle détecte tout ce que l'outil de SSR simple a identifié, et plus encore. Elle fournit également des commentaires plus nuancés par rapport à l'outil de SSR simple. La capacité à effectuer des vérifications croisées par rapport au contenu existant permettra à la fonction de SSR complexe d'éliminer certains des faux positifs produits actuellement par l'outil de SSR simple, comme indiqué dans le tableau ci-dessous.

TYPE D'ERREUR	FONCTIONNALITÉS D'IA GÉNÉRATIVE POUR LA SSR SIMPLE : REMARQUES ET LIMITATIONS	FONCTIONNALITÉS D'IA GÉNÉRATIVE POUR LA SSR COMPLEXE : REMARQUES ET LIMITATIONS	EXEMPLES
Contenu manquant	Peut signaler si le contenu cible est manquant	Peut confirmer si le texte cible est manquant intentionnellement car il n'était pas présent dans le contenu existant	Portuguese (Portugal) #4 Inconsistency: 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."
Contenu non traduit	Peut signaler le texte laissé en anglais et suggérer que cela pourrait être intentionnel (c.-à-d., texte relatif aux droits d'auteur)	Peut confirmer si le texte a été laissé en anglais intentionnellement	Bulgarian (Bulgaria) #1 Lingustic: Untranslated content Target Paragraph: 1 "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."
Traduction incorrecte	Peut identifier les écarts par rapport à la source	Peut identifier les traductions qui sont incorrectes car elles s'écartent du contenu existant, même si elles correspondent à la source	Portuguese (Portugal) #3 Lingustic: Inconsistent scale anchors Target Paragraph: 1 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.
Modifications d'eCOA manquantes		Fonctionnalité égale attendue pour les fonctions simple et complexe	Portuguese (Portugal) #1 Lingustic: Incorrect translation Target Paragraph: 4 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.
Incohérences dans les numéros de version	Peut signaler les incohérences dans les numéros de version entre la source et la cible, avec un avertissement précisant que les différences peuvent être intentionnelles	Peut confirmer si les différences de numéro de version sont intentionnelles en vérifiant le contenu existant	Korean (Republic of Korea) #3 Inconsistency: Version number mismatch Target Paragraph: 1 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.
Problèmes de mise en forme	Peut indiquer la mise en forme différente par rapport au rapport de l'écran principal source	Peut confirmer si la mise en forme s'écarte de la source intentionnellement, en fonction du contenu existant	Portuguese (Portugal) #1 Formatting/Layout Issues Target Paragraph: 5 In the source screen, the phrase 'past 7 days.' is bolded, while in the target screen only 'últimos' is bolded. The bolding should match the source, with 'últimos 7 dias.' in bold in the target.
Problèmes de morphologie	Peut détecter un genre inattendu pour les adjectifs, avec un avertissement précisant que cela peut être intentionnel (c.-à-d., que tous les objets peuvent avoir une forme féminine en raison d'une population de 100 % de patientes)	Peut fournir un contexte supplémentaire concernant le genre attendu, en vérifiant le contenu existant	Spanish (United States) #1 Lingustic: Inconsistent translation Target Paragraph: 7 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.
Problèmes d'étiquetage		Fonctionnalité égale attendue pour les fonctions simple et complexe	Thai (Thailand) #1 Formatting/Layout Issues Target Paragraph: 7 The target screen displays HTML tags (<u>) instead of rendering underlined text for 'นักศึกษาที่มีความต้องการ', 'แม่สี', and 'จำนวนผู้ดูแล'. 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.
Problèmes de saut de ligne		Fonctionnalité égale attendue pour les fonctions simple et complexe	French (France) #1 Formatting/Layout Issues Target Paragraph: 6 The target screen introduces a line break in 'habitueller' 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 'habitueller' appears on one line without a hyphen.
Problèmes d'utilisation des majuscules	Peut signaler des modèles d'utilisation des majuscules inattendus qui s'écartent de la source, avec un avertissement précisant qu'ils peuvent être intentionnels	Peut confirmer si l'utilisation des majuscules inattendue sur l'écran cible correspond au contenu existant	Hungarian (Hungary) #2 Lingustic: Inconsistent capitalization 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'.
Problèmes d'ancre d'échelle		Fonctionnalité égale attendue pour les fonctions simple et complexe	Portuguese (Portugal) #2 Formatting/Layout: Scale anchor placement Target Paragraph: 8 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.

Homme contre machine :

- L'IA générative est bien plus rapide que les humains pour évaluer le contenu de l'écran cible, pour une échelle donnée, en quelques secondes.
- Les spécialistes de l'assurance qualité de Lionbridge ont noté des gains de temps et d'efficacité grâce au processus d'IA générative, car l'interface SSR de l'outil Aurora Clinical Outcomes leur a permis d'accéder facilement aux fichiers existants, aux rapports sur les écrans principaux et aux rapports sur les écrans en langue cible simultanément.
- L'IA générative a détecté plusieurs problèmes que les réviseurs humains ont manqués, en particulier les sauts de ligne, l'utilisation des majuscules, les traductions incorrectes (y compris le texte en double) et les ancrages d'échelle. Ces problèmes peuvent être plus faciles à détecter pour l'IA générative, car l'outil utilise la reconnaissance optique de caractères (OCR) pour analyser les différences de texte dans le contenu d'image non modifiable, qu'un œil humain pourrait facilement manquer.
- L'outil SSR d'AIA générative a parfois détecté des erreurs de manière incohérente selon les langues. Par exemple, l'IA générative a détecté la traduction manquante d'« items » dans la plupart des langues, mais pas en espagnol. Cela est sans doute lié au fait que la seule différence était un accent (« ítems »).
- Les réviseurs humains étaient moins susceptibles de signaler de manière incorrecte les problèmes liés au contexte (c.-à-d., population de patientes, différences intentionnelles dans le nom de version, texte de droits d'auteur laissé intentionnellement en anglais, etc.).