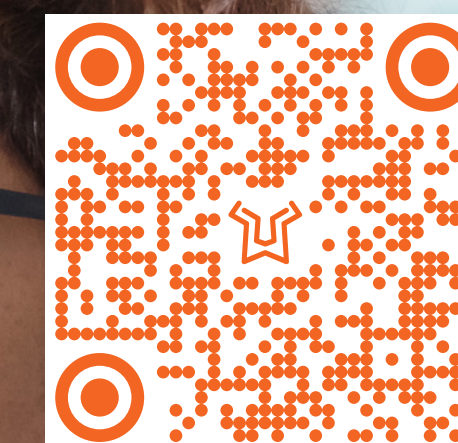


革新と適正利用のバランス： 知的財産のローカリゼーションにおける 生成 AI 利用についての著作権者の見解



LIONBRIDGE

著者: メリンダ ジョンソン、ミーガン ダフ、キャスリン ノルテ、ステファニー カサーレ

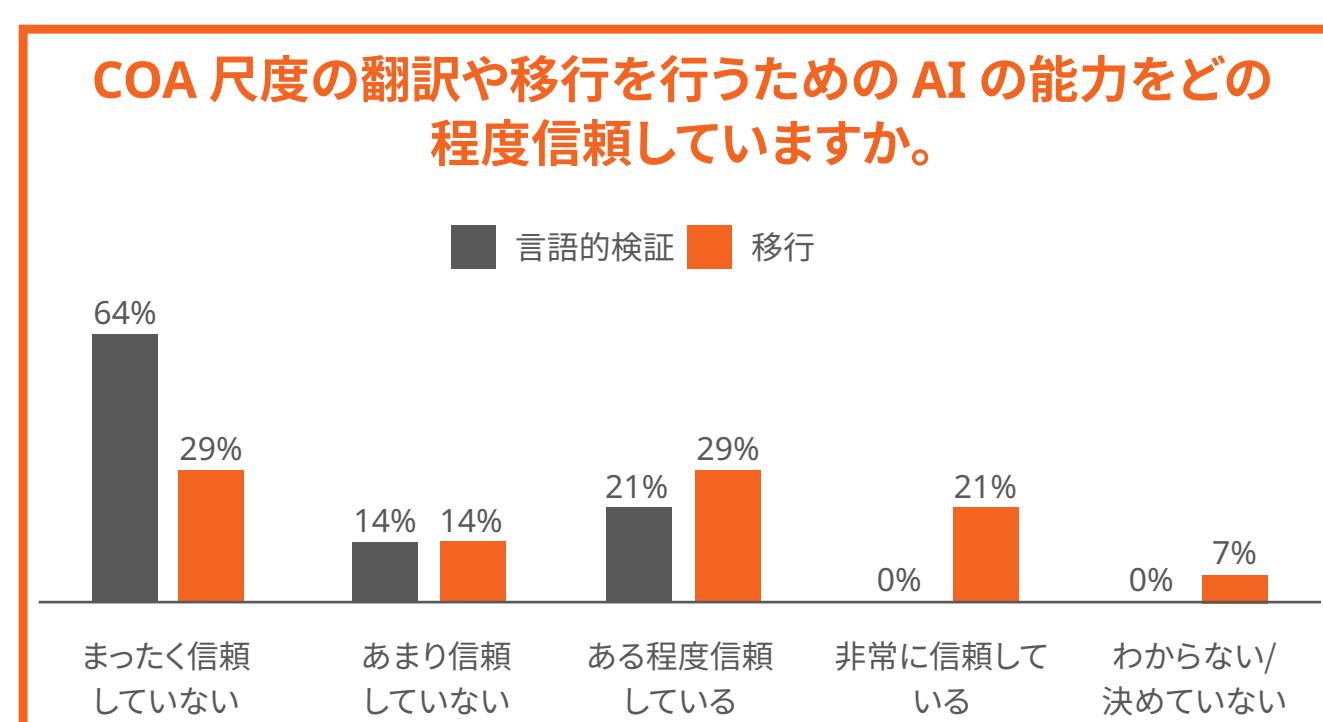
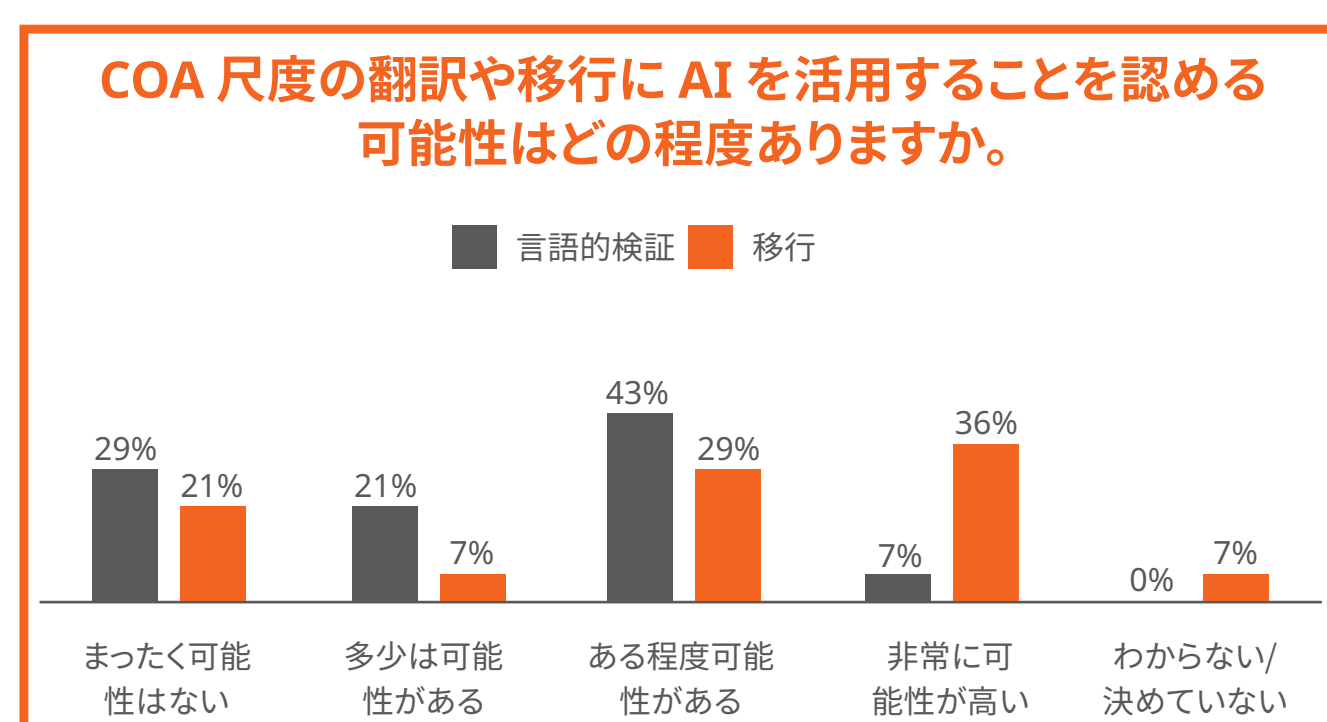
目的

臨床試験の運用に生成系人工知能 (生成 AI) を統合するという課題は、特に臨床アウトカム評価 (COA) に関する領域で、COA の知的財産 (IP) に対する生成 AI の適用について明確なガイダンスが不足しているため、対策の進展が妨げられています。この不明確さにより、COA の言語的検証や移行などという、患者の声を反映するうえで重要なプロセスにおける生成 AI の導入が滞っています。この調査は、こうしたプロセスでの生成 AI 利用に関する COA の著作権者の見解を収集・共有し、今後の導入戦略に活かすことを目的として実施したものです。

結果と解釈

1) 言語的検証および移行における生成 AI 使用の許容度と信頼度

a. 結果: 回答した COA 著作権者の 64% (9/14) は、言語的検証プロセスにおいて自らの COA 尺度の翻訳に生成 AI を使用することに対し、まったく信頼感を抱いていませんでした。しかし、言語的検証プロセスにおける自らの COA 尺度の翻訳に、生成 AI の使用を許可する可能性がまったくないとした回答は、29% (4/14) にとどまりました。eCOA 移行プロセスにおける生成 AI の使用については、信頼度と許容度の回答がより一致しており、生成 AI をまったく信頼していないとした回答が 29% (4/14) であったのに対し、使用を許可する可能性がまったくないとした回答は 21% (3/14) でした。

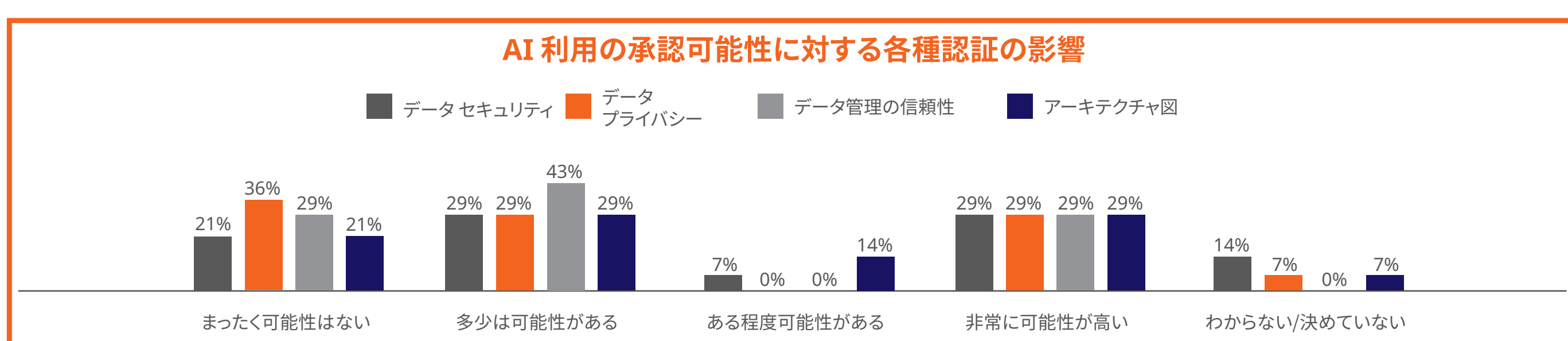


b. 解釈: AI の性能に対する信頼感という点では、著作権者の信頼度は言語的検証と移行の各プロセスで異なっていましたが、AI の使用を許可する可能性については両方のワークフローで一貫して肯定的であり、その程度には差があるものの (「ある程度」対「非常に」)、言語的検証については 50% (7/14)、移行については 65% (9/14) が、AI の使用を許可するという方向でした。自由記述形式の回答によると、全体として AI に対する信頼度は低い状況でしたが、丁寧な対話を経ることで、AI の使用を許可してもよいと考える回答者が大半を占めました。

2) 認証と信頼性

a. 結果: この質問では著作権者に、いくつかの認証や関連文書の有無が、AI の使用を認める可能性にどの程度影響するかを確認しました。取り上げたのは、1) データセキュリティに関する認証 (例: ISO27001:2022)、2) データ プライバシーに関する認証 (例: ISO27001:2019)、3) データ管理の信頼性に関する認証 (例: SOC2 Type 2)、4) AI アーキテクチャ図、の 4 種類です。これらの認証や文書の影響力については回答が分かれたましたが、とりわけデータ プライバシーに関する認証と、データ管理の信頼性に関する認証の影響力については、懐疑的な見方が最も多くみられました。一方で、これら 4 種類のいずれの認証・文書についても、それが適切に提示されれば、AI の使用を認める可能性が非常に高いと答えた回答者は 29% (4/14) でした。

b. 解釈: さまざまな文書類 (認証や図表)、データセキュリティ、プライバシー、および AI エンジンの動作の信頼性については疑念が見られるものの、こうした文書類を提示することで、少なくとも 29% の著作権者は、自身の IP を翻訳する際の AI 使用について前向きな姿勢が高まることが示されました。



まとめ

COA 尺度の著作権者は他のステークホルダーと同様、自身の知的財産をローカライズする際の AI の利用に関する明確なガイダンスが依然として不足している業界で、その対応策を模索している状況です。しかし、IP のローカライズ時の適切な AI 利用に対するステークホルダーの信頼感にはまだ改善の余地がある一方で、AI エンジンとやり取りする際に自らの IP の安全性を確保するための、議論やポリシー策定、ライセンス契約の改訂、文書類の再確認には前向きな姿勢が示されました。最後に、本調査は回答者数が限られていたことから、その結果には一定の制約があり、これらの結果を裏付けるには今後さらに詳細な分析が求められます。

手法

この目的で、以下の 3 つのトピックからなる 15 問の電子アンケートを、COA の著作権者 100 名に送付しました。設定した 3 週間の回答期間内に、14 名の著作権者から、この匿名アンケートへの回答が得られました。

- 1) COA の翻訳 (言語的検証) および eCOA への移行 (移行) における生成 AI 使用の許容度と信頼度
- 2) 認証と信頼性
- 3) 現行の生成 AI のアプローチ、方針、ライセンス

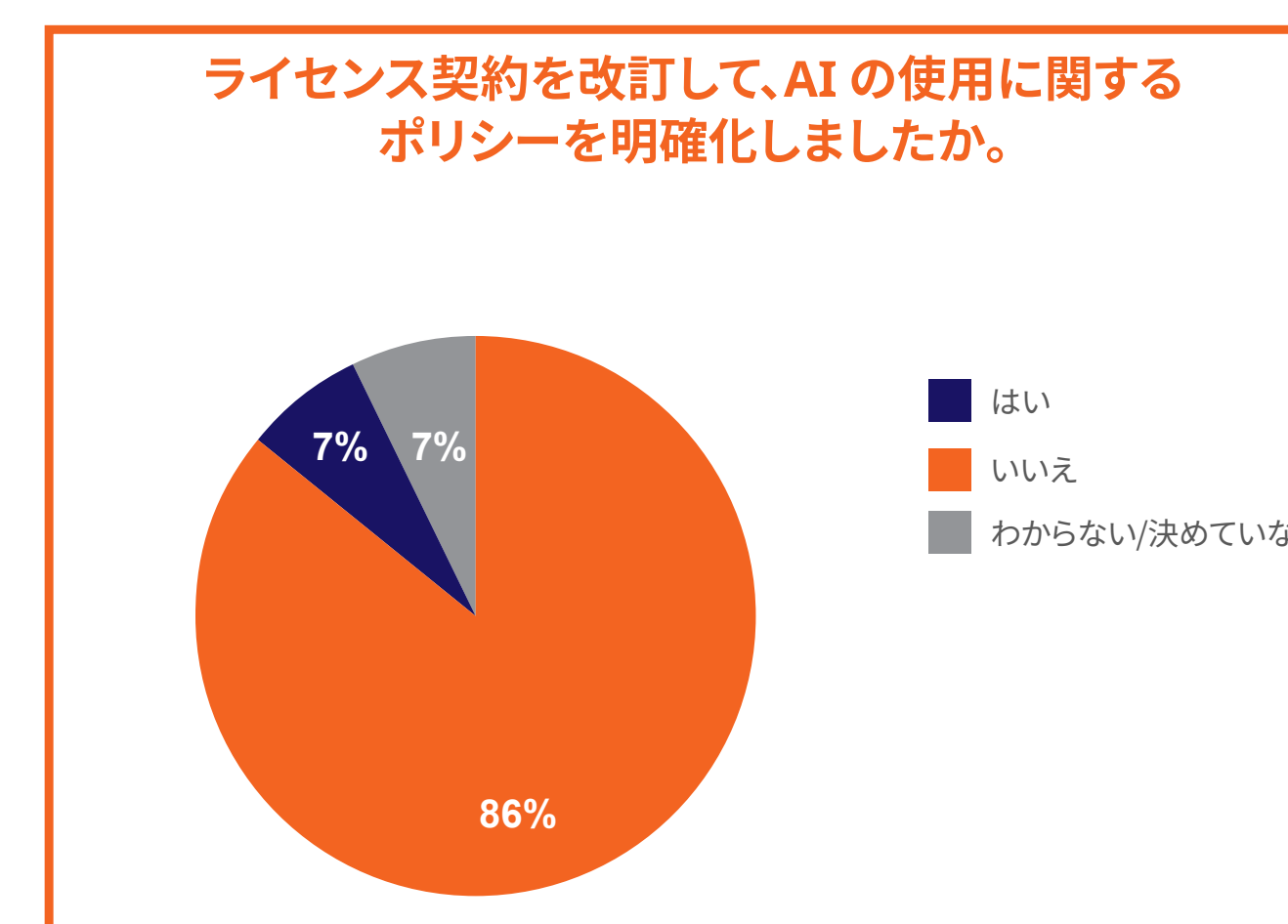
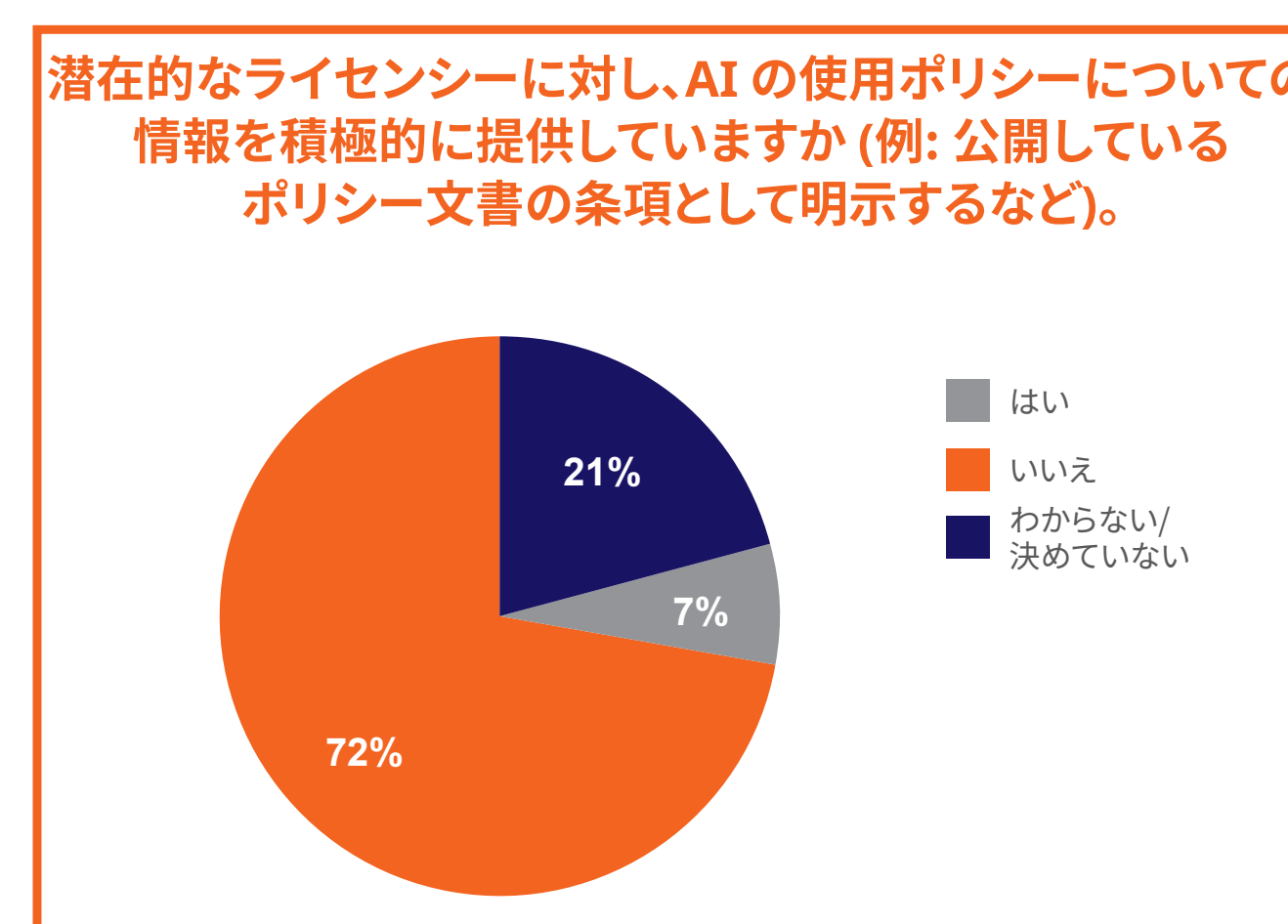
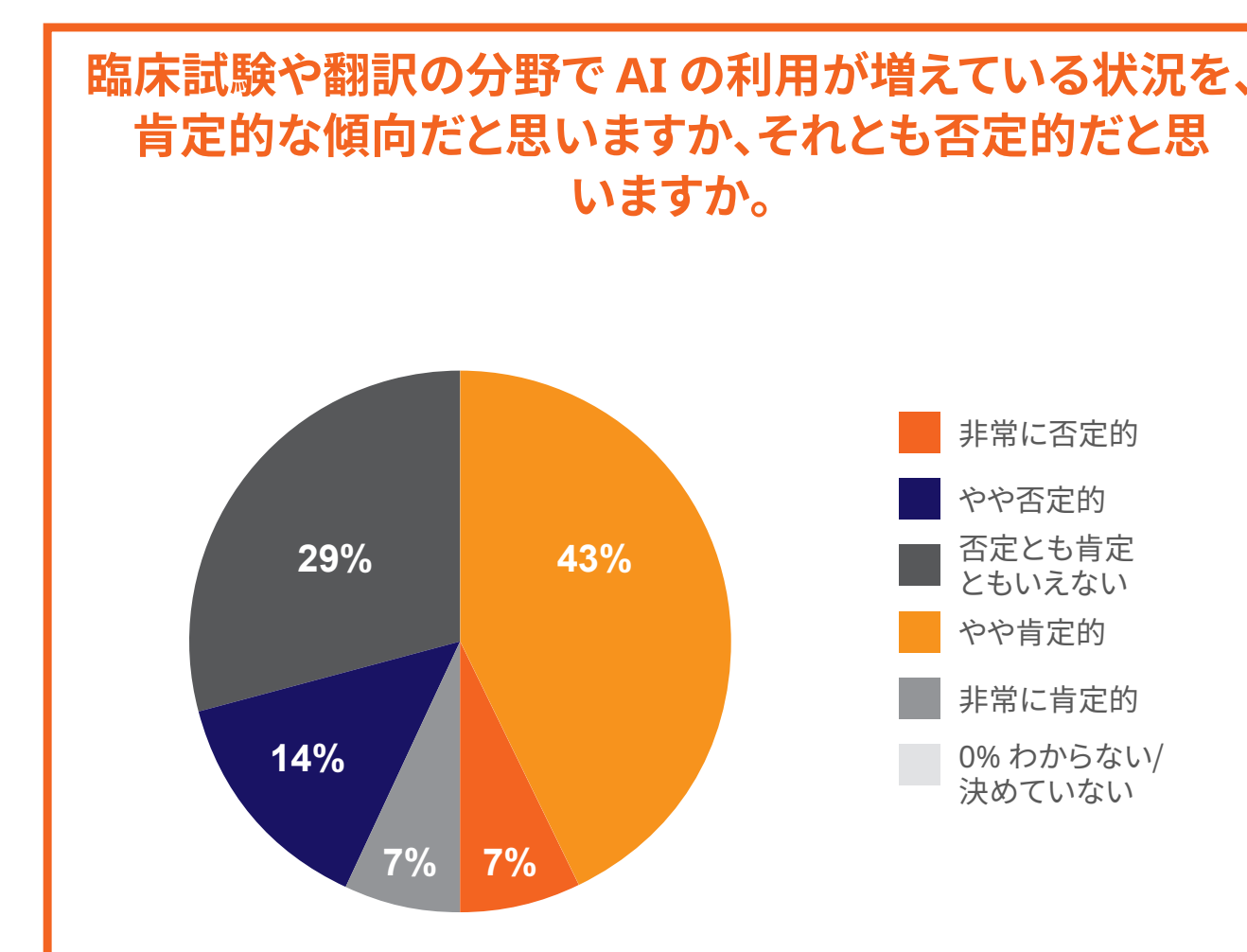
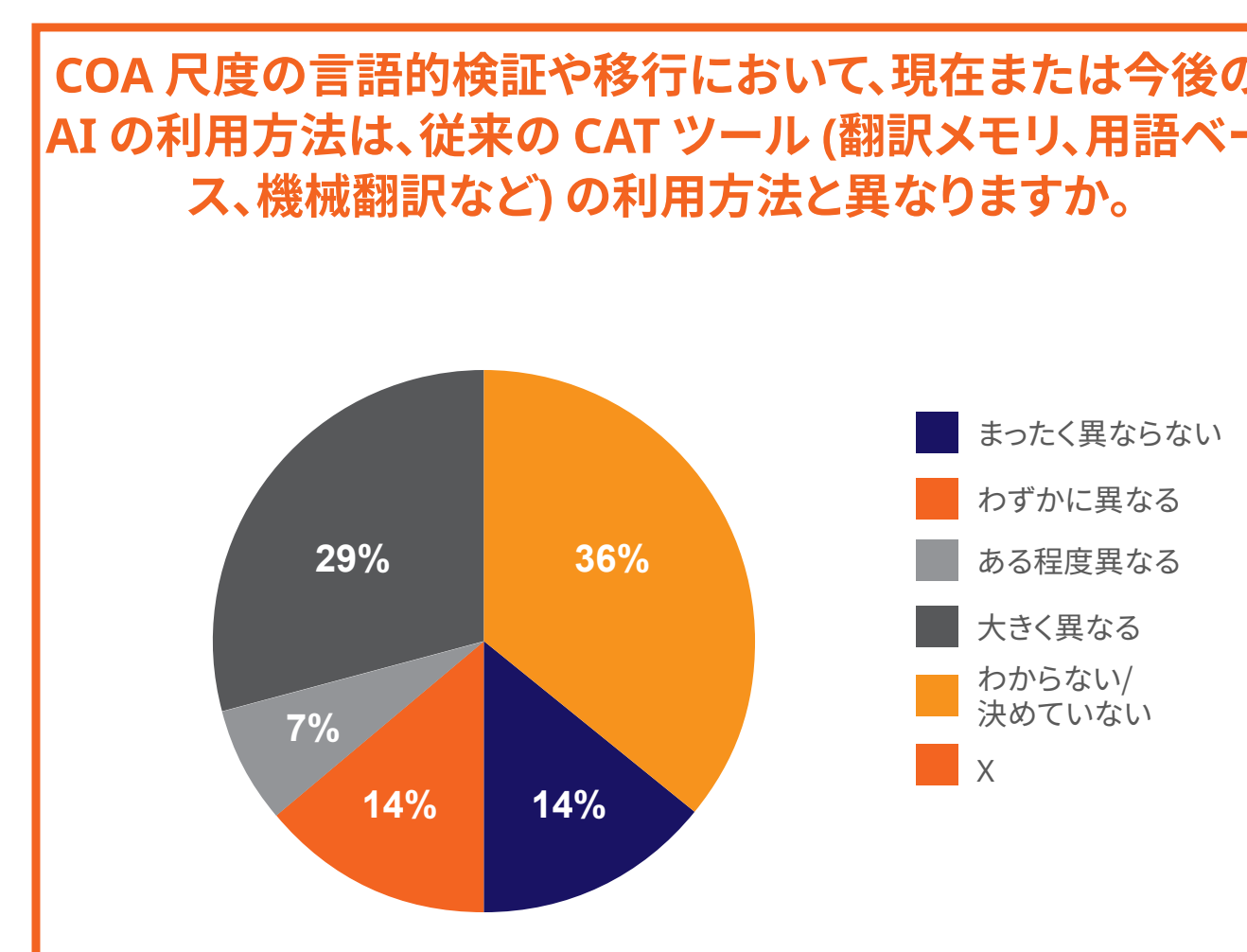
3) 現行の生成 AI のアプローチ

a. 結果: このテーマについては 4 問の質問を行いました。

1) 言語的検証と移行の各プロセスにおける AI の使い方が、CAT ツールの使い方とどの程度異なると思うかを尋ねたところ、最も多かった回答は「わからない/判断できない」で 36% (5/14) でした。一方、「大きく異なるだろう」と回答した人は 28.5% (4/14) で、さらに CAT ツールの場合と比べて「変わらない」または「わずかに異なる程度」と答えた人も 28.5% (4/14) にとどまりました。

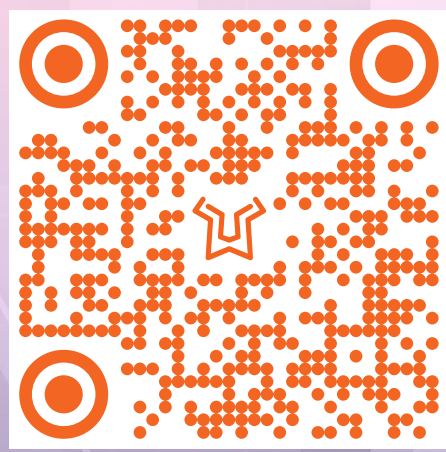
2) 臨床試験や翻訳における生成 AI の使用については、43% (6/14) が「やや肯定的」と回答し、それに次いで 29% (4/14) が「肯定的でも否定的でもない」と答えました。「非常に否定的」および「非常に肯定的」と回答した人はいずれも 7% (1/14) で、残る 14% (2 名) は「やや否定的」と回答しました。

3、4) AI の使用ポリシーを公開フォーラム上で積極的に公表しているかどうかを尋ねたところ、72% (10/14) が「いいえ」と回答しました。さらに 86% (12/14) は、AI 使用ポリシーを明確化するためのライセンス契約の見直しを行っていませんでした。



b. 解釈: このテクノロジーが従来の翻訳技術とどのように異なるのか、またその違いの大きさを踏まえてどのように使い方を調整すべきかについては、いくぶん迷いが見られました。また、業界内には AI の利用について幅広い意見が存在していました。しかしこのトピックにおいて、そしておそらくこの調査全体の中でも、特に際立っていた結果は、権利者が自らの IP (知的財産) に関する AI 使用ポリシーを潜在的ライセンシーに積極的に知らせていないという点であり、それはライセンス契約書の文書内でも同様でした。しかし自由記述回答によれば、この状況はまもなく変わるものと見込まれます。そのため本研究グループでは、ライセンシーが権利者とやり取りを行う際に情報が示されていない場合、現時点では意見や方針が確立されていないものとみなすことを推奨しています。





著者: キャスリン ノルテ、カロリナ エリソンド ヒメネス、ルパリ カダム、メリンダ ジョンソン

はじめに

電子臨床アウトカム評価 (eCOA) の翻訳、移行、およびスクリーンショット レビューは、画面上のコンテンツを編集できないという特性もあり、従来は主として人手による作業に大きく依存してきました。スクリーンショット レビュー (SSR) では、ターゲット言語の画面レポートをソース側のマスター画面レポートと照合し、必要に応じて旧来のコンテンツ (原本となる紙の質問票など) とも比較します。

ライオンブリッジが実施するスクリーンショット レビューには、Simple と Complex の 2 種類があります。Complex SSR が Simple SSR と異なるのは、ターゲット側の旧来のコンテンツ全体を対象に、明示的な逐語レベルの照合が必要になるという点です。当社では、従来のスクリーンショット レビューの手法を補完し、コストやレビュー期間 (レビュー ラウンド数) を削減し、品質を向上させ、最終的により良い患者アウトカムの実現につなげるための、生成 AI と OCR (光学的文字認識) 技術の可能性に着目し、検証を行いました。本調査は、当社独自の Aurora AI Clinical Outcomes™ ツールに搭載された、Simple および Complex の各スクリーンショット レビュー機能を活用して実施しました。

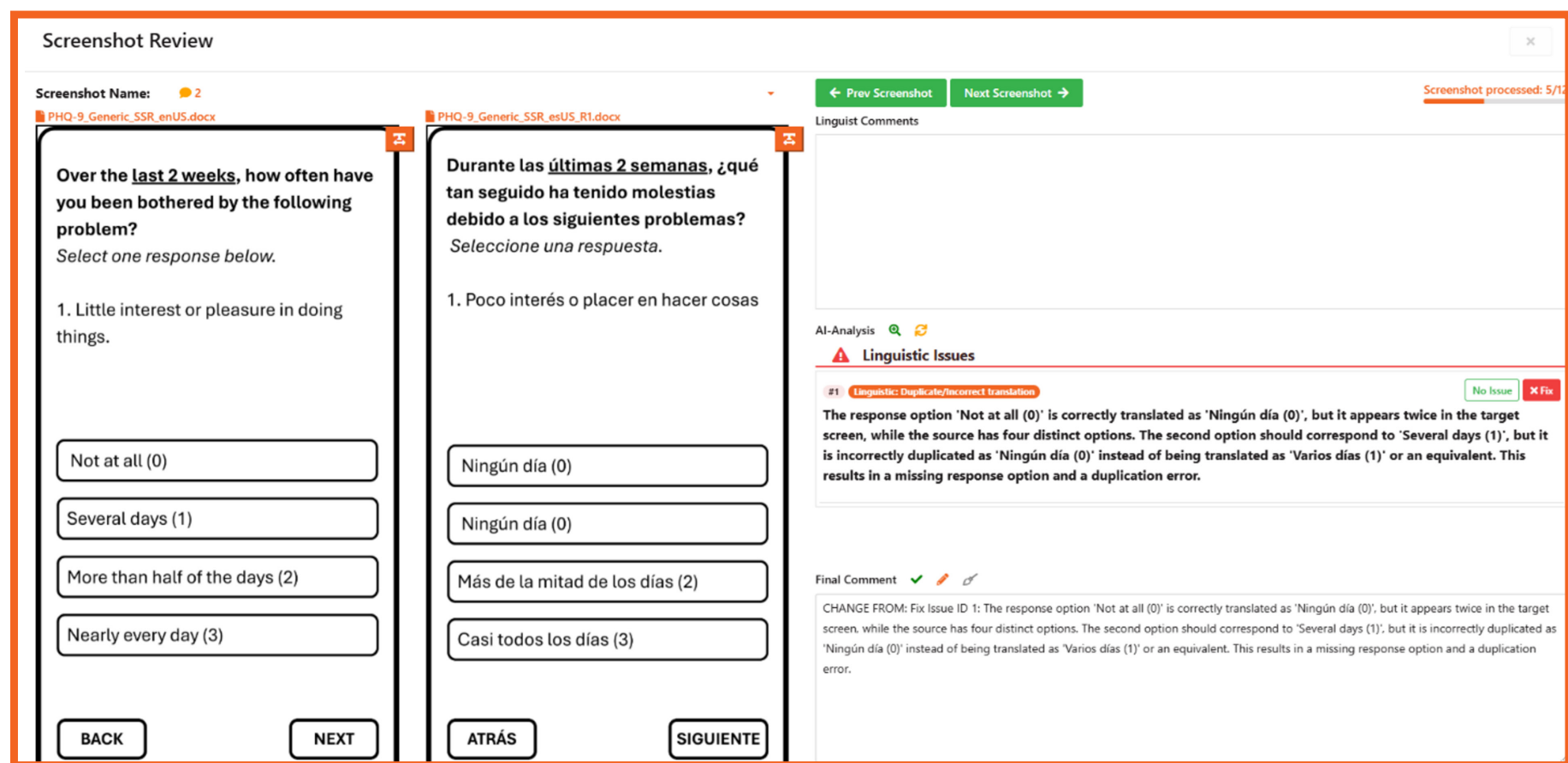
手法

ライオンブリッジでは安全な生成 AI エンジンを活用して、患者向けの 5 種類の電子臨床アウトカム評価 (eCOA) について、ターゲット言語版スクリーンショット レポートに対する品質保証フィードバックを生成しました。これらの eCOA は長さや複雑さのレベルがさまざまであり、テスト目的で 11 種類のエラーが意図的に組み込まれていました。すなわち、コンテンツの欠落、未翻訳コンテンツ、誤訳、eCOA 編集の内容の欠落、バージョン番号の不一致、書式の問題、語形の問題、タグの問題、改行の問題、大文字・小文字の問題、スケール アンカーの問題の 11 種類です。そしてさまざまな文字体系と言語系統を代表するターゲット言語を選択し、ブルガリア語 (ブルガリア)、ポーランド語 (ポーランド)、ルーマニア語 (ルーマニア)、ギリシャ語 (ギリシャ)、スペイン語 (アルゼンチン、メキシコ、米国)、フランス語 (ベルギー、フランス)、ポルトガル語 (ブラジル、ポルトガル)、ハンガリー語 (ハンガリー)、トルコ語 (トルコ)、韓国語 (韓国)、繁体字中国語 (台湾)、タイ語 (タイ) の合計 16 言語でテストを行いました。プロンプトは、eCOA に関する最新の業界標準および確立された実務プロセスに沿った適切な出力が得られるまで調整を加えました。また同時に、同じ eCOA スクリーンレポートを当社認定の言語担当者に送り、人間によるレビューも実施しました。その後、公平な第三者が AI と人間による両方の出力を検証し、正確さと完全性の観点から評価を行いました。

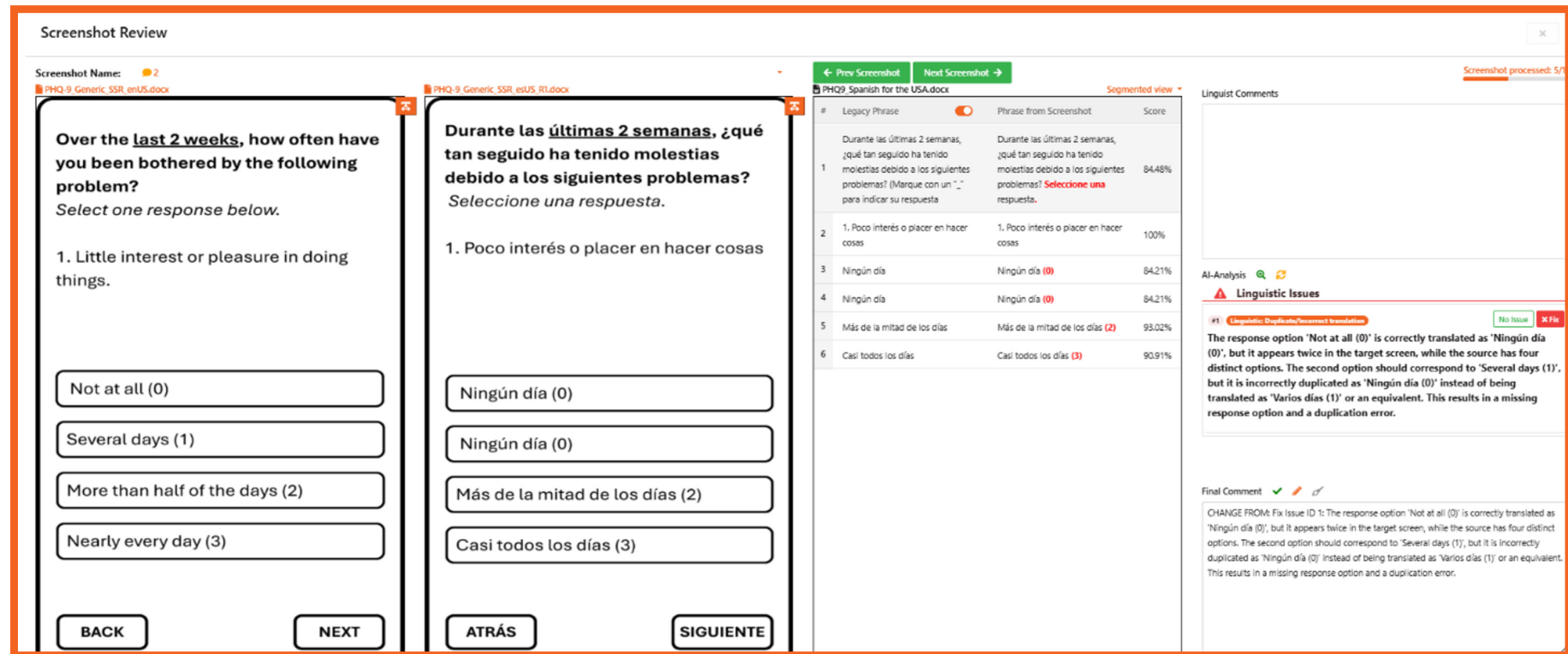
プロンプト設計の進化:

本調査の第 1 フェーズでは、Simple スクリーンショット レビュー向けの生成 AI プロンプトを使用しました。このプロンプトは、翻訳ワークフロー (翻訳/編集または順翻訳/逆翻訳) の後に、ソース側とターゲット側のスクリーン レポートを比較する用途に最適です。ツールのこの AI 機能が完成した後、当社ではより精緻な Complex スクリーンショット レビューに向けた生成 AI 機能の開発とテストに取り組みました。この機能は Simple 版より一段階進んだ処理が可能なため、移行作業を経たコンテンツにいっそう適しています。Complex 版の機能では、ターゲット側の画面上のコンテンツをソース側のマスター画面レポートと比較するだけでなく、ターゲット側の旧来コンテンツとも比較できます。

Simple スクリーンショット レビューのインターフェイス:



Complex スクリーンショット レビューのインターフェイス:



まとめ

この調査の結果、eCOA の翻訳や移行におけるスクリーンショット レビューでは、生成 AI と人間によるレビューを組み合わせることが最も効果的であることが示されました。プロンプトは、検出可能な問題の範囲を拡大できるよう、継続的に改良を進めています。ただし、この種のコンテンツは取り扱いに細心の注意が求められることから、現時点では最適な品質を確保するため、スクリーンショット レビューのプロセスに人間の関与を維持することが推奨されます。生成 AI は、初期段階のエラーを特定することで作業のスピード、効率、精度を大きく高めることができますが、一方でより微妙な判断や、AI が見落とす可能性のある問題の特定には、人間の専門技能が依然として欠かせません。この「人間参加型 (ヒューマンインザループ)」アプローチには、必要なレビュー ラウンド数を減らすことでスケジュールを効率化し、言語担当者による品質チェックの工数削減によってコストを抑えられるという効果があります。さらにとりわけ重要なのは、患者の声をこれまで以上に正確かつ質の高い形で反映させられるようになるため、より信頼性の高い患者中心の体験を実現できるという点です。

結果と解釈

Simple SSR で生成 AI 機能が出力した品質保証フィードバックを分析した結果、この機能は翻訳案件と移行案件のいずれにおいても、人間によるフィードバックと組み合わせで活用できる、非常に有用な QA チェック手段であることが明らかになりました。Simple 機能は特に、旧来のコンテンツとの照合が求められない翻訳案件やスクリーンショット レビュー案件で、言語担当者をサポートするために活用できます。Aurora AI 臨床アウトカム評価ツールでこの機能を利用することで、ターゲット側の画面に反映されていた 11 種類すべてのエラータイプについて、複数の発生例を検出することができました。ただし旧来のコンテンツを照合できないという特性から、一部のケースでは限界も見られました。

Complex SSR の生成 AI 機能は現在も開発中ですが (2026 年第 1 四半期にリリース予定)、これまでのテスト結果では、Simple SSR ツールが検出した内容をすべて拾えるだけでなく、それ以上の項目も検出できることが示されています。また Simple SSR ツールと比較して、フィードバックもよりきめ細かく返せるようになります。Complex SSR 機能では旧来のコンテンツとの照合が可能になることで、以下の表に示すような、現在 Simple SSR ツールで生成されてしまう誤検出の一部を排除できるようになります。

エラーのタイプ	SIMPLE SSR の生成 AI 機能: 注意点と制約	COMPLEX SSR の生成 AI 機能: 注意点と制約	例
コンテンツの欠落	ターゲット側のコンテンツが欠落している場合にそれを検出できる	ターゲット側のテキストが欠落している場合に、それが旧来のコンテンツに存在しなかったことによる意図的なものかどうかを判定できる	<div>Portuguese (Portugal)</div> <div>#4 [Secondary/Missing content] Target Paragraph: 1</div> <div>No Issue X Fix</div> <div>"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."</div>
未翻訳コンテンツ	英語のまま残されているテキストを検出し、それが意図的である可能性を指摘できる (例: 著作権表記の場合など)	テキストが意図的に英語のまま残されているのかどうかを確認できる	<div>Bulgarian (Bulgaria)</div> <div>Linguistic Issues</div> <div>No Issue X Fix</div> <div>"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."</div>
誤訳	ソーステキストとの不一致を検出できる	翻訳がソースと一致していても、旧来コンテンツに照らすと誤訳であるケースを検出できる	<div>Portuguese (Portugal)</div> <div>#3 [Linguistic/Inconsistent scale anchors]</div> <div>No Issue X Fix</div> <div>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.</div>
eCOA 編集の内容の欠落	Simple/Complex ともに同等の機能が予定されている		<div>Portuguese (Portugal)</div> <div>#1 [Linguistic/Inherent translation]</div> <div>Source Paragraph: 6</div> <div>No Issue X Fix</div> <div>Target Paragraph: 4</div> <div>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.</div>
バージョン番号の不一致	ソースとターゲット間のバージョン番号の不一致を検出できる。ただし、その不一致が意図的である可能性もある	旧来コンテンツを参照することで、バージョン番号の不一致が意図的なものかどうかを確認できる	<div>Korean (Republic of Korea)</div> <div>#1 [Secondary/Version number mismatch]</div> <div>✓ No Issue Fix</div> <div>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.</div>
書式の問題	ソースのマスター画面レポートと異なる書式を検出できる	旧来コンテンツに基づいて、書式の相違が意図的なものかどうかを確認できる	<div>Portuguese (Portugal)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout: inconsistent bolding/underlining]</div> <div>Source Paragraph: 7</div> <div>No Issue X Fix</div> <div>Target Paragraph: 6</div> <div>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.</div>
語形の問題 (一部の言語)	形容詞の語形が通常とは異なる性で用いられているケースを検出できる。ただし、それが意図的な場合もある (女性患者が 100% の患者群について、すべての形容詞に女性形が用いられているなど)	旧来コンテンツとの照合により、どの性の語形が使用されるべきかについて、追加の文脈情報を提供できる	<div>Spanish (United States)</div> <div>Linguistic Issues</div> <div>#1 [Linguistic/Inconsistent translation]</div> <div>Source Paragraph: 5</div> <div>Target Paragraph: 5</div> <div>✓ No Issue X Fix</div> <div>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.</div>
タグの問題	Simple/Complex ともに同等の機能が予定されている		<div>Thai (Thailand)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout: Underlining/HTML Tag Display]</div> <div>Source Paragraph: 7</div> <div>No Issue X Fix</div> <div>The target screen displays HTML tags (<u>) instead of rendering underlined text for 'หัตถ์หัตถ์หัตถ์', 'หัตถ์', and 'หัตถ์หัตถ์หัตถ์'. 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.</div>
改行の問題	Simple/Complex ともに同等の機能が予定されている		<div>French (France)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout: Line break]</div> <div>Source Paragraph: 6</div> <div>Target Paragraph: 4</div> <div>No Issue X Fix</div> <div>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.</div>
大文字・小文字の問題 (一部の言語)	ソースと異なる大文字・小文字の使い方を検出できる。ただしそれが意図的な場合もある	ターゲット側の大文字・小文字の使い方が通常と異なる場合、それが旧来コンテンツに基づくものかどうかを確認できる	<div>Hungarian (Hungary)</div> <div>#2 [Linguistic/Inconsistent capitalization]</div> <div>Source Paragraph: 6</div> <div>No Issue X Fix</div> <div>Target Paragraph: 6</div> <div>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'.</div>
スケール アンカーの問題	Simple/Complex ともに同等の機能が予定されている		<div>Portuguese (Portugal)</div> <div>#2 [Formatting/Layout: Scale anchor placement]</div> <div>Source Paragraph: 8</div> <div>No Issue X Fix</div> <div>Target Paragraph: 7</div> <div>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.</div>

人間 vs. 生成 AI:

- 生成 AI は、人間よりもはるかに高速にターゲット画面の内容を評価し、特定のスケールに属するターゲット画面すべてについて、わずか数秒でフィードバックを生成できます。
- ライオンブリッジの品質保証担当者は、生成 AI を利用したプロセスによって作業時間と効率が増えたと述べています。Aurora 臨床アウトカム評価ツールの SSR インターフェイスを使用することで、レガシー ファイル、マスター画面レポート、およびターゲット言語版の画面レポートを同時に、簡単に参照できるからです。
- 生成 AI は、人間によるレビューで見落とされていた複数の問題を検出しました。特に、改行、大文字・小文字の表記、誤訳 (テキストの重複を含む)、スケール アンカーの問題などは、生成 AI のほうが検出しやすい可能性があります。AI は編集不可の画像内テキストに含まれる不一致を解析する際に、人の目では見落としやすい部分も、OCR (光学的文字認識) を用いて読み取ることができるからです。
- 生成 AI による SSR では、さまざまな言語間でエラーの検出にばらつきが生じることがありました。たとえば、「items」の訳抜けについては多くの言語で検出できたものの、スペイン語では見逃されていました。これはおそらく、違いがアクセント記号 (ítems) のみであったためと考えられます。
- 人間によるレビューでは、文脈に依存する問題 (例: 女性患者集団に関する記述、意図的なバージョン番号の不一致、意図的に英語のまま残された著作権表記など) が、誤って検出される可能性は低くなりました。