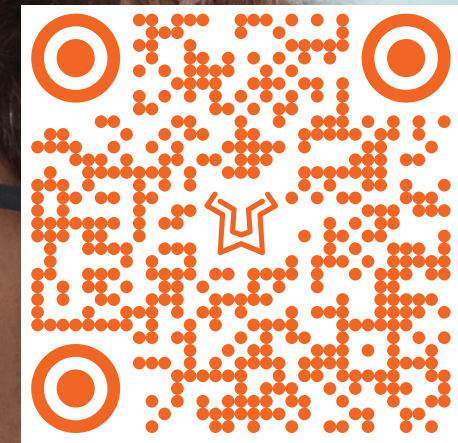


創新與所有權間的平衡

探討著作權所有人對其智慧財產本地化流程中使用 GenAI 的看法



LIONBRIDGE

作者：Melinda Johnson、Megan Duff、Kathryn Nolte、Stephanie Casale

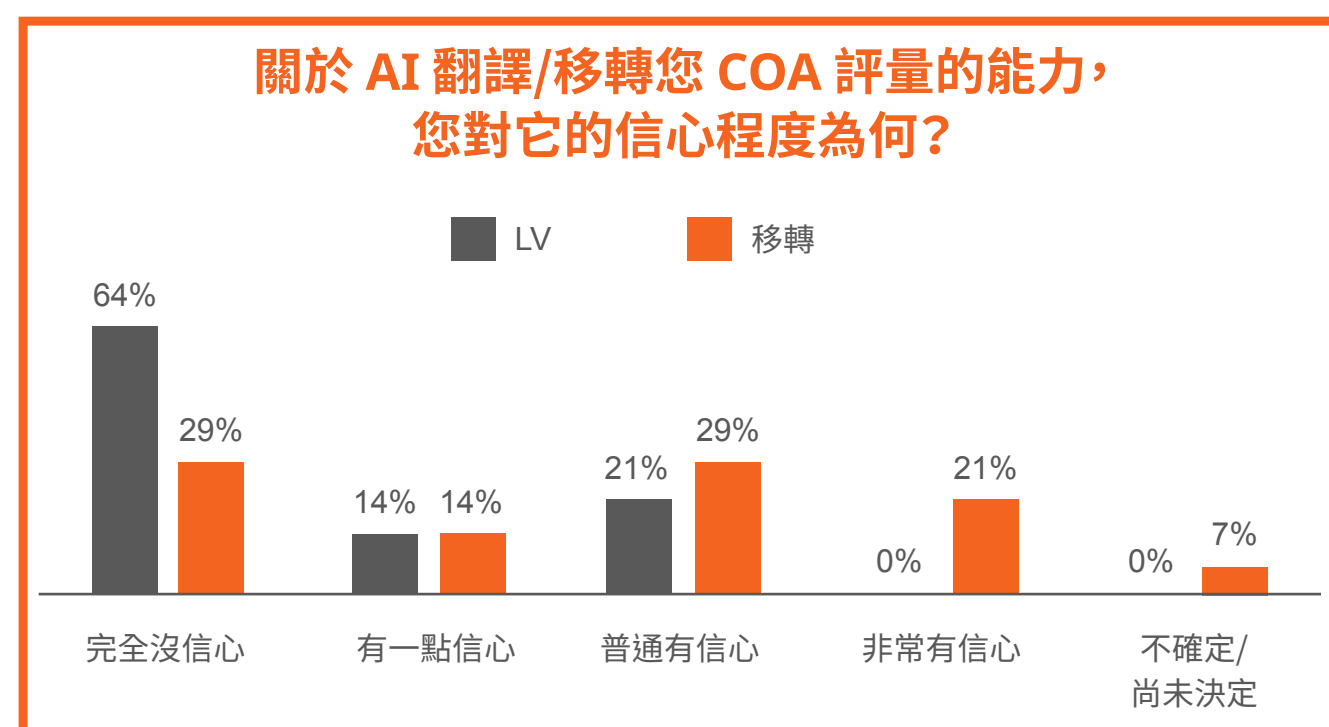
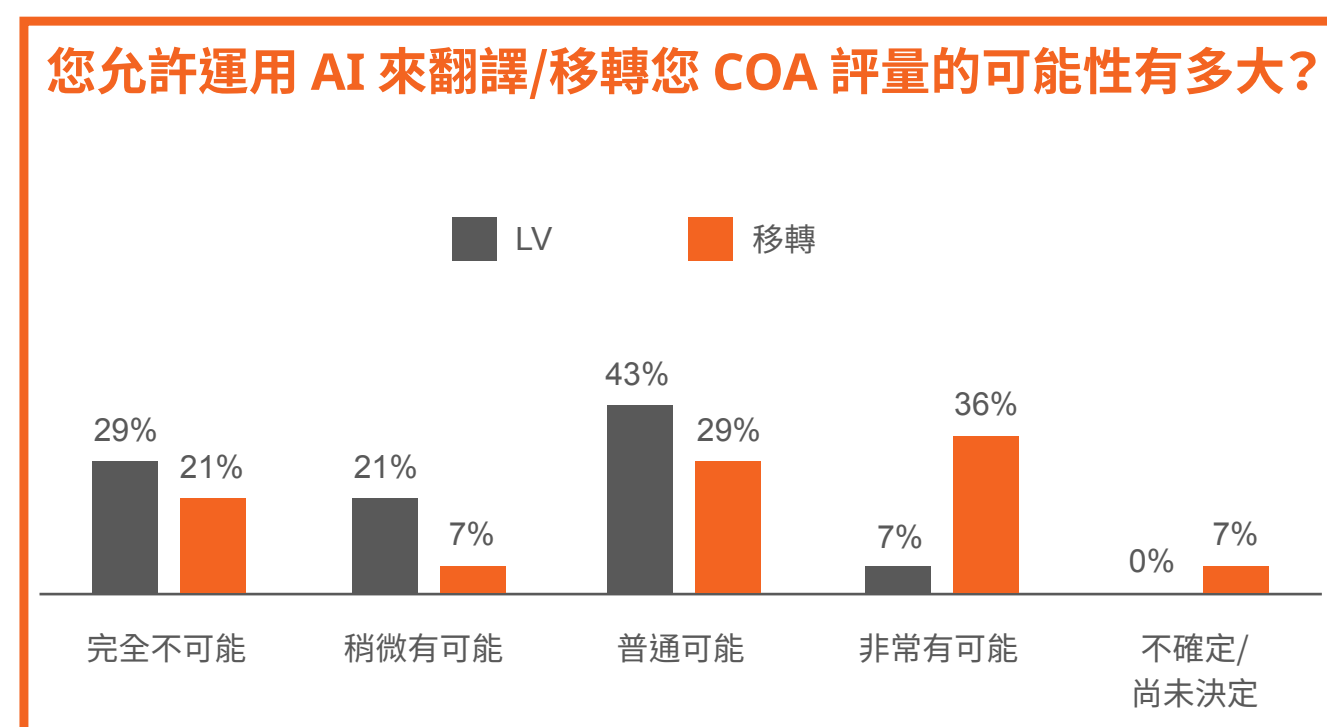
研究目標

關於將生成式人工智慧 (GenAI) 整合至臨床試驗作業中，尤其是在臨床結果評估 (COA) 這方面，往往會因為對 COA 智慧財產 (IP) 中應用 GenAI 缺乏清楚的指引而窒礙難行。在例如 COA 語言驗證和移轉等這些對納入病人心聲至關重要的領域採用 GenAI 的腳步，也因為這種不確定性而變得緩慢。本研究旨在收集與傳達 COA 著作權所有人對在這些流程中使用 GenAI 的看法，以供日後擬定部署策略時參考。

結果與解析

1) 對於在 LV 與移轉中使用 AI 的許可程度和信心程度

a. 結果：對於在語言驗證流程中使用 AI 翻譯他們的 COA 評量，回覆問卷的 COA 著作權所有人中有 64% (9/14) 回答「完全沒信心」。然而，只有 29% (4/14) 的受訪者「完全不可能」允許在語言驗證流程中使用 AI 翻譯他們的 COA 評量。對於在 eCOA 移轉流程中使用 AI，受訪者的信心程度跟許可程度就一致多了，「完全沒信心」跟「完全不可能」的比例分別是 29% (4/14) 和 21% (3/14)。

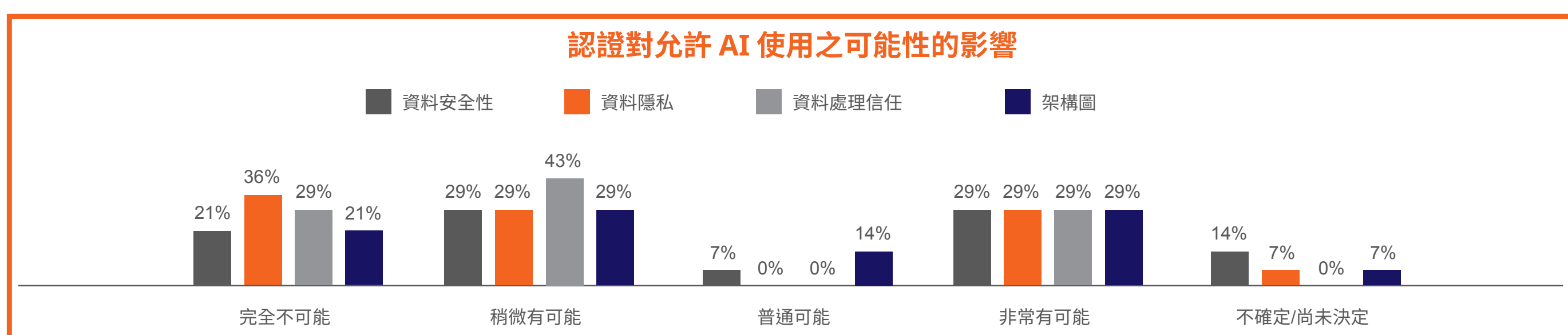


b. 解析：儘管著作權所有人對於在 LV 以及在移轉中使用 AI 的信心程度並不一樣，但對在這 2 種工作流程中允許使用 AI 的可能性則一致抱持正面態度，只不過程度不一樣（「普通」與「非常」），LV 為 50% (7/14) 而移轉為 65% (9/14)。根據「自由填寫文字回答」(Free Text Response) 來看，信心程度整體而言偏低。然而，經過細膩的對談後，回覆的受訪者大多數都願意允許使用 AI。

2) 認證與信任

a. 結果：我們詢問了著作權所有人以下幾種認證和/或文件對他們的影響 (例如允許使用 AI 的可能性)：1) 資料安全性認證 (例如 ISO27001:2022)、2) 資料隱私認證 (例如 ISO27001:2019)、3) 資料處理信任認證 (例如 SOC2 Type 2) 以及 4) AI 架構圖。不同認證/文件的結果不盡相同，其中最受質疑的是資料隱私及資料處理信任。然而在所有 4 種類型中，都有 29% (4/14) 的回覆者表示，若有這種文件存在的話他們「非常有可能」會允許使用 AI。

b. 解析：儘管回覆者對不同文件類型 (認證/圖表)、資料安全性、隱私以及 AI 引擎的功能運作方式等都多所懷疑，但至少 29% 的著作權所有人其接受程度確實會因這些文件而提高，變得更傾向同意在翻譯他們的 IP 時使用 AI。



結論

COA 評量的著作權所有人跟其他利害關係人一樣，也面臨業界在本地化其 IP 時缺乏 AI 使用相關指引的問題。然而，儘管這些利害關係人對於在本地化其 IP 時適當運用 AI 的信心仍有提升的空間，他們似乎也願意展開討論、建立政策、增補授權合約以及檢閱文件，來確保與 AI 引擎互動時他們 IP 的安全性。最後要提出一點，有鑑於本研究的回覆者數目並不多，因此還需要進一步分析來驗證這些結果。

研究方法

為了達成研究目的，我們將一份涵蓋下列三個主題、共 15 個問題的電子問卷調查表，寄送給 100 位 COA 著作權所有人。有十四位著作權所有人在為期三個禮拜的問卷調查時程內回覆了這份匿名填寫的問卷。

- 1) 對於在 COA 翻譯 (以下簡稱 LV) 與 eCOA 移轉 (以下簡稱移轉) 中使用 AI 的許可程度和信心程度
- 2) 認證與信任
- 3) 目前對 AI 的做法、政策與授權情況

3) 目前對 AI 的做法

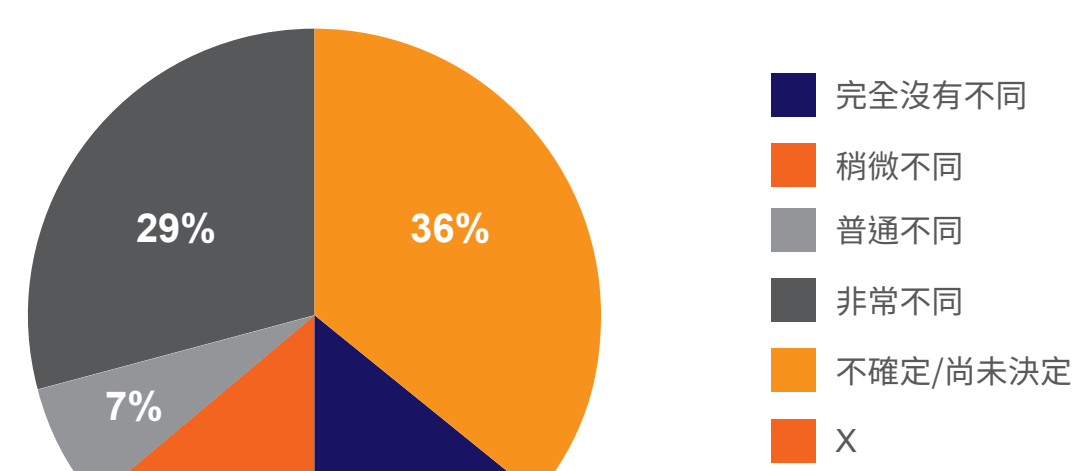
a. 結果：我們針對這個主題問了四 (4) 個問題。

1) 被問到相比於他們對 CAT 工具解決方案之使用的做法，與他們對在 LV 和移轉中使用 AI 的做法有何不同時，大部分的回覆者表示「不確定/尚未決定」(36% (5/14))。與此相較，有 28.5% (4/14) 的回覆者表示會「非常不同」。只有 28.5% (4/14) 的回覆者表示與 CAT 工具技術相比，他們對這個技術的做法「沒有不同」或只會「稍微不同」。

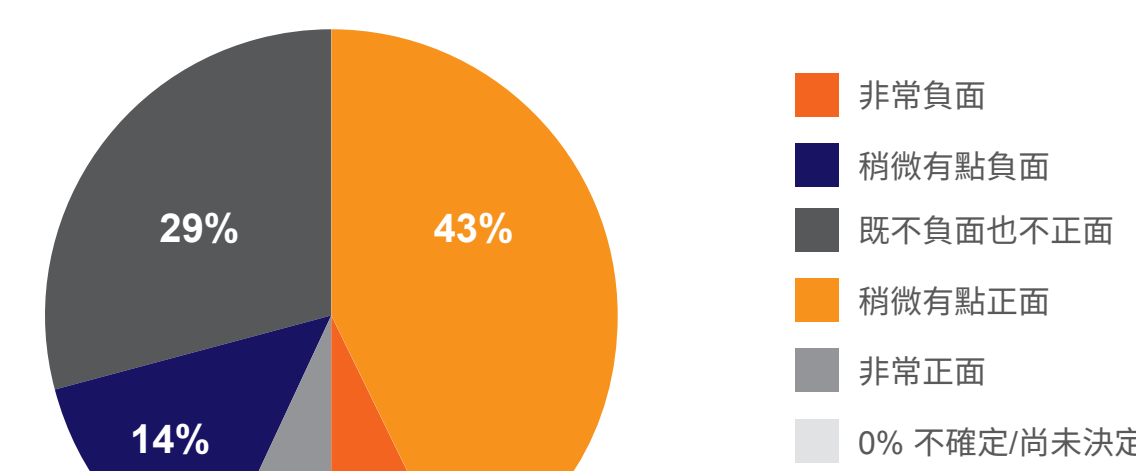
2) 43% (6/14) 的回覆者對於在臨床試驗與翻譯中使用 AI 覺得「稍微有點正面」，而有 29% (4/14) 覺得「既不負面也不正面」。「非常負面」和「非常正面」的結果都是 7% (1/14)，其餘 14% (2) 的回覆者則覺得「稍微有點負面」。

3、4) 被問到他們是否在公開論壇中主動提供 AI 使用政策，72% (10/14) 的回覆者表示「否」。此外，86% (12/14) 的回覆者並未調整其授權合約以清楚闡明他們對 AI 使用的政策。

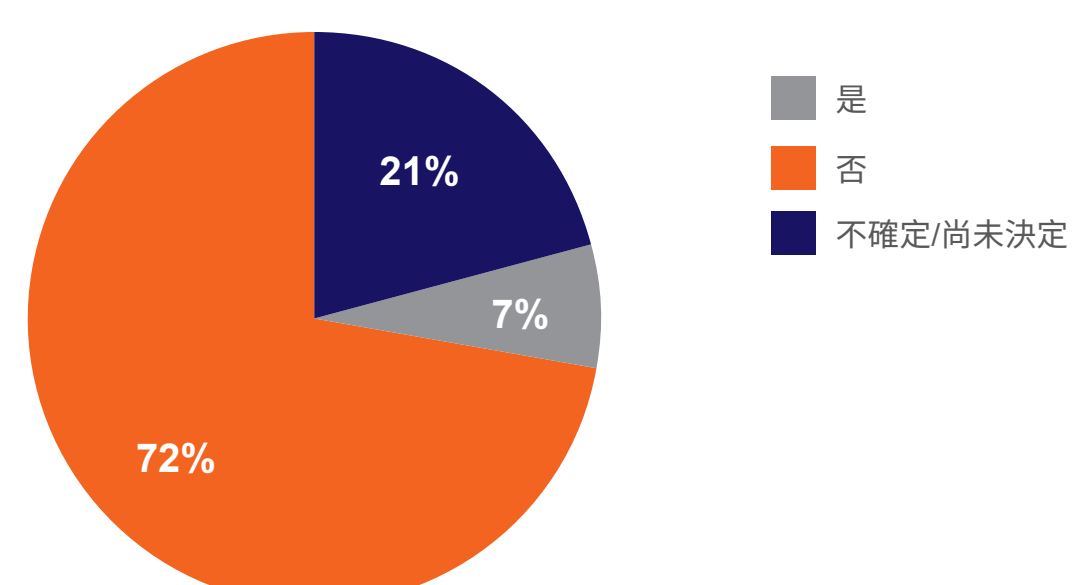
對於在 LV 或移轉中使用 AI 處理您的 COA 評量，您的做法會與/將會與之前對 CAT 工具解決方案之使用 (例如翻譯記憶庫/術語庫/機器翻譯) 的做法不同嗎？



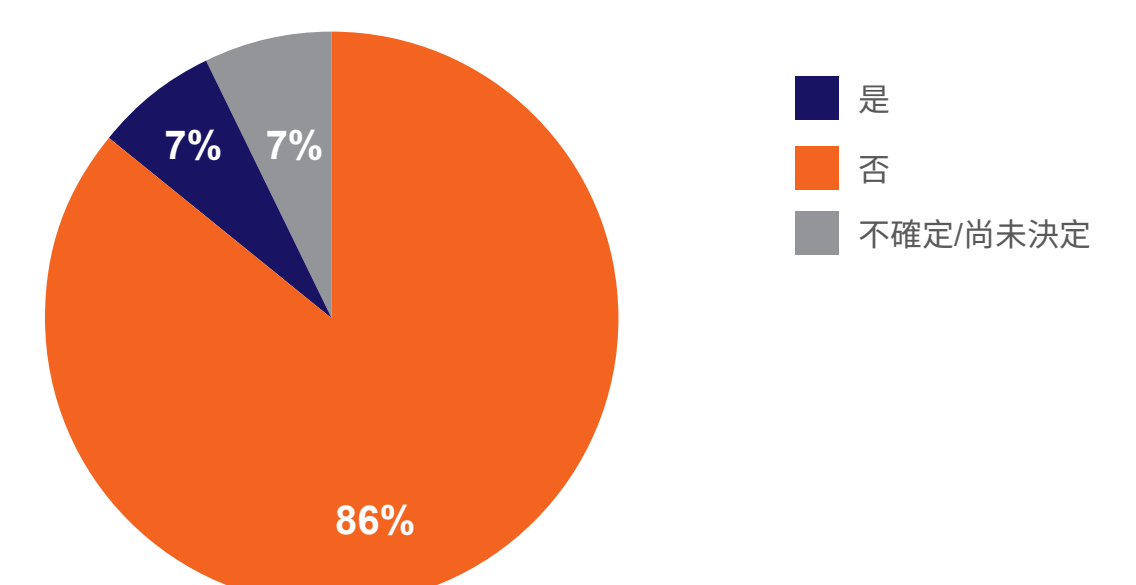
對於在臨床試驗與翻譯產業中日漸增加的 AI 使用，您覺得這是正面還是負面的趨勢？



您是否會主動地提供潛在被授權人您的 AI 使用政策資訊 (例如在可公開取得的政策中說明相關條款)？



您是否有調整您的授權合約，清楚闡述 AI 使用的相關政策？

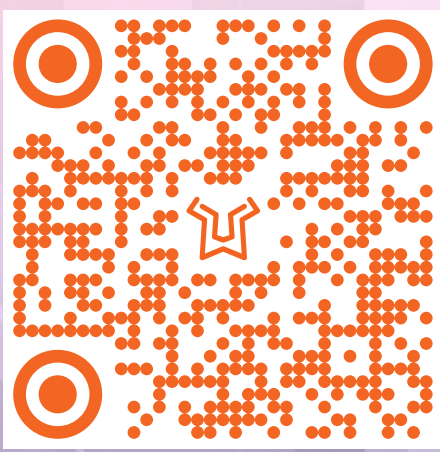


b. 解析：關於這個技術與翻譯中使用的舊有技術有何不同，以及如何根據這樣的差異調整相應的做法，似乎存在著一定程度的不確定性。業界對 AI 使用的意見也非常不同。但是，這個主題中 (甚至可能是整個問卷中) 最明顯的結果，就是著作權所有人鮮少積極主動地告知潛在被授權人使用其 IP 時的 AI 相關政策，使用在自己的授權合約文件中也缺乏相關條款。然而，根據「自由填寫文字回答」的結果，這個情況很快將有所改變，因此本研究小組也建議被授權人在與著作權所有人互動時，應將資訊的缺乏假定为是缺乏相關主張或政策，而非默示同意。



在 ECOA 本地化中以 GENAI 與人工進行螢幕截圖審閱之比較與評估

GenAI 會是改善審閱回饋的關鍵嗎？



作者：Kathryn Nolte、Karolina Elizondo Jimenez、Rupali Kadam、Melinda Johnson

簡介

電子臨床結果評估 (eCOA) 的翻譯、移轉和螢幕截圖審閱，傳統上都是高度依賴人工處理，這有部分是因為螢幕內容無法編輯的特性。在進行螢幕截圖審閱流程 (SSR) 時，會將目標語言的螢幕報告與來源主螢幕報告以及 (如適用的話) 舊有內容 (例如原本的紙本問卷) 相比較。

Lionbridge 會執行兩種螢幕截圖審閱：簡單 (Simple) 以及複雜 (Complex)。複雜 SSR 跟簡單 SSR 的不同之處，在於前者是要求與所有的目標舊有內容進行逐字逐句的詳細檢查。Lionbridge 體認且測試了生成式 AI (GenAI) 與光學字元辨識 (OCR) 技術的潛力，可以運用它們輔助傳統的螢幕截圖審閱方法、降低成本並縮短時程 (例如審閱次數)、改善品質，最終獲致更好的病患治療結果。在進行本研究時，我們同時使用了 Lionbridge 專有 Aurora AI Clinical Outcomes™ 工具中的簡單與複雜螢幕截圖審閱功能。

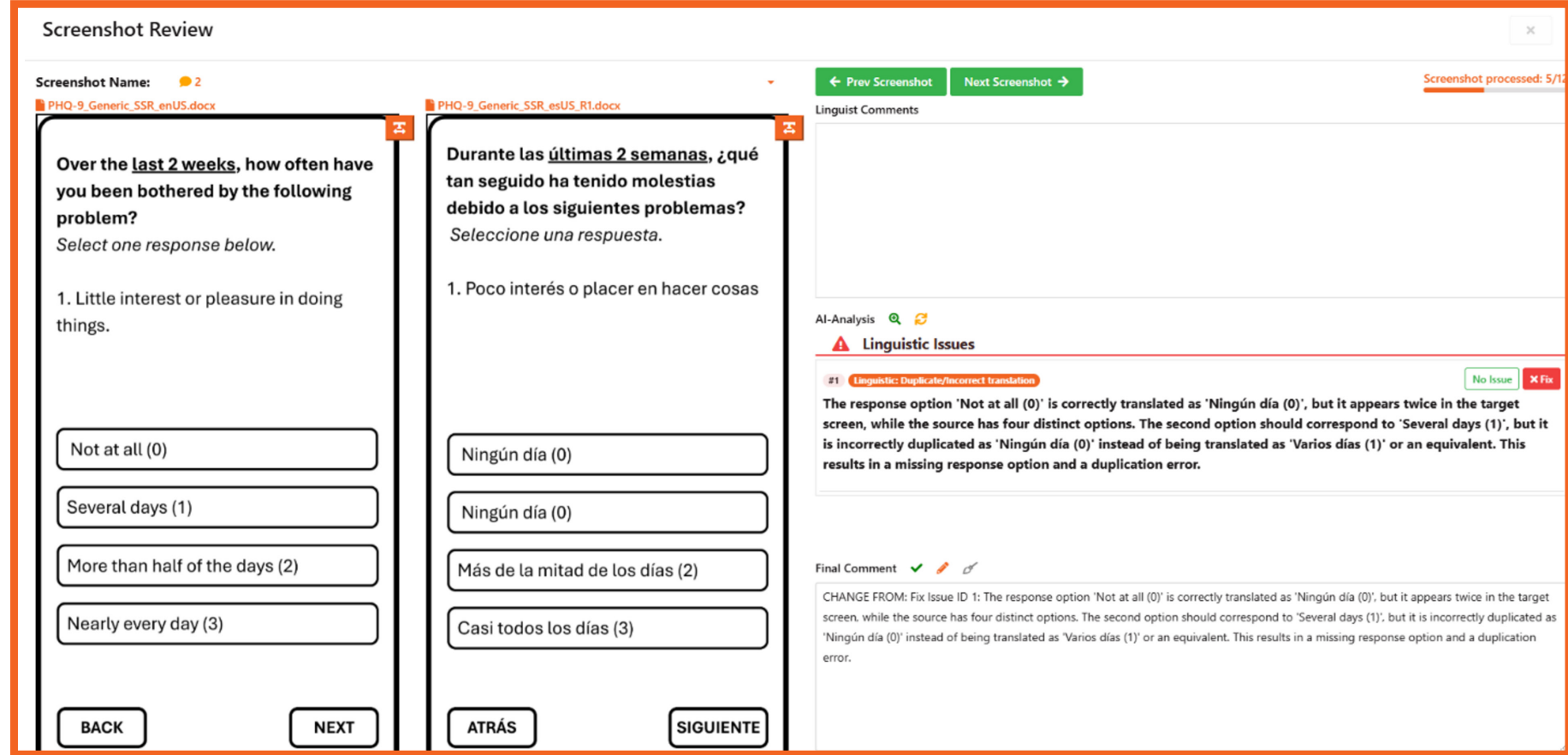
研究方法

Lionbridge 採用一個安全的 GenAI 引擎，針對 5 個病人用電子臨床結果評估 (eCOA) 的目標螢幕截圖報告生成了品質回饋。這些目標報告刻意加入了長度與複雜度各不相同的以下 11 種錯誤類型以供測試之用：遺漏的內容、未翻譯的內容、錯誤的翻譯、遺漏的 eCOA 校訂編輯、版本編號不相符、格式問題、構詞問題、標籤問題、換行問題、大小寫問題以及量表定位問題。測試的目標語言總共 16 種，涵蓋了多種字母和語系：保加利亞文 (保加利亞)、波蘭文 (波蘭)、羅馬尼亞文 (羅馬尼亞)、希臘文 (希臘)、西班牙文 (阿根廷、墨西哥、美國)、法文 (比利時、法國)、葡萄牙文 (巴西、葡萄牙)、匈牙利文 (匈牙利)、土耳其文 (土耳其)、韓文 (韓國)、繁體中文 (台灣) 以及泰文 (泰國)。提示經過持續自訂，直到能取得符合 eCOA 業界標準與既有實務做法的合適輸出為止。與此同時，我們也將同樣的 eCOA 螢幕報告同步傳送給通過 Lionbridge 審核的語言專家進行人工審閱。接著，我們再請公正的審閱人員驗證這兩類輸出 (AI 與人工)，並就正確性跟完整性評分。

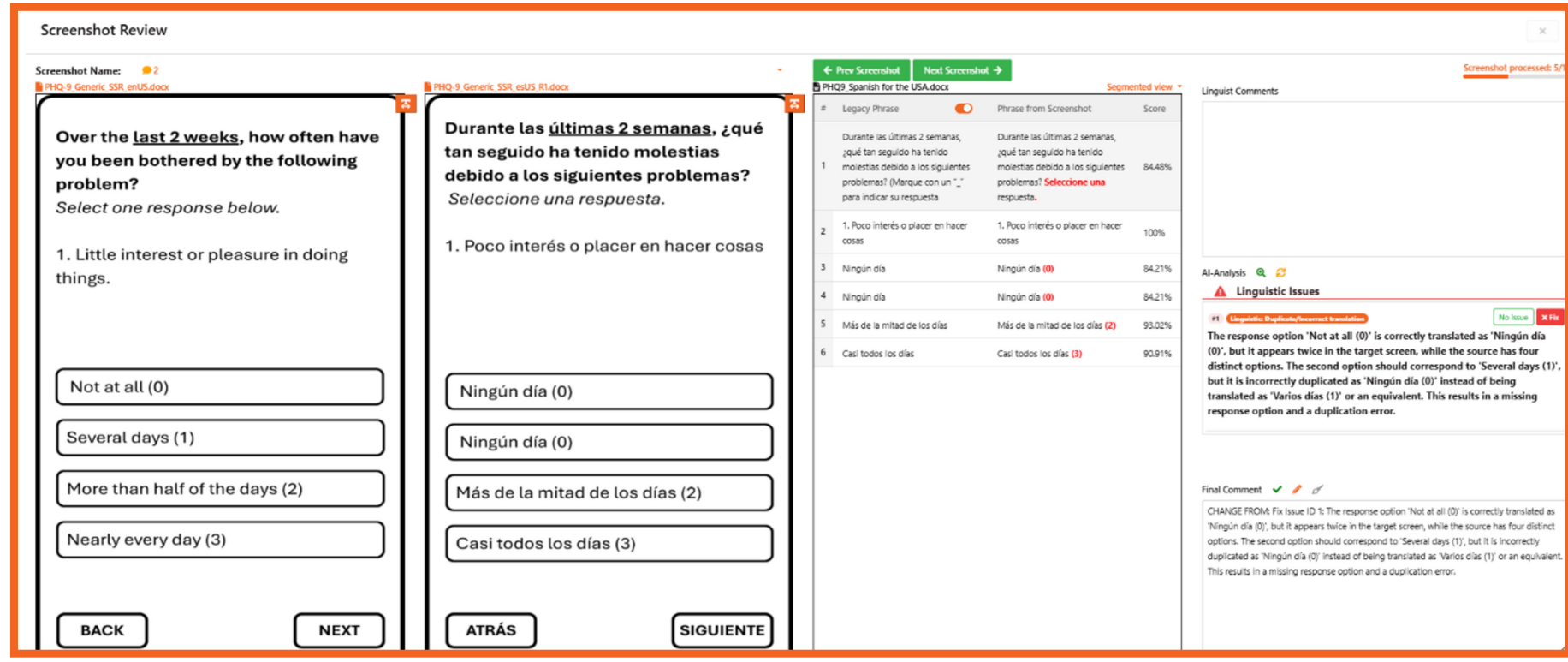
提示設計演進：

本研究的第一階段，使用了一個 GenAI 提示進行簡單螢幕截圖審閱，這個功能非常適合在完成翻譯流程 (翻譯/編修或是翻譯/回譯) 後，用來比對來源螢幕報告與目標螢幕報告。完成開發工具的這個功能後，我們便將注意力轉為開發與測試更細膩的複雜螢幕截圖審閱 GenAI 功能，這個功能比簡單版更上層樓，因此更適合用於已經完成移轉的內容。除了比對目標螢幕內容與來源主螢幕報告，複雜版功能也可以將目標螢幕內容與目標舊有內容相比較。

簡單螢幕截圖審閱介面：



複雜螢幕截圖審閱介面：



結論

我們的研究證實將 GenAI 與人工審閱相整合，是 eCOA 翻譯與移轉流程中進行螢幕截圖審閱最有效的一種策略。所使用的提示會持續一直精進，隨時間過去便可偵測出更多問題。不過目前還是建議在螢幕截圖審閱流程中維持有人類參與，確保這些高度敏感的內容能有最佳的品質。運用 GenAI 找出初步的問題，可以大幅提升速度、效率和正確性。人類的專業知識仍舊非常重要，能擬定細膩的決策並找出 AI 可能遺漏的問題。這種「人機迴圈」的做法有許多好處：減少所需的審閱次數進而縮短時程；減少語言專家花在品質檢查的時間進而降低成本；以及最重要的，強化我們的能力，能更精準地取得更高品質的病人心聲，進而確保提供更可靠、以病人為本的體驗。

結果與解析

分析簡單 SSR GenAI 功能的品質意見回饋後，我們發現它是個非常實用的 QA 檢查功能，可以搭配人工意見回饋一起應用於翻譯以及移轉專案上。「簡單」功能可以有效支援語言專家，尤其適合應用在不需要與舊有內容交叉比對的翻譯/螢幕截圖審閱專案。運用這個功能，Aurora AI Clinical Outcomes 工具成功地從目標螢幕截圖中，就所有 11 種錯誤類型分別找到多個實例。然而，由於它無法根據舊有內容進行比對檢查，因此在某些情況下會稍嫌不足。

複雜 SSR GenAI 這個功能雖然仍在開發中 (預計會於 2026 年 Q1 部署)，期中結果顯示它不但能揪出簡單 SSR 工具所能辨識出的任何問題，而且還能找出更多。另外，與簡單 SSR 工具相比，它也能交出更細膩的審閱回饋。能與舊有內容交互比對與檢查的能力，將可使複雜 SSR 功能得以排除目前簡單 SSR 工具會產生的一些誤報 (如下方表格所述)。

錯誤類型	簡單 SSR GENAI 功能：注意事項與侷限	複雜 SSR GENAI 功能：注意事項與侷限	範例
遺漏的內容	能標記出有遺漏的目標內容	可確認目標文字是否是因為不存在於舊有內容而刻意遺漏	<div>Portuguese (Portugal)</div> <div>#4 [Redundancy/Missing content] Target Paragraph: 1</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>
未翻譯的內容	可標記保留為英文的文字，並建議這可能是刻意保留的 (例如著作權文字)	可確認文字是否刻意保留為英文	<div>Bulgarian (Bulgaria)</div> <div>Linguistic Issues</div> <div>#1 [Linguistic/Untranslated content]</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>
錯誤的翻譯	能找出與來源文字有偏差的地方	可以找出因為與舊有內容有偏差的翻譯錯誤 (即使與來源相符)	<div>Portuguese (Portugal)</div> <div>#9 [Linguistic/Inconsistent scale anchors]</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>
遺漏的 eCOA 校訂編輯	簡單與複雜功能應會具備一樣的功能		<div>Portuguese (Portugal)</div> <div>#1 [Linguistic/Incorrect translation] Source Paragraph: 6</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>
版本編號不相符	能標記來源與目標版本編號不相符的問題，並會顯示警告指出這些不同可能是刻意為之	可藉由檢查舊有內容，確認版本編號的不同是否是刻意為之	<div>Korean (Republic of Korea)</div> <div>#1 [Redundancy/Version number mismatch]</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>
格式問題	可以標記不同於來源主螢幕報告的格式	可根據舊有內容，確認格式是否是刻意與來源不同	<div>Portuguese (Portugal)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout/Underlining/HTML, Tag Display] Source Paragraph: 7</div> <div>Target Paragraph: 5</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>
構詞問題	可偵測出乎意料的形容詞構詞，並會顯示警告指出這可能是刻意為之的 (例如因為病人族群 100% 全為女性，所以形容詞全都會是陰性)	可跟舊有內容比對，就預期的性別構詞提供額外的脈絡	<div>Spanish (United States)</div> <div>Linguistic Issues</div> <div>#1 [Linguistic/Inconsistent translation] Source Paragraph: 7 Target Paragraph: 5</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>
標籤問題	簡單與複雜功能應會具備一樣的功能		<div>Thai (Thailand)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout/Underlining/HTML, Tag Display] Source Paragraph: 7</div> <div>Target Paragraph: 5</div> <div>The target screen displays HTML tags (<u>) instead of rendering underlined text for ‘การเหนื่อยล้า’, ‘เหนื่อย’, and ‘การปวด’. 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>
換行問題	簡單與複雜功能應會具備一樣的功能		<div>French (France)</div> <div>Formatting/Layout Issues</div> <div>#1 [Formatting/Layout/Line break] Source Paragraph: 6 Target Paragraph: 4</div> <div>The target screen introduces a line break in ‘habituelles’ 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 ‘habituelles’ appears on one line without a hyphen.</div>
大小寫問題	能標記與來源不相符的大小寫模式，並會顯示警告指出這可能是刻意為之的	可確認目標螢幕中非預期的大小寫格式是否與舊有內容相符	<div>Hungarian (Hungary)</div> <div>#2 [Linguistic/Inconsistent capitalization] Source Paragraph: 6</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>
量表定位問題	簡單與複雜功能應會具備一樣的功能		<div>Portuguese (Portugal)</div> <div>#2 [Formatting/Layout/Scale anchor placement] Source Paragraph: 8</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>

人類與機器：

- GenAI 在評估目標螢幕內容上的速度遠超過人類，它能在數秒內，針對指定的量表為所有的目標螢幕生成審閱回饋。
- Lionbridge 的品質專家也發現採用 GenAI 的流程在時間與效率上多有助益，因為他們可以運用 Aurora Clinical Outcomes Tool SSR 介面，輕易地同時存取舊有檔案、主螢幕報告以及目標語言螢幕報告。
- GenAI 找出了多個審閱人員遺漏掉的問題，特別是跟換行、大小寫、翻譯錯誤 (包括重複的文字) 及量表定位的問題。這是因為對於無法編輯的影像內容，人眼比對很容易會漏失文字不一致處，但這個工具是使用光學字元辨識 (OCR) 來解析影像，因此 GenAI 會更容易找出這類問題。
- GenAI SSR 有時會無法一致地跨不同語言找到錯誤。舉例來說，儘管 GenAI 在大部分語言中都能找出遺漏翻譯「items」的問題，但卻沒能在西班牙文中揪出這個錯誤。這很可能是因為唯一的差別是重音符號 (「ítems」)。
- 在語境脈絡相關問題的標記上，審閱人員比較不會犯錯 (例如女性病人族群、刻意保留的版本編號不一致、刻意保留為英文的著作權文字等)。