



LIONBRIDGE

AI 及語言策略 在生命科學中的應用

不可不知的資訊



目錄

簡介 | 3

認識 AI 及大型語言模型 | 4

人機迴圈 AI 語言服務: 建立與培養信任 | 6

採用風險基礎方法取得可靠的結果 | 9

新臨床試驗設計: LLM 案例 | 11

MDR 與 IVDR: LLM 案例 | 13

語言 AI 應用探索與解決方案開發 | 15



是您聽命於 AI, 還是 AI 聽命於您?

人工智慧 (AI) 部署的成功與否, 取決於兩個因素: 控管和信任。

對受嚴格規範的生命科學領域來說, 控管已是他們司空見慣的工作; 但另一方面, 他們必須透過資料驗證來培養對技術的信任, 這對 AI 和大型語言模型 (LLM) 而言尤其重要。

生成式 AI 和 LLM 都是新興技術, 因此許多公司對它們是抱持著靜觀其變的態度。對這個受嚴格規範的產業而言, 它們的業務往往攸關性命, 因此確實需要謹慎小心。然而, 高速發展的 AI 在不少應用上展現了極大的潛力, 如果只是作壁上觀的話, 很可能會失去競爭優勢。有鑑於此, Lionbridge 通常會建議生命科學客戶在審慎小心之餘, 也要對創新抱持開放心態, 盡力在兩者之間取得最佳平衡。此外, 我們也認為, 對 LLM 產出之語言結果的信任是可以培養的。如果生命科學產業開始考慮在其語言策略應用這項技術, 這點就更顯重要。

在這份電子書中, 我們將說明有助解決信任鴻溝問題的方法, 以及大型語言模型在生命科學上的一些可能應用, 協助您善用它產出可靠且適宜的語言結果。

LLM 的進展正在重塑語言服務產業的面貌並帶來諸多挑戰, 身為語言服務供應商 (LSP) 的 Lionbridge 自然也深受其所影響。

許多客戶前來向我們求助, 希望我們能引導他們在對的時機使用對的方法來應用 LLM 和機器翻譯 (MT), 好在享有高成本效益卻又無需犧牲語言品質和法規遵循情況下達成目標。由於語言本質上是個視情境而定的複雜活動, 我們這份電子書的建議, 就是先專注在所欲取得的語言結果及使用情境, 接著再根據產品線或公司業務目標設計您的 AI 語言策略。



Camilla de Villiers

總經理
生命科學團隊
Lionbridge



Pia Windelov

副總裁
生命科學策略與
產品行銷
Lionbridge

認識 AI 及大型語言模型

人工智慧 (AI) 是指模仿人類智慧運作的方式，涵蓋了非常廣泛的可能應用以及多不勝數的先進工具。

AI 在近年造成轟動，經常被人以「極具突破性」來形容其發展，但是要透過五花八門的相關內容認識這個技術並了解其實際價值，往往不是件容易的事。以下這四個定義，有助您一揭這個主題的神秘面紗。

1. 人工智慧 (AI) 是電腦科學的一個分支，旨在建立系統來執行通常需要具人類智慧才能執行的工作。AI 系統可以根據規則運作，也可以使用龐大的資料集加以訓練。它們可執行的工作，包括理解下列事項：

- » 自然語言
- » 問題處理
- » 模式
- » 決策

2. 生成式 AI (GenAI) 是一種能根據提示和廣泛的多模態訓練來生成新內容 (通常是文字和/或圖片) 的 AI 系統或技術。它會判斷最可能的輸出，並以看似是由真人撰寫的方式呈現結果。

3. 大型語言模型 (LLM) 是一種以語言為主的生成式 AI。LLM 是使用由文本跟軟體程式碼組成的龐大資料集加以訓練，具備多種能力，包括：

- » 摘要
- » 翻譯
- » 預測
- » 根據從龐大資料庫學習得來的知識生成文本

4. 提示 (Prompt) 是個出發點，透過提供文字輸入給 LLM 或 AI 系統，以生成特定的回覆或內容。提示工程 (Prompt Engineering) 是指開發有效的一系列提示。作為 LSP，Lionbridge 業已欣然採用生成式 AI 技術來協助客戶拓展他們的商務內容。LLM 已經受到大多數產業的矚目，即使是對新興技術一向抱持觀望態度、往往很晚才應用於受嚴格規範內容的生命科學產業，也已注意到這個新技術的潛力。

深具潛力的 LLM 可望徹底改變產業界，促使人們以嶄新的方式向世界推出治療性介入措施和醫療保健解決方案。

LLM 的好處

目前的神經機器翻譯 (NMT) 典範將會走下舞台
而且很可能會被奠基於 LLM 的新典範所取代

隨著這個發展逐漸具體成形，您應該會看到以下結果：



翻譯成本降低



生產力提高，能生成更大規模的內容



翻譯品質提升 — 文字宛如是由真人撰寫



客戶體驗提升



進軍更多市場，享有新的商機



人機迴圈 AI 語言服務:建立與培養信任

一致性是成功研發治療性介入措施,並取得可預測之醫療結果的關鍵要素。但同時,它也是大規模應用 LLM 的主要挑戰,因為 LLM 的訓練往往是採用跨多個領域的龐大語料庫。因此,將 AI 使用案例應用到受規範內容的翻譯或語言修訂時,都必須要確保其語言結果的一致性與可靠性。要達到這個目標,必須具備以下條件:

- » 人工介入
- » LLM 的細微調整
- » 提示工程技巧
- » 能夠審閱針對特定領域自動產出的結果

此外,實行時也必須符合企業層級的法規遵循架構,才能充分發揮 LLM 在規模、效率和其他方面的好處。

產業界對 LLM 目前仍抱持著審慎的態度,而各監管機關尚未就 AI 提供一致的產業指引,可能是造成這種對 LLM 消極態度的原因。其中又以生物製藥及醫藥技術產業特別明顯,畢竟這攸關病人安全、資料完整性及生命倫理準則。

對生命科學領域來說,接下來的要務,是要在自動化及人工介入之間找到適切的平衡,進而對 LLM 衍生出的語言結果建立信任與信心。這不是件容易的事,因為這是個受法律高度規範的產業,面對越來越緊湊又難以預測的法規時程規定,相關人士還是得設法交出高品質的結果。除此之外,LLM 也還沒能產出符合 ISO 標準的翻譯。

Lionbridge 對這樣的挑戰並不陌生,我們為生命科學客戶大規模地提供神經機器翻譯 (NMT) 這個被視為是生成式 AI 前驅技術的服務,已有數年之久。

由於我們對 NMT 服務秉持同樣審慎的原則,堅持對所有受規範內容進行人工譯後編修,因此客戶無需犧牲品質,也能享有節省成本的好處。面對新一代的各式工具,我們也已做好踏上同樣成功旅程的準備。

生命科學語言服務中的 LLM 使用案例

簡言之,LLM 最明顯的使用案例就是語言翻譯。這個領域有許多應用這項技術的機會,例如大量的內容、眾多語言種類,以及極具挑戰性的法規時程等,因此非常適合大多數具有以下需求的新藥物及醫療器材:

- 需要取得重要資料,證明產品的有效性和安全性
- 需要以大量文件證明其法規遵循
- 需要提供強制性的定期報告,確保符合公開透明和安全性監督的要求
- 不同地區緊湊的審查時程及送件期限

隨著 LLM 持續進步,一些需要生成多語言內容、將其加以摘要或是針對特定用途或使用情境而修改的其他使用案例,也都可能會納入為標準服務項目。

以下是這類使用案例的一些例子:

- 將使用科學語言的內容轉換為淺白用語內容,以使用於公開揭露、投資者關係、說明或是供非技術性使用者、病人和其他目標對象使用的資訊。
- 將研究結果或科學計畫書以醫學用語或淺白用語進行摘要,供監管機關、試驗參加者、病人、使用者、倫理研究委員會以及其他目標對象使用。
- 直接從原文創作新的多語言內容,加快內容創作和翻譯的速度。這種應用會有助遵守嚴格的法規或其他商務關鍵時程。

- 評估機器翻譯 (MT) 產出內容的品質,找出品質不佳的段落。
- 針對找出的品質不佳 MT 段落,自動進行譯後編修/重新翻譯,節省許多時間。
- 針對從紙本資料擷取形式移轉至電子資料擷取形式的內容進行調適,例如病患回報結果 (PRO 到 ePRO 或 COA 到 eCOA)。

研究結果的總和有其風險存在,像是選取時存在偏見、對所欲呈現之結果的選擇偏見,以及讓讀者對跟藥物相關之成效或安全性主張有所誤會等。此外,由於 LLM 還在採用的初期階段,不同的 LLM 輸出驗證測試,其結果也可能大不相同。基於上述這些原因,監管機關可能永遠不會接受完全自動產出的結果摘要。因此,LLM 在這個使用案例中的附加價值,主要是在最佳化總和流程,而不是要取代人工作業。

人類如何有助彌補機器的不足之處

儘管生成式 AI 與 LLM 技術有著長足的進步,但在 AI 翻譯及其他語言工作流程中納入有人類參與其中的人機迴圈機制,仍舊非常重要。雖然 LLM 能提高翻譯效率及成本效益,但是在沒有監督的情況下,不能將最具風險或受嚴格規範的內容類型完全交由它們處理。再加上這個技術並無法獨立執行,因此人工介入也就不可或缺。LLM 的其中一個嚴重問題,是它們容易產生「幻覺」,這深切影響到它們的可靠性。所謂幻覺,是指生成的內容與輸入資料無關、是捏造的,或是與輸入資料不一致。人類的參與會有助降低 AI 信任鴻溝的問題,而機器則可以讓人類更有效率地工作並發揮遠超以往的聰明才智。所以,將 AI 與人類的長處結合在一起,將可產出符合您需求的語言成果。

LLM 的 3 大挑戰,以及人機迴圈模式如何有助解決這些問題

1. 展現一致性

LLM 為何表現不佳

LLM 技術最適合使用幾百字以內的提示,因此使用者必須將大且複雜的提示拆解成好幾個提示,這個長度限制往往導致分批輸出的語言結果前後不一致。在例如翻譯送審文件,或將數百頁的臨床試驗報告總結為 10 頁的結果摘要時,這個長度限制便會是個很大的問題。

人類能如何協助

可以由審閱人員審閱整個自動化輸出,確保語言一致性、格式符合規定、術語使用符合要求,並含有任何依法規必須納入的資訊。

此外,也可以由專家主導進行提示工程,根據動態的成效及品質參數持續微調輸入與輸出,減輕內文脈絡長度限制帶來的問題。

2. 運用語言資產

LLM 為何表現不佳

LLM 本身並沒有翻譯記憶庫 (TM)、詞彙表或術語清單,必須透過一連串提示分別加入這些內容,才能確保它使用您所指定的詞彙表、術語和語氣。

人類能如何協助

受過訓練且熟知如何執行這種工作的人類,可以根據專案類型在一系列的提示中加入多個詞彙表和指示,進而產生一致的語氣和術語。視目標使用者的健康識讀程度或技術知識而定,生命科學中的術語可能包含以下種類:

- 法規術語
- 特定疾病的術語
- 產品專屬詞彙表
- 不同的語氣或風格

基於這個原因,必須要有具生命科學相關經驗的語言專家,協助建立合適的術語清單。

3. 建立提示

LLM 為何表現不佳

能建立提示的是人類,而非 LLM。想要取得卓越的成效和語言結果,提示是第一個步驟也是必要條件,而 LLM 本身無法建立有效的提示。

人類能如何協助

技巧精湛的提示工程師能夠開發提示範本、將提示的再利用自動化,或是執行提示後處理作業,進而打造出最合適的工作流程。生命科學領域大多數的文件,往往會受例如調和 GxP 標準等規範所管制,而有預先定義好的特定結構、用途和內容要求。因此,可以為產品研發進程和產品生命週期內重複的內容類型,事先建立好提示範本。

採用風險基礎方法取得可靠的結果

AI 工具正大幅革新藥物研發的面貌，使人們能夠運用這些工具以創新的方法簡化與增進研究流程。

這些工具可以透過以下方式，改善臨床試驗的執行：

- 最佳化試驗參加者的遴選
- 加強試驗的監控
- 改善資料的收集
- 資料的管理與分析

AI 也有助於試驗設計，尤其是例如分散式臨床試驗 (DCT) 這類非傳統試驗，以及納入擷取自電子健康紀錄 (EHR)、醫療給付或其他來源之真實世界數據 (RWD) 的試驗。這些 AI 技術不但能提高臨床試驗的效率，也讓我們有機會打造更貼近病人需求的體驗，推動臨床試驗產業朝精準醫療的時代大步前進。

要使用生成式 AI 技術來取代或增進藥物或醫療技術之研發和行銷中的人工作業，就必須要能夠控管相關風險。監管機關會期望生命科學產業能實行風險基礎方法來開發、部署和監控 AI 技術，最終的目標是要能針對 AI 的使用情境與影響，主動實行適當的控管作為。

運用利益風險評估方法，提高實行的成功機率

對監管機關來說，藥物研發與評估的核心重點，是它們的利益風險剖析 (Benefit-risk Profile) 以及能否確實保護與推進公眾健康。正如已經使用於藥物及器材研發的其他技術 (例如電子資料擷取)，AI 應用同樣也得先滿足這些優先要務。

對於運用 LLM 的使用案例，我們建議採用類似利益風險評估的原則，並建立跨部門的語言策略，來判斷哪些內容類型、產品類型、預期用途和語言適合使用 LLM。

語言對全球醫療結果的影響往往遭到嚴重低估，只有在語言上遭遇困難時，人們才會驚覺它的重要性。這種作法使得產業界在運作上採用見招拆招的模式，在採購語言服務時也各自為政，缺乏對 LLM 之利益與風險的通盤了解。

然而，深具潛力的大型語言模型能為我們帶來很多機會，因此產業界應該敦促自己，策略性地去思考語言服務與成果間的關係。

以利益風險基礎方法，思考語言服務及 LLM

控管

語言策略 | 品質/工作流程控管
人機迴圈控管 | 風險管理控管

潛在利益

翻譯的數量越大、生命週期越長，節省的成本便越多

速度加快，能因應緊湊的時間壓力並加快監管程序的進行

提升敏捷度，因應複雜的新試驗設計在試驗的期中變更

能靈活彈性地支援規模更大、在多個國家/地區執行的多國臨床試驗

將具高度重複性的工作或工作流程自動化，減輕人員的負擔

語言一致性，有助正確地傳達產品使用方式、訊息、主張以及結果溝通

潛在風險

效用不佳的提示、「幻覺」或未採用適當的人機迴圈模式，可能會造成語言結果品質不良，進而傷害/危害到使用者和病人

對臨床決策或診斷意料之外/有害的影響

翻譯或總結錯誤/品質不佳，會對資料完整性/可靠性造成負面影響

沒有審慎考慮或控管的 LLM 應用，可能會造成在法規審閱或核准程序時遭到退件

忽略使用情境可能會造成意料之外或失敗的語言結果

揭露受保護的資訊 (個人/專利)

可靠的語言結果



新臨床試驗設計：LLM 案例

過去這十年來，生醫科技大幅躍進的革新，為人們帶來了不少創新且具突破性的治療方法。與此同時，藥物研發開發週期的縮短及日漸龐大的成本，也引發越來越多的擔憂。為了解決成本和時間方面的疑慮並因應其他現代藥物研發的挑戰，新的試驗設計也應運而生，像是由多個子試驗組成不同類型的總計畫書，以及能在計畫書執行期間進行預訂調整的調適性試驗設計等。

雖然傳統設計仍舊是標準作法，但新興的試驗設計提供了提高彈性與效率、加快病人招募速度以及降低研究成本等好處。不過，它們也同時在規劃、組織及統計分析上帶來了新的挑戰。

傳統試驗設計是以隨機、雙盲的方式比較同時進行的多個治療組，一直以來都是產出可靠且健全之臨床資料的黃金標準，但它們固有的限制也越來越受到人們的注意與檢視。這些限制包括：長時間的執行期、昂貴的成本，以及需要龐大的樣本數。

總計畫書或主計畫書 (Master or Main protocols) 是指一份總體計畫書中包含更多子試驗，一般可分為籃型 (Basket)、傘狀 (Umbrella) 和平台 (Platform) 試驗。對委託者、監管機關和病人來說，這類計畫書可說是種典範

轉移。這種類型的計畫書可以讓相關人士以同一個臨床基礎架構，對多個療法和/或疾病進行平行試驗。一般來說，跨國試驗計畫書的實行需要耗費大量的時間與人力。也因此，如果子試驗在例如以下事項上都一樣的話，那麼採用總計畫書的方式往往能大幅增進試驗執行的效率：

- » 試驗單位遴選
- » 病患篩選
- » 資料管理
- » 倫理或監督委員會

另一種現代試驗設計是調適性試驗，是指在試驗執行期間，試驗執行者可以根據所累積的試驗參加者數據來修改試驗。但這些變更必須要事先加以定義，而且必須在試驗期間進行期中分析，才得以進行例如樣本數調整或停止特定劑量等試驗期中調整。採用調適性試驗設計能為臨床藥物開發計畫注入高靈活彈性，同時亦可縮短時程、降低成本和減少受試病人數。

在新臨床試驗設計中應用 LLM 的可能性

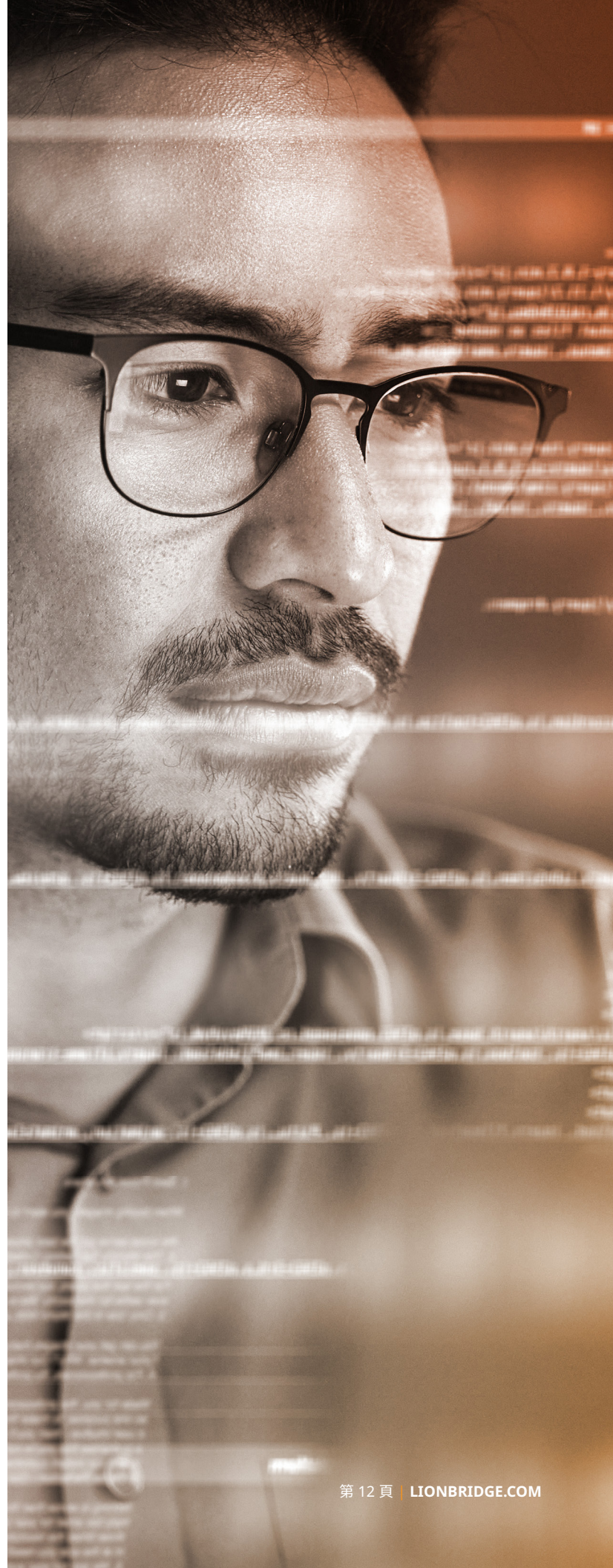
LLM 的設計目的,就是要提高語言結果的效率、速度和一致性,這對新試驗設計來說自然是非常好的應用機會。

如果在試驗計畫的開發階段便加以規劃,實行時也納入適當的控管及人工介入措施,LLM 其實深具潛力,能在新試驗設計的籌備與高效率執行上有所助益。

採用總計畫書或主計畫書這種設計時,由於要根據同一個計畫書提交多個子試驗,因此在初始臨床試驗申請 (CTA) 以及 CTA 修正期間,可能會需要提交大量的送審內容。因此,為了能依據同一個計畫書同步進行提交與審查,準備所有多語言文件時的翻譯效率和速度便很重要。如果子試驗間有一些重複的內容類型,LLM 甚至可以跳過這類文件的來源語言擷取步驟,直接為所有計畫書的語言建立多語言文件。

若是採用調適性試驗設計,由於可能會在試驗執行期間進行多次變更,因此會需要在非常緊湊的時間內完成新的或重複的翻譯。LLM 具有快速及高效率等好處,不但能增進這類調適性試驗設計所欲達成的敏捷執行力,也能滿足大規模跨國計畫在語言方面的需求。

如果希望 LLM 能帶來附加價值,重點就是要了解相關試驗設計如何管理語言事務,而且最好能在計畫書開發時便事先加以規劃。



MDR 與 IVDR:LLM 案例

AI 應用在醫藥技術上有很大的潛力，可以推動決策並增進診斷程序。或者，AI 也可以用來找出資料中較不明顯的趨勢走向，以嘉惠病人。一般談到 AI 的醫療技術應用時，經常會強調它在臨床決策支援軟體及影像增強分析上的潛力。但和藥物開發一樣，AI 也可以用來減輕人工作業流程、臨床試驗執行、送審程序和監管機構報告等方面的負擔。這個技術能協助醫藥技術公司更快、更聰明地處理研發 (R&D)、運作、監管單位送審、行銷、銷售到客戶支援等整個產品生命週期中的各項工作。

值得注意的是，在例如醫療器材法規 (MDR) 和體外診斷醫療器材法規 (IVDR) 等歐盟近日與目前法規改革的推動下，也為醫藥技術產業帶來不少運用 LLM 的機會。

這些新規範不但對效能、安全性和公開透明有著更嚴格的要求，法規審閱的時程也更加緊湊，對醫療器材製造商而言不啻為嚴峻的挑戰。他們必須要製作、翻譯、總結摘要以及更新內容，有時還得同時製作供專業及非專業人士使用的版本。其中有部分的內容，必須要經由代施查核機構驗證和/或上傳至 Eudamed (集中化的歐盟資料庫，用來收集有關醫療器材及其製造商的相關資訊) 公開揭露。此外，共有 24 個官方語言的歐盟是採多語言政策，這也代表製造商必須以有在其國內販售器材之成員國的所有語言提供特定報告。

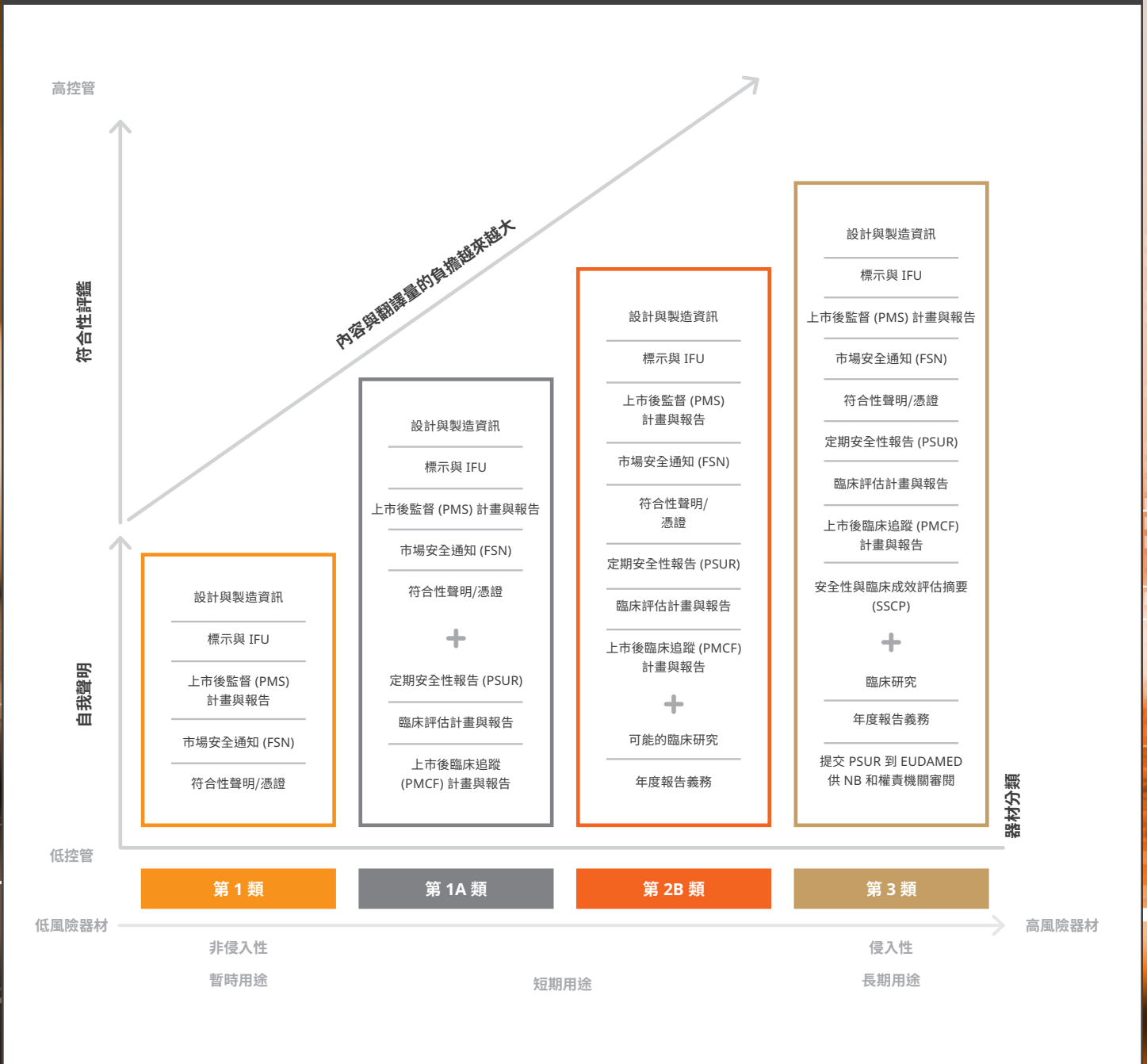
為了遵循新的歐盟架構，醫療器材和體外診斷醫療器材的製造商已經持續承受著極大的壓力，在這種情況下還要實行納入 LLM 的語言策略，難免會讓他們感到難以負荷。不過，製造商可以將策略領導事務託付給語言服務供應商 (LSP)，就自身產品和預期用途提供必要的專業知識，並在 LSP 的協助下了解語言技術和內容策略。

想要符合 MDR 及 IVDR 的語言要求，並充分利用與發揮 LLM 的價值，我們建議製造商應針對其器材或甚至是其產品系列的整個生命週期，根據器材風險剖析、預期用途、相關內容的類型以及目標使用者等要素，規劃與建立適當的語言策略。

比對內容也有助於判斷採用 LLM 的程度，以及從事翻譯和/或使用其他語言服務時所需的人工介入程度。

歐盟 MDR 的風險分類促使內容與翻譯量增加

法規要求



語言 AI 應用探索與解決方案開發

對全球研究結果以及醫療行銷來說，語言是成功的先決條件，大型語言模型也將深深強化語言服務的品質。這些技術可望協助我們打造最佳的語言工作流程和資產，也可以針對不同語言、目標對象和預期用途，生成與處理新的內容。

諸如 Lionbridge 等語言服務供應商 (LSP)，正積極快速地自行或是與業界合作探索和開發 AI 的使用案例。生命科學產業的其中一個特色，就是資訊與內容量極其龐大又不斷增加，內容類型也非常多樣，像是受法律規範或不受法律規範，或是醫學或淺白用語風格等。AI 不但正大幅改變語言服務產業，也將藉由使用 LLM 搭配其他語言資源，徹底革新未來支援醫療結果的做法。

LLM 可以為特定目標對象或市場生成或「重新合成」新的內容，而且不一定非得依賴傳統的來源檔案。只要有適當的指示和輸入，LLM 便能以不同風格產出不同的內容類型，來吸引特定的目標對象或媒體。然而，LSP 處理的經常是客戶的商務關鍵與機密內容，因此我們也必須要能妥善因應風險，並打造值得信賴的解決方案。

為了達到這個目標，Lionbridge 會持續設法深入了解客戶的內容與產品、法規要求以及預期用途。我們在生命科學產業中擁有穩固的關係以及深厚的專業能力，因此具有獨特優勢，能成為深受客戶信賴的合作夥伴，與客戶一同探索與實行採用 AI 技術的語言解決方案。我們也投入了大量的資源，開發一套健全又可靠、以 AI 為主的服務系列。時至今日，Lionbridge 已為超過 500 個客戶提供量身打造的 AI 解決方案。我們甚至還以 5 個重要評量，建立了獨特的 TRUST 架構。

TRUST

TRANSPARENCY (透明性)

在流程的每一步都可取得專案與資料的最新資訊。

RELIABILITY (可靠性)

Lionbridge 會盡全力處理您的解決方案與資料。我們擁有營運數十年累積的好成績，是值得您信賴的合作夥伴。

USEFULNESS (有用性)

我們只會提供您所需的解決方案，不會額外推銷或收費。

SAFETY (安全性)

我們設立了許多道防護機制來保障資料的安全性，讓您安心無憂。

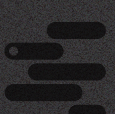
TIMELINESS (及時性)

Lionbridge 一向會準時交付您的 AI 專案。

探索由 LIONBRIDGE 提供、採用 AI 技術的生命科學語言解決方案

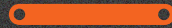
想探索有哪些安全、創新又適合您製藥或醫療器材活動使用的 AI 解決方案嗎？歡迎與 Lionbridge 合作，體驗我們 AI 及生命科學專家團隊的優質服務。我們精通 AI 事務、提供 TRUST 架構，並具備數十年的生命科學產業專業知識，能滿足您流程任何階段的語言或內容需求。

如需深入了解，請造訪
[LIONBRIDGE.COM](https://www.lionbridge.com)



關於 LIONBRIDGE

Lionbridge 與眾多品牌攜手合作，致力在全球各地達成「通情達意，暢行無阻」的使命。超過 25 年以來，我們已經為許多客戶與員工遍布全球的公司，提供超過 350 種語言的翻譯與本地化解決方案，協助他們順利獲得成功。Lionbridge 運用世界級作業平台，以及由全球各地充滿熱忱的專家所組成的網路，與各大知名品牌聯手打造豐富的文化體驗。我們憑藉著對語言的無窮熱忱，運用尖端的人類與機器智慧，打造出能讓客戶與其客戶之間產生共鳴、增進認同，並提高相互理解的優質內容。Lionbridge 總部位於美國麻州沃爾瑟姆市，在全球 24 個國家/地區設有解決方案中心。



如需深入了解，請造訪
LIONBRIDGE.COM



LIONBRIDGE

© 2024 Lionbridge. 著作權所有，並保留一切權利。

