

出國報告(出國類別：開會)

SNOMED CT 年度會議

服務機關：衛福部資訊處

姓名職稱：李建璋技監兼處長

派赴國家/地區：韓國/首爾

出國期間：113年10月18日至10月24日

報告日期：114年1月16日

摘要

台灣各家醫院的電子病歷系統多為各自開發，導致資料格式不一致，難以滿足智慧醫療發展的需求。為解決此問題並奠定國家長遠基礎，衛生福利部提出一綜合性次世代數位醫療平台策略，使台灣 HIS 能夠對接國際標準，實現資料統一、規則統一與應用程式統一三大目標。資料統一的重點在於：建立資料交換標準（FHIR）、醫療語彙（SNOMED CT）、檢驗檢查編碼（LOINC）與藥物編碼（RxNorm）等標準並導入醫院系統，並藉由訓練及工作坊持續推動臺灣在地應用與版本維運，以支持公共衛生研究及臨床醫療數據的標準化交換，實現與國際標準接軌的同時，也保持與本地實務的連貫性。透過三大核心基礎標準之建立，「車同軌、書同文」，實現台灣醫療數據可互通性，加速數據分析應用契機，並帶動產業發展。

其中，SNOMED CT（Systematized Nomenclature of Medicine - Clinical Terms）是全球最廣泛使用的醫學名詞編碼系統之一，由國際衛生標準發展組織（IHTSDO）維護與更新。今年，台灣以衛生福利部名義加入，首次成為這個國際重要組織的會員國，參與國際醫學名詞標準制定大會，由本人李建璋處長代表出席。雖然此次以觀察員身份參加，但透過加入 SNOMED CT 的國際標準推動組織，取得國際使用權並與現有國內標準協調推動，對台灣而言是一個重要的起點。

此次參加韓國的國際醫療名詞標準會議，深入了解各國在資料標準化與電子病歷統一方面的政策規劃與進展。會議中，美國、韓國、荷蘭、英國及阿根廷等國家分享了各自的政策方向，提供對全球趨勢的重要洞見。全球資料標準化與電子病歷統一的政策可分為三個主要內容：一是國家核心資料群的建立，確保不同機構之間的資料互通；二是國家個人健康資料入口網的成立，集中管理與查詢個人健康資料的系統；三是電子病歷的驗證與管理，確保系統的統一性與安全性。

台灣的電子病歷普遍以英文紀錄，這使得我們在推動國際標準化上具備一定優勢。透過發揮本地科技人才的能力並開發更多相關工具，台灣有望順利融入國際編碼標準的制定，並在全球電子病歷標準化進程中扮演關鍵角色。

目次

摘要I

壹、目的	1
貳、行程表	2
參、過程	3
肆、心得與建議	15

壹、目的

台灣各家醫院的電子病歷系統多為自行開發，導致資料格式不一致，難以支援智慧醫療的需求與發展。為解決此問題並奠定長遠發展的基礎，衛生福利部提出次世代數位醫療平台策略，旨在使台灣醫療資訊系統（HIS）全面對接國際標準，實現資料統一、規則統一與應用程式統一三大目標。

資料統一的核心在於推動以下標準的導入：資料交換標準(FHIR)、醫療語彙(SNOMED CT)、檢驗檢查編碼(LOINC)及藥物編碼(RxNorm)。透過訓練與工作坊，逐步促進標準的在地應用與版本維運，支持公共衛生研究及臨床醫療數據的標準化交換。不僅能實現與國際標準接軌，還可確保與本地實務需求的連貫性，進一步達成「車同軌、書同文」，實現台灣醫療數據的互通性，促進智慧醫療數據分析應用並帶動相關產業發展。

其中，SNOMED CT 作為全球應用最廣泛的醫學名詞編碼系統，由國際衛生標準發展組織(IHTSDO)負責維護。今年，台灣以衛生福利部名義首次加入該組織，成為會員國，並由李建璋處長代表參加在韓國舉辦的 SNOMED CT 國際醫療名詞標準會議。透過此次會議，與各國專家深入交流，學習資料標準化及電子病歷統一的政策規劃與實踐經驗，並探討如何將 SNOMED CT 與台灣現有醫療系統協調推動。雖然僅以觀察員身份參與，此舉已成功為台灣取得 SNOMED CT 的國際使用權，並為國內標準的推進奠定了重要基礎。

貳、行程表

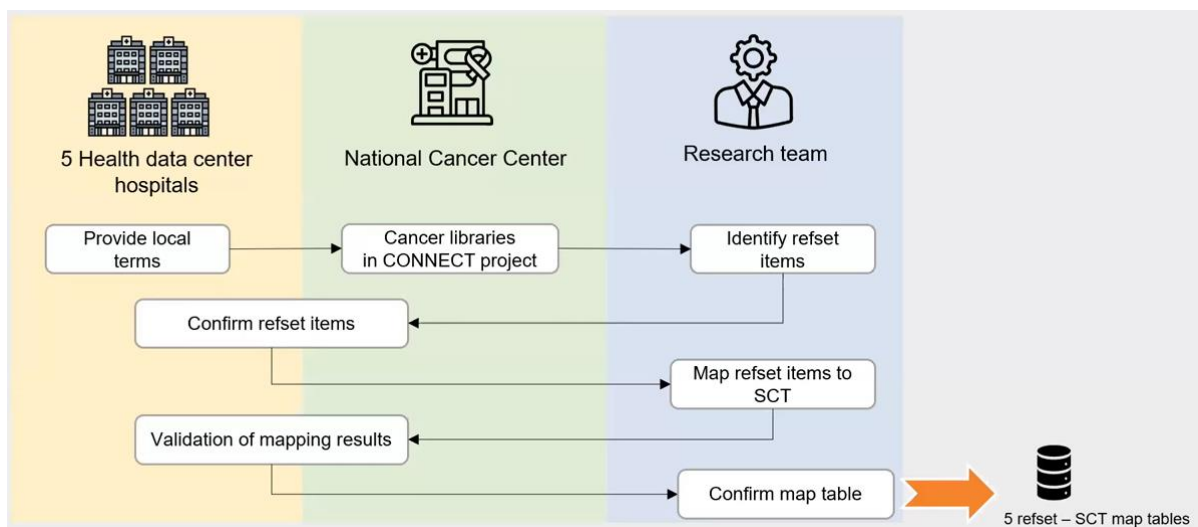
日期	地點(城市)	拜訪對象(行程)
10/18(五)	桃園機場	搭機抵達韓國首爾(移動日)
10/19(六) 至 10/23(三)	韓國首爾	October 2024 - SNOMED International Business Meetings
10/24(四)	韓國首爾	搭機返回台灣(移動日)

參、過程

1. Development of five cancer reference sets in SNOMED CT in South Korea -Sumi Sung (KR)

(1).內容

- 韓國首次使用 SNOMED CT 開發的乳癌、肺癌、大腸癌、胃癌和肝癌的癌症參考集。
- 使用國家癌症圖書館、CONNECT 癌症大數據平台、分析從 EHR 收集的自由文字資料並審查臨床實踐指南來定義癌症資料項目和值集。
- 將五種癌症的資料項和值集映射到 SNOMED CT 概念，並根據對應等級和目標概念的數量對映射結果進行分類。
- 如果我們找不到等效的 SNOMED CT 概念，我們將建立一個具有焦點概念和屬性值對的後協調表達式。



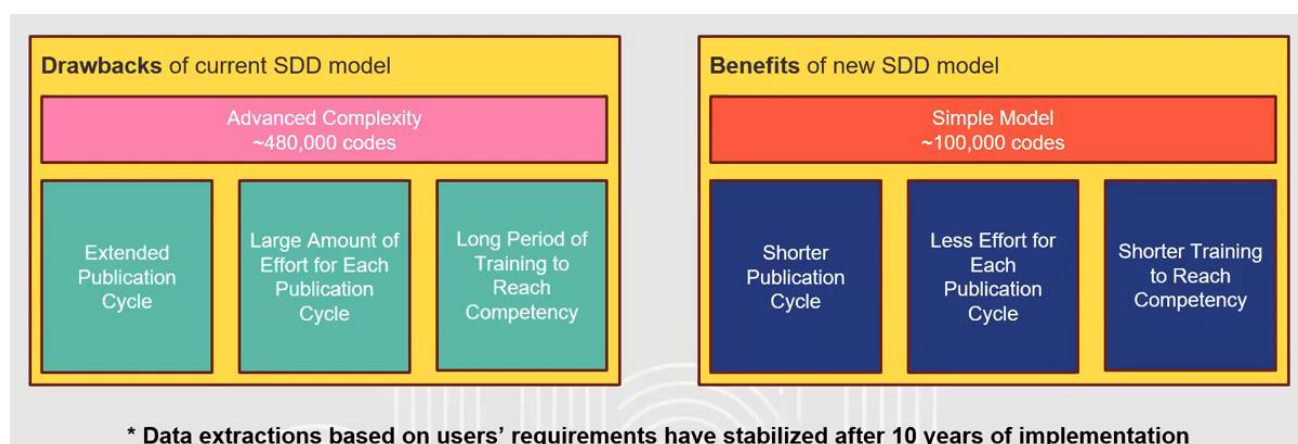
(2).結論

- 可參考韓國所開發的 23 個癌症參考集，將提供一種一致的方式來收集五種癌症的數據。這將有助於癌症患者的連續照護，並增強癌症大數據的互通性。

2. Enhancing the Model of the Singapore Drug Dictionary - Nicholas Seah (SG)

(1).內容

- 新加坡藥品詞典(SDD) 的最新版本是 SNOMED CT SG 的擴展，旨在滿足眾多使用例的需求，包括藥品整個生命週期（從主管部門審批和供應鏈到醫院開具處方和配藥等運營）的語義和功能互通性、藥物安全措施以及用於分析和研究的資料探勘。
- 全面的設計導致了過大的規模和複雜性，使資料庫的創作和維護變得複雜，並延長了發布週期。
- 新提出的模型透過重新分配描述的複雜性來解決這些挑戰，從而減少了約 80% 的概念和相關關係。
- 使用 ECL 表達式自動分配成員以對概念進行分組，以及從複雜的發布週期過渡到單步驟過程。



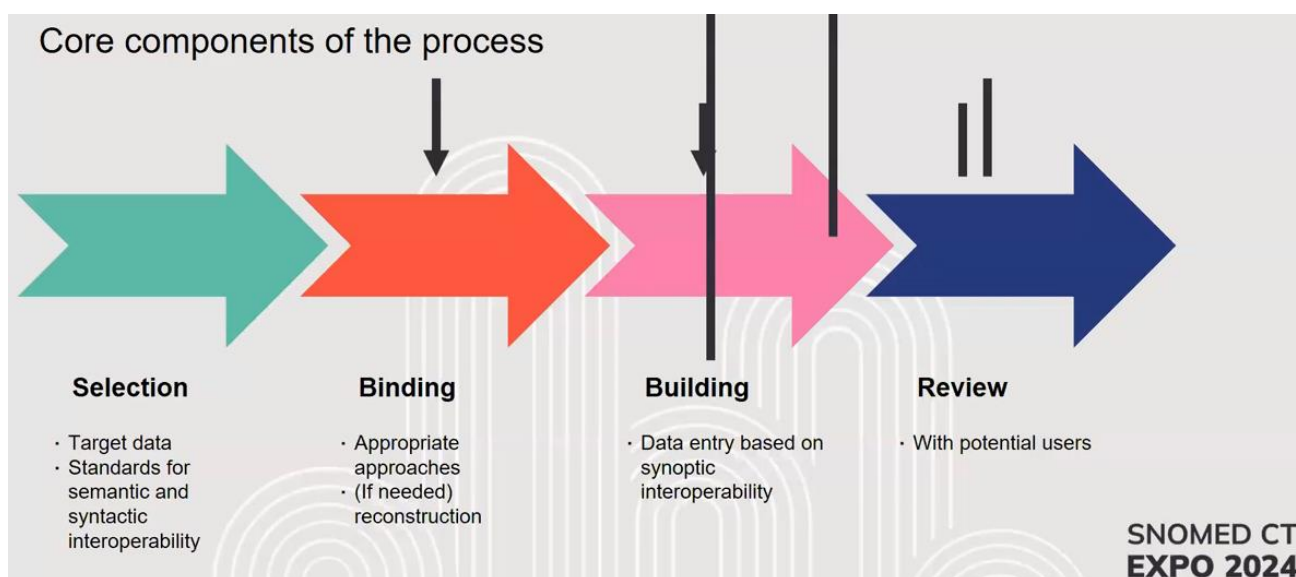
(2).結論

- 精簡的模型簡化了創作、維護和發布週期；由於概念的減少，本體生成和分類更有效率；隨著層次結構的減少和縮小，維護變得更簡單；由於規則更簡單，資料提取速度更快，可能不需要人工審查。

3. Toward Semantic and Syntactic Interoperability of Cancer Synoptic Report (KR)

(1).內容

- 病理報告對於診斷和治療至關重要，由於沒有資料元素的標準表示，互通性仍然是一個挑戰。
- 為了證明病理癌症報告中的互通性，作者將 SNOMED CT 與符合快速醫療互通性資源(FHIR) 的結構化資料擷取(SDC) 格式的國際癌症報告合作(ICCR) 病理癌症資料集綁定在一起。
- 使用美國國家醫學圖書館 (NLM) 提供的工具開發了 SDC 表格。 NLM 工具會為使用者建立一個電子資料輸入表單，並使用 SNOMED CT 編碼資料以符合 FHIR 的格式呈現已完成的表單。
- 未綁定的資料元素主要與「病理分期」相關，因為 ICCR 採用基於「國際抗癌聯盟 (UICC)」標準的癌症分期水準。



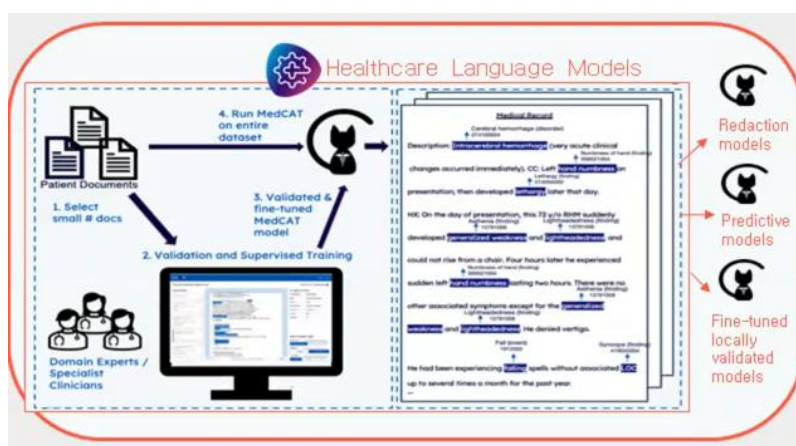
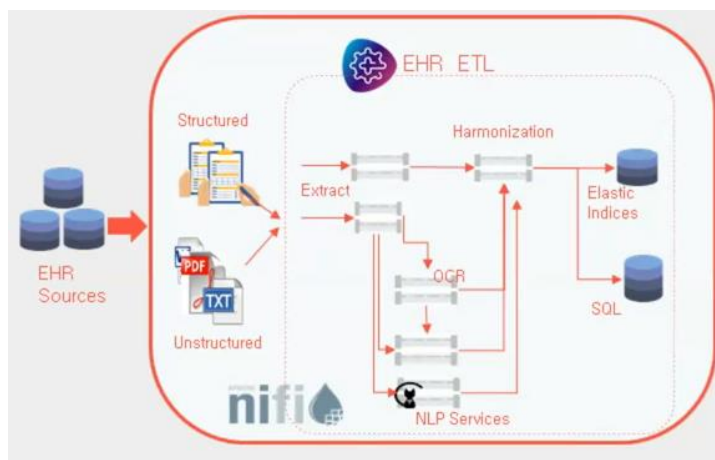
(2).結論

- UICC 概念尚未包含在 SNOMED CT 中，但不久的將來將會納入。本計畫證明了在病理癌症報告中實現語義和句法互通性的可行性和可處理性。

4. Multilingual support for systematic identification, linking and contextualisation of SNOMED CT terms from clinical narratives (UK)

(1).內容

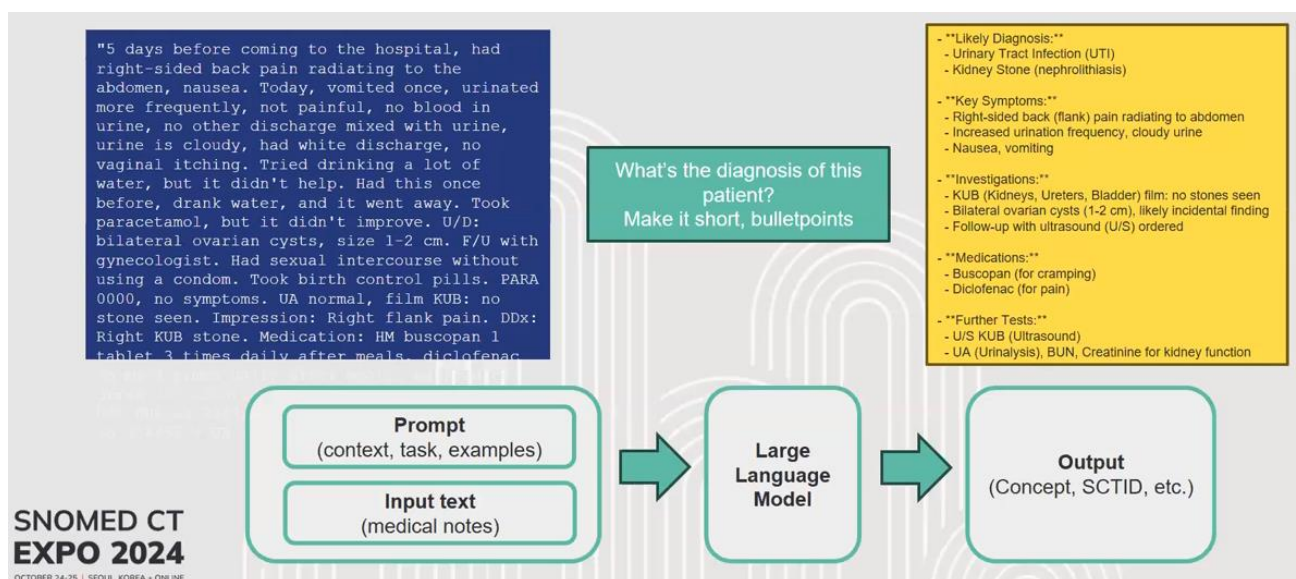
- 透過 CogStack 和 MedCAT，一系列工具支援從非結構化臨床敘述資料中識別、連結和語境化 SNOMED CT 術語及其在非英語語言地區的使用。
- CogStack 和 MedCAT - 一組用於 SNOMED CT 的自然語言處理 (NLP) 工具，其中臨床敘述來自英語地區。
- 講者介紹如何支援臨床資訊學和術語學家社群採用 SNOMED CT，並採用訓練、驗證和微調 NLP 模型的工作流程來識別/連結/情境化來自不同語言和衛生系統的 SNOMED CT 術語。
- 說明如何將 SNOMED CT 的持續版本快速無縫地整合到 CogStack NLP 管道中，從而加快本體中創作概念的使用及其在下游用例中的使用。



5. Healthcare Information Extraction Using LLMs and SNOMED CT: Sharing Thailand Experiences and Challenges - (TH)

(1).內容

- 該專案利用與 SNOMED CT 整合的大型語言模型(LLM)從雙語泰語醫療紀錄中提取資訊並將其對齊為結構化資料表示。
- 檢索增強生成(RAG)整合了 SNOMED CT 等外部知識庫，以提高 LLM 準確性。該項目使用標準化 SNOMED CT 術語評估 RAG，從 Ramathibodi 醫院的泰國紀錄中提取資訊，並將數據與 SNOMED CT 概念對齊。
- 分析的關鍵面向包括 LLM 從泰語紀錄中提取臨床資訊、將提取的數據映射到 SNOMED CT 概念以及為泰語擴展識別新術語的性能。



(2).結論

- 整合結構化術語可減輕 LLM 幻覺，同時建構泰國的多語言臨床數據。該計畫致力於編譯和映射泰語醫學術語至 SNOMED CT。未映射的術語要經過專家審查，才能在即將到來的國家擴展中作為新概念或後協調表達納入。

6. MedCAT 2.0 - Medical Concept Annotation Tool. Using large language models to extract SNOMED CT annotations from EHRs (UK)

(1).內容

- 用於回答醫學領域問題的大型語言模型(LLM)通常在訓練過程中缺乏透明度，因此其解決臨床問題的能力存在不確定性。
- 本研究利用 SNOMED CT 圖表中的結構化知識來評估大型語言模型(LLM)在醫療保健領域的表現。

MedCAT 1.0...	MedCAT 2.0...
<ul style="list-style-type: none">• utilises LSTM models for entity recognition and concept linking• models employ context free language models ("word2vec")• has to have some manual handling of misspellings ("ful remission" -> "full remission" SCT Id: 103338009)• has to manually handle multiple names of the same concept ("complete remission" -> "full remission" SCT Id: 103338009)	<ul style="list-style-type: none">• will be using a transformer base model architecture for both Named Entity Recognition (NER) and Linking (+L)• models will have contextual information (i.e. a contextual model would represent "shot" differently based on context)• aims to better handle misspelling as part of its architecture• will aim to be robust and handle multiple names without expert intervention

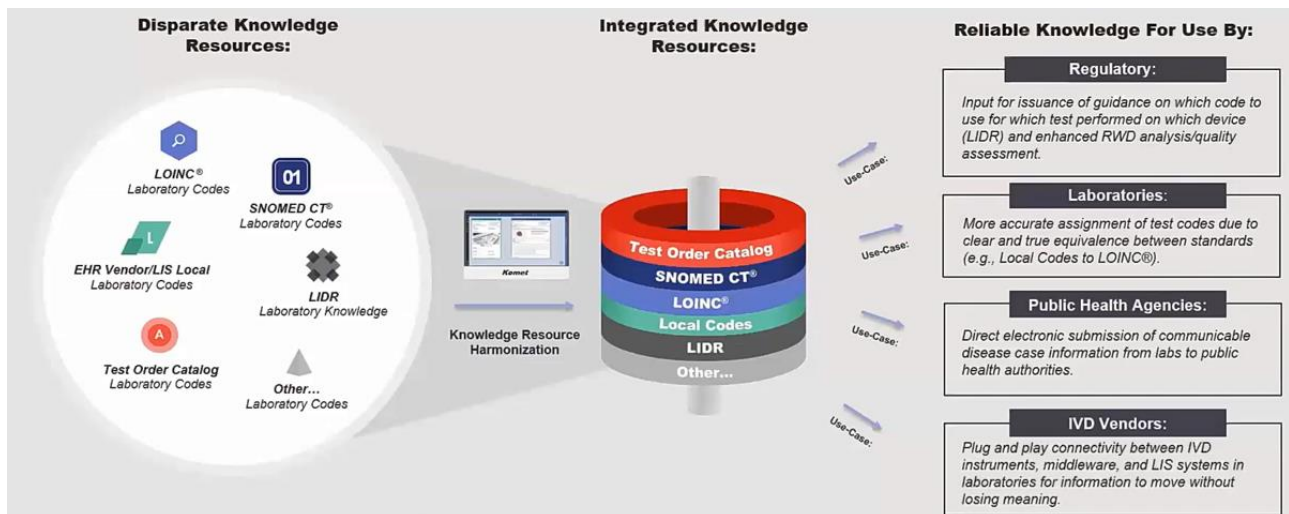
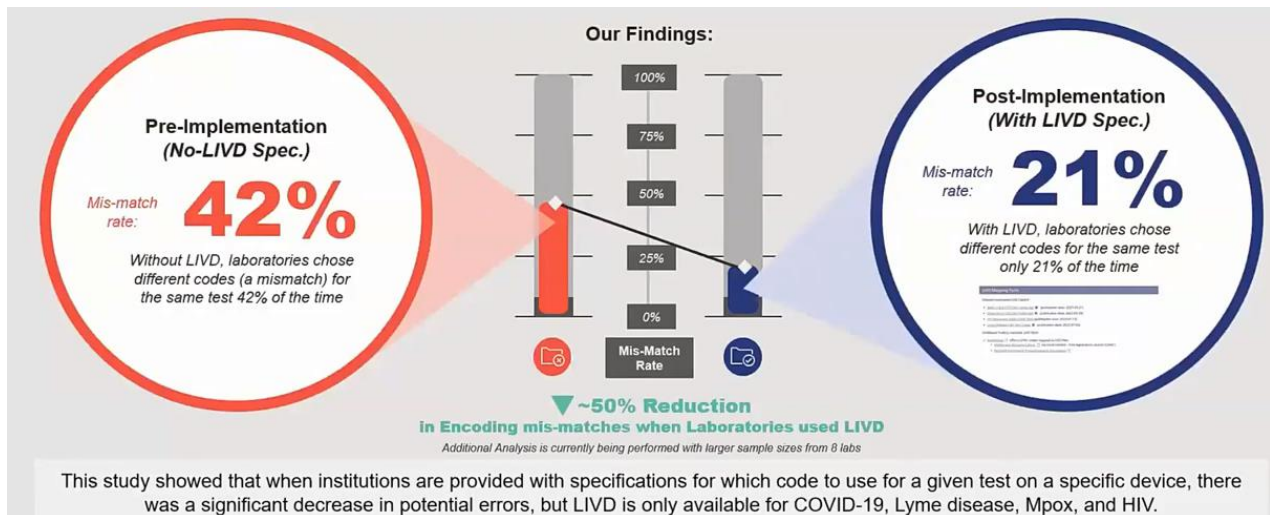
7. Modeling In Vitro Diagnostic Device Data using the LOINC SNOMED CT Collaboration Extension: A Path to Success - (US)

(1).內容

- 來自 IVD 設備的數據佔所有臨床決策的約 70%，但是，包括 IVD 設備數據在內的標準化數據採集仍然存在差距。
- 設備製造商和實驗室資訊系統(LIS)的編碼不標準化，捕獲標準化實驗室數據（包括 IVD 設備數據）仍然具有挑戰性。
- IVD 製造商、實驗室、電子健康紀錄(EHR)和 LIS 供應商使用的術語編碼實踐存在差異，因此在不丟失任何含義的情況下交換實驗室數據仍然存在問題。
- 目前，LOINC®、SNOMED CT®概念模型和其他互通性標準並未捕捉 IVD 設備資料

的含義、命名法和術語結構。本專案描述術語屬性要求來表示 IVD 設備的獨特組件。

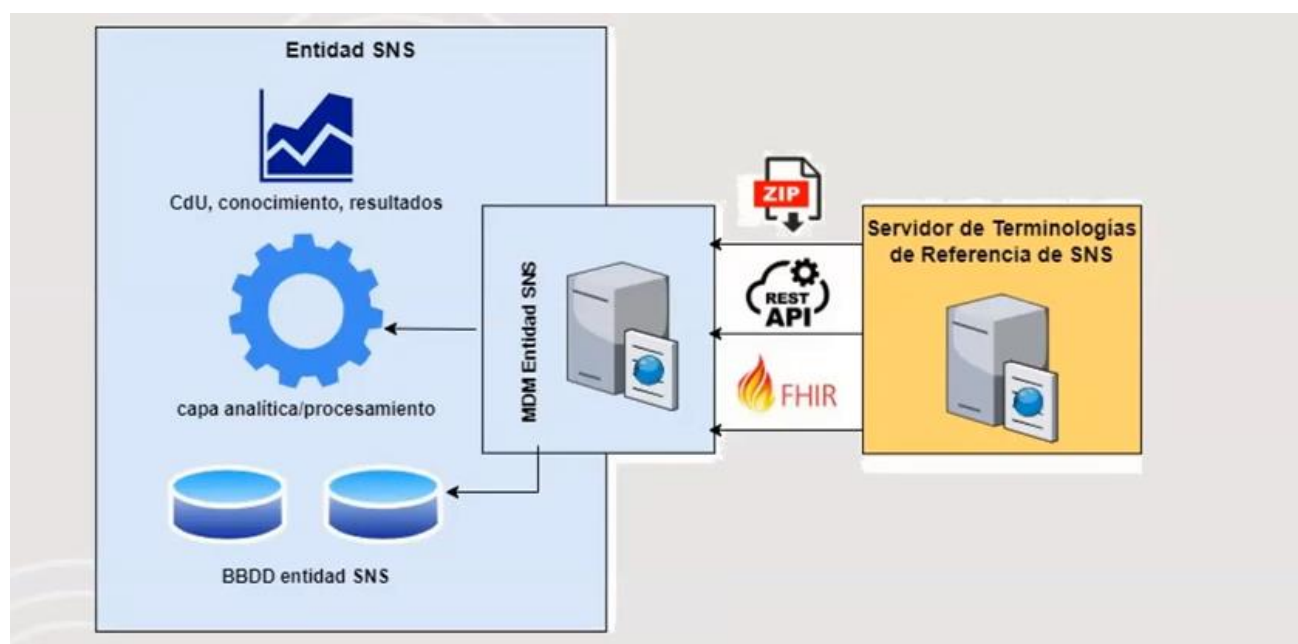
- 針對實驗室互通性資料參考(LIDR) 精選的 IVD 資料元素，以探索和擴展 SNOMED CT®/LOINC®共享概念模型的領域範圍約束，專案的最終目標是維護實驗室結果(包括 IVD 設備資料) 的資料完整性。



8. The Reference Terminology Server: Centralizing Medical Terminologies for Improved Healthcare System Integration in Spain (ES)

(1).內容

- 術語參考伺服器(strSNS)是西班牙國家衛生系統的基礎創新。strSNS 由西班牙國家參考中心管理，由衛生部啟動，使用 SNOMED International 的開源工具 Snowstorm 系統，向所有組織和使用者提供全國衛生系統使用的參考資源。
- 基於 snowstorm 構建，並透過 ElasticSearch，可以處理複雜的搜尋和管理大量數據，增強系統之間的整合。
- strSNS 不僅限於 SNOMED CT，還包括 ICD-10、ICPC-2 和 LOINC 等一系列醫療資源。此伺服器專為直接互動而設計，允許使用者透過 REST API 或 FHIR API 服務下載內容或查詢資料。



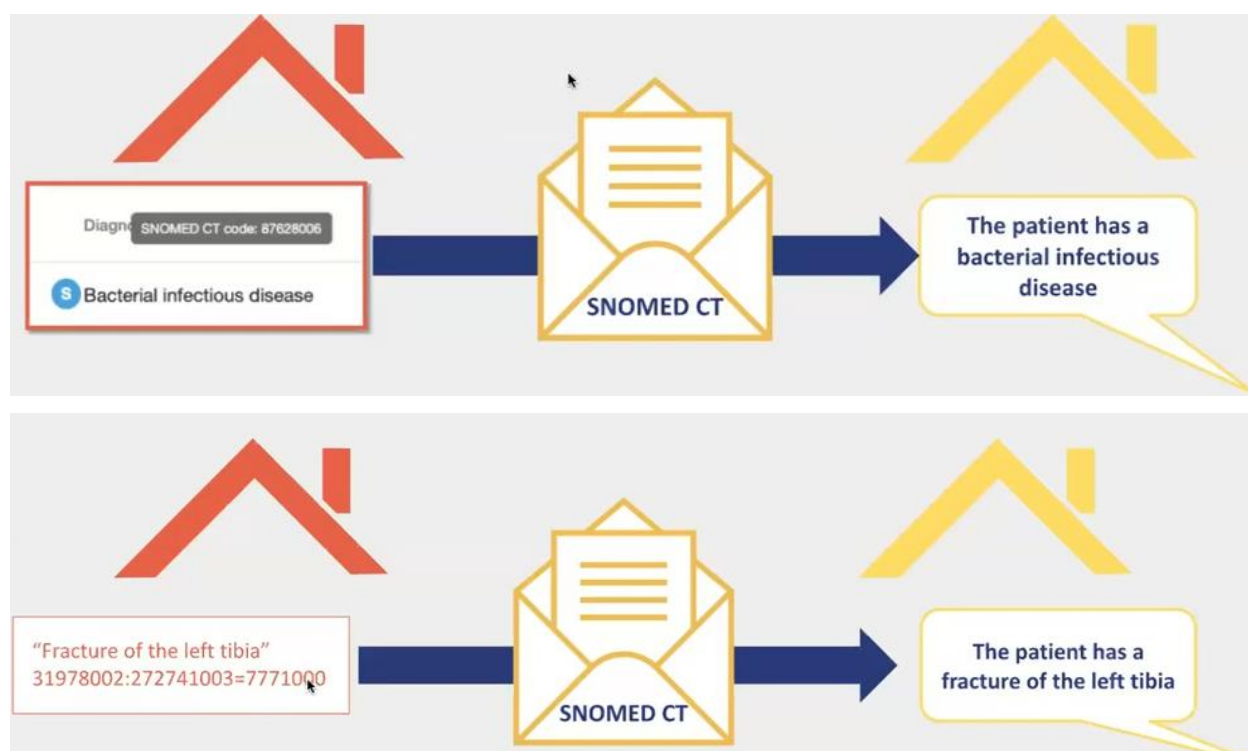
(2).結論

- 其自適應框架確保透過添加新目錄保持相關性，並透過無縫存取最新的醫學術語來支援資料驅動的醫療，從而改善工作流程和醫療保健服務。

9. Strategies for quality information exchange in healthcare

(1).內容

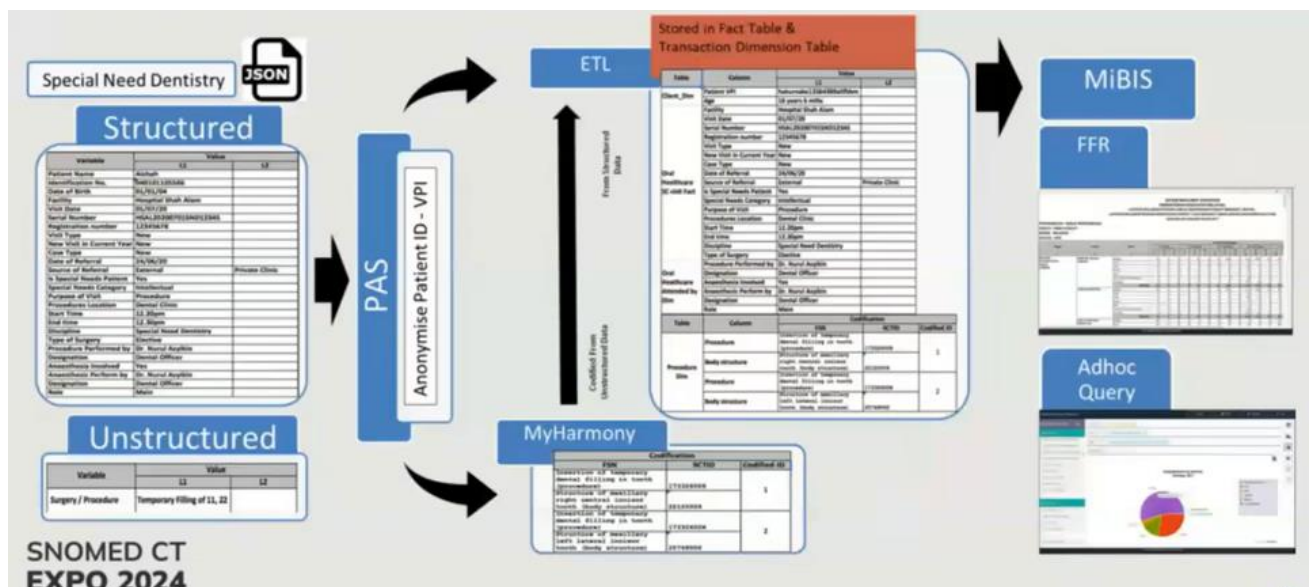
- 本議程深入探討了標準在促進有效資訊交換方面所發揮的關鍵作用。會議的很大一部分內容專門討論 SNOMED CT 及其在醫療保健資訊交換中的作用。
- 了解使用 SNOMED CT 的各種方法，包括共享預協調概念、後協調內容以及擴展和版本之間的共享技術。
- 議程中還探討了 SNOMED CT 如何與資訊模型的互通性，並涵蓋了 SNOMED CT 與 FHIR 資源和 FHIR 問卷的整合。
- 本教學最後將探討資訊交換架構的關鍵考量因素，讓學員全面了解推動醫療保健有效資料共享的因素。



10. Standardising Data Entry Using SNOMED CT Local Extension For Consistent Analysis and Expanding Data Dictionary - (MY)

(1).內容

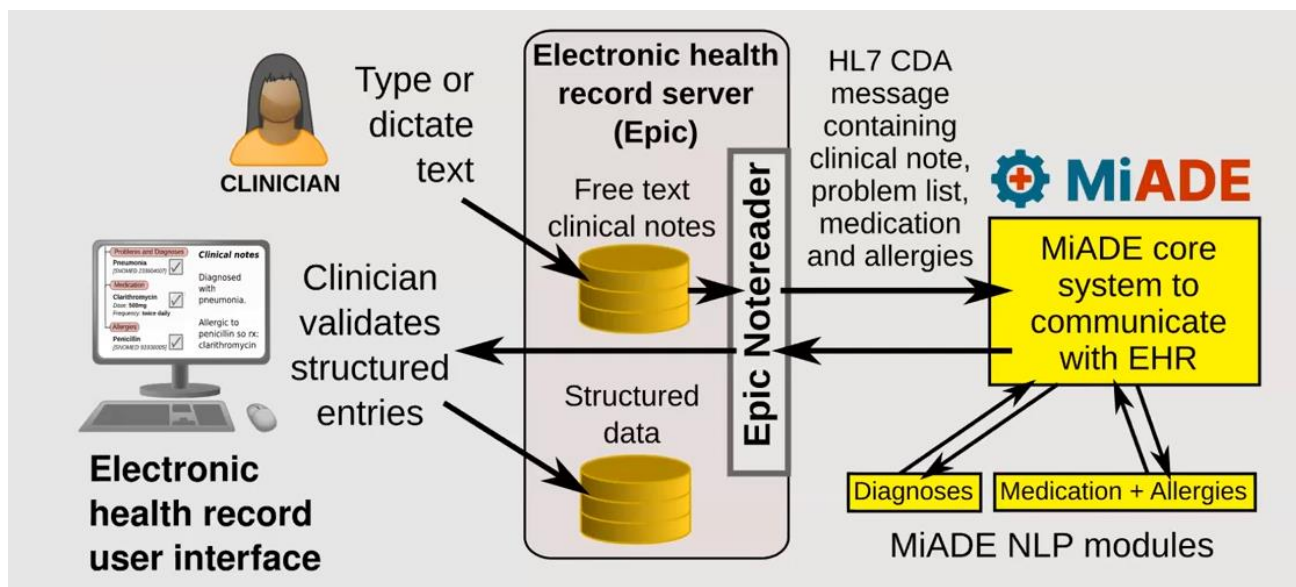
- 馬來西亞健康資料倉儲（MyHDW）的資料收集系統中收集的大部分資訊都是結構化的。
- 透過整合 MyHarmony，MyHDW 能夠透過心臟病學出院摘要協調和編碼數據來成功分析數據，並使用 SNOMED CT 建立儀表板或報告，從而實現非結構化分析。
- 在 2018 年，eReporting V2.0 Oral Health 被開發為另一個 MyHDW 的來源系統。此資料收集系統與 MyHarmony 集成，僅用於將自由文字與口腔健康參考集編纂成規範，以便及時產生靜態（固定格式報告）和動態（臨時查詢）報告，以適應預先存在的（所有口腔保健機構（包括馬來西亞各地的醫院和診所）BI 平台（結構化資料分析）。
- 本次演講中，講者展示馬來西亞團隊為標準化非結構化資料輸入所做的努力，以實現所需的輸出分析：固定格式的報告、臨時查詢，進而擴展馬來西亞健康資料字典 (MyHDD) 。



11. Feasibility study of ‘MiADE’ point of care natural language processing (NLP) system: methodology and initial results (UK)

(1).內容

- 有效的病患照護依賴於良好的醫療紀錄，但當今許多醫療資訊仍然是非結構化，阻礙了臨床使用和研究。
- 醫療資訊 AI 資料擷取器 (MiADE) 基於開源 CogStack MedCAT 命名實體提取系統，以及基於規則的截面檢測演算法，是講者團隊開發的創新自然語言處理 (NLP) 系統。
- 無縫整合到電子健康紀錄 (EHR) 系統中。 MiADE 即時將非結構化文字轉換為結構化數據，並將 SNOMED CT 概念呈現給臨床醫生進行驗證和改進。
- MiADE 在倫敦大學學院醫院 (UCLH) 的臨床評估研究中實施，並於 2024 年 2 月 26 日向有限的使用者群體上線。

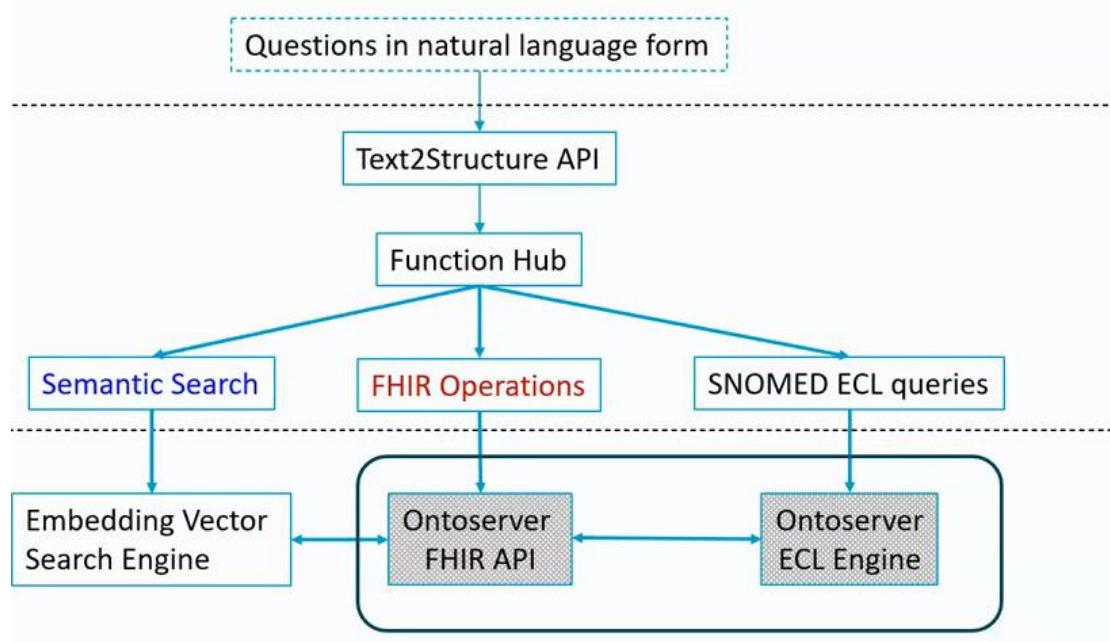


12. NL2ECL: Transforming Natural Language Queries to SNOMED CT Expression Constraint Language - (AU)

(1).內容

- 在醫療資訊學領域，醫療數據的互通性對於有效的溝通和分析至關重要。SNOMED CT 作為一套綜合的臨床術語，是此項努力的基石。然而，要充分發揮其潛力需要熟練其表達約束語言（ECL），這對非技術使用者來說是一個障礙。
- 本研究提出了一種將自然語言查詢轉換為 ECL 表達式的系統，提出了一種彌合這一差距的新方法。透過採用先進的自然語言處理技術、生成式人工智慧、深度學習技術，結合對 SNOMED CT 概念的語義理解，此系統 NL2ECL 方便使用者進行直覺查詢，實現與 SNOMED CT 術語的無縫互動。

Recap: OntoGPT workflow



(2).結論

- 本套工具可以使醫療保健專業人員能夠更好地存取 SNOMED CT，而且還為”改善臨床資訊學的互通性和語義查詢”開闢了途徑。

肆、心得與建議

台灣各家醫院的電子病歷系統多為各自開發，資料格式不一致，難以支援智慧醫療需求。為解決此問題，衛生福利部推動次世代數位醫療平台策略，旨在對接國際標準，實現資料、規則與應用程式的統一。策略重點包括導入資料交換標準（FHIR）、醫療語彙（SNOMED CT）、檢驗檢查編碼（LOINC）及藥物編碼（RxNorm），並透過訓練與工作坊推動在地應用與版本維運，支持公共衛生研究及臨床醫療數據的標準化交換，達到國際接軌與本地實務平衡，促進數據互通及產業發展。今年，台灣首次以衛福部名義加入國際衛生標準發展組織（IHTSDO），並派李建璋處長參加 SNOMED CT 國際醫學名詞標準會議，透過此次會議取得 SNOMED CT 國際使用權，並為國內標準推動奠定重要基礎。

在醫療實務中，不同醫師或醫療機構可能會使用不同的術語來描述同樣的事件，這會造成電腦系統無法準確識別它們實際上指的是同一件事。例如心臟病發作，醫師 A 使用 "心肌梗塞"（Myocardial Infarction, MI），醫師 B 使用 "心臟病發"（Heart Attack），醫師 C 使用 "冠狀動脈阻塞"（Coronary Artery Occlusion），這三個名詞在臨床上指的是同一種情況，然而，沒有標準化的編碼系統，這些名詞會被電腦視為不同的資訊，導致診斷紀錄和資料統計的錯誤或重複，透過 SNOMED CT 這樣的標準化醫學名詞系統，這些不同的術語會被對應到一個統一的編碼。心臟病發作、心肌梗塞和冠狀動脈阻塞可能都會對應到同一個 SNOMED CT 編碼，讓電腦系統能夠準確識別它們是相同的醫學事件，避免混淆。

這次參加韓國的國際醫療名詞標準會議，了解了各國在資料標準化與電子病歷統一方面的政策規劃與進展。會議中，美國、韓國、荷蘭、英國及阿根廷等國家分享了各自的政策方向，可對全球趨勢有了更深入的認識。

首先，美國政府在資料標準化的推動上擁有最完整的規劃，涵蓋了標準的制定、測試、工具開發、法規完善以及認證機制，形成了一套系統性的推動策略。目前，全球只有少數國家如荷蘭、德國、愛沙尼亞、澳洲、新加坡及韓國等，正依照美國的模式來推動全國性政策。其中，韓國與荷蘭在標準化領域已走在前列。英國目前僅以倫敦為中心，採用 Oracle 商業化的電子病歷系統作為統一標準，尚未擴展至全國。而阿根廷則僅提及其首都地區的醫院發展情況。

全球資料標準化與電子病歷統一的政策大致可分為三個主要內容，一是國家核心資料群的建立：確保不同機構之間的資料互通；二是國家個人健康資料入口網的成立：集中管理與查詢個人健康資料的系統；三是電子病歷的驗證與管理：確保系統的統一性與安全性。

至於我國的進展，我們除了已經完成了「台灣核心資料群」的定義與執行指引外，也開發了利用人工智慧的半自動化資料轉換工具，讓不同醫療資訊系統之間的資料可以無縫轉換。目前，我們正規劃推動全國性的個人健康資料入口網，與現行的健康存摺不同，此入口網將採用國際資料標準，確保與各大醫院的電子病歷系統無縫接軌，有助於提升醫療效率及醫師的連續性照護，進而保障民眾健康。

除了資料的整合，我們也發展了規則圖書館，這可用來進行跨院的醫療品質管理與監測，提升行政效率。例如，健保的事前審查、藥物規範的處方提醒，以及醫院的預審，均可透過規則圖書館的應用來實現。此外，規則圖書館還可推廣實證醫學，進一步提升醫療品質。

我們的未來規劃還包括建立以SMART on FHIR 為標準的應用程式市集。目前，除了美國已有國家級的規劃外，其他國家尚未推動類似計畫。如果我們能在兩年內順利完成這些計畫，台灣將在次世代電子病歷的轉型中進入國際領先的行列。

最後，我們都知道，資料的規模與計算能力已成為國家實力的象徵。在醫療領域，若要創造龐大的資料庫，首要任務是實現不同醫院資訊系統之間的互通，其次是資料內容的標準化。唯有如此，我們才能生成高品質的大數據，推動智慧醫療與資料經濟的發展。

台灣的電子病歷普遍以英文紀錄，這使得我們在推動國際標準化過程中具備一定優勢。只要充分發揮台灣科技人才的能力，開發更多相關工具，我們將有機會順利融入國際編碼標準的制定中，並在全球電子病歷標準化的過程中扮演關鍵角色。