

基於多模態迭代檢索之本地衛教問答系統： 整合視覺、語音與專業反饋

**Multimodal Iterative Retrieval-Augmented Local Health Education Q&A system:
Integrating Vision, Speech and Professional Feedback**

指導教授：蔣榮先 合作醫師：劉冠宏、邱裕桓 學生：李乃仁

大綱

現狀與面臨瓶頸	01
---------	----

解決方案	02
------	----

系統介紹	05
------	----

系統流程圖	06
-------	----

UI展示	07
------	----

效能驗證	14
------	----

未來展望	15
------	----



臨床衛教的實務挑戰

在繁忙的臨床現場，護理師需重複解答病患關於飲食、術後照護的各種疑問。面對資訊焦慮的病患，需一個能提供：

- **即時性**：24/7 不間斷的基礎衛教諮詢。
- **一致性**：確保每位病患獲得標準且正確的指導。
- **精確性**：結合病患個人數據的針對性回覆。

AI現狀與面臨瓶頸



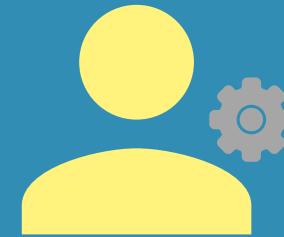
可信力不足

LLM 常產生「幻覺」，在零容錯的醫療環境中，無法確保參考資料的準確性。



缺乏專業評估

醫療問答缺乏「黃金標準」，難以由系統自動化且精確地評估回覆品質。



無法個人化

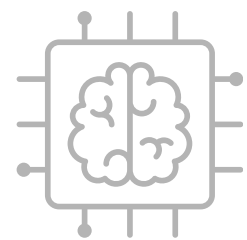
AI無法根據個人生理狀況給每位病人個人化的衛教問答。

瓶頸：可信力不足

解決方案一

強化系統可信力與來源可追溯性

策略：本地端語言模型與追溯機制



LLM + RAG

架設本地端語言模型，採取 RAG 結合多輪檢索，增加回答醫療準確度。



可驗證的來源文獻機制

透過 Prompt 強制模型輸出引註來源，並在 UI 介面精確展示文獻頁碼與內容片段，提供醫護人員快速校核。

瓶頸：缺乏專業評估

解決方案二

產學合作評估與自動化

策略：人機協同評估體系

建立Ground-Truth

與成大醫院專科醫師合作，針對臨床衛教建立標準題庫與答案集，使自動化評估具備專業基準。

導入 Ragas 與專業指標

參考 Amazon 公開之評估框架，並導入 Ragas 自動化評估指標，使模型優化有跡可循。

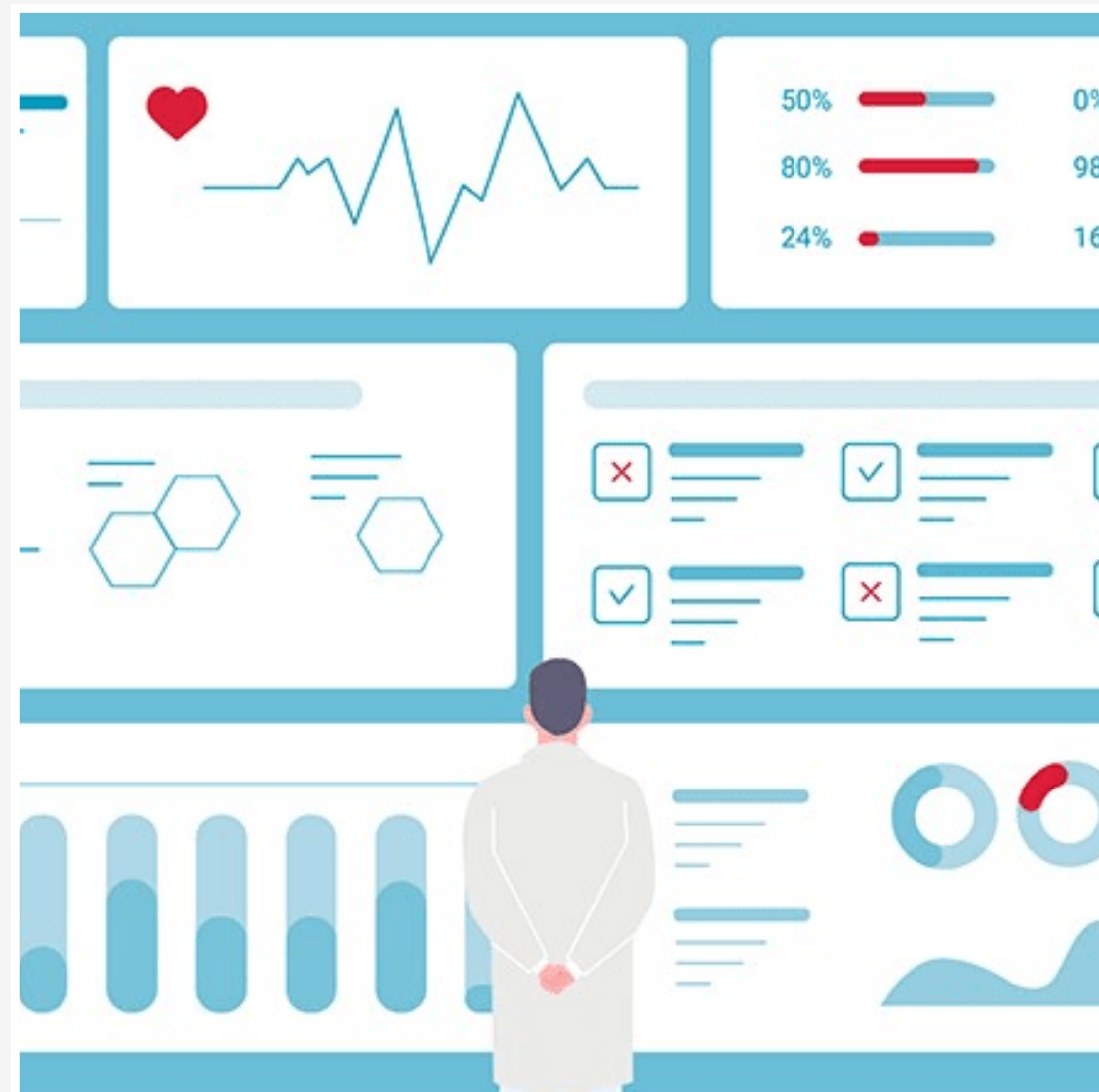


瓶頸：無法個人化

解決方案三

結合病患特徵的衛教問答

策略：針對病患數據精準分析



動態生理數據整合

系統將自動提取病患最新的檢驗值，並將其加入 Prompt 框架中。

精準化對話

LLM 根據注入的病患生理狀況，生成「針對性」的衛教指引。

系統介紹

核心架構：

前端

文字 / 語音輸入
輸出區塊渲染
文獻/個人數據管理
鏡頭監控

HTTP / SSE 串流

後端

多模態混合檢索
Database
API路由

特色：

精準檢索與重排

混合檢索：結合「語意向量」與「詞頻」檢索。
多模態整合：提取guideline中的圖片與表格。

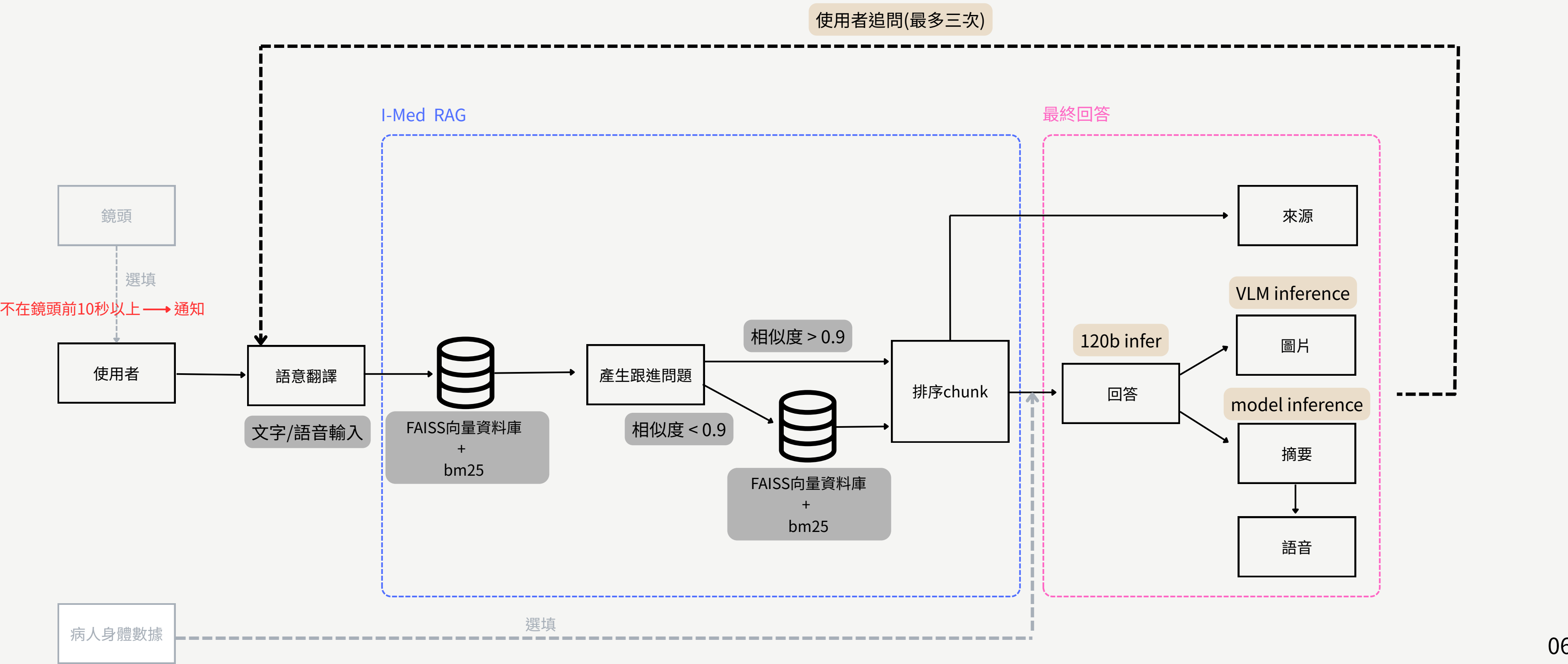
病患個人化

輸出多元：回答涵蓋摘要、全文、語音、圖片、來源」
個人數據：考慮病患個人化數據產生回應。

病患行為分析

鏡頭監控：前端偵測病患是否在鏡頭前。
行為數據化：記錄病患網站上的行為並分析。

系統流程圖



UI展示

病人個人化問答

病患身份

00000001

最新檢驗值

日期：2026-04-22

Hb

11.5 g/dL g/dL

Ca

9.5 mg/dL mg/dL

K

3.7 mmol/L mEq/L

ALB

4.0 g/dL g/dL

P

2.1 mg/dL mg/dL

✓ 已自動注入查詢

Nephrology AI

腎臟科智慧諮詢系統

病患身份

(未選擇)

未選擇病患，AI 不考慮個人檢驗值

參考文件

上傳

✓ KDIGO-2025-IgAN-IgAV...

✓ KDIGO-2026-Anemia-i...

✓ PIIS008525382300766...

已選 3 / 3 份

本次對話

追問進度

0 / 3

回饋紀錄

對話

鏡頭偵測病人情況

未啟動

腎

腎臟科 AI 諮詢

請在下方輸入問題，或選擇常見問題開始諮詢。

AI 將依據已選文件提供參考資訊。

常見問題

尿毒症的早期症狀是什麼？

腎臟病患者要注意哪些飲食習慣？

急性腎損傷的常見原因是什麼？

腎臟病患者可以服用哪些止痛藥？

輸入問題，或按麥克風錄音

問題輸入(打字/語音)

自由選擇參考文獻

記錄每個問題最多只能追問三次

紀錄使用者問答回饋

Nephrology AI

腎臟科智慧諮詢系統

病人身份

(未選擇)

未選擇病患，AI 不考慮個人檢驗值

參考文件

上傳

☒ KDIGO-2025-IgAN-IgAV...

☒ KDIGO-2026-Anemia-i...

☒ PIIS008525382300766...

已選 3 / 3 份

本次對話

追問進度 0 / 3

對話

腎

腎臟科 AI 諮詢

請在下方輸入問題，或選擇常見問題開始諮詢。
AI 將依據已選文件提供參考資訊。

常見問題

尿毒症的早期症狀是什麼？

腎臟病患者要注意哪些飲食習慣？

急性腎損傷的常見原因是什麼？

腎臟病患者可以服用哪些止痛藥？

輸入問題，或按麥克風錄音

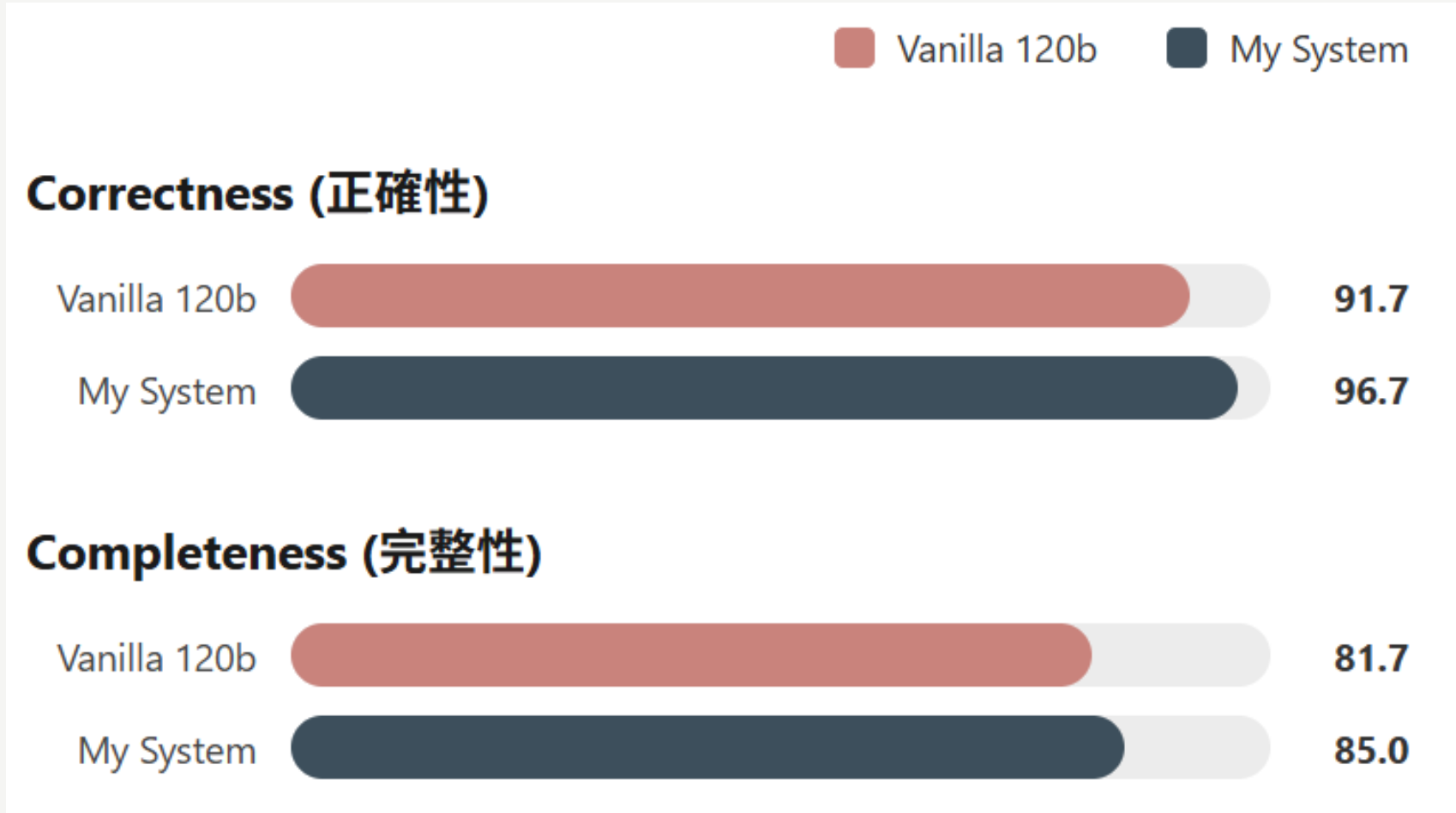
未啟動

開啟監控

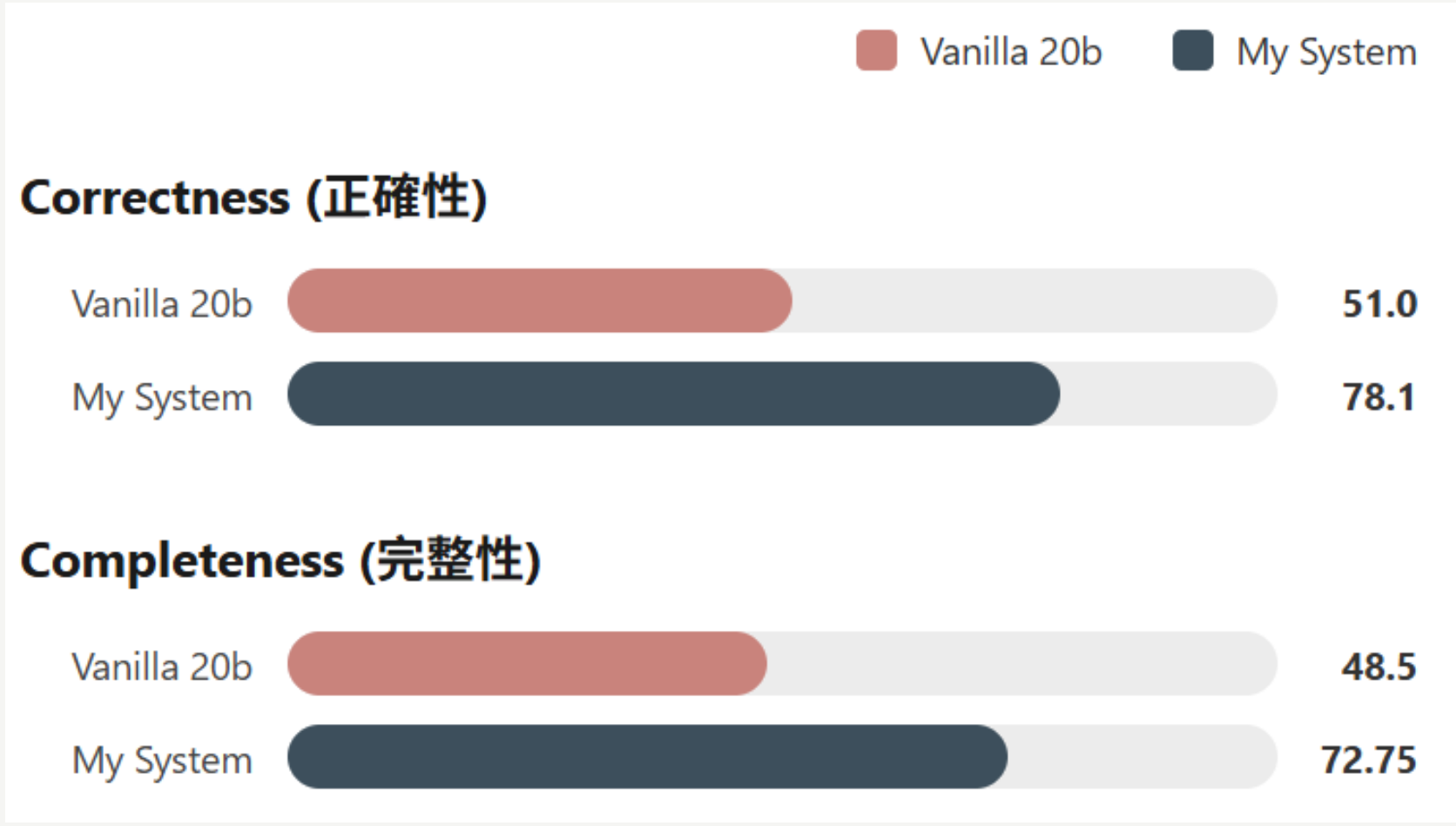
效能驗證：系統評測與臨床實測成果

技術指標對比

120b：



20b：



效能驗證：系統評測與臨床實測成果

Ragas 自動化量化評估

評估指標	平均分數
Faithfulness	0.988
Context recall	0.878

臨床醫療人員實測



未來展望

1. multi-hop: knowledge graph：針對RAG技術可以考慮以multi-hop加強。
2. 鏡頭偵測本地化：為保護病患隱私，鏡頭偵測之後可以繼續修改成只在本地運行。
3. 實時監控：建立api定期打入病患數據到系統，讓系統可以根據最新數據進行回答。
4. 整體流程加速：目前需要約一分鐘產生最終答案，未來可以嘗試在不失準確度的情況下增加速度。
5. Parsing pdf 速度：目前使用Marker來parsing pdf，但因醫院本地運算資源不足，無法發揮Marker效能，未來可以嘗試從實驗室接Marker的api到醫院。

REFERENCE

1. Guangzhi Xiong, Qiao Jin, Xiao Wang, Minjia Zhang, Zhiyong Lu, Aidong Zhang. (2024). Improving retrieval-augmented generation in medicine with iterative follow-up questions. arXiv. <https://arxiv.org/abs/2408.00727>
2. Baur, D., Ansorg, J., Heyde, C.-E., & Voelker, A. (2025). Development and evaluation of a retrieval-augmented generation chatbot for orthopedic and trauma surgery patient education: Mixed-methods study. JMIR AI, 4, e75262. <https://doi.org/10.2196/75262>
3. Amazon Web Services. (2025). Evaluate knowledge base–prompt and knowledge base with Claude 3.7 Sonnet. In Amazon Bedrock User Guide. Retrieved June 11, 2026, from <https://docs.aws.amazon.com/bedrock/latest/userguide/model-evaluation-type-kb-prompt-kb-sonnet-37.html>

醫院部屬

