

NTU AI Chatbot - PRD

一、專案簡介

本專案目標為開發一套 NTU AI Chatbot，部署於 NTU AI 社群平台（如 Discord、Facebook、Instagram、Web 等），提供社員與外部對 AI 社有興趣者一個即時互動的資訊管道，涵蓋社團簡介、課程安排、費用規則、專案組報名條件等核心資訊。

二、目標使用者

- 任何對 NTU AI 社有興趣的人
- 任何想要了解 NTU AI 社相關資訊的人

三、主要目標

- 提供即時、準確的 NTU AI 社資訊
- 降低幹部負擔，強化對外說明與服務品質
- 支援多平台與未來擴充空間（雙語系、API、檔案解析等）

四、功能規劃

MVP（優先開發功能）

功能模組	說明
Chatbot 主體	能透過 LLM + RAG 回答常見問題 如：社課時間、社費、如何加入等
FAQ 知識庫	內建固定回答邏輯，涵蓋常見問題集 如：沒有基礎可以加入嗎？社費是多少？
檔案向量處理	對社課課表（CSV）、守則（PDF）、講義（PDF）等檔案，使用繁體中文支援的嵌入模型（如： bge-large-zh 或 text-embedding-3-large ）轉為向量，儲存於資料庫
平台串接	Discord、Facebook Messenger、Instagram
語言支援	支援繁體中文與英文（以使用者輸入語言為基準）

延伸功能（後續迭代）

功能模組	說明
社課自動更新系統	自動抓取課表 / 管理者上傳資料後自動同步知識庫
使用者身分分級	根據是否為社員給予不同回答層級（例如：免費參與 vs 需繳費）
Web 服務入口	製作 Web 端聊天入口，包含資料上傳介面
行為記錄與統計	紀錄使用者提問熱區，回報常見資訊缺口

五、技術架構建議

層級	建議技術
前端整合	Python / Node.js
向量儲存	FAISS / ChromaDB / Qdrant (視部署環境選擇)
語言模型	本地部署 (如 Ollama + MediaTek Breeze 2 or BreeXe) / OpenAI gpt-4o / Gemini 2.0 Flash
向量模型	text-embedding-3-large (OpenAI) 或 bge-large-zh
資料處理	LangChain 或 Haystack 作為 RAG 框架
檔案處理	支援 PDF、CSV，自動內容擷取與分段處理

六、平台與部署策略

項目	說明
初期平台	Discord、Facebook Messenger、Instagram
未來擴展	Web
部署	可使用 GCP / AWS / Azure，或使用社內 Linux Server (具備 Docker) 若使用社團 Linux Server，建議配合可用本地 Ollama + GPU

七、建議開發時程

週次	任務
1	互相了解、明確需求確認、知識庫初稿整理、工作分配
2	向量資料流程建立、FAQ 系統整合測試
3	Chatbot 初版部署、RAG 整合、回應語言切換功能
4	對話準確性調整、修正回答邏輯
5	Discord / IG / Facebook 上線內測、FAQ 評估迭代、錯誤辨識加強
6	後期調整、準備分享內容

八、參考資料

1. [NTUAI 大小事](#)
2. [NTUAI 各站連結](#)
3. [MediaTek Research Foundation Models 聯發創新基地基礎模型](#)