

能源資產績效管理知識 GPT 模型程式委託開發

一、委託目標與需求

目前電力市場和能源政策包含大量複雜的規範和資料，資料分散於 POXA Info 網站的多個 Blog 文章和多份電力政策相關資訊之 PDF 文件中，這使得相關人員在尋找和解析資訊時面臨挑戰，難以快速了解市場的機會與風險。其中 POXA Info 網站¹每週都會分析電力交易市場的相關資料與數據，並會發布一份報告，總結市場活動，報告會包含電價、不同公司參與度和新規則等相關資訊。本委託案將規劃開發能源資產績效管理知識 GPT 模型程，簡稱 POXAGPT：一個專門針對電力交易和能源相關資訊的聊天機器人服務，POXAGPT 將透過整合多個資料來源（POXA Info 及台電能源局規範 pdf）並提供問答功能，以簡化資訊檢索和解析的過程，提升使用便利性(Usability)及擴大服務對象。

POXAGPT 服務之主要目標包含：

1. 提供一個能夠自動回答使用者問題的聊天機器人。
2. 可快速檢索和解析電力交易和能源政策相關的規範和資訊。
3. 能自動生成 Podcast 的腳本和講稿，以利電力新知資訊之傳播與推廣。

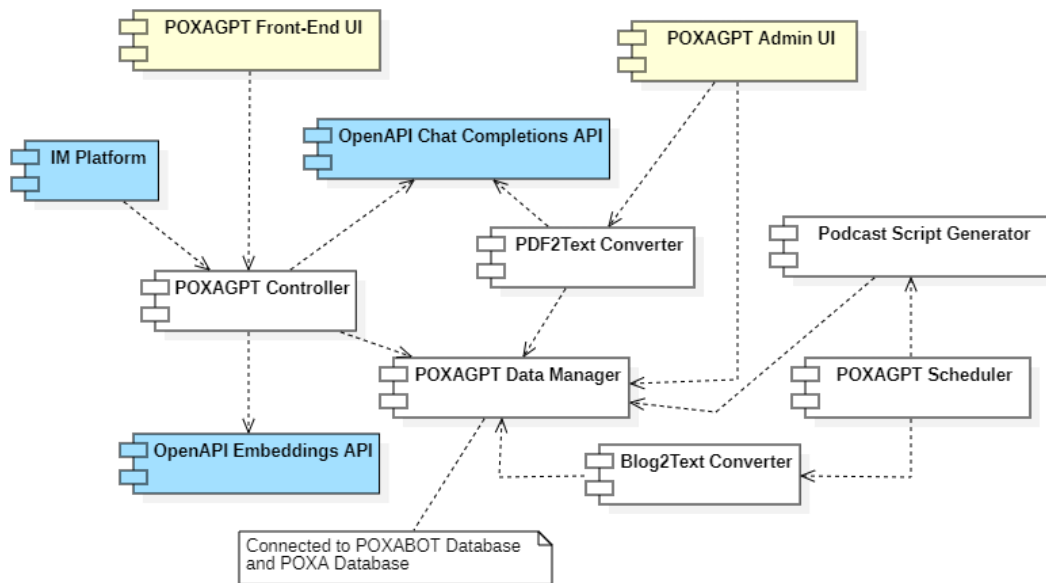
二、功能需求

1. POXAGPT 可整合多個資料來源，包括 POXA Info 網站內容、PDF 文件、MongoDB 資料庫等。
2. POXAGPT 能提供準確和相關的回答，涵蓋從市場價格到政策規範的廣泛問題。
3. POXAGPT 能提供 RESTAPI 介面，以便與其他服務(如 Line Bot、Slack Bot 和網頁)進行整合。
4. POXAGPT 應有管理者設定功能，以加入新的資料(如新的電力規範 PDF 文件)。

三、系統架構

POXAGPT 服務之預期系統架構如下圖所示，前端模組標示為黃色，後端模組標示為白色，外部模組則標示為藍色。

¹ <https://info.poxa.io>



1. POXAGPT Front-End UI：是 POXAGPT 的主要聊天/問答介面，可讓使用者提問並顯示回覆資訊。
2. POXAGPT Admin UI：是系統管理者介面，允許管理員新增新的電力法規或其他技術資訊之 PDF 文件到系統中。
3. POXAGPT Controller：核心控制器，負責處理使用者查詢和調用其他模組，如直接呼叫 PoxaBot Data Manager 進行資料查詢，亦會外部調用 OpenAI Embeddings API 以實現 Retrieval Augmented Generation (RAG)；透過結合 POXAGPT Data Manager 與 OpenAI 之 Function Calling 機制，以能進行資料查詢(如查詢指定時間之 dReg 商品價格)；以及調用 OpenAI Chat Completions API，以生成一般性的對話回覆資料。
4. POXAGPT Data Manager：資料管理模組，負責從對 Blog 和 PDF 轉換出的 JSON 資料進行讀寫操作，亦即負責與資料庫溝通，以能儲存和檢索轉換後的文本資料。此外，除了 Blog 與 PDF 轉換之文本資料庫外，POXAGPT Data Manager 亦可銜接 POXA Info 系統本身的資料庫，以利 POXAGPT Controller 實現 Function Calling。
5. Blog2Text Converter：Blog 內容轉換器，負責將 Blog 網頁內容轉換為文本，或取得原始的 JSON 資料，再轉換為適當的 JSON 格式 (例如過濾掉不相干的資訊)，最後再透過 POXAGPT Data Manager 將轉換後的資料儲存到資料庫。
6. PDF2Text Converter：PDF 文件轉換器，負責將包含法規或技術內容之 PDF 文件轉換為文本，然後轉換為適當的 JSON 格式，最後透過 PoxaBot Data Manager 將轉換後的資料儲存到資料庫。目前規劃運用 PyPDF2 套件²來擷取 PDF 內文，並透過 OpenAI Chat Completions API

² <https://pypi.org/project/PyPDF2/>

以進行文本校正。

7. Podcast Script Generator 用於生成 Podcast 腳本，根據 Blog 資料生成相關的 Podcast 腳本。
8. POXAGPT Scheduler：負責管理定期任務，安排和執行周期性工作，如每週使用 Blog2Text Converter 進行 Blog 內容的轉換與儲存，及定期調用 Podcast Script Generator 生成 Podcast 腳本。

POXAGPT 服務系統預計以 Python 為主要開發語言，後端模組 API 規劃以 FastAPI 框架建置，前端規劃運用 jQuery/Bootstrap 套件或 react 框架實現使用者介面。

四、實施步驟(工作項目)

1. 開發基本問答功能：

- 設計與實作 POXAGPT Front-End UI。
- 設計與實作 POXAGPT Data Manager，並先預載範例資料，以能成功提供基本問答效果。
- 設計與實作 POXAGPT Controller，以能串接前端、Data Manager 與 Open AI 之 API。Controller 會運用 OpenAPI 之 Chat Completions API，以生成合適的回覆資料。

預期產出：

完成 POXAGPT 系統之基本雛型，能夠根據預載之範例資料，理解使用者查詢與檢索相關資訊，並生成合適之回覆內容。

2. 開發資料轉換功能與完整問答功能：

- 設計與實作 Blog2Text Converter 和 PDF2Text Converter。
- 設計與實作 POXAGPT Scheduler。
- 設計與實作 Podcast Script Generator。
- 實現 Retrieval Augmented Generation (RAG)機制，能先擷取出合適的 Blog 或 PDF 之資料，以做為 LLM 文本生成的基底資料。
- 實現 Function Calling 機制，能夠透過問答介面進行資料庫查詢與回覆資料生成。

預期產出：

完成 POXAGPT 系統之核心功能，能夠將 Blog 和 PDF 內容轉換為結構化資料，並自動化任務執行與產生 Podcast 腳本。此階段之 POXAGPT 應能根據資料庫(包含既有 POXA Info 資料庫)之內容，較廣泛地理解使用者查詢與檢索相關資訊，並生成合適之回覆內容。

3. 開發服務管理與進階功能：

- 設計與實作 POXAGPT Admin UI。

- 設計與實作 RESTAPI 介面，以能讓各種前端介面或 IM 平台服務 (如 Line 或 Slack) 進行銜接。
- 提供離線版本 (安裝 MediaTek DaVinci 或 TAIDE-LX-7B 等繁體中文語言模型) 之相關規劃與 POC (Proof-of-Concept) 展示。

預期產出：

完成 POXAGPT 系統之完整功能與進階需求，以提供使用者管理介面與 Podcast 腳本生成功能，並支援可供外部服務調用的 REST API。此階段亦將提供離線版本 (不銜接 OpenAPI 之相關 API) 之 POC 展示。

4. 進行系統測試與發展相關文件：
 - 進行完整的端到端測試 (End-To-End Testing)。
 - 進行完整的效能測試 (Performance Testing)。
 - 撰寫軟體安裝文件與使用說明文件

五、產出物

1. POXAGPT 之基本功能：
 - 使用者聊天/問答介面 (POXAGPT Front-End UI)。
 - 問答控制器 (POXAGPT Controller)。
 - 資料處理和儲存模組 (POXAGPT Data Manager)。
2. POXAGPT 之核心功能：
 - 資料轉換工具 (Blog2Text Converter 和 PDF2Text Converter)。
 - 任務自動化執行工具 (POXAGPT Scheduler)。
 - Podcast 腳本生成工具 (Podcast Script Generator)。
3. POXAGPT 之服務管理與進階功能：
 - 管理員介面 (POXAGPT Admin UI)。
 - REST API 介面。
 - 離線版本 POC。
4. 設計文件、測試報告與操作手冊：
 - 系統架構、API 與資料設計文件。
 - 功能測試與效能測試報告。
 - 軟體安裝文件。
 - 使用說明文件

以上功能若引用第三方函式庫，須確保該第三方函式庫可合法商業使用。

六、交付時程

交付時程對應上述產出物

交付日期	交付項目
第一期	● 基本功能程式碼與設計文件。

(2024/9/31 前)	
第二期 (2024/12/31 前)	● 核心功能程式碼與設計文件。
第三期 (2025/2/28 前)	● 服務管理與進階功能程式碼與設計文件。 ● 系統測試報告、軟體安裝文件與使用說明文件。

七、驗收方式

第一期：可正確回答測試案例：

- Q1: 最新的 dReg 商品價格？以及目前參與容量？
- Q2: 本週市場情況摘要？
- Q3: 本週頻率變化？
- Q4: 本週是否有台電新的公告？
- Q5: 光儲的參與規則？
- Q6: 光儲無限套娃是啥意思？
- Q7: 我有 1MW 的光電案場，可以蓋多大的儲能案場？收益大概如何？
- Q8: 目前 E-dReg 的投報率如何？你建議投資嗎？

第二期：可正確回答測試案例：

- Q1: 跟我介紹一下一下 E-dReg 的規範？
- Q2: 幫我說明目前 sReg 價金的計算方式？
- Q3: E-dReg 的升降載是怎麼計算的？
- Q4: 成為合格交易者的資格是啥？請幫我整理申請流程。
- Q5: 即時備轉的 SBSPP 要如何計算？
- Q6: 這個網站是誰做的？團隊成員是誰？（客製化問題）

第三期：完成底下的進階需求：

- 提供銜接雲端 LLM 服務(使用 OpenAI 或 Gemini)之完整 PoxaBot 服務。
- 提供可供其他 IM 平台服務串接之 RESTful API。
- 提供離線版本(安裝 MediaTek DaVinci 或 TAIDE-LX-7B 等繁體中文語言模型)之相關規劃與 POC (Proof-of-Concept)展示。

八、經費

經費含 OpenAI GPT API 之預估費用 2 萬元，其估算方式：

- OpenAI GPT API 若採用 gpt-3.5-turbo 模型，每 1000 個 token (字詞單位)約為 0.002 美元 (以匯率 31.5 計算約 0.063 台幣)，單日接收到之問題數以 200 估算，回覆所需之 token 以 2000 估算 (以匯率 31.5 計算約 0.126 台幣)，每月花費約為 756，一年約為 9,072，兩年則為 18,144 元。

九、保固期

驗收合格後提供免費保固12個月。

十、智慧財產權歸屬要求：

承接方需確認相關軟體已取得合法使用，並未涉及侵權。委託之產出智慧財產權屬於委託方，未經委託方同意，不得擅自進行揭露或重製。

十一、 其他需求

受委託方需要每兩週至少一天至工研院中興院區或遠距進行進度報告、系統架構 review、或其他相關交辦事項。

建議廠商：

國立臺灣海洋大學 資訊工程學系

馬尚彬 教授

信箱：albert@ntou.edu.tw、shangpin.ma@gmail.com

電話：02-24622192 ext 6625