

<Cloud Computing Final Project Proposal>

Group 21

Group Member: 114522019,114522095,114522020,114522090,114522071

1. Motivation and Goal

對於現今快節奏的生活中，我們需要同時處理來自不同地方的資訊和任務。可能需要打開 A 程式來查看天氣預報，再切換到 B 程式搜尋資訊，然後用 C 程式來記錄日常開銷。這樣在多個程式間來回切換的操作長久累積起來會浪費大量的時間並降低效率。

此外，對於正在忙碌的使用者，例如在通勤途中或準備餐點時，以及無法輕鬆使用鍵盤或觸控螢幕的人，如年長者或行動不便者，傳統的觸碰與文字輸入方式往往成為一項明顯的障礙。因此，需要一個更聰明、直觀和集中的解決方案，以簡化日常生活的使用。

我們的目標是開發一個「個人助理機器人」，可以整合日常生活中常見的需求，並提供一個以語音輸入為核心的使用介面。希望能簡化用戶的操作流程，讓取得資訊和日常管理變得更即時、直觀和方便。

2. Related Work/Market Survey

隨著人工智慧與雲端運算技術的快速發展，AI 個人助理正逐漸從單純的語音控制工具，演進為能夠理解語意、學習使用者行為並主動提供資訊的智慧系統。這些系統不僅能執行語音指令，更強調與使用者的互動體驗與個人化學習能力。近年來，相關研究與產品的發展主要聚焦於以下四個方向：

A. 語意理解與自然語言搜尋

傳統的關鍵字式搜尋已難以滿足使用者的複雜需求。近年的研究重點轉向語意搜尋與檢索增強生成。這類技術透過語意向量嵌入與大型語言模型的結合，使系統能理解自然語言問題、自動判斷使用者意圖，進而檢索最相關的資料並生成自然流暢的回覆。

B. 人機互動

現今的智慧助理，如 Google Assistant、Amazon Alexa 與 Apple Siri，已能整合語音輸入與行為指令，提供多樣化的互動體驗。然而，它們在多語言理解、跨應用整合及個人化推薦等方面仍存在侷限。近年出現的新一代 LLM，如 ChatGPT、Claude 與 Gemini，更進一步強化了跨模態能力，支援語音、文字與圖像等多種輸入形式，以打造更自然且流暢的互動體驗。

C. AI 搜尋與雲端整合

企業與開發者愈來愈傾向運用雲端 AI 平台來部署智慧搜尋型應用。以 AWS 為例，其提供的 Amazon Bedrock、OpenSearch、Kendra、Lambda、EventBridge、DynamoDB 等服務，使開發者能建構具擴充性與安全性的 AI 搜尋系統，實現從「語音輸入 → 語意理解 → 文件查詢 → 回覆生成」的一體化流程。

D. 研究與市場差異

現有的 AI 助理多著重於通用應用，對個人化資料搜尋（例如「查詢我的支出」）或主動推播（例如「自動提醒明日天氣」）的功能仍相對有限。因此，本專案的特色與創新之處在於：

- 結合 AI 搜尋與個人資料分析
- 整合語音輸入與自然語言理解
- 利用 AWS 雲端服務實現彈性部署與自動化推播

E. 相關市場產品比較

為進一步說明本系統功能定位，以下列出相關應用之市場參考：

- 語音輸入
 - **Translatotron:** 是一個語音對話的模型，相比於傳統的語音對話模型，會將聲音轉成文字再給 AI 進行回覆，Translatotron 則是直接將語音當成輸入並輸出語音，回覆的效率更快之外也可以保留原始使用者的個人特質與情緒等等。
- 天氣預報
 - **Open-Meteo:** Open-Meteo 是一個免費、開源且無需金鑰的天氣資料服務，支援多種氣象模型（ECMWF、ICON、GFS 等）。
- 記帳
 - **Moze:** Moze 是一款個人記帳 App，提供多裝置同步、信用卡/分期/預算/多幣種，並加入 AI 語音、截圖、拍照記帳 與台灣電子發票整合等功能。
- AI 搜尋
 - **ChatGPT:** 由 OpenAI 開發的語言模型，能進行自然對話、撰寫文章、程式輔助、翻譯與資料整理等多種任務。它基於 GPT 架構，具備強大的理解與生成文字的能力。ChatGPT 支援多語言互動，且可根據使用情境調整語氣與內容風格。
 - **Gemini:** 由 Google DeepMind 推出的大型語言模型，前身為 Bard，整合了 Google 搜尋與多模態能力。Gemini 可同時理解文字、圖片與程式碼，擅長資訊查詢與邏輯推理。它強調與 Google 服務緊密結合以提升生產力。

- **Claude:** 由 Anthropic 開發的 AI 助手，基於 Constitutional AI 理念設計，注重安全性與倫理性。Claude 擅長長文本閱讀、摘要與創意寫作，並能保持穩定且符合人類價值觀的回應。它的回答通常語氣溫和、條理清晰，適合知識與文件處理任務。

3. Users

主要用戶：

- 忙碌的上班族和學生：上班族或通勤族可以在準備出門時用語音詢問天氣，或在處理其他事情時快速紀錄開銷。學生族群則可以利用 AI 搜尋快速查詢資料並同時管理個人預算。
- 科技愛好者：喜歡嘗試新科技，並使用工具來優化日常工作流程的用戶。

次要用戶：

- 年長者或行動不便者：對於不熟悉電子產品或不方便打字的用戶，語音數入提供了一個低門檻的操作方式，讓他們也能輕鬆體驗到科技帶來的方便。

4. Explanation of product features

A. 語音輸入

- 語音輸入的功能可以提供使用者非常多的便利性，不需要拿起手機即可與助理對話。
- 透過 Amazon transcribe 建立此功能，可以即時的將使用者所講的話轉成文字檔在進行下一步的 AI 回覆。

B. 天氣預報

- 使用者只需輸入地名系統即能即時回傳該地區的氣象資訊，包括目前氣溫、體感溫度、濕度、風速、天氣狀態，以及未來三天的預報。
- 可直接部署於 AWS Lambda + API Gateway，並搭配 AWS EventBridge 設定排程，自動每日推播天氣資訊。

C. 記帳

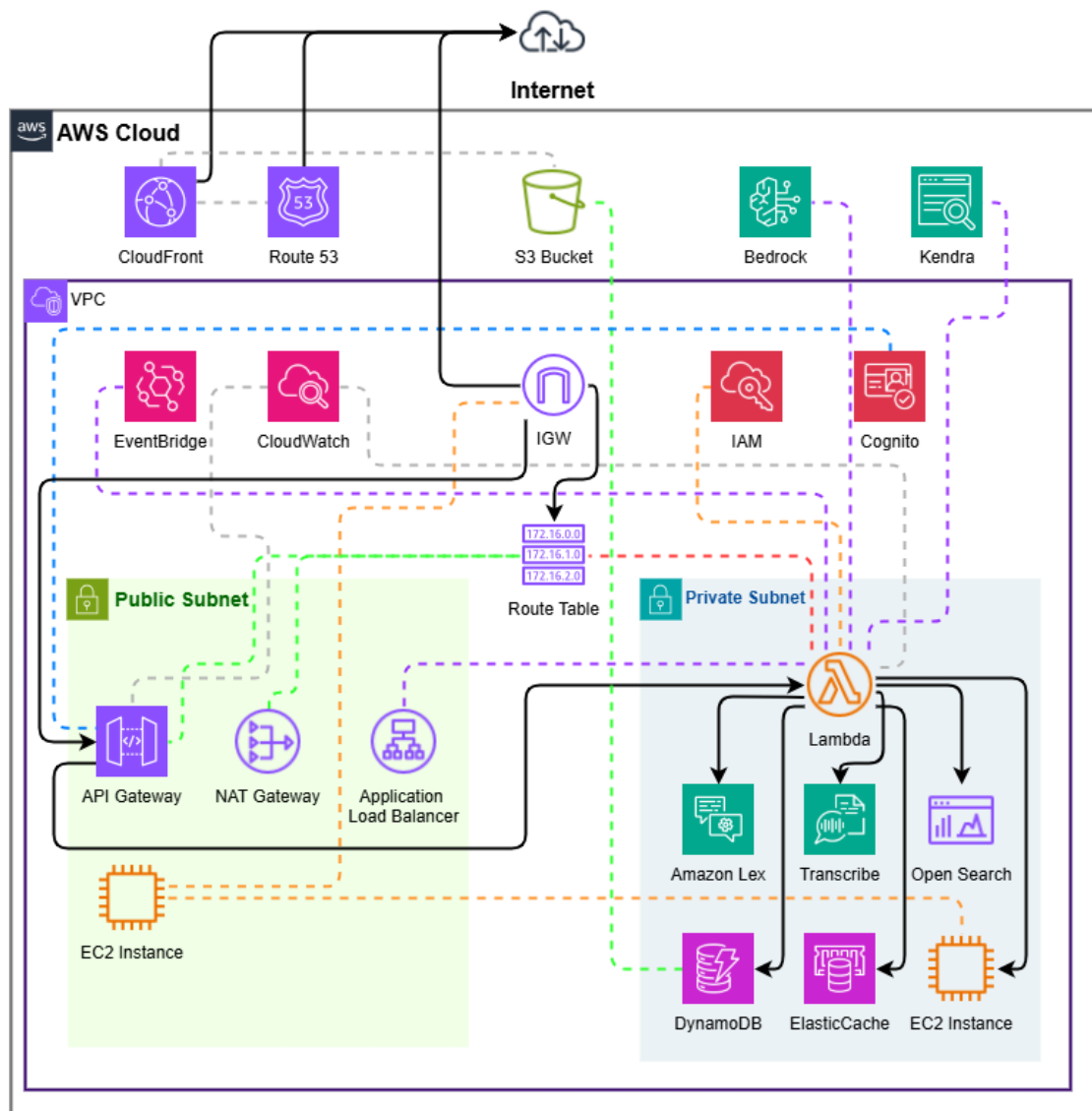
- 使用者可透過語音說出如「我剛剛花了 200 元買午餐」或「幫我記一下昨天的交通費 50 元」，系統將自動辨識支出金額、類別及日期，並自動記錄到雲端帳本中。
- 所有記帳資料皆儲存在 AWS 雲端（如 AWS DynamoDB 或 RDS），確保使用者可跨裝置即時同步資料，無論是在手機、平板或電腦上

操作，資料皆一致且安全。

D. AI 搜尋

- 同時搜尋網路公開資料與內部文件資料庫，結合語義分析技術進行內容比對，提供更準確與完整的回答。此功能可即時從外部知識與內部知識中擷取資訊，實現跨來源智慧搜尋與自動摘要整合。
- AI 助理可將使用者的個人資料（如記帳紀錄、天氣偏好）與即時外部資訊整合分析。透過語義搜尋與資料匹配，使用者能以自然語言查詢個人化結果，達到生活資訊與個人資料的連動。
- 使用自然語言理解，能判斷查詢意圖並轉化為資料檢索指令。當使用者以口語輸入問題時，助理可自動解析時間、類別與條件，從資料庫擷取相關內容後生成清晰、語意化的回答。

5. The solution architecture



6. Mind Map

