

# RAG PDF Reading Assistant

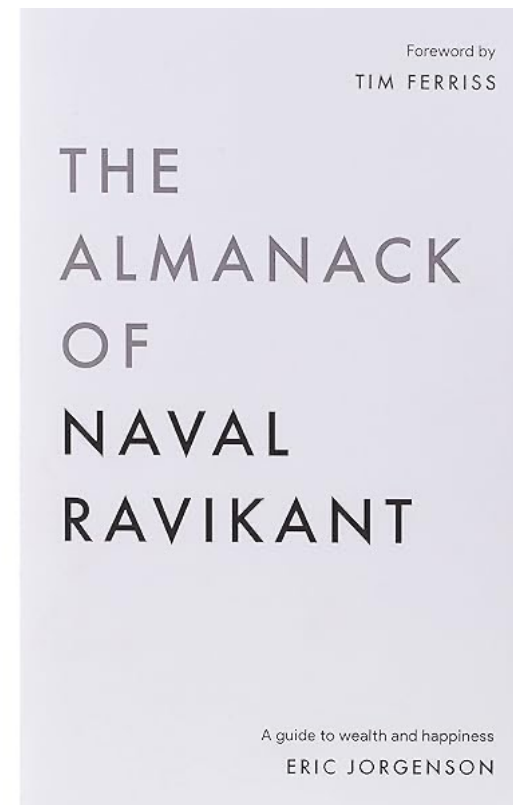
報告人：李盛廉

# Background

- 專案目標：建立一個基於PDF的 AI 諮詢機器人，能夠以PDF內容價值相關的語氣和原則回答問題。
- PDF: 《The Almanack of Naval Ravikant》
- Tools:
  - LangChain (框架)
  - Pinecone (向量資料庫)
  - OpenAI (LLM & Embedding)
  - Cohere (Rerank 模型)。

# Data

- The Almanack of Naval Ravikant.PDF
  - 由其推特金句和訪談精華而成
  - 涵蓋多個主題：
    - 創造財富
    - 快樂生活
    - 清晰思考
    - 建立長期思維...



# Text Processing

- Read PDF
- 文本切分策略
  - RecursiveCharacterTextSplitter。
  - chunk\_size=1000 and chunk\_overlap=100。
  - 因為這本書很多章節是由推特金句和訪談精華組成且橫跨多個主題，因此使用較小的Chunk與overlap。

# Embedding and Storage

- Model: text-embedding-3-large
- 維度: 3072
- Uploaded vectors to Pinecone

# 檢索

- gpt-4o 進行Query Expansion
- Pinecone 查詢回傳各query TOP 15 相關結果
- 以Cohere Cross-Encoder重新依與query的相關性排序  
Pinecone回傳結果，並取前5個。



# 生成

- Model: gpt-4o
- 自訂prompt
  - 根據PDF回答
  - 風格控制
  - 引用要求
- 輸入：自訂prompt + 檢索回傳的相關資料

```
=====
RAG Bot is ready (enter 'exit' to end the chat)
=====
Thinking...

Answer:
要擁有更好的人生，可以考慮以下幾點：

- **選擇幸福**：幸福是一種選擇和技能。你可以通過改變內心來提升幸福感，而不是依賴外部環境的改變。[Page 134]
- **培養良好習慣**：用好的習慣取代壞習慣，並與積極向上的人交往。你的習慣和你周圍的人會影響你的長期幸福。[Page 140]
- **內心的平和**：追求內心的平和，而不是試圖解決所有外部問題。放下對問題的執著，才能獲得內心的平靜。[Page 143]
- **降低身份認同**：減少對不重要事物的關注，不參與政治，遠離不快樂的人，珍惜生命中的時間。[Page 133]
- **活在當下**：不要過多地擔心未來或過去，接受當下的現實，這樣才能找到內心的平和。[Page 4]
- **避免嫉妒**：嫉妒是幸福的敵人，專注於自己的生活，而不是與他人比較。[Page 77]

這些方法可以幫助你逐步提升生活質量，並找到內心的平和與幸福。
```