

# FUNCube AI 產品說明書

## 產品概述：

FUNCube AI 是一款強大的 AI 智能助理平台，集成大型語言模型（LLM）、嵌入模型（Embedding Models）、向量資料庫（Vector Databases）、音訊轉錄模型（Transcription Models）等先進技術，幫助企業輕鬆打造高效、安全且具備彈性的 AI 解決方案。

### 產品核心功能

1. 模型靈活切換：根據業務需求自由選擇不同大型語言模型，可雲端運行或地端部署，實現最佳效能與成本平衡，確保數據安全與隱私。
2. AI 代理：類似 ChatGPT 的智能助手，可訪問工具執行網頁抓取、文件摘要等任務。
3. API 整合：開發者可透過完整的 API 介面，實現工作區管理、資料嵌入、即時對話等多樣操作，高效串聯業務流程。
4. 客製化外觀：可更換品牌標誌、歡迎訊息、頁尾連結，塑造企業專屬品牌形象。
5. 工作區對話紀錄：支援 CSV、JSON、OpenAI fine-tune 格式匯出，方便分析與保存。
6. 嵌入式對話元件：可嵌入網站，即時提供 AI 客服服務，並支援使用限制及安全設定。

## 技術架構與支持模型：

技術架構 FUNCube AI 採用 LLM（大型語言模型）+ RAG（檢索增強生成）雙引擎驅動，並整合多模型支持與向量資料庫，形成以下完整技術架構：

1. LLM 大型語言模型：生成式 AI 回答
2. RAG 檢索增強生成：基於知識庫提高準確性
3. 嵌入模型：將文本轉為向量，實現快速檢索
4. 向量資料庫：儲存和檢索語義向量，提高查詢效率
5. 音訊處理模型：提供即時語音轉文字（STT）與文字轉語音（TTS）功能，適合即時對話或語音客服場景
6. 語音轉錄模型：專門處理音訊檔案轉錄為文字，用於長音檔或會議記錄等場景

## 7. API 整合：無縫連接外部系統

### 支持模型類型

1. 嵌入模型 (Embedding Models)：文字轉向量，用於檢索增強生成 (RAG)
  - 本地模型：Ollama、LM Studio、Local AI、LiteLLM、OpenAI API相容模型
  - 雲端模型：OpenAI (語音轉文字與文字轉語音)、Azure OpenAI、Cohere、Gemini、Voyage AI、Mistral AI
2. 大型語言模型 (LLM)：生成對話與內容
  - 本地模型：Built-in、Ollama、LM Studio、Local AI、NVIDIA NIM、KoboldCPP、Oobabooga Web UI、LiteLLM
  - 雲端模型：OpenAI、Azure OpenAI、AWS Bedrock、Anthropic(Claude)、Google Gemini、HuggingFace、DeepSeek、Mistral AI、Perplexity AI、OpenRouter、Novita AI、Together AI、Fireworks AI、Groq、Cohere、AWS Bedrock、APIpie、OpenAI API相容模型
3. 音訊處理模型 (Speech Models)：語音轉文字與文字轉語音
  - 本地模型：Whisper (語音轉文字)、TTS 模型 (文字轉語音)
  - 雲端模型：OpenAI
4. 語音轉錄模型 (Transcription Models)：專門處理音訊檔案轉錄為文字，用於長音檔或會議記錄等場景
  - 本地模型：Whisper
  - 雲端模型：OpenAI
5. 向量資料庫 (Vector Databases)：支援高效儲存與檢索
  - 本地資料庫：LanceDB (內建)、Chroma、Milvus、Weaviate、QDrant
  - 雲端資料庫：Pinecone、Zilliz、AstraDB

### 安全性與帳號管理

1. 單用戶模式
  - 全權控制平台，僅限特定人員使用。
  - 可開啟密碼保護。

## 2. 多用戶模式

- 支援角色權限管理，分為：
  - 全域管理員：系統全權管理
  - 管理者：可管理工作區，部分系統設定除外
  - 一般用戶：僅限使用授權工作區

## 3. 資料隱私

- 內建 LanceDB，本地資料不外流，適用於資料安全需求高的企業。
- 支援匿名數據回報功能，可隨時關閉。

# 實務應用案例

## 1. 教育業

- 校務問答、學習成果分析

## 2. 製造業

- SOP 查詢、員工培訓、維修助手

## 3. 零售業

- 商品查詢 Chatbot，提高銷售與顧客滿意度

## 4. 人資管理

- 員工入職、請假制度、FAQ 智能問答