NARLabs 財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

AI TalenTech Labs
AI 人才科技實驗室

TAIWAN AI RAP × G-RapX

國家算力創新落地應用

G-RapX 主權 AI 天地通



從算力到能力，建立政府主權 AI 落地架構

① 可介接國網中心算力 → ② 主權 AI／地端 LLM → ③ G-RapX Agent 平台 → ④ 機關自主發展 Workflow 與 Skill









應用案例


會議紀錄


公文編修


談參稿編寫


審案自動化


智能客服

算力、模型、平台與人才能力形成完整閉環

G-RapX 系統全景圖

政府機關、G-RapX 與國網中心的分工

機關內網



政府機關

使用者與業務系統

- 公文、檔案、會議系統
- 業務系統資料庫

HTTPS／grx- 金鑰

回應



G-RapX

機關 AI 應用平台

- 網頁操作與 API Playground
- 多模態解析、Skill、Workflow
- 金鑰、額度與稽核
- SQLite 與檔案庫

國網中心

HTTPS 443／RAP Key

回應



國家高速網路中心

TAIWAN AI RAP

- RAP API
- 超級電腦 GPU 算力
- 大型語言模型
- 對話／多模態、Embedding、Rerank、語音

G-RapX 主權 AI 天地通：政府機關專用 Agent 解決方案



- 直連國家高速網路中心
- 支援工作流自動排程
- 支援 Skill Hub 技能庫
- 支援 MCP 外掛功能
- 地端 LLM 及 Agent

https://app.g-rapx.com

AI 能力層：模型路由、檔案解析、檢索

1 模型自動分類與路由

國網模型上架後自動歸類為對話／多模態、Embedding、Rerank、語音轉文字、語音合成，並送到對應端點。用 **GET /v1/models** 可取得清單與每個模型的類型、是否看圖、是否支援工具呼叫。

GET /v1/models

```
{
  "id": "xxx",
  "type": "對話／多模態 /
          Embedding / Rerank /
          語音轉文字 / 語音合成",
  "vision": true | false,
  "tools": true | false
}
```

模型數量會隨國網上下架變動，現場以 /v1/models 回傳為準。

2 檔案解析

多格式檔案解析

PDF → 文字頁 直接抽字

掃描頁 → 先做 OCR

圖表頁 → 才交給多模態模型

轉逐字稿 並轉為繁體

解析 → 抽音軌與影格 後分別處理

3 RAG 檢索

以 bge-m3 向量檢索，再用 BGE-Reranker 精排，最後交給對話模型依據段落作者並標註出處。檢索段數與生成模型都可調整。

使用者問題／解析後內容

→ **bge-m3** 向量檢索

→ **BGE-Reranker** 精排

→ **對話模型生成** 依據段落作者並標註出處

檢索段數（可調整）

生成模型（可調整）

4 模型適配

國網模型上下架、串流與推理模型，的格式差異，由 G-RapX 持續適配；介接端的程式不必跟著改。

G-RapX 持續適配

- 模型上下架適配
- 串流格式適配
- 推理模型格式差異適配

介接端的程式不必跟著改

國網模型池（動態上下架）

對話／多模態 · Embedding · Rerank · 語音轉文字 · 語音合成 · ...

應用層：機關同仁在網頁上用的功能

對話與知識庫

上傳公文、掃描檔、錄音，
建立專案知識庫，
回答標註出處。

上傳資料 | 知識庫 | 對話問答

公文 (PDF/DOC)　掃描檔 (JPG/PNG)　錄音 (MP3/WAV)

拖曳檔案至此或點擊上傳

請問本案的辦理依據？

依據「○○要點」第3點規定……

資料來源 1

○○要點.pdf
第3點

Workflow v2 工作流

步驟、條件分支、逐項處理、
人工簽核、批次與排程；
內建公文回覆、會議紀錄等範例。



內建範例

- 公文回覆
- 會議紀錄
- 申請案件處理
- 問卷彙整
- 資料整理
- .....

範本與技能

Word／Excel 範本自動套版；
技能庫收錄公文摘要、陳情回覆、
簽呈格式檢核等。

文件範本 | 技能庫

Word 範本自動套版　Excel 範本自動套版

○○機關公文範本

- 公文摘要
- 陳情回覆
- 簽呈格式檢核
- 其他技能 .....

內部問答機器人

綁定指定文件、
以通行碼保護的同仁問答助理。

機關內部問答助理

請輸入通行碼

••••••••••　登入

請問本機關差勤規定？

依本機關差勤管理要點規定……

資料來源

○○機關差勤管理要點.pdf
人事室 Q&A.docx

請輸入問題…

三步接入

G-RapX
政府專屬 AI 服務

1 取得 grx- 金鑰

向機關的 G-RapX 管理員申請。
金鑰**只顯示一次**，
請存進你們系統的機密設定。



金鑰**只顯示一次**
請妥善保存於系統的機密設定

2 設定 base_url

指向機關的 G-RapX：
https://<G-RapX 主機>/v1。
認證標頭為
Authorization: Bearer grx-…



base_url

https://<G-RapX 主機>/v1

Authorization

Bearer grx-....

3 用現有方式呼叫

OpenAI SDK 直接可用；
沒有 SDK 的語言用一般 HTTP POST 送 JSON 即可。

OpenAI SDK（範例・Python）

```python
from openai import OpenAI
client = OpenAI(
    base_url="https://<G-RapX 主機>/v1",
    api_key="grx-...",
)
response = client.chat.completions.create(
    model="...",
    messages=[{"role": "user", "content": "..."}]
)
```

HTTP POST（範例）

```
POST https://<G-RapX 主機>/v1/chat/completions
Authorization: Bearer grx-...
Content-Type: application/json
{
  "model": "...",
  "messages": [{"role": "user", "content": "..."}]
}
```

已經接過 OpenAI 或 Azure OpenAI 的廠商，改的就是**網址**與**金鑰**兩個設定值。

相同介面 • 低調整成本 • 快速接入

API Playground：試接台就是活文件

開放 × 安全 × 可實作
讓 AI 真正服務公共任務

1 網址

機關的 G-RapX 主機加上 /api，貼上 grx- 金鑰、按「連線驗證」即可使用。

https://您的機關 G-RapX 主機/api

2 六個分頁

對話、逐字稿、語音合成、Embedding、Rerank，以及「公文參考撰擬」情境。

對話 | 逐字稿 | 語音合成 | Embedding | Rerank | 公文參考撰擬

https://您的機關 G-RapX 主機/api

G-RapX API Playground

試接台就是活文件
探索能力、立即實作

API 設定

API 網址 https://您的機關 G-RapX 主機/api

grx-金鑰 ••••••••••••••••••••

連線驗證



對話 | 逐字稿 | 語音合成 | Embedding | Rerank | 公文參考撰擬

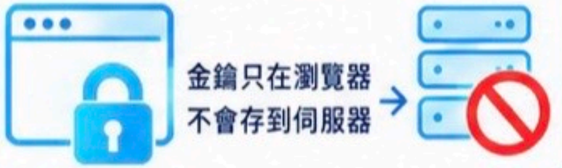
3 六種語言

curl、Python、JavaScript、C#、Java、PHP，程式碼依畫面上的欄位即時產生，複製即可執行。

curl | Python | JavaScript | C# | Java | PHP

4 金鑰不落地

金鑰只留在瀏覽器，不會被頁面存到伺服器。



5 現場示範：貼一份來文 → 摘要 → 複製 C# 程式碼

1 貼一份來文

來文內容
(貼上機關來文)

2 摘要

產生摘要
(由模型回傳)

3 複製 C# 程式碼

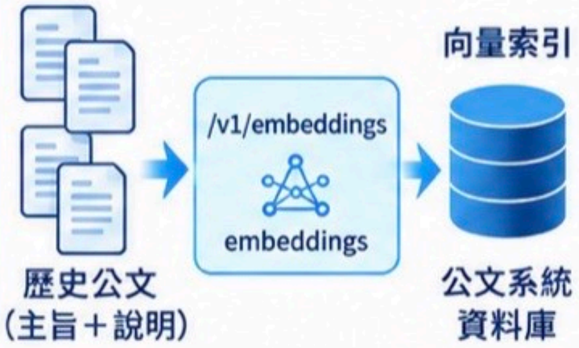
C#
1
2
3

複製程式碼
即可執行



情境一：從歷史公文找相近案例，撰擬新稿

1 建立公文庫索引



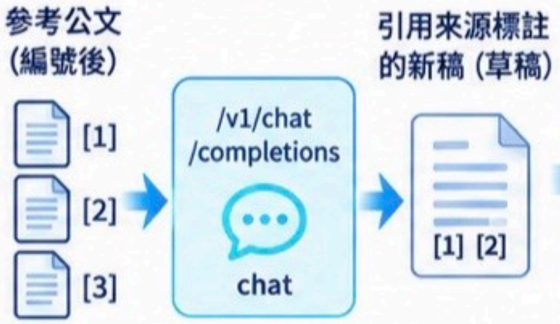
歸檔時把「主旨＋說明」送 /v1/embeddings，向量存進公文系統自己的資料庫。

2 查詢相近公文



承辦人用白話描述需求；向量比對取前 5 份，再用 /v1/rerank 留下最相關的 3 份。

3 組成新稿



把參考公文編號後交給 /v1/chat/completions，要求以 [1][2] 標註來源；不確定處留白。

4 人工核稿與產出



產出一律是草稿，須承辦人核稿；來源對照表由程式依檢索結果產生，不依賴模型自述。

歷史公文 → 向量索引　取前 5 份 → 重排 3 份　生成引用新稿　人工核准 才能定稿

情境二：來文摘要到回覆函稿

來文

PDF / DOCX / 其他格式

內建工作流的四個步驟

1 來文摘要

chat

2 研擬回覆要點

3 產出回覆稿 (JSON 欄位)

```
{
  "主旨":"..."
  "說明":"..."
  "辦理建議":"..."
}
```

4 套用公文回覆範本 產生 DOCX

DOCX

DOCX 範本

進階版本

來文 PDF 智慧解析

PDF

依公文規章檢核

主管核定

輸出 DOCX 與 PDF

DOCX PDF

在公文系統裡怎麼做

收文畫面

公文系統 — 收文

來文文號

來文主旨

來文內容

AI 摘要

呼叫

/v1/chat/completions

擬稿畫面

公文系統 — 擬稿

受文對象

主旨

擬稿內容

產生初稿

呼叫

/v1/chat/completions

提示詞可沿用工作流程的文字。

工作流與 API 說明

工作流本身在 G-RapX 網頁執行。

API

由外部系統直接觸發工作流的任務層 API 仍在規劃中。

規劃中（尚未提供）

目前 /v1 提供的是模型端點，工作流是網頁功能。

情境三：會議錄音到會議紀錄與決議追蹤

運用 AI 將會議錄音轉為結構化紀錄，並產出可追蹤的決議事項，提升會議管理效率

1 audio/transcriptions

MP3

會議 MP3 送
/v1/audio/transcriptions

長音檔切段與簡轉繁
由 G-RapX 處理。

繁

2 chat

錄音轉逐字稿

00:00:12 主席：
今天會議開始……

00:01:35 委員A：
建議先進行……

00:03:20 委員B：
補充說明……

00:05:10 主席：
以上作成以下決議……

3 整理成正式紀錄

逐字稿交給對話模型整理出
決議事項、列管事項與發言要點，
承辦人確認後套用會議紀錄範本。

會議紀錄（範例）

一、會議資訊

二、發言要點

三、決議事項
1.
2.
3.

4 決議追蹤表

從會議紀錄抽出決議與待辦，
一件一列，含負責人、期限與狀態，
輸出 Excel。

會議決議事項追蹤表（範例）

序號	決議/待辦事項	負責人	期限	狀態
1	推動系統優化			進行中
2	研議作業流程			待辦
3	提出評估報告			已完成
4	辦理教育訓練			進行中
5	追蹤辦理情形			待辦

從會議到行動
讓決議落地執行

內建範本：正式會議紀錄、公文回覆草稿、訪談整理報告、會議決議事項追蹤表等六種。

- 正式會議紀錄（範例）
- 公文回覆草稿（範例）
- 訪談整理報告（範例）
- 會議決議事項追蹤表（範例）
- 會議議程範本（範例）
- 會議簽到表範本（範例）

標準化範本
快速套用

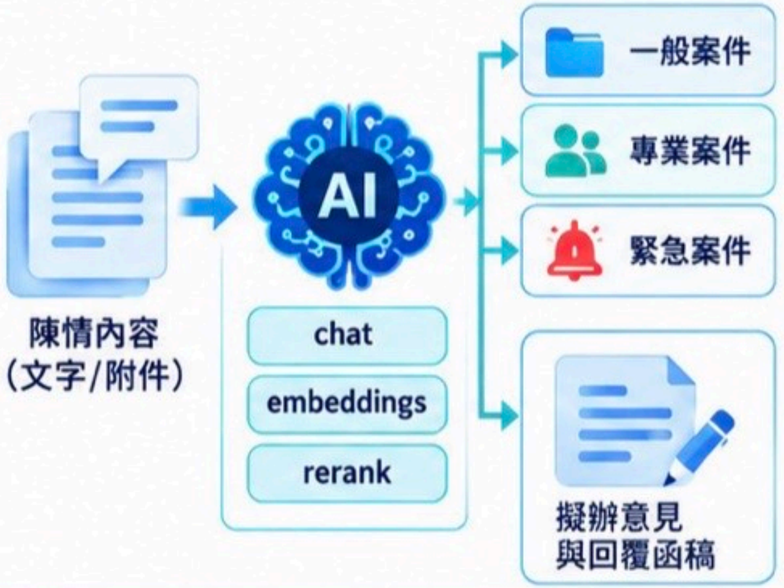
結構化輸出
提升可讀性

銜接後續作業
強化執行力

情境四：陳情分流、法規問答、簽呈檢核

01 陳情案件分流

解析陳情內容、判斷類別與緊急程度，分別產生擬辦意見；回覆函稿語氣同理、於法有據。



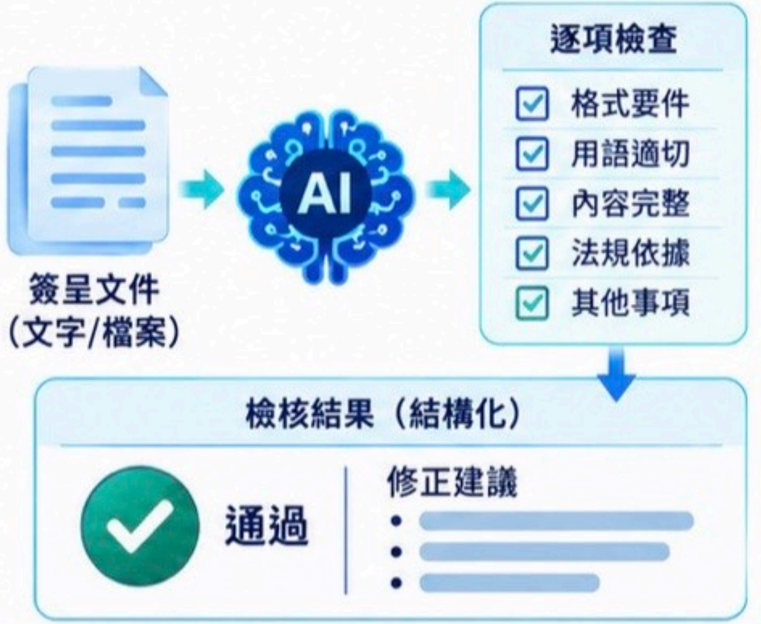
02 法規與作業指引問答

問題先檢索機關的法規與指引，合併證據後才作答，回答附依據。



03 簽呈格式檢核

逐項檢查簽的格式與用語，輸出通過與否及修正建議的結構化結果，適合陳核前自我檢查。

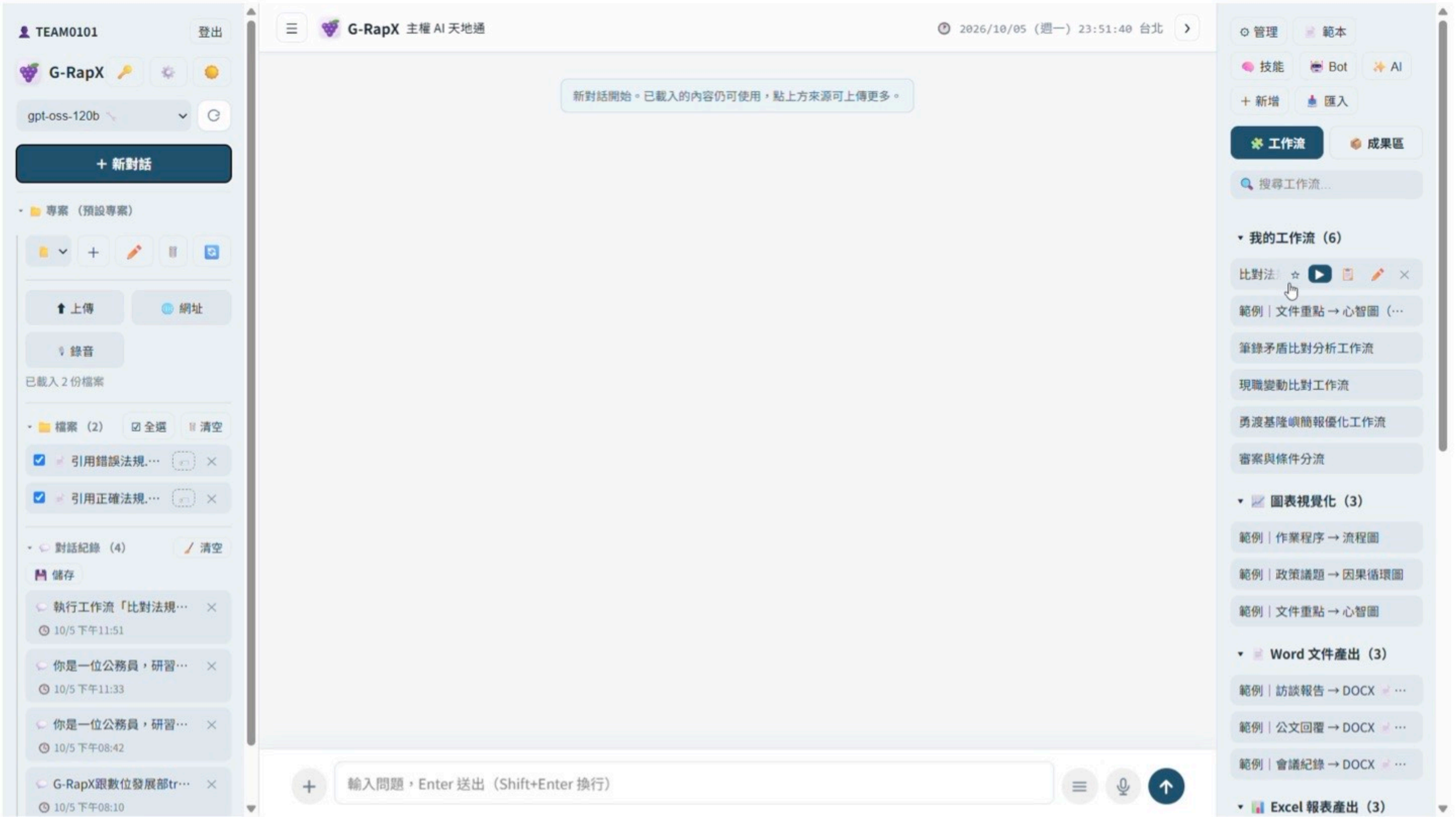


技能以 SKILL.md 開放格式撰寫

內容就是可重用的提示詞與檢核清單，可匯出後移植到你們的系統。



範例：機關使用 G-RapX 設計審案 Agent Skill 工作流



範例：公文系統呼叫 G-RapX API 直連國網中心 AI 算力



政府 AI Agent 導入加速計畫

AI Agent Accelerator（AAA）

八週實作陪跑，完成第一個可上線的 AI Agent



Powered by G-RapX

AI TalenTech Labs
AI 人才科技實驗室

NARLabs 財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

政府不缺 AI 工具，缺的是「導入最後一哩」

- 不知道從哪個流程開始
- 個人會用，組織卻沒有穩定流程
- 展示很多，真正上線與衡量很少



AAA 把 AI 體驗，轉化為可運作的組織能力

找對場景 ····· 做出 Agent ····· 共同實作 ····· 上線擴散





交付的不只是一個 Agent，而是一套可持續發展的方法

八週，交付一個可驗證的完整案例



1. AI 成熟度診斷
2. 導入診斷與建議報告
3. 一個工作流程的 AI Agent
4. 一個共同發展的 Skill
5. 測試、SOP 與下一階段 Roadmap

AAA 是導入方法，G-RapX 是落實方法的平台



下一步：選定第一個真實場景

1. 選出值得驗證的工作流程
2. 指定業務與技術對接窗口
3. 確認八週試點的目標與衡量方式



從一個成功案例開始，建立政府持續創造、治理與擴散 AI Agent 的能力

AI TalenTech Labs
AI 人才科技實驗室

NARLabs 財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing