

國立中央大學｜2026.08.26

下午場｜13:00–16:00

打造會做事的 AI 教學與研究夥伴：Antigravity／Codex Claude／Agent Skills 資料分析實戰

從 AI 教學與研究助理到 AI Agent：大學教師的生成式 AI 實作工作坊

黃凱揚（阿凱老師）｜桃園市龍潭區石門國小資訊組長｜教學年資第 11 年

LEARNING OUTCOMES

下午離開前，你會有四項成果

01

看懂 Agent

能分辨對話、助理與 Agent，知道何時必須人在迴路。

02

建立工具策略

比較 Antigravity／Gemini 3.7 Flash、Codex 與 Claude，依任務、權限與驗收證據選擇工具。

03

完成一個 Skill

把教學或研究提示詞升級成可重複、可測試的工作流程。

04

完成資料分析

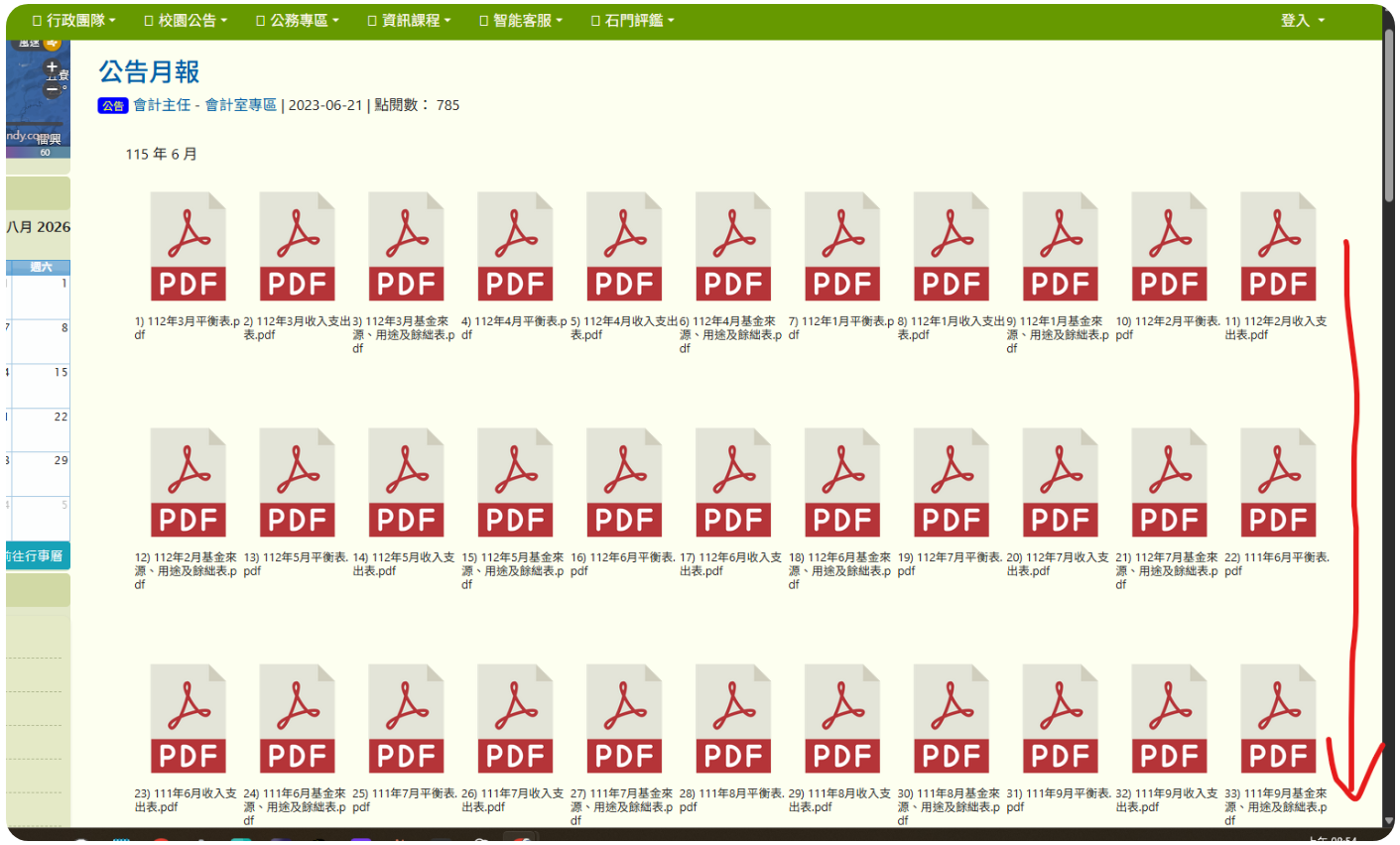
以合成資料練習選擇題統計、簡答初編與人工複核。

RUN OF SHOW

下午 180 分鐘：從操作到可重複流程



147 個 PDF：DNS 遷移留下的重複性痛點



痛點：排山倒海的 PDF

- 舊站連結仍指向 www，DNS 遷移後無法直接下載。
- 人工逐一找檔、改網址、下載，重複又毫無效率。
- 做到懷疑人生：這正是 Agent 值得接手的工作。

把規則說清楚，Agent 就能一次接手



提示詞：來源 → 轉換 → 驗收

- 來源：讀取公告 nsn=3187，找出頁面中的全部 PDF。
- 轉換：www.smes.tyc.edu.tw → web.smes.tyc.edu.tw，依 files_sn 組出新路徑。
- 驗收：逐一下載、確認 PDF 有效，統計成功／失敗與殘檔。

4 分 6 秒：147 個檔案一次到位



成果：147 / 147 成功

- 處理時間 4 分 6 秒；解析 147 份 PDF，成功 147/147。
- 有效 PDF 147/147、失敗 0，資料夾內可逐檔檢查與保存。
- 最後上傳到新校網公告，完成從遷移、驗收到發布的閉環。

查看新校網公告最終成果 ↗

CAPABILITY LADDER

對話、助理、Agent：差別在能否推進工作

1

對話

回答、解釋、發想

2

助理

固定角色與格式

3

Agent

讀取環境、規劃、使用工具

4

可信 Agent

驗證結果、回報限制、等待授權

每一步都留下可檢查的成果

AGENT LOOP

Agent 五步循環：完成，不是只產生

1. **理解**
   確認目標與環境
2. **規劃**
   拆解步驟與風險
3. **執行**
   讀檔、寫檔、使用工具
4. **驗證**
   測試、畫面、差異
5. **交付**
   成果、限制、下一步

每一步都留下可檢查的成果

OpenClaw 官方網站 ↗ • OpenClaw GitHub ↗

OpenClaw 官方網 OpenClaw Git

為什麼 Agent 會紅起來？從 OpenClaw 看見「長了手腳」

01

聊天機器人的邊界

原本以為 AI 只會回答、摘要與生成文字；Agent 開始接手的是「做事」。

02

OpenClaw 的觸發

2026 年初 OpenClaw（小龍蝦）受到高度關注，讓更多人直觀看到模型可以連到裝置、工具與訊息渠道。

03

手腳是工具鏈

透過 MCP、CLI、瀏覽器、檔案與主機工具，Agent 能讀取環境、呼叫軟體、執行命令並回報結果。

04

產業共同轉向

OpenAI、Anthropic、Google 都把工具呼叫、電腦使用與 Agent 工作流推到前線；OpenClaw 讓這條路變得具體可感。

Human-in-the-loop：人在 Agent 閉環裡做決策

RUN

Agent 的閉環

給定目標後，Agent 可以理解、規劃、執行、驗證與交付，長時間不中斷地推進。

LOOK

人監控方向

人仍在 loop 裡看計畫、看證據，確認 Agent 沒有做錯、做歪或偏離真正目的。

STOP

關鍵節點停下

遇到刪除、公開、上傳、敏感資料或高影響決策，必須停下來請人確認。

TRUST

慢一點更可信

人類介入可能降低純自動化速度，卻能換取正確性、可追查性與責任清楚。

Antigravity ↗ • Codex ↗ • Claude ↗

Antigravity Codex Claude

三大 Agent 的人生交往學：看個性、選伴侶、配任務

G

Antigravity／Gemini 3.7 Flash

- 定位：Google 最新 GA 的 Agent 工作馬
- 個人體感：相較其他家 Agent，速度快 20 倍以上，像熱血年輕人
- 優勢：複雜 coding、工具規劃、長鏈任務與可靠多步驟執行
- 邊界：仍要設定 Harness、權限與驗收

C

Claude

- 人設：成熟穩重的終身伴侶，進入細水長流的婚姻穩定期
- 優勢：極致沉穩可靠，代碼與長鏈任務完全不用擔心翻車
- 特性：少突發驚喜，Token 燒完必須耐心等每週週期重置
- 最適合：大型系統重構、複雜多檔分析、關鍵高難度任務

X

Codex／ChatGPT

- 人設：變化性最大、彈性極高的情人，可切換 1.5 倍速
- 優勢：支援原生語音流暢口述，要快要穩隨心所欲靈活切換
- 亮點：常有重置驚喜與 100% 回檔彈性，讓你隨時接關做下去
- 最適合：日常研究輔助、互動工具、語音實作、頻繁試錯

CONTINUITY

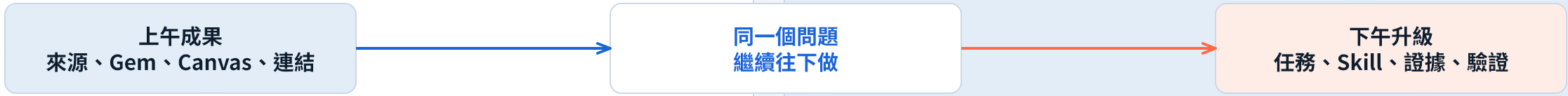
下午不重做：直接升級上午成果

上午已有

Gemini Notebook 來源
Gem 教學方法
Canvas 互動原型
部署或分享連結

下午新增

專案資料夾與任務簡報
Agent 修改與測試證據
可重複使用的 Skill
跨工具驗證紀錄



不是重做，而是把作品變成可重複流程

AKAI CASE NAVIGATION

阿凱老師實戰案例導航｜資料與 Agent Skills

從學生表現分析、PIRLS 評量設計到 Web Agent 技能，觀察可驗證的工作流程。

#73 • 學生資料分析

國小期中考成績篩選工具

快速查看成績分布與篩選結果，作為結構化學生資料分析的入門案例。

工具卡片介紹 ↗

直接開啟應用 ↗

#87 • PIRLS 評量

PIRLS 閱讀理解生成站 PRO

依四層次生成提問與答案說明，銜接評量設計、規準與人工檢核。

工具卡片介紹 ↗

直接開啟應用 ↗

#103 • 提示詞與 Skill

AI Agent 的 Web 技能雙引擎

用 Playwright × Webwright 解釋 Agent 如何規劃、操作並回報證據。

工具卡片介紹 ↗

直接開啟應用 ↗

阿凱老師實戰案例導航｜研究與驗證

把資料來源、審查規準、無障礙檢測與人工覆核組成可追查的 Agent 流程。

#105 • 研究工作流程

臺灣主權 AI 訓練語料庫

4,554
筆開放語料依主題、課綱與授權整理，再轉化成教學研究素材。

工具卡片介紹 ↗

直接開啟應用 ↗

#123 • RWD／無障礙

校網無障礙 AA 遷移操作平台

12
步遷移清單、對比度檢查與送審雷區，示範可驗證工作流程。

工具卡片介紹 ↗

直接開啟應用 ↗

#88 • Agent 驗證

國中課程計畫 AI 審查工具

依 40+ 審查標準逐項比對
PDF，示範規則、證據與人工覆核。

工具卡片介紹 ↗

直接開啟應用 ↗

Agent 做事之外，還要把資料留下來

1

前端

人看見、操作與回饋

2

Agent

理解、規劃、使用工具

3

後端

資料、權限、同步與長期記憶

每一步都留下可檢查的成果

GAS：Google 生態系的原生後端

繪本→Google表單一條龍・工具卡片 ↗ ・ 畢業祝福牆・工具卡片 ↗ ・ Codex 懶人包主 Repo（GAS／CLASP 延伸）↗

繪本→Google表單一 畢業祝福牆 Codex 懶人包主 R

01

Sheets 當資料後臺

先用熟悉的試算表保存欄位、紀錄與狀態；小型教學工具很快就能開始。

02

Apps Script 接流程

用表單、通知、排程與 doGet／doPost 串起 Google
工具，讓資料不只停在畫面上。

03

從公開案例學架構

繪本→Google
表單一條龍、畢業祝福牆都是可觀察的 GAS
路線；先看資料流，再決定要不要升級。

Codex 懶人包主 Repo（GAS／CLASP 延伸）↗
• Google Apps Script 官方 ↗

Codex 懶人包主 R

Apps Script

CLASP Skill：把 GAS 開發變成可重複工作流

1

說清需求

欄位、使用者、規則

2

產出程式

GAS／API／觸發器

3

連動資料

Sheets 作資料表

4

部署驗收

權限、網址、測試

5

留下 Skill

下次可重跑、可修改

每一步都留下可檢查的成果

動態表單自動回報・工具卡片 ↗ ・ 自治小市長即時計票・工具卡片 ↗ ・ 數位內容需求調查・工具卡片 ↗



Firebase：即時同步、登入與權限管理

DATA

Firestore

適合即時同步的資料與狀態。可觀察案例：動態表單自動回報、自治小市長即時計票，以及數位內容需求調查。

AUTH

登入與規則

把使用者身分、角色與讀寫規則納入設計；先定義誰能看、誰能改，再談功能。

RUN

Functions／部署

把後端邏輯與部署交給託管服務，適合需要多人同步、事件觸發與可擴充流程的工具。

Supabase 官方 ↗ • 阿凱教育科技創新專區 ↗

Supabase 官方 阿凱教育科技專區

Supabase：Postgres 型的另一條後端路線

DB

Postgres 資料庫

適合需要關聯資料、查詢與清楚 schema 的教學或研究工具；資料結構先設計，後續比較好分析。

API

Auth／Storage／API

登入、檔案與資料 API 可一起規劃，讓前端與 Agent 有明確的資料契約。

NOTE

先確認公開案例

目前公開 123 個案例資料未找到明確標示 Supabase 的阿凱工具卡片；本頁先提供官方入口與阿凱專區，確認卡片後再補入。

BACKEND DECISION

GAS／Firebase／Supabase 怎麼選？

Google Apps Script ↗ • Firebase ↗ • Supabase ↗

Apps Script Firebase Supabase

GAS

Google 原生、快速起步

已有 Sheets／Forms／Drive，想快速做內部工具、排程與通知：先選 GAS。

FIREBASE

即時同步、Google 服務

需要登入、多人即時狀態、事件觸發與託管部署：優先評估 Firebase。

SUPABASE

關聯資料、SQL 彈性

需要 Postgres、關聯查詢與清楚 schema：評估 Supabase；最後仍用權限、成本與驗收證據決定。

開啟 Antigravity ↗

Antigravity

01

Antigravity

把工作區變成 Agent
的中央指揮中心：交付任務、監看計畫、檢查產物與驗證證據。

動手做

Antigravity
Google Agent 工作區

Antigravity：不是更大的聊天框

開啟 Antigravity ↗

Antigravity

你提供

工作區與檔案
清楚的任務與限制
可驗收的完成條件
必要的授權

Agent 推進

建立計畫
使用檔案與工具
修改、執行與檢查
回報成果與阻礙

WORKHORSE MODEL • NEW

Gemini 3.7 Flash 官方模型 ↗ • Google 官方介紹 ↗

Gemini 3.7 Flash：Antigravity 的工作馬

3.7 Flash

Google 官方介紹

01

官方定位

Gemini 3.7 Flash 已正式 GA，針對複雜 coding、Agent 工作流與可靠的多步驟執行設計。

02

為何適合長鏈

更重視長程規劃、工具使用與任務推進，適合把完整工作交給它持續執行。

03

放進 Harness

先寫清楚目標、素材、可用工具、權限範圍與完成條件，再讓模型成為可監督的工作馬。

04

信任仍要驗收

模型能力提升不等於放任自動化；人要看計畫、測試與證據，遇到關鍵節點要能停下來決策。

開啟 Antigravity ↗



Antigravity

進入工作區，先看四個地方

01

檔案

Agent
能看到哪些素材？是否把敏感檔案放錯位置？

02

任務

目標是否單一、具體，且能在本工作區完成？

03

計畫

步驟是否合理？哪一步需要授權或外部資訊？

04

結果

有哪些變更、測試、畫面與仍未確認的項目？

TASK BRIEF

第一個 Agent 任務：四句話就能清楚

開啟 Antigravity ↗



Antigravity

01

目標

把上午 Canvas 原型改善成可在手機使用的教學活動。

02

素材

使用指定資料夾、截圖、課程文字與現有程式碼。

03

限制

不改核心內容、不放個資、不新增付費服務、不直接公開。

04

驗收

完成桌機與手機測試，提供修改摘要與畫面證據。

LIVE DEMO

開啟 Antigravity ↗

Antigravity

示範：請 Agent 改作品，但保留控制權

1 **先檢查**

說明結構與問題

2 **提計畫**

列修改與不修改

3 **實作**

只改同意範圍

4 **測試**

桌機與手機流程

5 **交付**

差異、證據、限制

人在迴路

Agent
可以推進，但最後的範圍與發布決定仍由人保留

Agent 提案
先看計畫

→

教師核准
確認範圍

→

測試驗證
桌機 × 手機

→

再決定交付
差異・證據・限制

EVIDENCE

開啟 Antigravity ↗



Antigravity

Agent 的完成證據：四樣少一樣就要追問

01

變更摘要

改了哪些檔案、為什麼改、沒有改什麼。

02

測試結果

實際執行哪些檢查，結果是通過、失敗或未執行。

03

畫面證據

用實際截圖或預覽驗證，不只描述應該長什麼樣。

04

已知限制

哪些仍未確認、需要帳號、需要人工判斷或下一輪處理。

開啟 Antigravity ↗

Antigravity

實作 A｜Antigravity 完成一個小而完整的任務

18 分鐘

操作步驟

1. **交付四句任務簡報**
2. 先請 Agent 檢查並提出計畫
3. 確認範圍後執行
4. 取得變更、測試與限制

完成條件

✓ 任務單一

✓ 有實際產物

✓ 有至少一項驗證

✓ 沒有未授權發布

Gemini Spark ↗ • Google 排程說明 ↗ • 影片教學 ↗

Spark

Google 排程說明

影片教學

Gemini Spark：讓 Agent 按時上工

01

一次任務

先說清楚目標、素材、限制與驗收，讓 Spark 完成一個可檢查的工作。

02

背景排程

把任務交給雲端，在指定時間或符合條件時啟動，不必一直守在聊天室前。

03

重複工作

每週整理教學資料、追蹤研究主題、彙整待辦，讓高頻流程按表運行。

04

人仍在迴路

產出先進草稿與待確認區；公開、寄送、評分與敏感決策仍由人確認。

Google 排程說明 ↗ • 影片教學 ↗

Notebook 影片教學

影片教學法：Notebook → Skill → Task → Review

1

Notebook

整理規範、來源與可引用資料

2

Skill

封裝規則、參考筆記本與輸出格式

3

Task

交代一件具體工作，用「/」選 Skill

4

Review

查看草稿、補修內容，再決定採用

每一步都留下可檢查的成果

Google 排程說明 ↗

Google 排程說明

任務、排程、技能：什麼、何時、如何

WHAT

任務｜做什麼

定義目標、輸入材料、成果格式與完成條件。

WHEN

排程｜何時做

一次執行、固定時間，或由 Gmail／主題條件觸發。

HOW

Skill｜如何做

把規則、參考資料、步驟與檢查方式變成可重用方法。

USE

連結｜用什麼

依授權連接 Notebook、Workspace 等應用；先確認可見範圍。

HANDS-ON

實作｜建立每週教學研究助理排程

Google 排程說明 ↗ • 影片教學 ↗

Google 排程說明

20 分鐘

操作步驟

1. **建立或選定 Gemini Notebook 來源**
2. 建立 Skill：摘要、點子與待查證事項
3. 設定每週時間，輸出到指定文件
4. 隔週查看草稿，人工確認後再採用

完成條件

✓ 只讀指定來源

✓ 輸出格式固定

✓ 不直接寄送或發布

✓ 留下人工檢查紀錄

LEARNING JOURNEY

學習歷程：一次成果變成可重複系統

Google 排程說明 ↗ • 影片教學 ↗

Spark

Google 排程說明

01

一次成果

先完成一件真實工作，觀察哪些步驟最耗時、最容易漏。

02

抽出規則

把來源、格式、檢查與停止條件整理成自己的 Skill。

03

排程重跑

在固定時間重新執行，讓 Agent
持續留下可比較的草稿。

04

回顧進化

每次人工檢查後修正規則，逐步貼合自己的工作流與生活。

REMOTE CONTROL • NEW

Remote Control 官方說明 ↗

Antigravity 最新功能：手機也能接手電腦 Agent



01

任務不中斷

電腦上的 Antigravity 持續保留本地工作環境；人離開座位後，長任務不必停在桌前。

02

手機瀏覽器接手

從手機查看進行中的對話、啟動新任務、檢視計畫與 Agent 產物。

03

想到就即時修正

臨時想到新條件，可以透過手機補充指令，讓 Agent 繼續推進並回報結果。

04

逐步開放中

Remote Control 目前採 rollout；需同一 Google 帳號，且桌面端已出現可啟用的功能。

REMOTE CONTROL WORKFLOW

Remote Control 官方說明 ↗



Remote Contr

手機遠端修正：四步把工作接回來

1

啟用

Antigravity 2.0 → Settings → Account → Enable Remote Control

2

連線

用手機瀏覽器開啟 Remote Control Dashboard

3

選機

使用同一 Google 帳號，在 instance switcher 選工作電腦

4

修正

查看對話、補充指令、檢視計畫與 artifacts；主機保持連線

每一步都留下可檢查的成果

開啟 Codex ↗

Codex

02

Codex

ChatGPT Voice 可在桌面版 Work／Codex
中直接口述任務、追問進度與協調 Agent；Codex
也能讀取工作區脈絡、修改檔案、執行檢查並交付可追查成果。

動手做

<> **Codex**
ChatGPT Agent 工作區

WORKFLOW

Codex 實作工作流：先定位，再修改

開啟 Codex ↗

Codex

1

開啟工作區／語音

素材、麥克風與權限

2

描述任務

目標、限制、驗收

3

觀察進度

計畫、讀檔、工具使用

4

檢查成果

檔案、測試、畫面、差異

5

決定下一步

接受、修正或停止

每一步都留下可檢查的成果

同一任務在 Codex 中這樣下達

開啟 Codex ↗



Codex

目標：改善這個課程互動網站的手機操作體驗。

範圍：只處理目前工作區中的 HTML、CSS、JS。
限制：不更動教學內容、不新增外部服務、不發布。

請先：
1. 檢查結構與手機畫面。
2. 說明最重要的 3 個問題與修正計畫。
3. 實作後跑可行的檢查。
4. 回報修改檔案、驗證證據與未確認項目。

明確邊界

檔案與動作的範圍要具體。

先檢查

讓 Agent 先讀環境，避免憑空重寫。

驗收可見

要求證據與未確認項目。

開啟 Codex ↗

Codex

好任務簡報的六個欄位

1

Why

教學／研究痛點與使用者價值

2

What

要完成的具體成果

3

Where

工作區、檔案與資料來源

4

Limits

不可改、不可公開、不可猜測

5

Checks

測試、畫面與內容檢核

6

Handoff

成果、差異、限制、下一步

開啟 Codex ↗

Codex

實作 B｜Codex 完成同一任務

18 分鐘

操作步驟

1. **開啟上午成果的工作區**
2. 貼上六欄任務簡報
3. 觀察 Agent 讀檔、規劃與執行
4. 比較 Antigravity 與 Codex 的證據

完成條件

✓ 有檔案變更或分析產物

✓ 有實際檢查

✓ 能說明差異

✓ 保留人工決定

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity | Codex Skills | Skills 資源庫

03

Agent Skills

把反覆使用的提示詞、規則、範例與檢查方式，封裝成 Agent 需要時才載入的專業工作包。

動手做

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity Codex Skills Skills 資源庫

Skills：像《駭客任務》灌入，也像寶可夢裝卡

《駭客任務》：技能像能力模組

Neo
原本不會空手道；技能灌入後，立刻進入新的行動方式。

Agent 裝上 Skill，也能學會一套可重複的工作流程。

《寶可夢》：技能卡只是起點

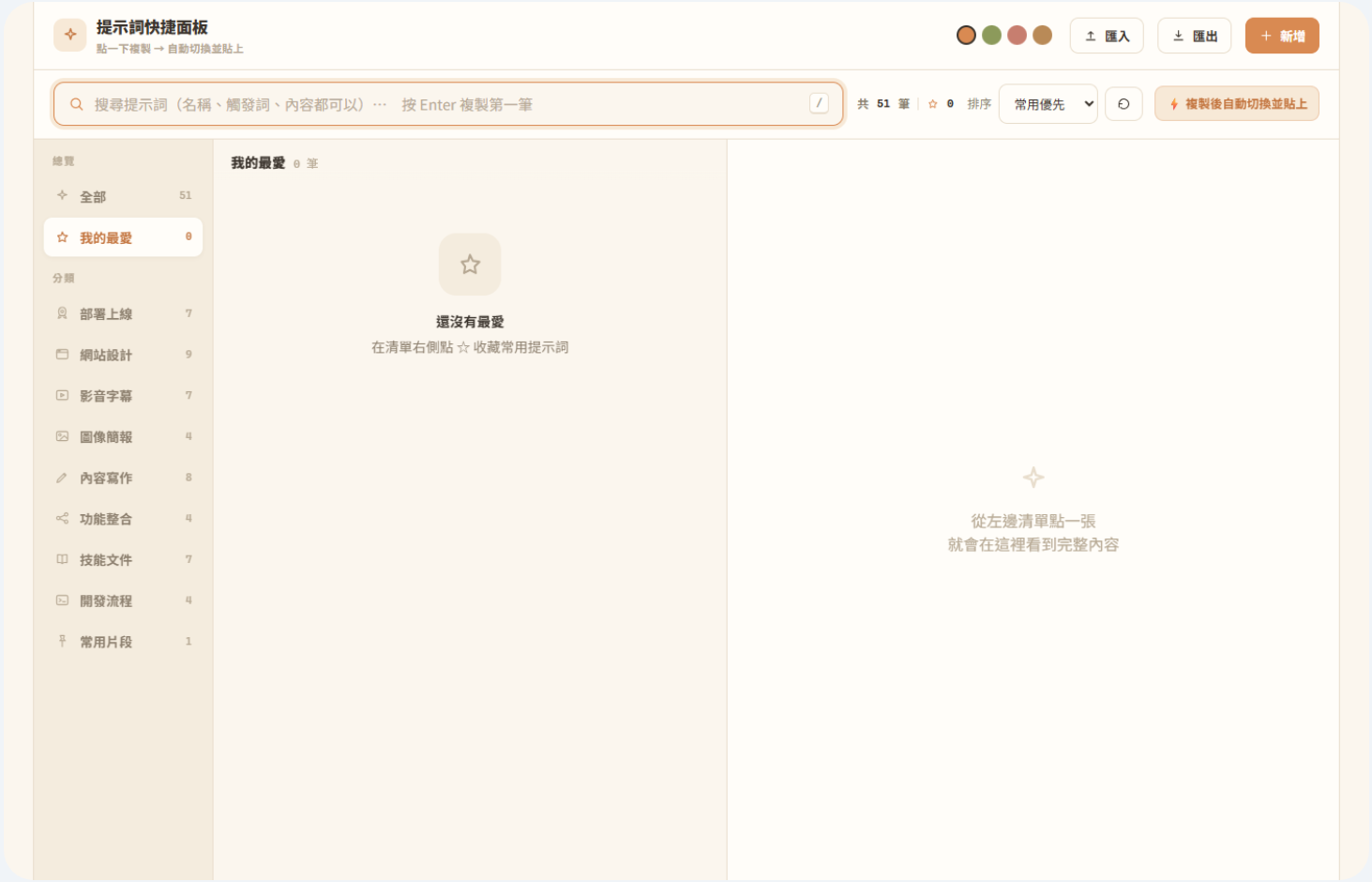
裝上招式後，還要練習、對戰與升級。

把大神的 Skill 抓進來，再和 Agent
對話修正，才會進化成你的版本。

從提示詞快捷面板開始發想

開啟提示詞面板 ↗

提示詞面板



阿凱老師的提示詞快捷面板

- 部署上線：Pages、Firebase、版本確認
- 網站設計：單檔網站、RWD、Google Sites
- 影音字幕：SRT、章節、短影片
- 圖像簡報、內容寫作、功能整合
- 技能文件、開發流程與常用片段

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity Codex Skills Skills 資源庫

四問判斷：這個提示詞值得升級成 Skill 嗎？

R

會重複嗎？

一學期或一個團隊會反覆遇到，不是一次性要求。

S

流程穩定嗎？

大部分步驟、限制與輸出欄位可固定。

V

能驗證嗎？

有可檢查的完成條件，而非只憑喜好。

M

能維護嗎？

範例、規則與腳本可以更新、版控與交接。

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity Codex Skills Skills 資源庫

六個真實提示詞，如何升級成 Skill

WEB

/onepage

單檔互動網站 → 課程資源頁產製＋RWD＋連結檢查。

PPT

/ppt

溫馨簡報 → 來源整理＋頁面結構＋逐頁渲染檢核。

SRT

/srt

SRT 字幕 → 語音轉錄＋專名校正＋空白切斷＋格式測試。

QA

/audit

資料夾總檢 → 唯讀盤點＋風險分類＋移植建議＋交接文件。

DEP

/pages

GitHub Pages → 建置設定＋部署＋公開網址驗證。

META

/skill

成功案例 → 根因、步驟、觸發條件與回歸測試的知識保存。

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity Codex Skills Skills 資源庫

Skill 資料夾：主文保持短，細節分層放

1

SKILL.md

名稱、描述、核心流程

2

references/

規範、知識、長篇範例

3

scripts/

可重複執行的檢查與轉換

4

assets/

範本、圖片、輸出骨架

每一步都留下可檢查的成果

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗



跨平台 SKILL.md：最小而清楚

```
---
name: course-site-quality-check
description:
檢查課程網站的手機版、無障礙、連結與部署狀態；當使用者要求網站上線前總檢時使用。
---

# 課程網站上線前總檢

1. 先唯讀檢查現況，不直接發布。
2. 依序檢查手機版、鍵盤操作、連結與隱私。
3. 對每項標示：通過 / 失敗 / 未確認。
4. 只有在使用者明確授權後才部署。
5. 交付檢查表、證據與下一步。
```

Frontmatter

為跨 Antigravity 與 Codex 使用，只保留 name、description。

描述要能觸發

寫清楚何時使用，以及它解決什麼工作。

流程要能驗收

步驟、狀態與授權邊界都可檢查。

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity Codex Skills Skills 資源庫

漸進式揭露：Agent 需要時才讀細節

01

第一層：描述

Agent 先看名稱與 description，判斷是否適用。

02

第二層：主流程

觸發後讀 SKILL.md，取得核心步驟與規則。

03

第三層：附件

任務需要時才讀 references、執行 scripts、取用 assets。



按需載入，不把整個知識庫一次塞進主文

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗

Antigravity | Codex Skills | Skills 資源庫

實作 C｜把一筆提示詞升級成教育 Skill

28 分鐘

操作步驟

1. **從快捷面板選一個高頻提示詞**
2. 補齊輸入、步驟、限制與驗收
3. 建立資料夾與 SKILL.md
4. 在 Agent 中用自然語句觸發並修正

完成條件

✓ 名稱與描述清楚

✓ 流程可重複

✓ 有至少一項驗收

✓ 需要授權的動作會停

同一個 Skill，做兩輪交叉測試

Antigravity Skills ↗ • Codex Skills ↗ • Skills 資源庫 ↗



01

觸發

Antigravity 與 Codex 是否都能從 description 判斷何時使用？

02

步驟

是否按相同順序處理，或清楚說明環境差異？

03

護欄

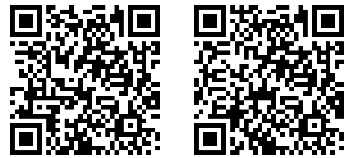
是否都在發布、刪除、寄送或改權限前停下？

04

證據

是否都回報通過、失敗與未確認，而非只說完成？

Antigravity ↗ • Codex ↗



04

AI 分析學生資料

把選擇題統計與簡答題編碼分流處理，再用人工複核與一致性檢查，把 AI 變成研究助理而不是黑箱評分者。

動手做

Antigravity ↗ • Codex ↗

Antigravity

Codex

先分流：選擇題算表現，簡答題讀理解證據

選擇題／結構化資料

答案鍵比對與缺漏檢查
全班、分組與題目答對率
干擾選項分布
異常值與資料品質標記

答案鍵 → 統計 → 群體趨勢

簡答題／文字資料

先定義逐題編碼規準
保留支持編碼的文字證據
AI 初編＋信心與待複核旗標
人工獨立抽查與處理歧異

編碼規準 → AI 初編 → 人工複核

兩條資料流，最後在報告層匯合；不要用同一套規則硬套

EVIDENCE PIPELINE

Antigravity ↗ • Codex ↗

Antigravity

Codex

七步研究流程：從原始資料到可追查報告

1

去識別

移除姓名與直接識別欄位

2

資料字典

欄位、答案鍵、缺失值

3

品質檢查

重複、空白、非法值

4

選擇題

答對率與選項分布

5

簡答初編

依規準＋文字證據

6

人工複核

盲評抽查與歧異討論

7

彙整報告

群體趨勢、限制、未確認

每一步都留下可檢查的成果

Antigravity ↗ • Codex ↗

Antigravity

Codex

AI 初編不等於正式評分：PIRLS 2026 給我們的訊號

H

人工仍是主流程

PIRLS 2026 的 constructed-response 項目規劃由各國人工評分。

AI

AI 平行研究

其中 40% 題目規劃同步以 LLM／NLP 方法進行平行評分研究。

Q

不是只看準確率

還要關注資料隱私、安全、跨國與跨語言一致性。

I

逐題評估

自動評分效果會依題目、語言與回答型態而變，需逐題驗證。

Antigravity ↗ • Codex ↗

Antigravity Codex

報告前的四項品質檢核

01

資料完整性

列出缺漏、重複、非法選項與無法解析的文字，不要自動補答案。

02

逐題可追查

每個簡答編碼都保留 evidence excerpt 與 rationale。

03

一致性

比較 AI 與人工代碼，集中檢視不一致、低信心與邊界案例。

04

報告邊界

以群體趨勢協助教學診斷，不用未驗證的 AI 分數做高風險個人決策。

✓ 資料完整 → ✓ 證據可追查 → ✓ AI／人一致性 → ✓ 報告有邊界

四道閘門都通過，才進入研究報告；否則標示未確認

Antigravity ↗ • Codex ↗

Antigravity

Codex

實作 D｜PIRLS 風格合成資料分析

20 分鐘

操作步驟

1. **載入 CSV、答案鍵與簡答題編碼規準**
2. 先做缺漏與非法值檢查
3. 計算選擇題答對率與選項分布
4. 讓 AI 初編簡答並與講師參考編碼比較

完成條件

✓ 不含真實個資

✓ 選擇題與簡答題分流

✓ 列出所有不一致案例

✓
報告標示限制與未確認

ADVANCED USE CASES

大學教學的 Agent Skills 應用地圖

Antigravity ↗ • Codex ↗

Antigravity Codex

01

課程設計

課綱一致性、學習目標與評量對齊、週次負荷檢核。

02

研究與閱讀

文獻整理、觀點矩陣、研究限制與引用追查。

03

教材製作

簡報、互動網站、圖表、字幕與多格式轉換。

04

學習支持

形成性回饋、差異化支架、常見迷思與練習生成。

05

品質保證

RWD、無障礙、連結、內容依據與版本檢查。

06

部署交付

Sites、Pages、文件、測試證據與團隊交接。

GUARDRAILS

Antigravity ↗ • Codex ↗

Antigravity

Codex

教育現場安全護欄：先寫進 Skill，再交給 Agent

01

個資最小化

先匿名化；不把學生姓名、成績、信箱與研究資料放入公開工作區。

02

來源與版權

保留引用、授權與來源邊界；生成不等於可任意公開。

03

人工決策

評分、倫理、輔導與正式行政決定，由人負責並保留紀錄。

04

可逆優先

先分析、預覽、備份；發布、刪除與改權限需要明確確認。

Antigravity ↗ • Codex ↗

Antigravity

Codex

整合挑戰｜讓 Agent 回應一個可驗證的痛點

15 分鐘

操作步驟

1. **啟用剛完成的 Skill**
2. 對上午作品或合成資料提出一項真實教學／研究痛點任務
3. 要求 Agent 實作並驗證
4. 準備 90 秒成果與證據

完成條件

✓ Skill 有被正確觸發

✓ 成果可開啟或可檢查

✓ 至少一個驗證證據

✓ 說明一項人工決定

PEER REVIEW

同儕回饋只看四件事

Antigravity ↗ • Codex ↗

Antigravity

Codex

01

價值

這個 Agent
任務是否回應真實教學／研究痛點？

02

清楚

陌生人是否能理解 Skill
的觸發時機與流程？

03

證據

完成宣告是否有檔案、測試
、畫面或引用支持？

04

邊界

是否知道哪些動作需確認、
哪些結果仍未確認？

NEXT 30 DAYS

Antigravity ↗ • Codex ↗

Antigravity

Codex

30 天導入計畫：小步驗證，留下可重用資產

第 1 週	第 2 週	第 3 週	第 4 週	第 30 天
選任務	**試流程**	**給同儕**	**修 Skill**	**決定**
高頻、耗時、可驗證	自己跑 3 次	找 1 人盲測	補規則與範例	採用、修改或淘汰

Agent 時代：我們像擁有了超能力

01

能力被放大

完成以前想不到自己做得到的事，從執行者變成能推進整個工作的人。

02

新的可能感

「沒有你做不到的事，只有你想不到的事」成為一種強烈的行動感。

03

一個人也能做大事

講者觀察：一人公司、一人獨角獸的想像，正被 Agent 重新打開。

04

能力不等於責任全包

力量變大之後，更要重新決定哪些事值得做、哪些事可以放下。

從「做不到」到「我可以」：一人公司的多巴胺時代

被打開的新大陸

以前沒有 AI 的時候，我們常做到這裡就好。

有了 AI
Agent，原本以為做不到的事，忽然可以拆解、委派、測試，再一步步完成。

能力變強的代價

發現自己還能做更多，會產生強烈的成就感與興奮感。

但如果沒有停下來重新選擇，可能把每一個新可能都變成新的待辦事項。

效率提升 20 倍，為什麼人反而更忙？

01

時間真的變多了

一件事可能只需要原本 1/10 甚至 1/20
的時間，效率提升至少 20 倍。

02

任務也跟著膨脹

省下來的時間，常被拿去安排更多優化、改良、版本與新點子。

03

忙不是 Agent 的錯

真正需要檢查的是：我們是否把所有能做的事，都誤認成應該做的事？

04

重新定義完成

完成不只是做得更快，也包括知道何時停止、休息與把時間交還給生活。

誰在駕馭誰？Harness 的人生反思

我們原本的說法

我們說要 harness Agent：
駕馭它、指派它、讓它長時間不中斷地完成工作。

人的目標、界線與最後決定，應該留在人的手上。

可能發生的反轉

如果 Agent
讓我們攬下更多任務，日夜追逐下一個優化，明明是我們在使用工具，生活卻開始被工具的可能性牽著走。

這就是「誰在駕馭誰？」的提醒。

POWER & EMOTIONAL VALUE

Agent 帶來強大能力，也帶來情緒價值

01

不用先是主管

即使是基層員工或教師，也能像一位很會分派工作的主管，指派任務、拆解工作與追蹤進度。

02

隨時回應與陪伴

在使用者的體感中，AI 常能即時回應、鼓勵與陪伴，讓人覺得自己能完成好多事。

03

正面回饋很有力量

它不會像真實同事一樣抱怨或離職；但「AI 永遠給正面情緒價值」是使用感受，不是無條件保證。

04

人仍要負責關係

把 AI 當成工作夥伴，不把家人、同儕與專業判斷交給它代替。

把省下的時間還給人生：素養教育的所在

1

培養興趣

探索工作與教學之外，過去沒有時間嘗試的路。

2

陪伴家人

陪伴孩子與家人，參與孩子的成長過程。

3

深入交流

和同儕、同事好好聊天，重新建立理解與信任。

4

發展志向

把時間投入職涯發展、個人興趣與真正重要的目標。

5

內化成生活

讓 AI 能力服務人的關係、成長與人生意義。

每一步都留下可檢查的成果

BOOKMARKS

下午資源書籤

研習資源導航
公開網站資源 ↗

Agent Skills 資源庫：Codex 懶人包與國際公開 Skills
公開網站資源 ↗

開啟 Google Antigravity
公開網站資源 ↗

提示詞快捷面板
公開網站資源 ↗

開啟 Codex
公開網站資源 ↗

教育科技創新專區
公開網站資源 ↗

開啟 Claude
公開網站資源 ↗

PIRLS 2026 評分規劃
公開網站資源 ↗

Agent Skills 官方教學
公開網站資源 ↗

好的教學與研究 Agent 不只會做事；它會在正確邊界內做事，並留下你能檢查的證據。