

低階應用  
但大眾關注焦點



# 運用生成式AI輔助撰寫研究計畫及論文： 心態、技巧、學倫

孫以瀚

2026.05.11

陽明交通大學 新竹光復校區

個人意見，仍請查看遵守期刊或國科會規範

**Helped Greatly by AIs**

本演講即是運用 AI 的範例，展現人的主導

面對新事物，找類比

針對研究人員(已具備核心技能)

主要針對實驗科學

PPT 可提供(細節慢慢看)，歡迎分享出去

# Outline (Mine)

- 面對AI的心態 (General)
- 如何使用AI (General)
  - 避免資訊外洩
  - 如何提問、使用付費版、選擇模型、設定角色
- 如何使用AI輔助撰寫研究計畫、論文
- 使用AI輔助寫作的學倫問題

# Take Home Messages

- 要努力用AI (The single most important powerful research tool)
- 不要怕學生用，要帶著學生一起學習用AI
- 充分發揮AI功能，研究的每一步都用AI幫忙
- 驗證AI生成的內容
- 不用在意論文內容是否AI寫的，只在意內容是否可靠，判斷是否有價值
- 遵守現有規範

## I. 面對AI的心態 (General)

# Generative (生成式) AI 的特性與問題

- 以大數據資料庫透過大型語料模型(LLM，如Transformer)訓練而成
  - 根據上文來預測下一個最可能出現的字詞，基於統計規律的再創造(主流、平均化、典型化)
  - 幻覺 (Hallucinations): 內容可合理但不一定真實
  - Many AIs: different training database and algorithm => 各有千秋
    - 公開可取得的網路內容、授權取得的資料、人類產生或標註的資料、使用者互動資料
  - Potential info leakage (chat info goes into learning database)
- (Gemini + ChatGPT)

# Hallucination by ChatGPT

## ChatGPT: references on TNBC and EGFR research :

Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011;121(7):2750-2767. doi:10.1172/JCI45014

Ueno NT, Zhang D, Targeting EGFR in Triple Negative Breast Cancer. *J Cancer Sci Ther* 2014; 6(6): 236-243. doi: 10.4172/1948-5956.1000262

 wrong

Baselga J, Albanell J. Mechanism of action of anti-HER2 monoclonal antibodies. *Ann Oncol.* 2001;12 Suppl 1:S35-S41. doi:10.1023/a:1011199018579

Yardley DA, Kaufman PA, Brufsky A, et al. Treatment patterns and clinical outcomes for patients with de novo versus recurrent HER2-positive metastatic breast cancer. *Breast Cancer Res Treat.* 2021;186(1):107-117. doi:10.1007/s10549-021-06105-5

Hynes NE, MacDonald G. ErbB receptors and signaling pathways in cancer. *Curr Opin Cell Biol.* 2009;21(2):177-184. doi:10.1016/j.ceb.2008.12.010

Moulder SL, Yakes FM, Muthuswamy SK, Bianco R, Simpson JF, Arteaga CL. Epidermal growth factor receptor (HER1) tyrosine kinase inhibitor ZD1839 (Iressa) inhibits HER2/neu (erbB2)-overexpressing breast cancer cells in vitro and in vivo. *Cancer Res.* 2001;61(24):8887-8895.

Siddiqui S, Chopra R. EGFR: A Potential Target for the Treatment of Triple-Negative Breast Cancer. *Chemotherapy.* 2017;62(3):177-184. doi:10.1159/000452984

 totally false

# 破壞性創新 Disruptive Innovations

能源：動物、水車、風車、蒸汽引擎、電力

交通：輪子、火車、飛機、自駕車

資訊傳播：紙、印刷、網路、電報、電話、手機、智慧手機

生物醫學：抗生素、recombinant DNA, PCR, CRISPR/Cas9

★ Google Search、Google Map、ChatGPT

- Free and open for all. => 有利平等，提高baseline
- 使用門檻低(自然語言)
- 付費版(功能升級) => affordable
- 阻擋限制都無效

# Competition: Human + AI >> Human

- 會使用新工具者佔優勢 (AI提升效率、品質)
- 用得好才贏 (ChatGPT)
- 人與人的競爭，避免被善用AI的人取代 (人+AI >> 人)
- AI 是新的selection pressure (Gemini) => 必然有淘汰 (殘酷的現實)
- AI 拉高了「地板」，人決定了「天花板」(Gemini)

# GenAI 是受人指揮的工具，人是主導

- AI 是**工具、助理**（類比）
- 要**善用**（擅長、做好事），也要了解工具的侷限
- 助理/工具不完全可靠，會犯錯，**需要檢驗**
- 老闆要為助理/工具犯錯**負責**
- AI 是**被動**的，有問才有答，只動口不動手，生成內容須透過人（篩選、判斷）才能呈現
- **人是主導（掌控工具）!**

## 要充分發揮AI 的功能

- 善待問者，如撞鐘，叩之以小者則小鳴，叩之以大者則大鳴 《禮記 學記》
- AI 的能力強大，要充分利用，如果只用來修改文法，是錯失大好機會
- 不要只把AI 當工具/助理，要當成是見多識廣且無私的資深學者，凡事都找AI 討論
- 從構想最初期就讓AI 參與

# 使用工具導致核心能力進階

核心能力 流失 => 進階

## 書寫與文字(Gemini)

- 蘇格拉底（透過柏拉圖的記載）曾強烈反對「書寫」。他認為如果人們依賴文字記錄，就會停止鍛鍊記憶力，最終將導致「靈魂的健忘」，而且文字是被動的，無法像真人辯論那樣產生智慧。（柏拉圖（Plato）對話錄《斐德羅篇》（Phaedrus））
- 我們將記憶「外包」給書籍，反而釋放了大腦去進行更高階的思考

## 計算機 (Claude)

- 複雜計算能力下降了，但學生能處理更複雜的應用問題
- 數學教育轉向概念理解和問題建模

## 維基百科/Google vs. 博聞強記 (Gemini)

- 單純知識記憶提升為
  - 資訊查核與辨偽
  - 知識連結

# 使用AI是最重要的核心技能(老師、學生)

用同樣的工具 => 不同的水準 (珠算 12級 → 10段)

- 不要擔心學生使用，要鼓勵/要求學生使用。
- 學生用AI做作業，可以有不同程度的產出。同學之間要比較、競爭，沒有標準答案，只有更好的答案。
- 作業範例: 如果諾貝爾獎要頒給NO作為signaling molecule，該頒給哪些科學家? 他們的得獎文章是哪些? 為何重要? 超越其他人的理由是什麼?
- 作業、考試沒有標準答案，老師負擔增加 (用AI幫忙改作業)
- 必然有人偷懶跟不上時代，會被淘汰
- 老師: 接受 not everyone can make it (每個人要為自己負責)
  - 只能盡力告訴學生要學AI

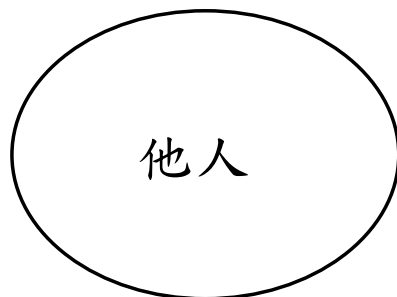
## 系統性的隱憂

- 錯誤/假內容的擴散
- 驗證的惰性 (當 AI 的準確率超過人類專家) => 更依賴 AI
- AI 放大個人差異，促進優勝劣汰，加大社會階層落差
- 評量失真 (教育、學術圈)

# 不同思考角度



- 要用、要善用
- 人為主導
- 避免被淘汰
- 遵守當下規範
- 人要負責



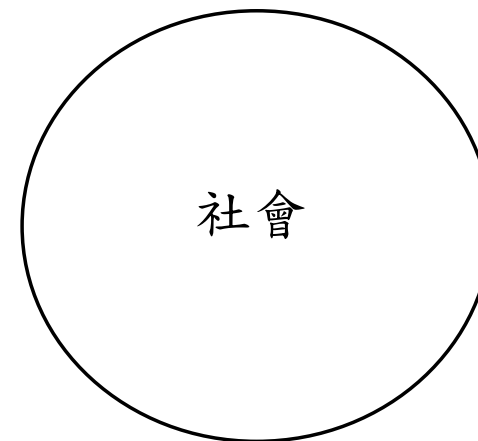
學生

- 鼓勵用、要求用、帶著一起用

- 偷懶就會被淘汰

其他研究者

- 只管內容是否可靠
- 不用管是否偷懶



- 即使擔憂，也無法阻擋，只能適應
  - 網路、手機
- AI developers 在設法改善

## AI輔助閱讀文獻，但人仍需深讀原典

科學文獻的摘要或文學名著的故事架構類似骨，原典的細節是肉(精華)

- *Romeo and Juliet*: 男孩女孩相戀，雙方家庭世仇，兩人殉情

- AI 將文獻摘要是「收斂」的，目的是縮減資訊

- 閱讀原文是「發散」的，可能因聯想而激發靈感

(Gemini)

- AI善於摘要，抓出重點，建立全貌與架構

- 精華藏在細節裡：研究方法、研究數據、詮釋與結論的不確定性

- 透過 journal club 學習深讀

- 找篇文章，讓學生及AI分別評論，帶領學生討論分析AI的評論（不但深讀論文，也向AI學習）

# Summary I: 面對AI的心態

核心心態：人為主導，工具為輔

- Human + AI (協作、互動) > Human
- 人要為結果負責

競爭本質：不是人與 AI 的對決，是人與人的競爭

- 工具決定優勢，用得好才贏
- AI 拉高了能力的基本水準(地板)，但卓越程度(天花板)仍由人來決定

風險控管：警惕幻覺與驗證義務

能力轉型：核心能力提升

(Gemini)

## II. 如何使用 AI (General)

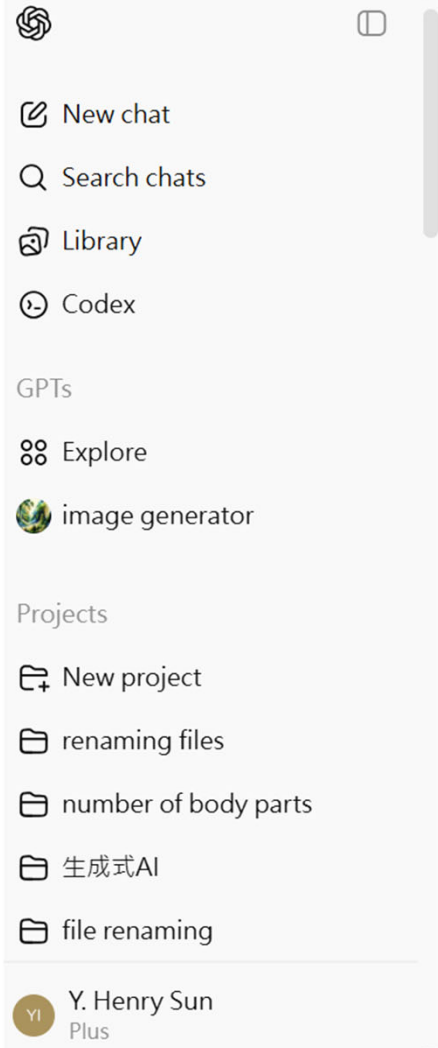
## 強烈建議使用付費版本 (不要輸在起跑點)

- ChatGPT (GPT-5.5)
  - Gemini (3.1 Pro)
  - Claude (Opus 4.6)
  - Grok (4.3)
  - Perplexity (GPT-5.5)
  - DeepSeek (V4 Pro)
- 付費(\$20 = NT650/月)
- 免費版
- 物超所值! 絕對建議!
  - Use >1 paid AIs.

不斷快速進步中!


不斷推出新model

# Keep chats out of training (ChatGPT)



The sidebar navigation menu includes the following items:

- New chat
- Search chats
- Library
- Codex
- GPTs
- Explore
- image generator
- Projects
  - New project
  - renaming files
  - number of body parts
  - 生成式AI
  - file renaming
- Y. Henry Sun Plus



ChatGPT 5 Thinking 5.2

Share

## How to keep our chats out of training

- Turn off training for your account: ChatGPT → Settings → Data Controls → Improve the model for everyone → Off. New chats won't be used to train models. [OpenAI Help Ce... +1](#)
- (Optional) Turn off Chat history in the same menu. When history is off, new conversations also aren't used for training and won't appear in your sidebar. [OpenAI Help Ce... +1](#)

## What OpenAI says about your data

- Consumer ChatGPT (Free/Plus/Pro): You can opt out of training via Data Controls; you can export or delete your data at any time. [OpenAI Help Ce... +1](#)
- Business/Enterprise/EDU & API: By default, OpenAI does not use your inputs/outputs to train. You own your data; enterprise plans add retention controls. [OpenAI Help Ce... +2](#)
- Privacy Policy basics: It covers what is collected and your choices, and it

Ask anything

+ Thinking 5.2



ChatGPT can make mistakes. Check important info.

# Keep chats out of training (ChatGPT)

Also Gemini, Claude, Perplexity, Co-Pilot

The image shows a sequence of three overlapping screenshots from the ChatGPT interface, illustrating how to access the 'Data controls' settings. The first screenshot on the left shows the main navigation menu with 'Settings' highlighted by a red arrow. The middle screenshot shows the 'Settings' menu with 'Data controls' selected, also indicated by a red arrow. The final screenshot on the right shows the 'Data controls' page, where the 'Improve the model for everyone' toggle is set to 'Off' and highlighted with a red arrow. Other options on the page include 'Remote browser data' (On), 'Shared links' (Manage), 'Archived chats' (Manage), 'Archive all chats' (Archive all), 'Delete all chats' (Delete all), and 'Export data' (Export).

GPTs

- Explore
- image generator
- yihenrysun@gmail.com
- Upgrade plan
- Personalization
- Settings
- Help
- Log out

Y. Henry Sun  
Plus

Settings

- General
- Notifications
- Personalization
- Apps & Connectors
- Schedules
- Data controls
- Security
- Parental controls
- Account

Data controls

- Improve the model for everyone: Off
- Remote browser data: On
- Shared links: Manage
- Archived chats: Manage
- Archive all chats: Archive all
- Delete all chats: Delete all
- Export data: Export

# Temporary chat

- Open AI logo
- New chat
- Search chats
- Library
- Codex
- GPTs
  - Explore
  - image generator
- Projects
  - New project
  - renaming files
  - renaming files
  - number of body parts
- Y. Henry Sun Plus

ChatGPT 5 Thinking ▾



ChatGPT 5 Thinking ▾



## Temporary Chat

This chat won't appear in history, use or update ChatGPT's memory, or be used to train our models. For safety purposes, we may keep a copy of this chat for up to 30 days.

Ask anything

+ ⌚ Thinking ▾



# How reliable is zero-retention and not-use-for-training?

Corporate promise is reliable (business reputation)

## But still has risks

- Software bugs
- Human access for safety/abuse/legal issues
- Hackers

Anthropic 全新旗艦 AI 模型意外流出，  
網路安全風險太高急尋防禦者測試  
[TechNews 編輯台](#) 2026.3.27

敏感機密資料不要上傳

- 與產學合作廠商的機密數據、合作者的未發表數據
- 涉及病人個資的臨床數據
- 研究的 idea, data?

# NotebookLM by Google

The screenshot displays the NotebookLM interface. On the left is the 'Sources' panel, which includes a search bar for finding new sources online, a 'Deep Research' feature, and instructions on how to add sources. The central area is a chat window titled '對話' (Conversation), which contains a red annotation: '建立新的筆記本(自訂、具體、可靠資訊範圍) => 沒有幻覺' (Create a new notebook (custom, specific, reliable information range) => no hallucinations). Below the chat is a 'Studio' panel with various tools like 'Voice...', 'Mind Map', 'Report', etc. At the bottom, there is a disclaimer: 'NotebookLM 提供的資訊未必正確，請查證回覆內容。' (Information provided by NotebookLM may not be accurate, please verify the content of the response.)

來源

對話

工作室

+ 新增來源

試用 Deep Research 取得深度報告和新聞來源！

在網路上搜尋新來源

新增來源即可開始使用

上傳來源

上傳來源即可開始使用

0 個來源

建立新的筆記本(自訂、具體、可靠資訊範圍) => 沒有幻覺

NotebookLM 提供的資訊未必正確，請查證回覆內容。

# NotebookLM by Google

Mammary Gland Evolution and MPSUs

數據分析 共用 設定 PRO

來源

對話 重新整理

→ Castano Leon 2010 Bilateral Familial ...

Info source  
快速驗證 來源導覽

reported refer to the relationship between polythelia and  
454 CASE AND RESEARCH LETTERS  
congenital renal and urinary tract malformations, such as  
adult autosomal dominant polycystic disease, unilateral  
renal agenesis, ectopic kidney, cystic renal dysplasia,  
and congenital dysplasia of the pelviureteric  
junction.<sup>4,6,8-10</sup>

The case of our patient is notable for the presence of  
bilateral familial supernumerary nipples at a distance  
from the typical location and apparently with no  
associated

→ systemic disease or congenital malformation.  
Nevertheless, when polythelia is detected, a detailed  
family history should be taken, a careful physical

The sources provide extensive data regarding the prevalence of  
supernumerary mammae (polythelia and polymastia) in humans, but note  
that the occurrence rates vary significantly depending on the study,  
ethnicity, and sex 1 ...

## Incidence of Supernumerary Mammae (Polythelia/Polymastia)

Supernumerary nipples (polythelia) or accessory breast tissue (polymastia)  
are relatively common developmental abnormalities, resulting from the  
failure of the embryonic mammary ridge (or "milk line") to fully regress 2  
...

The reported incidence of these condit in humans generally ranges  
from **0.2% to 6%** of the population, depending on the cohort studied:

開始輸入... 151 個來源

Define polythelia  
and polymastia.

What is the  
"one-half rule"?

Which mammals show  
nipple number asymme

# Agentic AI

- **Agentic AI: execution** in multistep task (human not in control at every step)

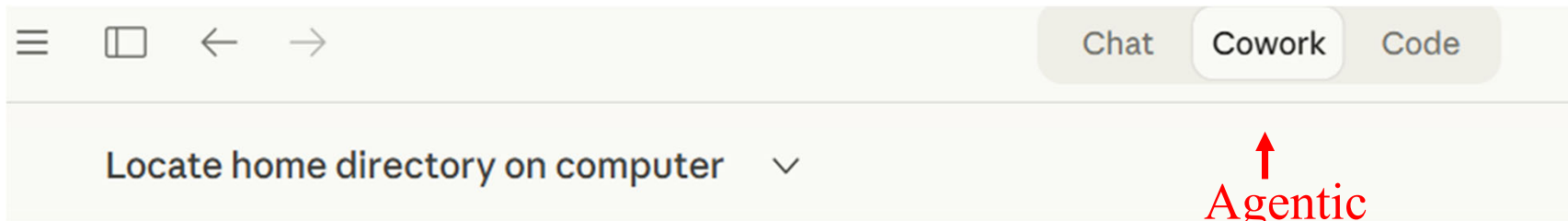


李弘毅: 解剖小龍蝦 — 以 OpenClaw 為例介紹 AI Agent 的運作原理  
<https://www.youtube.com/watch?v=2rcJdFuNbZQ>

Claude Code 比較安全

# Claude Pro

Download the Claude app => run on computer, not on mobile phone



↑  
Agentic  
(executions)

Cowork 採取權限導向設計——你選擇哪些資料夾和連接器 Claude 可以存取，它在隔離的 VM 中運行，網路存取透過白名單控制。在執行重大動作前，Claude 會展示計劃並等待你的批准，而刪除檔案一律需要你明確許可。  
(Claude)

# 使用AI的技巧

- **提問** (問對問題、試著用不同方式提問、問AI不怕丟臉，可以重複問)
  - 非**學**無以致疑 非**問**無以廣識 (清·劉開《問說》)
- 角色扮演
- 對AI提供的回答要做判斷，調整問題，**追問**
- **問多個AI，比較、互相批評、辯論**
- **AI幻想，不易重複發生 (重複問、問不同AI)**
- **要求資訊來源要附連結，方便驗證**

# Context 比提問更重要

## 提問 Prompts

- Chain-of-Thoughts (think step-by-step)
- Now embedded in generative AIs

## 提供 Contexts (情境/脈絡)

- 餵入相關文檔、對話歷史、參考範例、

## 關鍵技術

- 長上下文窗口 (Long Context Window)
- 檢索增強生成 (RAG; Retrieval-Augmentation-Generation)
  - 讓 AI 在回答問題之前，先去查閱參考資料
  - 解決了傳統 AI 模型最大的兩個痛點：幻覺、知識過時

# Many New AI tools for research

- SciSpace
- Jenni AI
- PaperPal
- AvidNote
- ResearchRabbit
- NotebookLM
- , , ,

## Summary II. 如何使用AI

- 用付費版本
- 保護敏感資訊 (別完全相信AI)
- 問、追問
- 比較多個AI、AI 辯論
- 自設可靠資訊範圍 (domain-specific knowledge): NotebookLM
- 提供 Context

### III. 如何使用AI輔助撰寫研究計畫、論文

# Acceptable and un-acceptable human/AI contributions

Simple prompt. **Copy, paste and submit.**

- Please generate a research grant proposal on cancer biology.



Prompt with explicit and clear instructions. **Copy, paste and submit.**

- Based on my recent publications and CV, please follow the same line of thoughts to write a one-year research proposal with novelty and significance, following these guidelines.



Prompt with explicit and clear instructions. **Discuss with AI multiple rounds to revise until satisfaction.** Submit. **Human inputs and decisions**



類比：大教授的研究計畫由博後撰寫，在教授的研究脈絡下，與教授來回討論，教授審閱後以教授的名義提出申請。

# Ask AI to help on manuscript

1. Please evaluate the manuscript for **significance and novelty**. Please also suggest how to modify it.
2. What are the **main points** or contributions of this work? How to make these points stronger?
3. Is the **organization** appropriate? How can I improve?
4. What are the **weaknesses**? How to fix?
5. Are there essential **controls** or data missing? What more experiments are needed?
6. I am planning to submit it to the **journal** XXX. Is this appropriate? Any other journals can be considered?
7. Please help to revise the manuscript.
8. Please suggest **potential reviewers** that may be more friendly.
9. Any important **references** that I missed?

# How to ask AI to help? useful prompts (ChatGPT)

You are a grants editor. Summarize the last 5 years of {topic}: (i) problem, (ii) mechanisms/tech, (iii) **resolvable gaps in 3–5y**, (iv) 5 fundable angles aligned to {FOA}, **each with a one-sentence hypothesis and measurable endpoint**. Return bullets with citation placeholders.”

Act as a **hostile reviewer**. List **top 10 criticisms** by category (Significance, Innovation, Approach, Investigators, Environment). For each, **propose fixes** (experiment, control, analysis, prelim data).

Rewrite these aims to (i) **one-sentence objective**, (ii) **2–3 aims each with hypothesis, rationale, innovation, approach, go/no-go, outcomes**, (iii) a **convergence statement**. Max 1 page.

收斂、結語

Audit this Approach for blinding, randomization, inclusion/exclusion, sample-size justification, replication, key biological variables, authentication, data sharing. Output a table of gaps + suggested text.

Convert Abstract to **lay level (grade 8)** without changing claims; keep 150–200 words.

# 研究計畫常有的缺點

## 研究主題與重要性（核心思想問題）(modified by AIs)

- 文獻回顧的廣度或深度不足
- 缺乏可被驗證或推翻的核心科學假設
- 失焦: 目標過於分散與不連貫，彼此之間邏輯聯繫薄弱
- 科學與臨床意義(重要性)闡述不足
- 研究優勢不明，未考慮國際競爭對手

## 研究設計與策略(方法論問題)

- 未說明在諸多可能的方法或策略中如何選擇
- 缺少適當的對照實驗
- 實驗設計不足以獲得明確答案
- 統計方法不恰當
- 缺乏統計功效與樣本量依據

## 研究可行性與管理（執行能力問題）

- 時程與資源評估不切實際
- 缺乏風險評估與備選方案

AI can help at every step!

# 範例：如何要求嚴厲批評文獻

王介立醫師 **FB** 2026.2.9 FB (文長，僅摘要，請看原文，順便點讚追蹤) **批判文獻是否可靠 (自我檢驗)**

批判性評估 (四大面向；每點都用「問題 → 文內證據 → 風險/影響 → 我需要看到什麼才放心」)

## (1) 假說與理論邏輯

- 問題：假說是否「事後補寫」？是否存在撒網式測量、再回頭挑顯著結果？
- 檢查：是否預先註冊主要終點？是否明確界定 primary vs secondary？是否大量生物標記/多終點？
- 你必須輸出：「最可能的合理機轉」與「最可疑的敘事跳躍」各 1-3 點 (各自附文內依據)。

## (2) 受試者與外推性

- 問題：受試者是否與宣稱對象一致？(健康人 VS 病患；一般人 VS 運動員；訓練程度、性別比例、年齡)

## (3) 研究設計與安慰劑

- 問題：隨機化與分派隱藏是否充分？

## (4) 測量工具與實務/臨床意義

- 問題：測量工具的信度/變異度如何？(test-retest reliability、CV、最小可偵測差異)

## 統計與資料完整性

- 樣本數與 power：是否事前估算？是否明顯 underpowered？
- 多重比較：有沒有校正？若沒有，哪些結果最像假陽性？

## 紅旗清單 (Red Flags, 越短越尖銳越好)

- 以條列列出所有你認為會讓可信度下降的點 (設計/統計/敘事/COI/外推)。

## 反向解釋

- 請提出 2-4 個「不需要補充品也能解釋結果」的替代解釋 (例如：期待效應、訓練/飲食控制差、學習效應、測試日狀態差、回歸平均等)，並指出文內有哪些線索支持/反對。

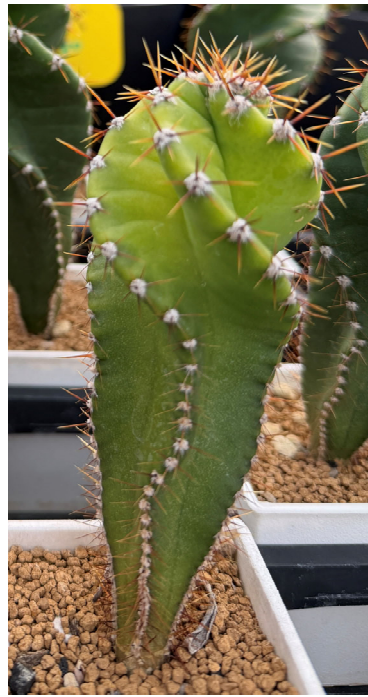
## Asking AI for help **at an early stage**

In a biological system (development or disease), **how do I experimentally demonstrate** the incomplete penetrance of phenotype is due to threshold effect? Please give some examples.

I am trying to find an experimental system where I can analyze data to validate my ideas on the threshold effect. I would like to explore the single cell RNAseq data available from published articles. Please **give me several examples** where the cells exist in two distinct population (e.g WT vs. disease state; differentiation state A to B; physiological state A to B), which are distinguished by a clear ON/OFF expression of a molecular marker gene. These examples should be backed by scRNAseq data. Please provide the reference paper and its DOI link.

# Questions from Nature

Q: How to explain?



# AI 可以如何幫忙

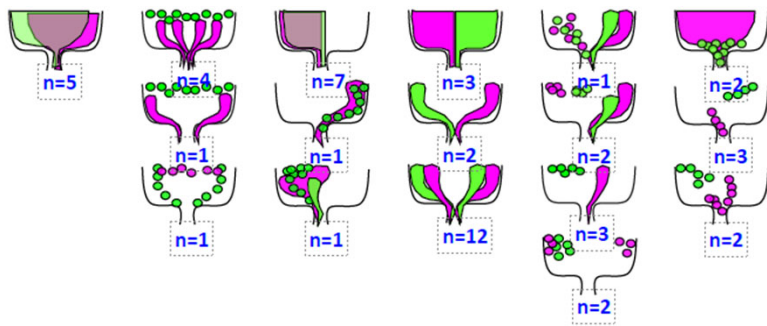
- 整理文獻，找出知識缺口，評估重要性、創新性
- 釐清思緒、整理架構、抓出重點
- 彌補缺失、完善套路 “routine” (common research strategy, experiments need to do, required controls, sample size, questions reviewers often ask)
- 改善表達
- 幫助聯想 (跨領域)
  
- Steady path, but not exciting/innovative.
- Follow AI (routines) => 80分 (but **everyone** gets 80分) (地板提升)
- To get to 90分 (ahead of others), need **human innovations** (天花板由人決定).

類似的觀點 (簡立峰):

<https://www.youtube.com/watch?v=XHTfeCk0GwQ>

# Data analysis

## A. Summary of Twin-spot MARCM results



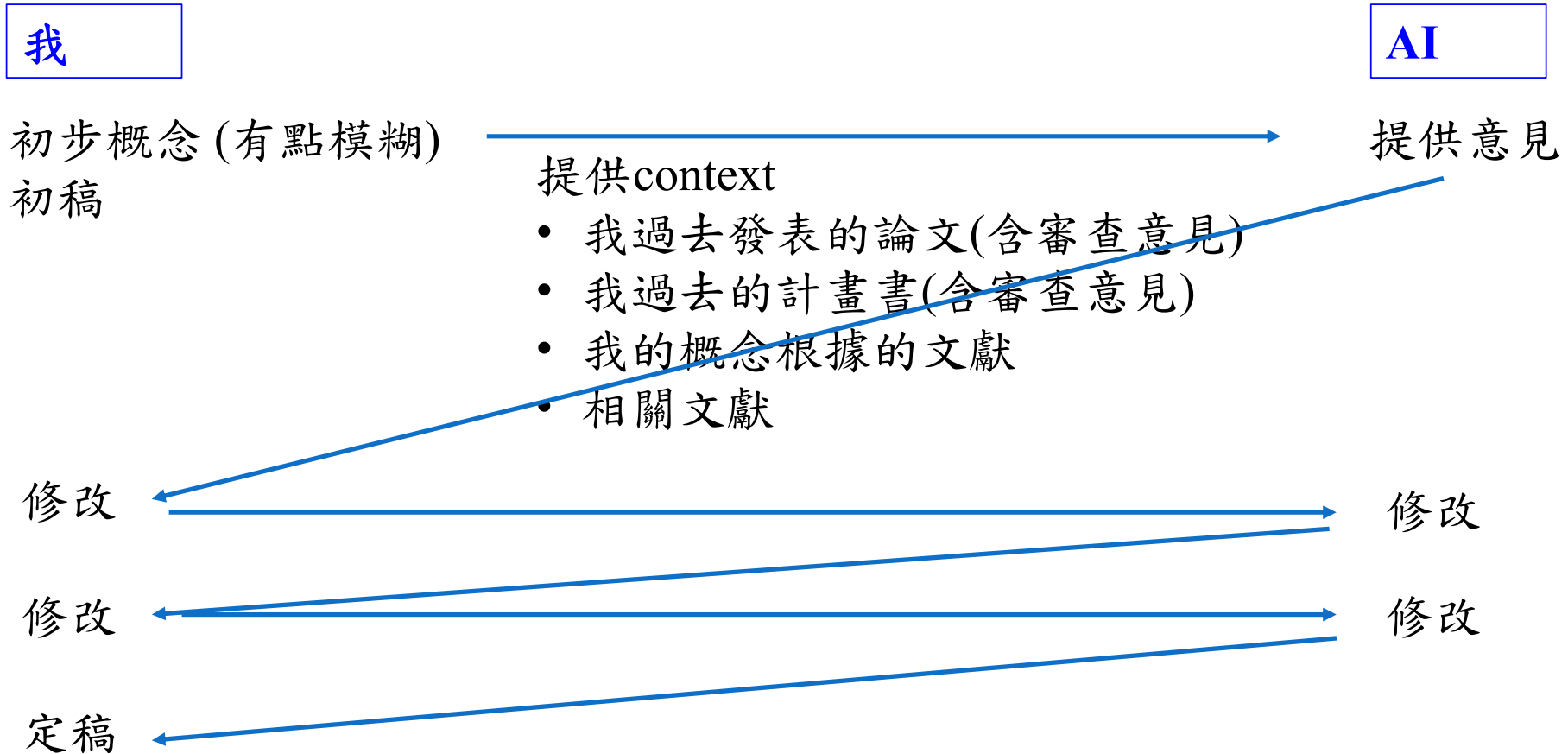
Q: this is a summary of the twin-spot MARCM results from our experiment. We induced the clones at different developmental times and examined the morphology (SG or WG), spatial distribution and number of cells. The twin-spot clones are marked by green and red. Their spatial distribution, morphology and numbers are depicted. The number for each class is indicated. **What conclusions can you make?**

非常見數據型態，刻意不做很多說明

Pattern	Observation (Green / Red)	Conclusion
<b>Symmetric SG</b>	SG / SG (e.g., Col 1)	Progenitor produced only Surface Glia.
<b>Symmetric WG</b>	WG / WG (e.g., Col 4)	Progenitor produced only Wrapping Glia.
<b>Mixed (Asymmetric)</b>	SG / WG (e.g., Col 2)	<b>Key Finding:</b> Single division produced distinct glial types.
<b>Stem Cell Mode</b>	Large Clone / Single Cell	Progenitor underwent self-renewal vs. differentiation.

**Excellent!**

# Reciprocal interactions between AI and human



絕非單一指令可以完成

# AI 的角色超越工具或助理

AI 輔助撰寫計畫或論文，角色已經超越工具或助理，更像是與資深同事或合作者(來自不同領域)的互動

- 腦力激盪、來回辯證、盲點偵測、釐清架構、抓出重點、提供意見、改善表達，甚至提出新的思路

但

- AI 只是對我的提問做被動的回應，不具研究意圖與責任能力
- 最終是由我做取捨判斷，由我定稿，由我承擔全責
- 如果是人，有足夠貢獻，應列為共同作者
- 但AI無法作為共同作者，亦無私心，只能揭露其貢獻並致謝

“作者感謝大型語言模型（ChatGPT-4）在研究構思階段擔任了\*\*批判性對話對象（Critical Conversational Agent）\*\*的角色。透過與 AI 的遞迴式討論，作者得以釐清論述架構並發現邏輯盲點。然而，所有最終的科學判斷、論點採納及文字定稿，均由作者本人全權負責。”(ChatGPT)

# Disclosure of AI usage

## 使用 ChatGPT 必須註明

- 註明於研究方法(**工具**; ~本抗體委託XX公司製備)或致謝(**助理**; ~感謝XX 協助)

### *Elsevier: Generative AI policies for journals*

‘Declaration of Generative AI and AI-assisted technologies in the writing process’ (above References section)

*“During the preparation of this work the author(s) used [NAME TOOL / SERVICE] in order to [REASON]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.”*

### **My suggestion:**

The **conception and writing** of this grant/paper is **assisted** by AI (Gemini 3 Pro and ChatGPT 5.2 Thinking). *The author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.* **Our complete discussion history with AIs demonstrating my commanding role is available upon request.**

# Some journals require listing the full prompts

## *Science Journals (AAAS)*

*“The full prompt used in the production of the work, as well as the AI tool and its version, should be disclosed.”*

Thorp and Vinson (2023) Change to policy on the use of generative AI and large language models

<https://www.science.org/content/blog-post/change-policy-use-generative-ai-and-large-language-models>

*“For transparency and reproducibility, the full prompt used to steer GPT-4.1 is provided in Appendix A.”*

## *Journal of Medical Internet Research (JMIR)*

*“keep the complete original prompts/responses/transcripts on file and submit all ChatGPT conversations used in the preparation of the manuscript as a Multimedia Appendix”*

Leung et al., (2023) Best Practices for Using AI Tools as an Author, Peer Reviewer, or Editor. *J Med Internet Res* 25:e51584

## 使用生成式AI應遵守的規範

- 生成式AI不可作為作者: 工具缺乏人格，無法對作品承擔法律責任
- 使用生成式AI要註明 (Methods or Acknowledgment)
- 生成式AI產生的圖像或影片不可使用 (包含示意圖、摘要圖)
  - 除非專門討論生成式AI的文章，須經編輯許可
  - 分析數據的統計圖表(e.g. 曲線圖、柱狀圖)需要驗證確保正確
- 不可用作審查: 違反審查保密協定
- 敏感資訊不應上傳給生成式AI

### 非生成式 AI (e.g. AlphaFold)

- 視為方法/工具，應詳細說明使用設定條件 (可重現性 Reproducibility)

# 使用生成式AI的規範

- 科技部 人工智慧科研發展指引
- 中央研究院 使用生成式AI之參考指引
- 國立台灣大學 針對生成式AI工具之教學因應措施
- 國立成功大學 生成式AI於教學研究的學術誠信指引
- 國立陽明交通大學 生成式AI用於學術研究之參考指引
- 國立清華大學 生成式人工智慧倫理聲明
- 臺灣學術倫理教育學會 人工智慧技術對學術倫理的影響及因應建議
  
- Committee on Publication Ethics (COPE) Council. (2023, February 13). *COPE position - Authorship and AI*. <https://doi.org/10.24318/cCVRZBms>
- *Science journals' guidelines for AI usage*
- Springer Nature: <https://www.springernature.com/gp/policies/editorial-policies>
- Elsevier: The use of generative AI and AI-assisted technologies in writing for Elsevier <https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier>
- Wiley: Using AI tools in your research [https://www.wiley.com/en-gb/publish/article/ai-guidelines/?\\_gl=1\\*14gb8gk\\*\\_gcl\\_au\\*MzMzMyNzQ2MDk4LjE3NzQ5NDA2OTE](https://www.wiley.com/en-gb/publish/article/ai-guidelines/?_gl=1*14gb8gk*_gcl_au*MzMzMyNzQ2MDk4LjE3NzQ5NDA2OTE)

## Summary 3 如何使用AI輔助撰寫研究計畫

- 理清思路、去蕪存青、找出重點、整理架構
- 來回討論
- **AI as a critical reviewer**
- 不同AI 提供意見的多樣性
- 人要對AI提供的內容正確性、架構、邏輯、取捨做判斷
- 應遵守之規範
- 註記AI的使用
- 保存所有與AI互動的紀錄 (以證明自己的實質創作性貢獻)

## IV. 使用AI輔助寫作的學倫問題

- 我用AI，別人怎麼看我？
- 別人用AI，我怎麼看待？

# 我們在擔心什麼？

## 如果研究者遵守規範

- 為AI產生的內容確保其準確性、真實性與完整性
- 為AI產生的內容負責
- 沒有機密敏感資訊外洩
- 避免不當研究行為（如捏造、篡改或抄襲）
- 揭露AI之使用
- 那使用AI還有什麼問題？
- 為何我可以負責任的使用，卻擔心別人會不負責任的使用？（偷懶、錯誤）

## Gemini: Asymmetry of Trust

I know I am the "Executive Chef", but others may just microwave the food.

## The key of our worry: Is human in control?

保存所有與AI互動的紀錄 (包含時間)，以證明自己的腳色、貢獻

## 我們擔心的是AI代寫，不是AI輔助寫作

- “Use my **58-page ChatGPT prompt** — a complete, automated AI academic writing tool — **to generate entire research papers**, from Chapter 1 to Chapter 5, in just a few hours.”
  - Provide key information: research topic, case study, target population, data collection methods, type of study (qualitative, quantitative), referencing style
  - **Data generated by AI?**
- ThesisAI: “**One-prompt drafting** → generates structure, sections, and inline citations in seconds”
  - Possible to generate **review** articles, **not for primary research paper**.
- =買論文、代寫
- 永遠有人投機取巧，也必然有人僥倖獲利（自古皆然，於今尤烈）
- 如果涉及虛假數據或文獻 => 違反學倫
- 涉及學位論文：學生能力考核

## AI 可能涉及的學倫問題

- 抄襲 (是否偷?)
  - 研究結果是否真實 (是否騙?)
  - 虛假數據、虛假文獻、錯誤示意圖
- } 人要負責
- 工具是中性的，可以為善，可以為惡，人是操縱者，人要負責，與工具無關。
    - e.g. Image processing by Photoshop (也可造假)
  - 問題不在AI，但AI讓甄別困難許多
  - 今天偵測不出造假、抄襲，未來也許可以! 別冒險!

# 虛假圖像的案列

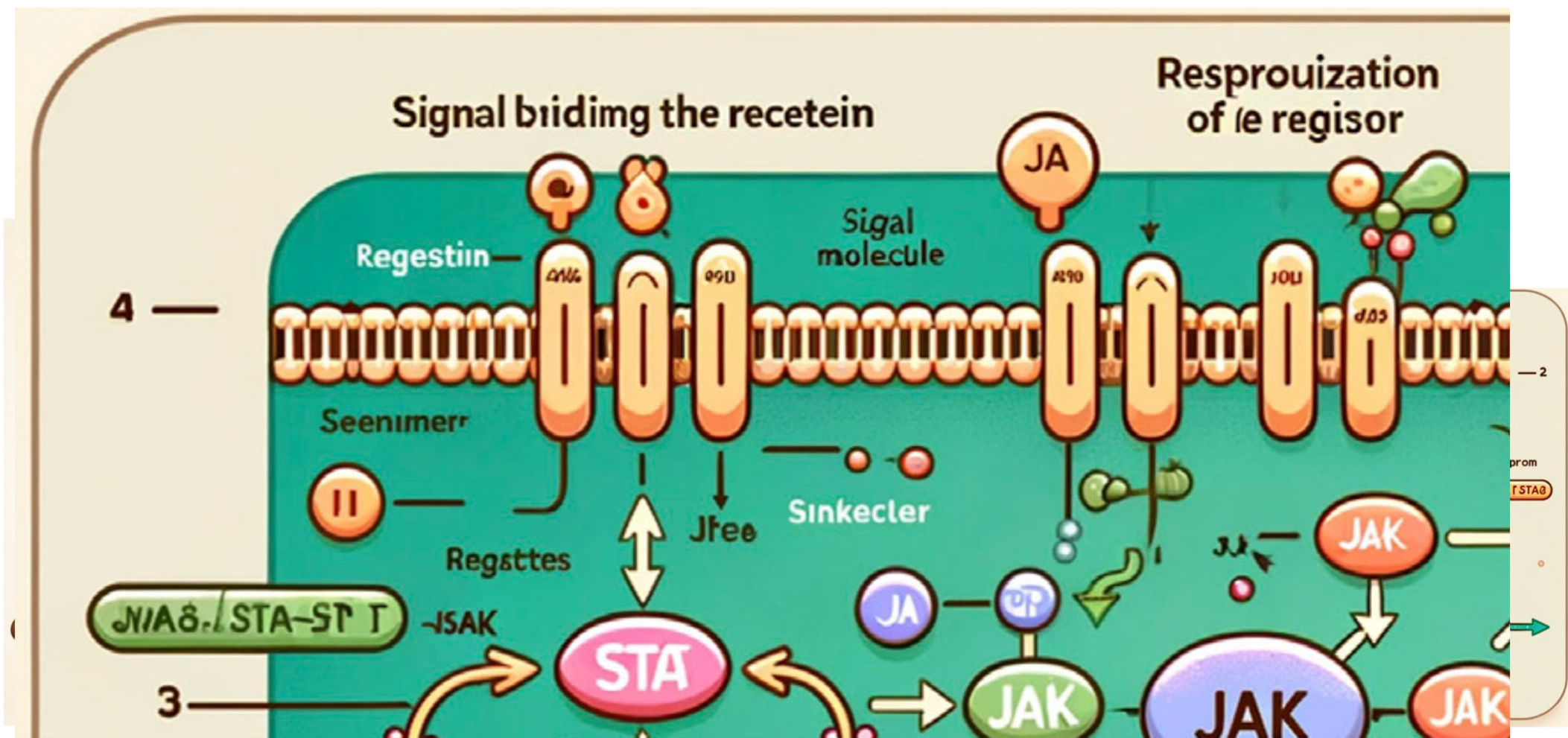


FIGURE 1 Spermatogonial stem cells, isolated, purified and cultured from rat testes.

FIGURE 2 Diagram of the JAK-STAT signaling pathway

# 虛假文獻的案例

**University of Hong Kong professor steps down from associate deanship after AI-generated references scandal**

*used non-existent publications generated by AI in its references*  
(Kelly Ho, HK Free Press, 2025.12.18)

辭去副院長職務(非撤職)

論文虛構書目、引用錯漏 政大博士生認錯

2026/03/14 05:30 記者林曉雲

校方已先行將該論文自資料庫下架，並啟動學術倫理審議程序進行調查。

尚未有以學術不端處理之具體案例

# 如何避免虛假文獻、圖像？

- 如何避免虛假文獻：

- 要求AI提供資訊來源(文獻)及連結及doi，方便檢驗
- 請AI(或另一AI)檢驗(請核實本論文所列參考文獻之真實性，並檢驗文獻內容是否與本文之引用語境相符)
- 仍須人工謹慎檢驗 (Google Scholar, PubMed, <https://consensus.app/> , <https://scite.ai/> )
- 提供真實的文獻庫
- 提示詞範例 (Gemini)：「請針對『生成式 AI 對高等教育評量方式的影響』提供 5 篇相關研究摘要。限制要求：
  - 每篇文獻必須包含：標題、作者、年份、期刊名稱及有效的 DOI 連結。
  - 如果你無法在真實資料庫中找到匹配的文獻，請直接回答『找不到』，**嚴禁自行編造**。
  - 請優先引用來自 Google Scholar 或 PubMed 已收錄的文獻。」

- 如何避免虛假圖像：

- 放大後，仔細檢驗(尤其文字) => 修正錯誤，但**投稿時應註明**

# AI 產生的虛假內容（幻覺）：如何處理？

虛假數據：影響研究結論，造假，嚴重學術不端

虛假文獻、示意圖：性質上屬於造假（學術不端）

如何處理：

- 如為論文核心論述基礎時 => 嚴重學術不端（即使非蓄意，但影響嚴重）
- 如不涉及論文核心論述基礎時
  - 雖不涉及虛假研究結果，但影響文章的可信度
  - 建議行政處理（要求修正、退稿、撤稿）
  - 類比：新購儀器未校正，產出錯誤數據，用於論文發表。應為錯誤負責，但未蓄意欺騙。

# 有辦法分辨AI與人類的作品嗎？

AI 生成的內容似乎有某種風格，能否分辨？

- 分辨工具發展中，只能推測，**無法100%證明**      <= 以此做為審查依據，站不住腳
- OpenAI Classifier 已於 2023 年永久下架：“連製造矛的人都造不出盾”(Gemini)
- 軍備競賽：道高一尺，魔高一丈
- 我不預期能發展有效可靠的辨別工具

## 需要分辨嗎？

- 內容真假，作者須負責
- **誰寫的並不重要**（見後章）

# How to tell the use of AI?

**Step2 :** The value of  $L$  is found by the principle of balance.

**Step3 :** Substituting Eq. (5), with Eq. (6) into Eq. (4), we obtain a polynomial expression that depends on the Jacobi elliptic function  $\mathcal{Z}(\chi)$ . By equating the coefficients of  $\mathcal{Z}^l(\chi)$ ,  $\{l = 0-7\}$  equal to zero, we obtain a system of equations. We solve this system to find the unknown parameters. The solutions of Eq. (5) are represented in Table [1] based on the values of the parameters  $s$ ,  $c$  and  $r$ :

→ Regenerate response

→ As an AI language model, there is no access to the specific database details of any particular research study. However, in general, a well-designed database for a hydroponics system should include the following:

顯然未經檢驗

**Retracted:** “the use of an LLM in the drafting of this work was **not declared** by the authors upon submission, **in violation of IOP Publishing’s ethical policy**”

<https://iopscience.iop.org/article/10.1088/1402-4896/aceb40>

# 使用AI未揭露，是否學倫問題？

使用AI未揭露，**未騙(造假)未偷(抄襲)**，違反的是

- 與期刊或國科會的約定 => **行政處分** (修改、退稿、撤稿、禁投)
- 透明原則(transparency): 如果牽涉研究成果實質內容，**必須揭露**
  - 如: 以AI生成或模擬數據、程式碼
  - 屬於「數據源頭」的問題，而非「表達形式」的問題。使用工具/方法、詳細條件/參數**必須詳細揭露**，才符合研究再現性的要求。這就像我們必須交代實驗試劑來源。
  - 但如未揭露，**非屬違反學倫**，而是審查端應**退稿**
    - 類比: 沒把實驗條件或工具寫清楚 => 退稿
    - 很多計畫/論文也沒交代清楚實驗條件

# 要求揭露只是過渡期反應

## Gemini

- 聚合酶連鎖反應 (PCR)
- 基本序列比對工具 (BLAST)

早期都需詳細揭露使用條件，現在已默認

## Gemini

在學術監管史上是一種非常典型的\*\*「安撫性儀式」(Pacification Ritual)\*\*。

當一個新事物（如 AI）打破了既有的規則平衡，機構（期刊、大學、資金方）的第一反應往往不是「解決問題」（因為還不知道怎麼解決），而是「建立儀式」，以此來宣告：「我們注意到這件事了，我們在控制局面。」

10% 人使用

防衛心態

要求揭露

100% 人使用

接受、默認

無需揭露

## 要求揭露造成道德困境

使用者:

- 揭露是否讓審查人覺得我偷懶? 是否對我不利?

審查人:

- 看起來像用了AI幫忙, 是否該追究?
- 是否用AI幫忙審稿?

53% reviewers already using AI to assist reviewing (Frontiers, 2025)

國科會

- 很多人(申請人、審查人)可能用AI但未揭露, 是否該追究?
- 追究起來, 行政負擔很大

我的看法: 揭露意義不大, 且無法求證(除少數太懶的案例), 造成大家的困擾, 建議無須深究。

## 他人幫助，但我主導

我是一名研究員，我碰到研究上的問題，請生成式AI幫我想解決方法，然後我再去讀懂這個方法，並判斷此方法是否真的有幫助解決問題...那這樣是否可以稱為我想出的方法呢？

我查文獻/問老師/與同事討論，找到可能的解決方法，判斷是否可以幫助解決問題

AI 提供有效率的幫助，最終判斷取捨還是你自己

你的貢獻：問題的形成、篩選與判斷、後續執行

# 原創性 (originality)?

我研究的idea來自別人/AI，這樣的研究成果算我的原創嗎？

幾乎所有研究進展都奠基於前人的研究，要如何界定原創性？

- 讀文獻、聽演講，獲得啟發、靈感 => citation (give credit)
- 合作者、學生、博後、AI 提出具體構想 => coauthor

牛頓：「如果我看得比別人遠，是因為我站在巨人的肩膀上。」

- 只要你誠實標註了「肩膀」（引用前人與 AI），剩下的那部分「你自己跑完的最後一哩路」，無論多短，都是你的原創。(Gemini)

**智識的功勞 (intellectual credit) (Claude)**

- 我站在哪些人的肩膀上？（誠實引用，包含AI）
- 我的研究加入了什麼？（知識的進展）

類似觀點: Nakazawa, E., Udagawa, M., & Akabayashi, A. (2022). Does the use of AI to create academic research papers undermine researchers' originality? *AI*, 3(3), 702–706.

## 審查人不必在意計畫是誰寫的或創意來自誰

- 政府資助研究，要的是成果，不論由誰提出或做出（只要credit歸屬合乎學術規範）。
- 要成果，所以要看的是計畫的可行性及執行力。
- 在高度分工合作的研究中，部分的創意、執行、修正、撰寫，可能來自不同的合作夥伴（學生、博後、助理、合作者、廠商、AI），而非來自PI，但PI綜規全局，協調各方，統整、取捨、決策、管控，無疑是主導。
- AI只是研究團隊中的新成員，且不承擔名分與責任

### Gemini (我跟AI辯論的結果)

學術界必須放棄對「來源」的執著：接受創意和內容可能來自多元化輸入（包括AI）的事實。

# 未來可能得用AI幫忙審查

- 撰寫計畫的成本（時間、精力）低
  - ⇒ 投機性的計畫書數量暴增
  - ⇒ 審查負擔重 ⇒ 限制可投計畫數目
- 計畫品質趨中，鑑別度降低（審查人：一則以喜，一則以憂）
  - ⇒ 審查難度高，需驗證真偽、鑑別創新
  - 利用 AI 幫助審查？ (internal, “within-firewall” AI systems)
  - 整個學術出版及審查制度都面臨危機

# 用AI 輔助審查?

*Frontiers* journals: AIRA (Artificial Intelligence Review Assistant)

- 為了減輕審查人的「**行政與查核負擔**」
- 揪出學術不端與防弊：除了傳統的文字抄襲偵測外，AIRA 還能自動掃描論文中的實驗圖片，偵測是否有被局部複製、竄改或不當修圖的痕跡。
- 利益衝突排查 (Conflict of Interest)：自動交叉比對作者群與被邀請的審查人之間，是否有未揭露的潛在關係（例如過去曾共同發表過論文或隸屬同一機構）。
- 合規性與倫理檢查：確認論文是否齊全了必要的聲明，例如數據可用性聲明 (Data availability)、人體或動物實驗的倫理批准，以及受試者同意書。
- 語言與基礎品質：評估英文寫作的品質，並揪出是否有使用具爭議性或不當的關鍵字。
- **不做接受 (Accept)或拒絕 (Reject)論文的決定**

(Gemini)

# AI Review Trap by ICML

## ICML Desk Rejects 497 Papers: AI Review Trap Explained 2026.03.19

- **International Conference on Machine Learning**
- **desk-rejected** 497 papers—nearly 2% of all submissions
- catching **506 reviewers** (795 reviews; about 1% of all reviews) violating the conference's AI usage policy
- **rejecting papers authored by these reviewers**
- unfairly harms co-authors who weren't involved in the violation
- ICML watermarked submission PDFs with invisible **prompt injection** instructions that LLMs would follow, embedding detectable phrases in AI-generated reviews.
- enforcing academic integrity or **unethical entrapment**
- **Opt-in**: reviewers knew the rules, explicitly agreed not to use AI, then broke their commitments.
- Only works once.
- Root cause: reviewer overload

AI 學者用 AI 作弊

結果被 AI 技術抓到退稿

<https://www.facebook.com/ProgressBarTW>

<https://byteiota.com/icml-desk-rejects-497-papers-ai-review-trap-explained/>

## Summary 4 使用AI輔助寫作的學倫考量

- AI產生的內容是否真實、抄襲，人要負責驗證
- 工具是中性的，可以為善，可以為惡，人是操縱者，人要負責
- 無法肯定的分辨AI與人的作品
- 目前規範要求揭露，但**未揭露非屬違反學倫**

# SUMMARY

- AI 是破壞性創新，勢不可擋
- 人+AI >> 人; 不善用AI者會被淘汰(殘酷的現實)
- AI 在研究的每一步都可以幫忙
- AI 拉高了地板，天花板的提升要靠人的創意
- 工具是中性的，使用AI，人要為內容負全責(相同的學倫原則)
- 研究是多元分工合作，AI 是研究團隊中的一員
- 審查計畫的重點在於創新性、可執行、成功機會，計畫由誰寫的不重要
- 目前使用AI必須揭露，其實無效且造成困擾
- 目前不能用AI審查，未來可能會放寬

# Acknowledgments

感謝周倩、張雯、呂俊毅、彭麗春、周美吟、連正章、張雯、周雅惠  
提供寶貴意見

感謝生成式AI（主要是 Gemini 3 Pro, ChatGPT-5.2 Thinking，但也包含 Claude Sonnet 4.5, Grok 3 Pro, Perplexity Sonar, DeepSeek V3.2）在本演講構思  
撰寫階段，**透過遞迴式討論，回應我各項問題與對我想法提出批評與建議，  
並與我來回辯論，讓我得以釐清論述架構並發現邏輯盲點。所有最終的論述  
判斷及文字定稿，均由本人全權負責。** [My complete discussion history with  
AIs is available upon request.](#)

**While  
INTELLIGENCE  
is becoming  
artificial,  
STUPIDITY will  
always remain  
ORIGINAL**