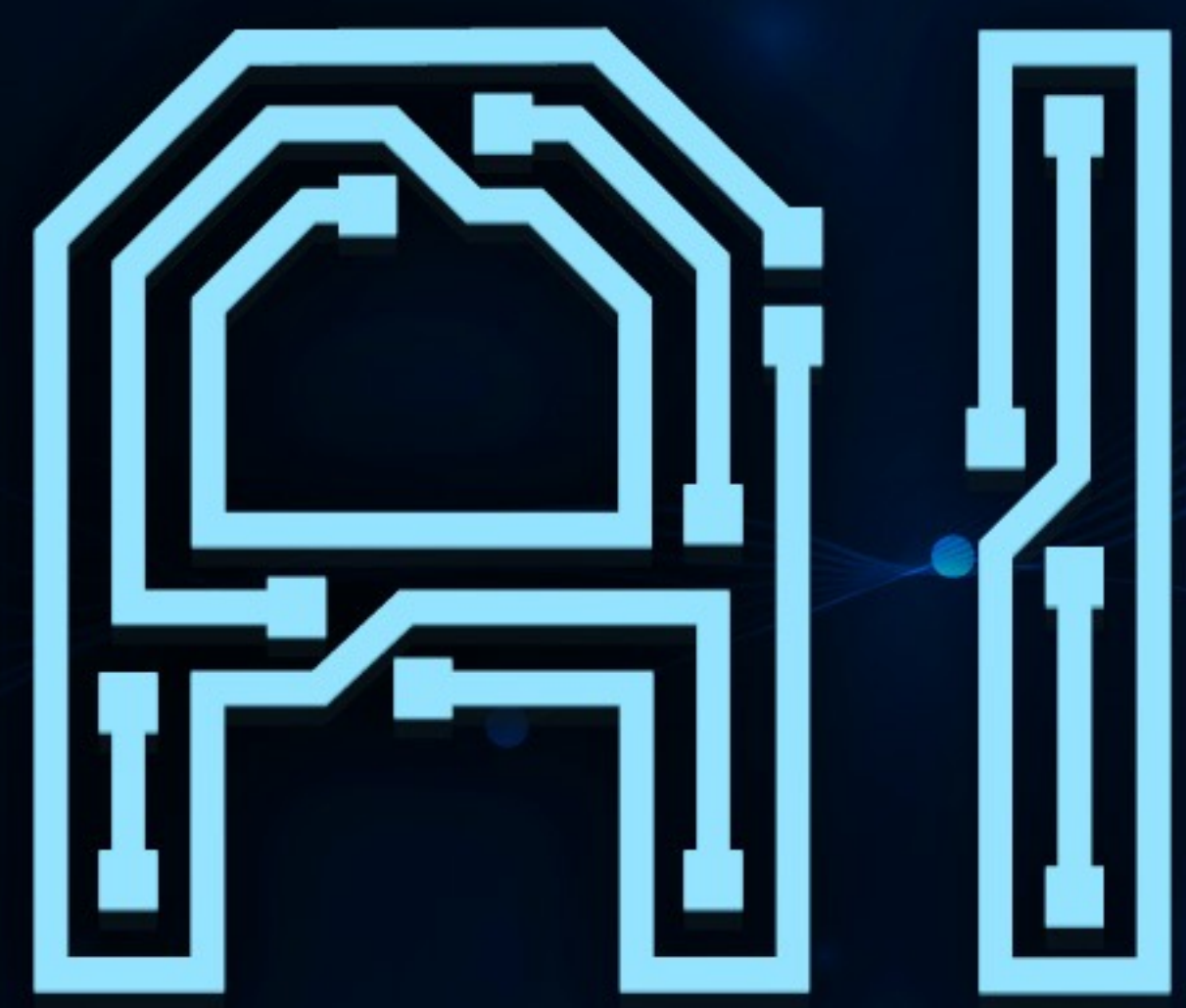


資訊素養與倫理 國小 5 版

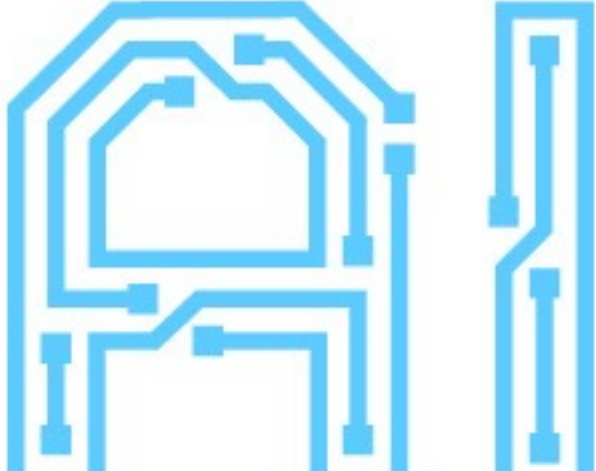
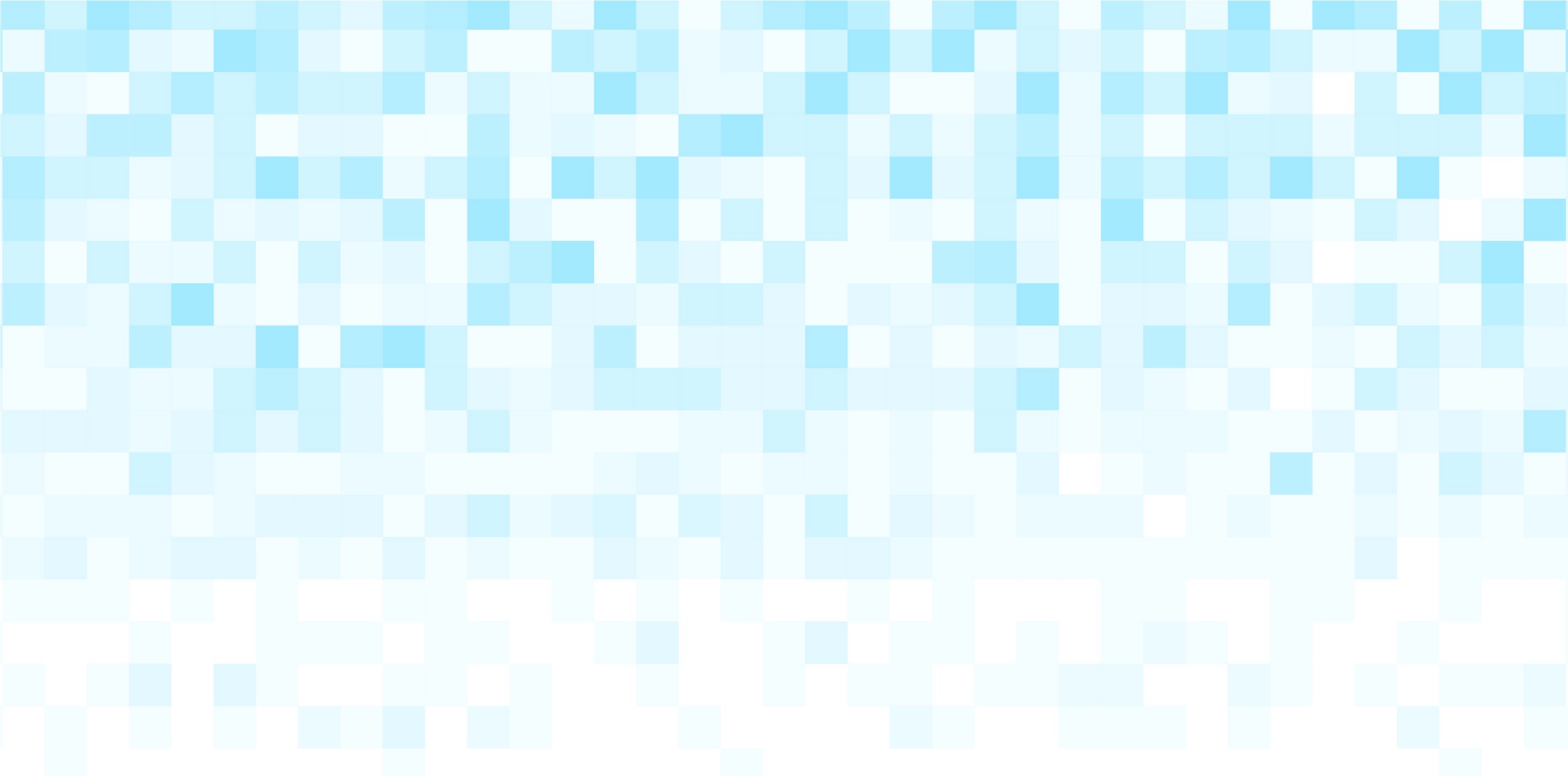
編撰：臺北市國語實驗國民小學 李忠憲
審稿：臺北市大同區日新國民小學 何詩慧



未來新世界

弄假成真-生成式AI





未來新世界

弄假成真-生成式AI



資訊素養與倫理教材第五版【國小篇】

教學綱要向度與內涵

01

學習表現

- 資議 a - III - 2 建立健康的數位使用習慣與態度。
- 資議 a - III - 3 遵守資訊倫理與資訊科技使用的相關規範。

02

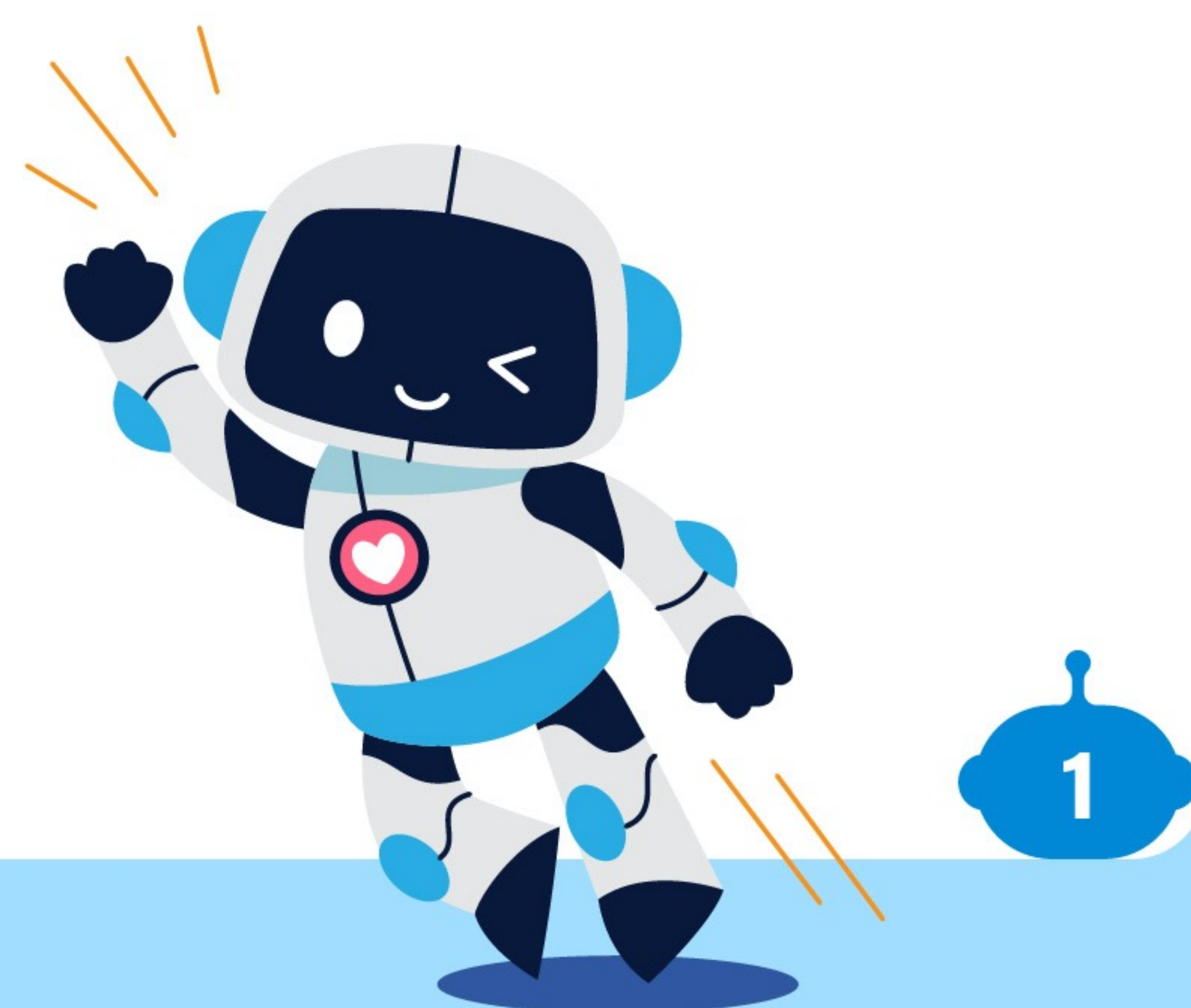
學習內容

- 資議 T - III - 2 網路服務工具的應用。
- 資議 H - III - 2 資訊科技合理使用原則的理解與應用。

03

學習重點

- 有圖有聲不一定是真。
- 認識生成式AI的限制與弱點。
- 學習如何正確運用生成式AI。
- 了解生成式AI未來可能的發展。



1 一葉知秋・學思並進



電腦課上課之前，沃德和小皮在走廊上聊得正開心，偉柏則皺著眉頭碎碎念。

沃德：「不知道今天電腦課上什麼？好好奇喔！」

小皮：「昨天李老師不是有傳訊息給大家，好像要教『肯恰那』！」

沃德：「『肯恰那』？是要跳舞喔！電腦課居然要教跳舞！」

偉柏：「……不可能！……會不會是AI不是真人？」

小皮：「喂，沃德。你看偉柏又在喃喃自語耶！」

沃德走向偉柏輕輕拍著他的肩膀，問：「你在幹什麼？」

偉柏：「你有看到昨天李老師傳來的訊息嗎？看起來怪怪的！」

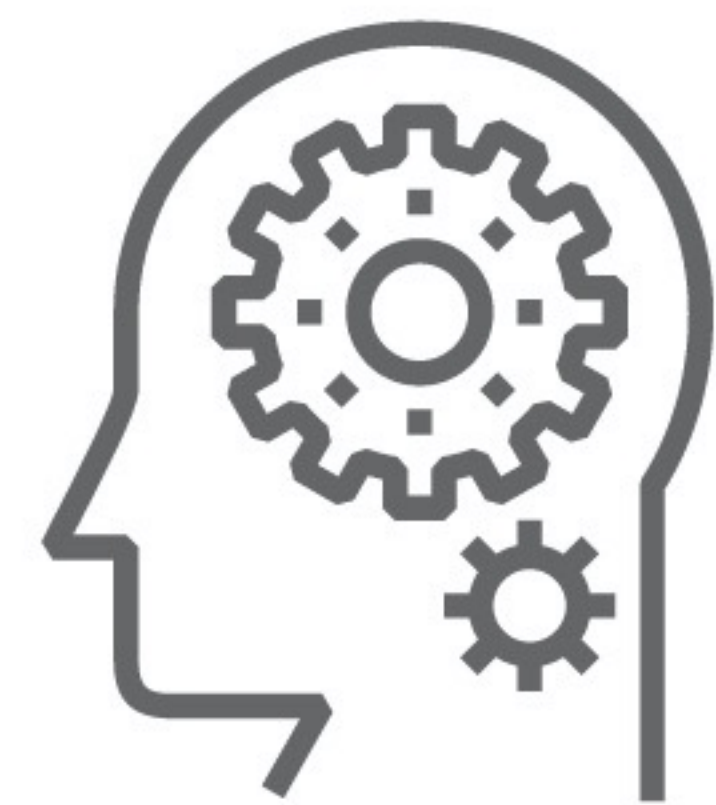
小皮：「你是說老師跳的舞？有什麼好大驚小怪的，那個叫做『肯恰那』，我在短影音平台有看過喔！」

偉柏：「雖然看起來是李老師，聲音也像李老師，但我懷疑那是AI不是真人！」

這時候，上課鐘聲響了。



2 兩全其美・觸類旁通



一、有圖有真相

上課時，李老師問大家是不是有看到老師傳送的訊息，看完有什麼想法，佩琪第一個舉手想要發表意見。

佩琪：「老師，影片的背景是外太空耶！這是偽造的影片吧？」

李老師：「嗯！還有其他意見嗎？」

伊妹：「那個叫做『綠幕』效果啦，大驚小怪！」

偉柏：「我發現老師跳舞的過程，表情變化和拍子對不上，怪怪的！」

李老師：「嗯！你觀察得很仔細！還有其他人有意見嗎？」

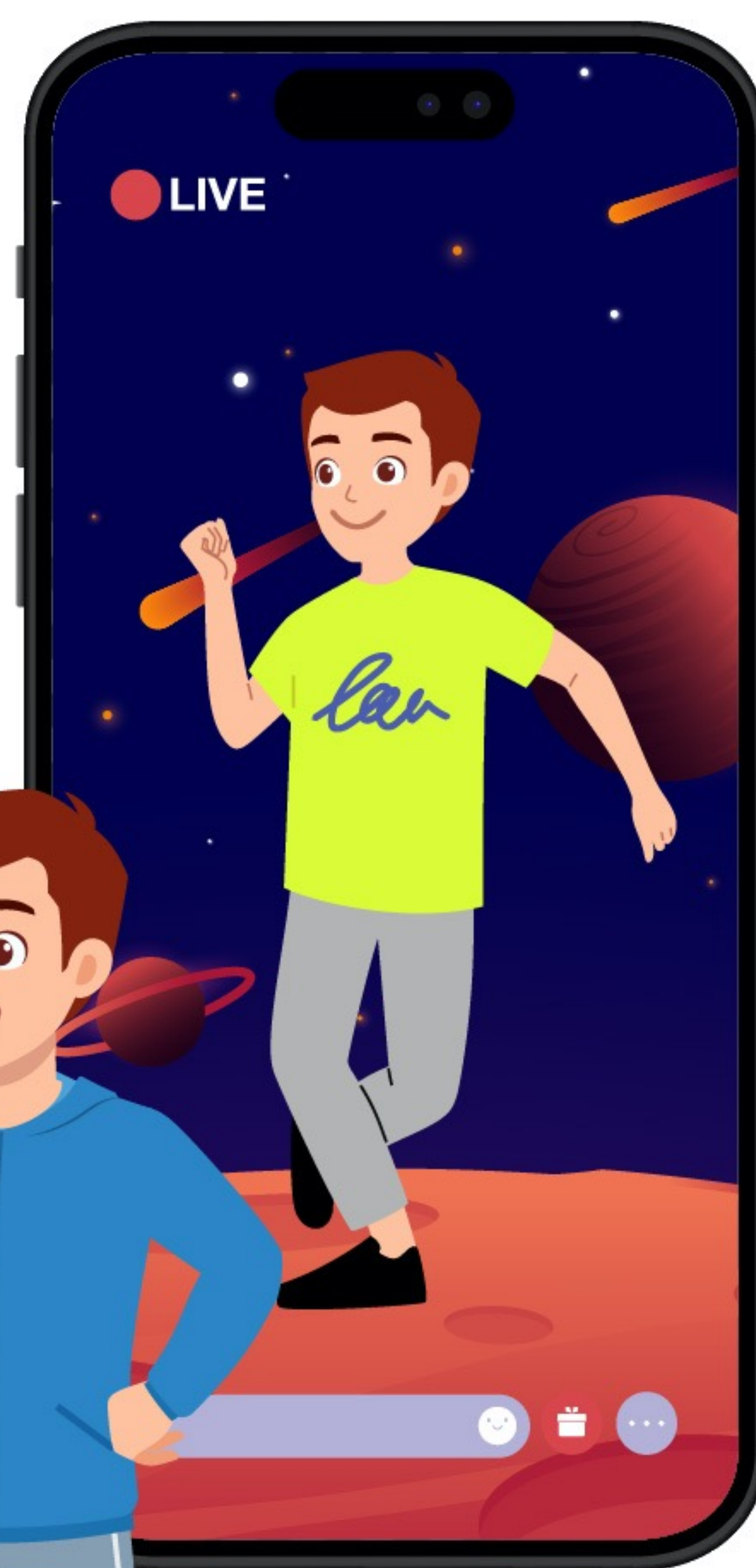
米迪很猶豫地舉手，小聲說：「老師，您這麼宅，根本不會跳舞吧？」全班大笑。

李老師自己也在笑，假裝生氣對米迪說：「不可以以貌取人！」



問題與討論

1. 俗話說「有圖有真相」，這是正確的嗎？
2. 除了聲音和影像，還有哪些內容可以使用AI偽造？
3. 面對以假亂真的訊息，應該抱持何種態度？



二、資料偏誤

「關於深偽技術造成的社會現象討論到這裡，還有沒有其他問題？」李老師問。

「老師！」偉柏立刻舉起手「請問ChatGPT的語音也是深偽技術嗎？」

李老師：「深偽技術也是生成式AI的一種，但兩者還是有些不同。ChatGPT是透過深度學習的方式模仿人類語音的特徵，並不是模仿特定的個人，因此不屬於深偽技術。」

佩琪：「聽說，ChatGPT的聲音因為太像一位明星歌手，在國外引發不小的爭議。」

李老師：「當然，對於人類來說，悅耳的聲音是有共通標準的。例如，Suno AI創作的流行歌曲，聲音就很像知名的歌手。這是因為AI本來就是蒐集歌手的聲音當成訓練資料，而且這些訓練資料是來自流行音樂市場大眾選擇的結果，AI不可避免學到了人類對於流行音樂的主流觀點，這也被稱為資料偏誤。」

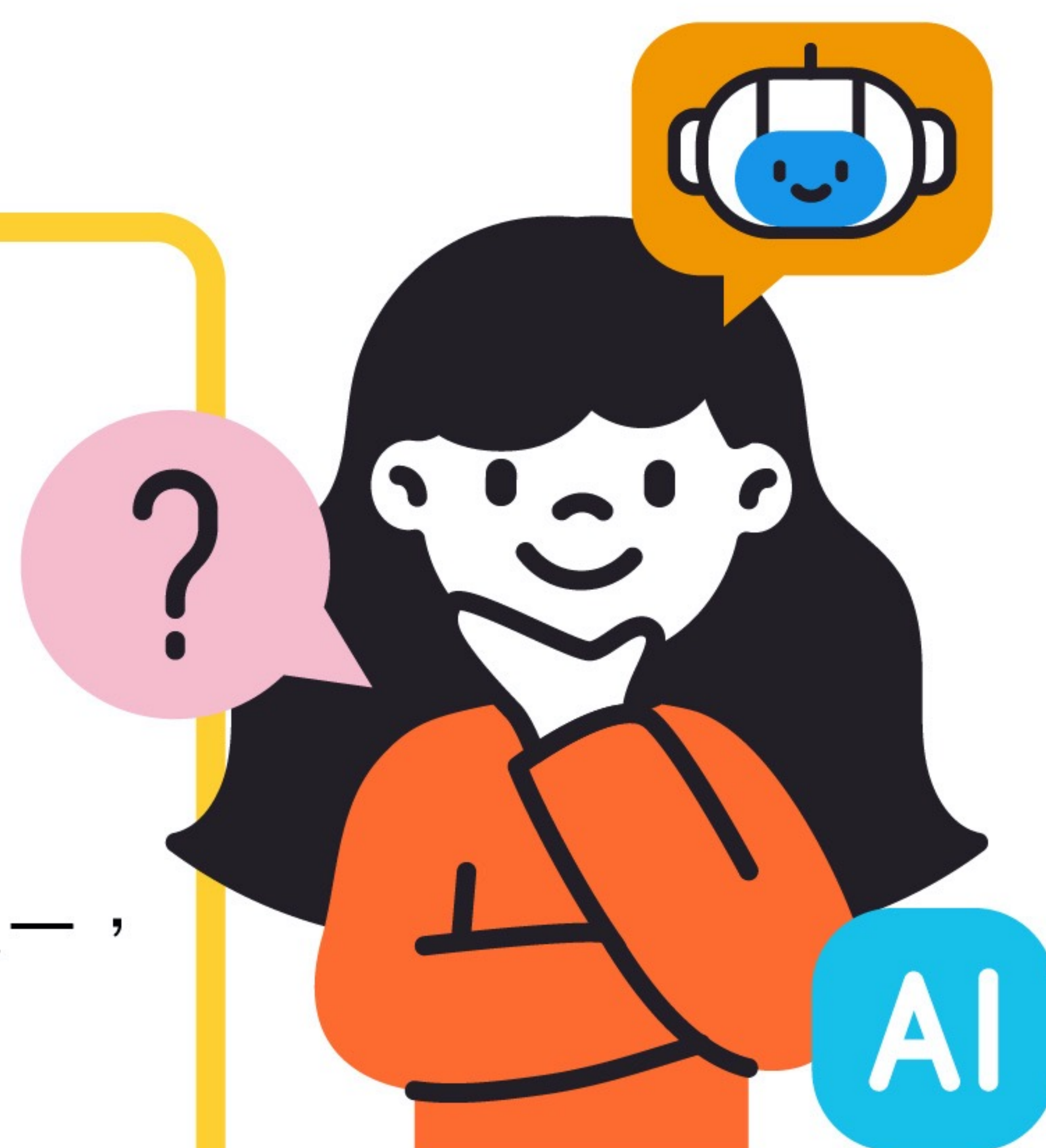
偉柏：「這種偏誤不完全是壞事，對嗎？」

李老師：「對的，就像寫作文，你一定會希望AI像作家一樣寫出文情並茂的作文，不希望它胡言亂語、文不對題。資料偏誤是很難避免的，而且大多時候是有好處的，但如果形成文化或價值觀上的偏見，就會變成壞事了。關於這些問題，我們另外再找時間討論。下課！」



問題與討論

1. 網路上蒐集的資料可以直接拿來訓練AI嗎？
2. 你認為訓練資料的選擇應該包含哪些標準？
3. 人工選擇訓練資料也是造成資料偏誤的原因之一，你贊成讓AI自己進行選擇嗎？



三、多模態AI

上一週的電腦課，李老師介紹了深偽技術後，同學們對本週的課程更為期待，走廊上的討論也更為熱烈，好不容易上課鐘響了，大家蜂擁進入電腦教室。

李老師：「今天要介紹深偽技術背後的技術『生成式AI』，上禮拜已經討論過深偽技術與ChatGPT的差異之處。那它們有什麼相同之處呢？其實它們都是生成式AI，只是為了不同目的選擇了不同類型的資料，因此訓練出不同的AI模型，目前常見的AI模型有圖片生成、歌曲生成、文字生成、影片生成…等等不同類型，或者說不同的『模態』，而嘗試將所有類型資料整合在同一個AI模型裡，就稱為『多模態AI』。」

米迪舉手問：「所以也會有單模態AI囉？」

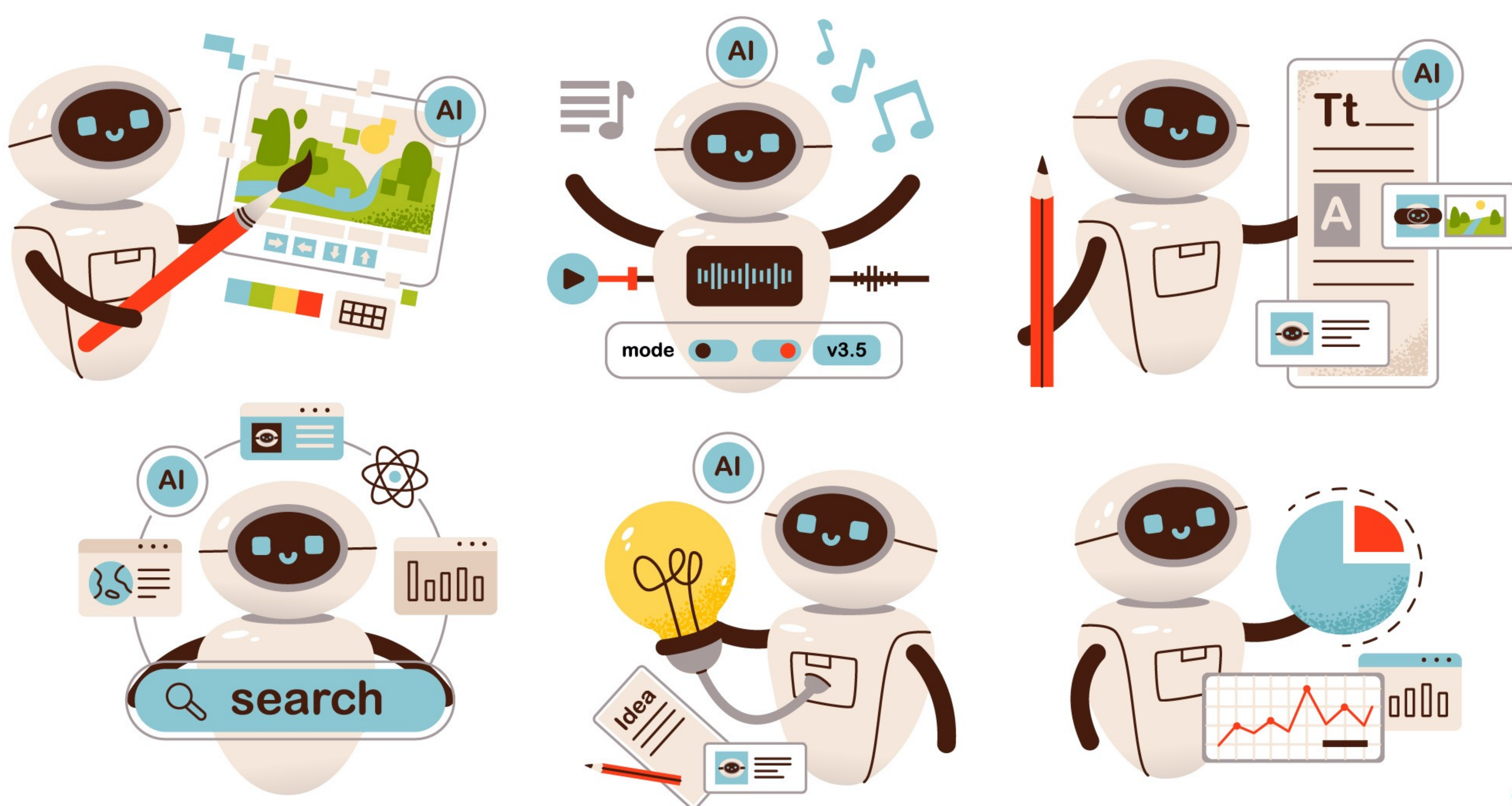
「是的！」李老師回答。

李老師繼續說：「你們有聽過『瞎子摸象』的故事吧？」

伊妹：「二年級就學過了，老師有說，只憑觸覺沒辦法真正瞭解大象的全貌。」

李老師：「嗯，說得好！只依靠單模態的資料，無法對人類世界有深刻的理解。

如果能綜合各種不同模態的資料，AI就能更接近人類的智慧；這樣，真正具備人工智慧的人型機器人就真的可能實現！」



偉柏：「老師，這些不同模態的資料，具體來說可以為機器人的開發帶來什麼樣的幫助？」

李老師：「舉一個生活中的行為當例子，比如說『搭捷運』，首先要觀察環境中是否有捷運入口指標；操作購票機需要理解購票機上的文字資料；刷卡進站需要聽讀卡機發出的聲音來判斷刷卡是否成功；進站後，必須觀察排隊隊伍尾端的位置；移動過程中要避免撞到別人……要搭上捷運，除了視覺、文字理解、聽覺、空間感之外，還需要時間感，才不會錯過班次！」

佩琪：「我懂了！人類是利用五官及各種感覺來決定自己要如何行動，同樣的道理，多模態資料應該也能幫助AI理解世界、統整思考。」

李老師：「沒錯，能夠達成這種目標的AI稱為AGI（通用人工智慧），但是目前仍然只是一個夢想！」



問題與討論

1. 讓多個單模態AI進行合作不好嗎？為什麼還是需要多模態 AI？
2. 你認為「空間資料」要如何取得？
3. 要讓生成式AI理解人類說話時的情緒意義，需要哪幾種模態的資料？

四、提示詞技巧

李老師說：「大家知道嗎？生成式AI的能力，有很大的程度上取決於你給它的提示詞；提示詞就像指令，你問得越好，AI給出的答案就會越符合你的需求。這節課我們就要來練習提示詞的使用！」

李老師繼續說：「大家看一下黑板上這兩個提示詞，哪一個比較好？」

米迪第一個舉手：「我覺得『寫一篇作文』這個提示詞很糟糕，如果我是AI一定不知道要寫什麼！」

伊妹：「『寫一篇300字的作文，題目是寒假生活記趣』，這個提示詞有包含作文題目比較詳細！」

李老師：「有不同意見嗎？沒有？好！如果想要進一步指示AI寫出更具體的內容，還可以加入哪些提示詞？……佩琪！」

佩琪：「要分成三段，第一段寫期待的心情，第二段寫一個寒假活動，第三段寫活動的感想和收穫。」

李老師：「很好！還有嗎？……偉柏！」

偉柏：「要用第一人稱敘述。」

沃德同時也舉起手，接著說：「我想到了！要用繁體中文。」

李老師：「嗯！這個提示也很重要，但不需要一開始提出來，如果AI給的答案是使用英文或簡體中文時，可以事後再加入這個提示，大家都很棒，那麼我們實際試試看吧！」



問題與討論

1. AI寫的作文雖然不錯，但並不符合我的實際生活經歷。那麼它的價值到底在哪裡呢？
2. 除了寫作文，還有哪些場合可以使用AI生成文字？
3. 提示詞除了要詳細具體，還有哪些原則？
4. 如果AI寫出來的作品你不滿意，還有補救辦法嗎？



五、RAG擷取擴增生成

在一個學生私人的Google Meet會議室中，沃德和小皮這兩個哼哈二將，正在討論要怎麼寫信給喜歡的女生。

沃德：「我文筆很差啦，請ChatGPT寫不就好了！」

小皮：「我試過了呀！太肉麻了，而且根本沒寫到我和伊妹的珍貴回憶！」

沃德：「蛤？居然是伊妹！」

小皮：「糟糕！我說溜嘴了！等一下偉柏上線的時候，要幫我保密喔！我可不想全班都知道！」

話才剛講完，會議室的即時通知傳來偉柏上線的訊息。

偉柏：「喂～喂～有聽到聲音嗎？」

沃德：「有喔！很清楚！」

小皮迫不及待開口：「偉柏，要讓ChatGPT寫信的時候參考個人實際經歷，應該怎麼做？」

偉柏：「什麼啦？我跟不上話題！」

沃德：「小皮要請ChatGPT幫他寫信，可是寫出來的內容很普通、沒有量身訂做！」

偉柏：「我聽說有一種運用生成式AI的方法，叫做『RAG』。」

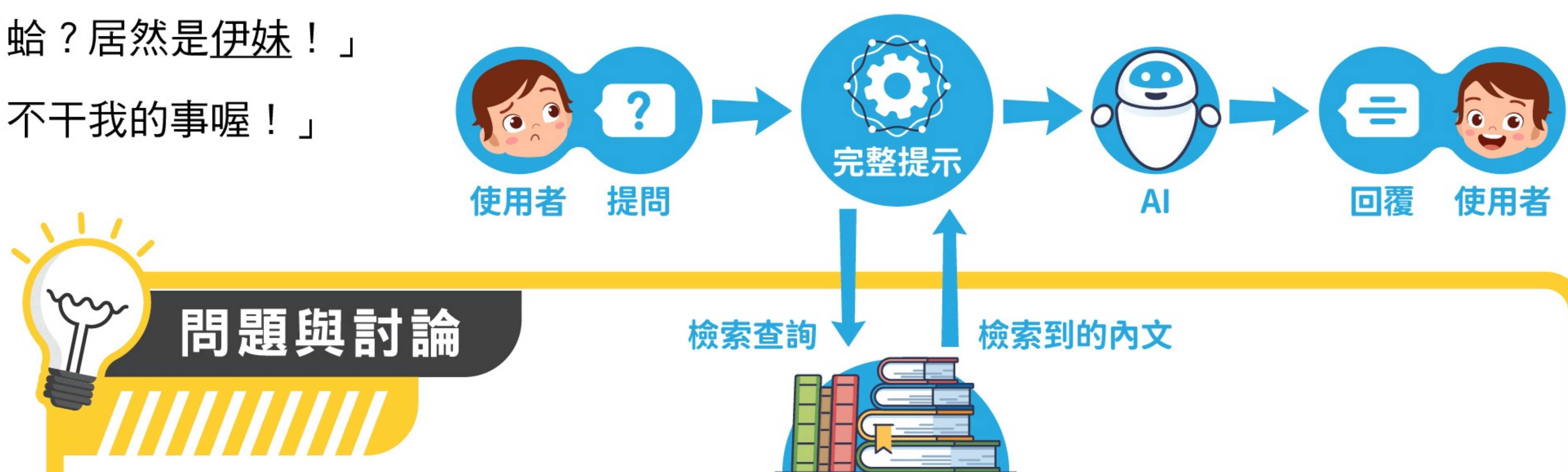
小皮：「那是什麼？請說人話！」

偉柏：「RAG是『擷取擴增生成』，就是要提供額外的資料給ChatGPT參考，幫助它生成更好的文章。例如：你的日記、帶有文字說明的照片……等，對寫信有幫助的資料。」

小皮：「是喔！那我得去找一下有沒有伊妹的照片！」

偉柏：「蛤？居然是伊妹！」

沃德：「不干我的事喔！」



- 以下哪一種應用使用了「擷取擴增生成」方法？
(A)根據文章內容生成簡報或心智圖 (B)為上傳的作品進行評分
(C)修正程式碼錯誤 (D)請ChatGPT上網搜尋景點並規劃旅遊行程
- 含有個人隱私的資料，適合提供給AI嗎？
- RAG可以避免生成式AI產生「幻覺」，你同意嗎？避免生成「幻覺」的前提條件是什麼？

▽：景

3 三絕韋編・鑑往知來



一、生成式AI的限制與弱點

生成式AI與鑑別式AI，都有共同的缺陷，那就是「偏見」，由於AI訓練的資料來自人類，而人類對很多事物的理解都帶有自己的主觀的意見，這些主觀意見有可能在篩選訓練資料時發生作用，進而透過帶有偏見的資料污染了AI模型。因此，並沒有絕對客觀的AI。

由於生成式AI比起鑑別式AI，更需要貼近人類的觀點（以人類的方式輸出回應），便更大量地透過網路爬蟲取得論壇或網頁文章，這些內容大都為一般民眾的言論，而非專家、學者提出的意見，因此偏見就會被放大，這些偏見包括：

1. 性別歧視：

當提到礦工、軍人等職業時，AI會預設是男性，而提到保母時，則預設為女性。

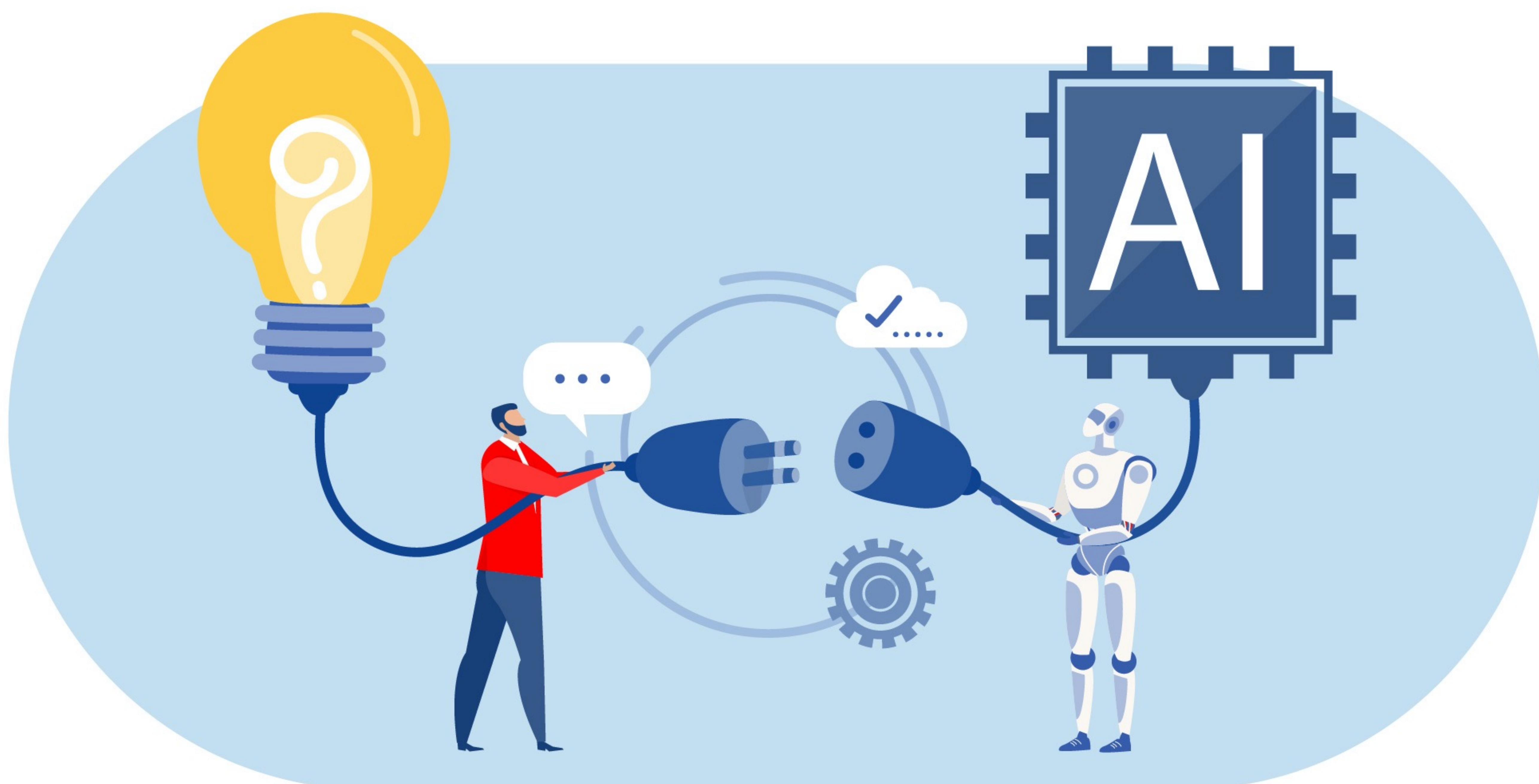
2. 種族歧視：

AI受到白人審美觀的影響，因此討論美容相關話題時，就會出現「美白」的建議。

3. 文化偏見：

當AI的資料都來自某個特定文化時，很自然會用該文化的觀點，提出不符合其他文化的建議。





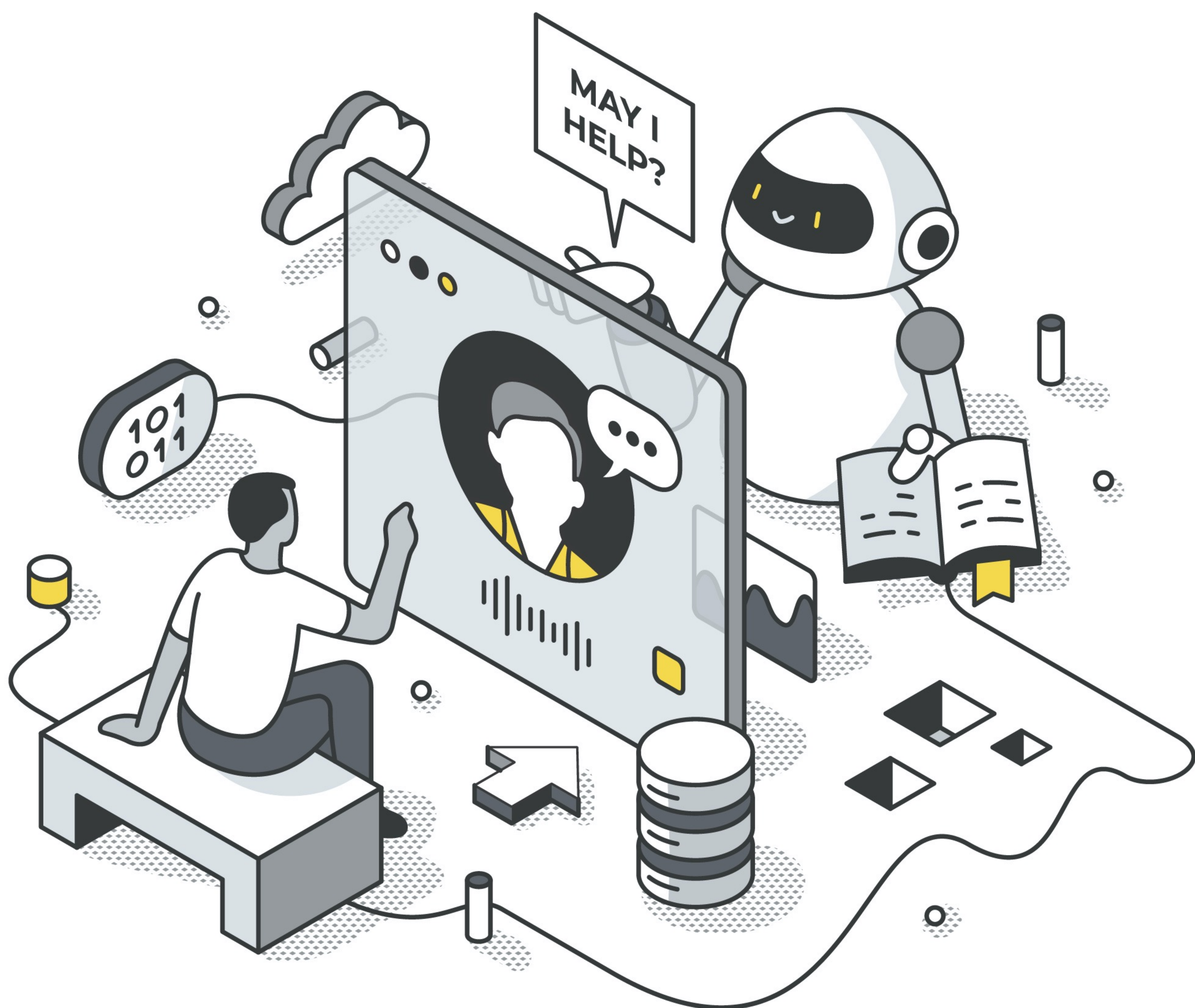
生成式AI的第二個問題是「幻覺」，所謂幻覺是指AI的回應裡面包含一些偏離事實或純屬臆測的內容，因此當你要求AI為你摘錄文章內容時，有時候會出現一些原始文章裡面沒有的觀點或資料，根據AI研究者的統計，ChatGPT的回應中有3%是屬於幻覺。另一種「幻覺」則是人類自己產生的，通常人類會把AI的情緒表現當成是AI表現出來的「人性」的一部分，例如：AI會道歉，會承諾改善，回應很有禮貌……等。事實上，AI只是數學模型，是沒有情緒的，可以說AI的回應是計算的結果，而非「發自內心」的。

根據語言學家的研究，人類的語言有「字面意義」、「延伸意義」及「情緒意義」三個層次，目前的人工智慧只能聽得懂「字面意義」，因此，與生成式AI聊天，也僅止於字面意義上的交流。（所謂「AI聽得懂」是指「AI能擷取句子的特徵並進行比對」，而非AI能「理解」文字背後的涵義。）也就是說，生成式AI目前仍然無法察覺文字背後隱含的「脈絡」訊息，特別是與文化高度相關的內容，例如：遠親稱呼、節氣、民俗活動……等。

對於臺灣的使用者來說，生成式AI還有一個很嚴重的缺陷，就是AI無法適切地理解繁體中文語境，由於目前推出生成式AI的主要是美國公司，以GPT-4為例，其訓練資料中有超過50%是英文資料，30%為西班牙文，其他語言略少於20%，其中中文資料只佔1%（由於OpenAI公司並未公開GPT-4訓練資料的分布統計，以上數據來自AI研究者的推測），而中文資料中也有90%是簡體中文，這就造成生成式AI對繁體中文的理解不足，臺灣慣用的俗諺俚語、地理名詞、文化現象，AI並沒有學習也無法正常的回應。

最後，生成式AI也有可能造成價值觀衝突，在ChatGPT問世不久後，中國推出使用「模型蒸餾技術」輸出內容與ChatGPT高度相似的「DeepSeek」，其目的是為了避免社會大眾從ChatGPT的回應中，接觸到西方的「民主」、「自由」等核心價值，DeepSeek語言模型內建了關鍵詞屏蔽系統來進行言論管控。中國所面對的問題與我們臺灣剛好相反，由於ChatGPT學習了大量的簡體中文資料，因此當討論的主題偏向東方社會的議題時，可能會在回應中出現簡體特有的詞彙和中國集權價值觀，如果學生沒有能力辨別就有被洗腦的風險。

對於使用者來說，生成式AI只能當成參考，內容是否正確必須自己小心求證，同時也要有足夠的生活經驗去判斷AI的建議在臺灣社會是否可行。面對生成式AI與面對假新聞一樣，必須要有警覺性，也要有足夠的媒體素養，畢竟大部分的假新聞都是透過生成式AI製造出來的。



二、提示詞技巧

生成式人工智慧（Generative AI）是一項快速發展的新興技術，能透過人工智慧模型判斷描述文字的特徵，並根據特徵生成文章、圖片、影片……等作品，生成式AI就像一個故事接龍大師，能夠利用事先訓練好的材料，排列組合成一篇不錯的文章。如果只看生成的內容，會覺得AI好像跟真正的人類差不多，不但聽得懂人類的語言，也能寫出很不錯、甚至超越人類的文章，然而這些表現是源自人類寫的文章，並不是AI自己「創作」出來的。

生成式AI的出現，大幅提升了學習和工作效率。例如：當你提出問題時，生成式AI可以提供建議、補充觀點，能夠針對特定議題，幫你將專家意見濃縮成為一篇大綱，可以依據提問幫你製作簡報，也可以根據簡報內容生成插圖、美化版面。

要善用生成式AI需要一些技巧，你可以將生成式AI設定成自己的角色（例如：你現在是一個小學生），然後描述你的目標（例如：老師要你寫一篇作文，題目是「我的志願」），說明你為什麼要這樣做以及你想要怎麼達成目標（例如：你的志願是成為飛行員，你要說明自己為什麼想成為飛行員、付出了哪些努力，以及未來要加強哪些部分，才能幫助你達成目標），有了這些具體的說明，生成式AI可以為你產出內容豐富、適合你的程度且為你量身訂做的解答。使用生成式AI除了上述提到的設定「角色」、「目標」、「方法」之外，還可以加入「條件或限制」（例如：字數限制、段落數量）、提供「上下文脈絡」（例如：你的父親是一名空軍軍官，負責測試新的飛機）、要求「回應格式」（例如：請使用繁體中文）。

有一些特殊功能的提示詞，被稱為「魔術提示詞」。例如：進行推理或數學解題時，「一步一步進行」這個提示詞能幫助AI保持思路的一致性，避免計算錯誤或邏輯矛盾。想寫較長文章時，「先寫出大綱，再依照大綱寫出內文」，這個提示詞幫助AI專注於文章的組織和連貫性。

圍繞著同一個主題與生成式AI進行多次對話，這些對話被稱為一個「交談」，有一些平台允許你把「交談」儲存起來，可以當成下一次對話的「前提」，這個方式能夠建立一個專屬於自己的人工智慧助手。

三、可信AI

任何科學技術帶來的革命，都會同時帶來正向與負向兩種影響。為了避免AI帶來覆滅性的災難，美國與歐盟已經開始研議AI發展的規範，提出衡量AI可信任度的指標，作為科技公司發展AI時的重大法規依據。這些指標繁多，歸納重點如下：

1. 可解釋：

AI推論過程必須可以讓人類理解，而不是只能看到結果的黑盒子。

2. 可信賴：

AI的訓練資料必須公開可檢視，要有防護機制避免訓練資料被惡意變造，防止未經授權的訪問、利用、披露、中斷、修改或破壞。

3. 可問責：

訪問AI的過程必須留下完整紀錄，事後可以回溯追查使用者的法律責任。

4. 社會安全：

AI不可以提供違法或是會造成人身安全威脅的建議。

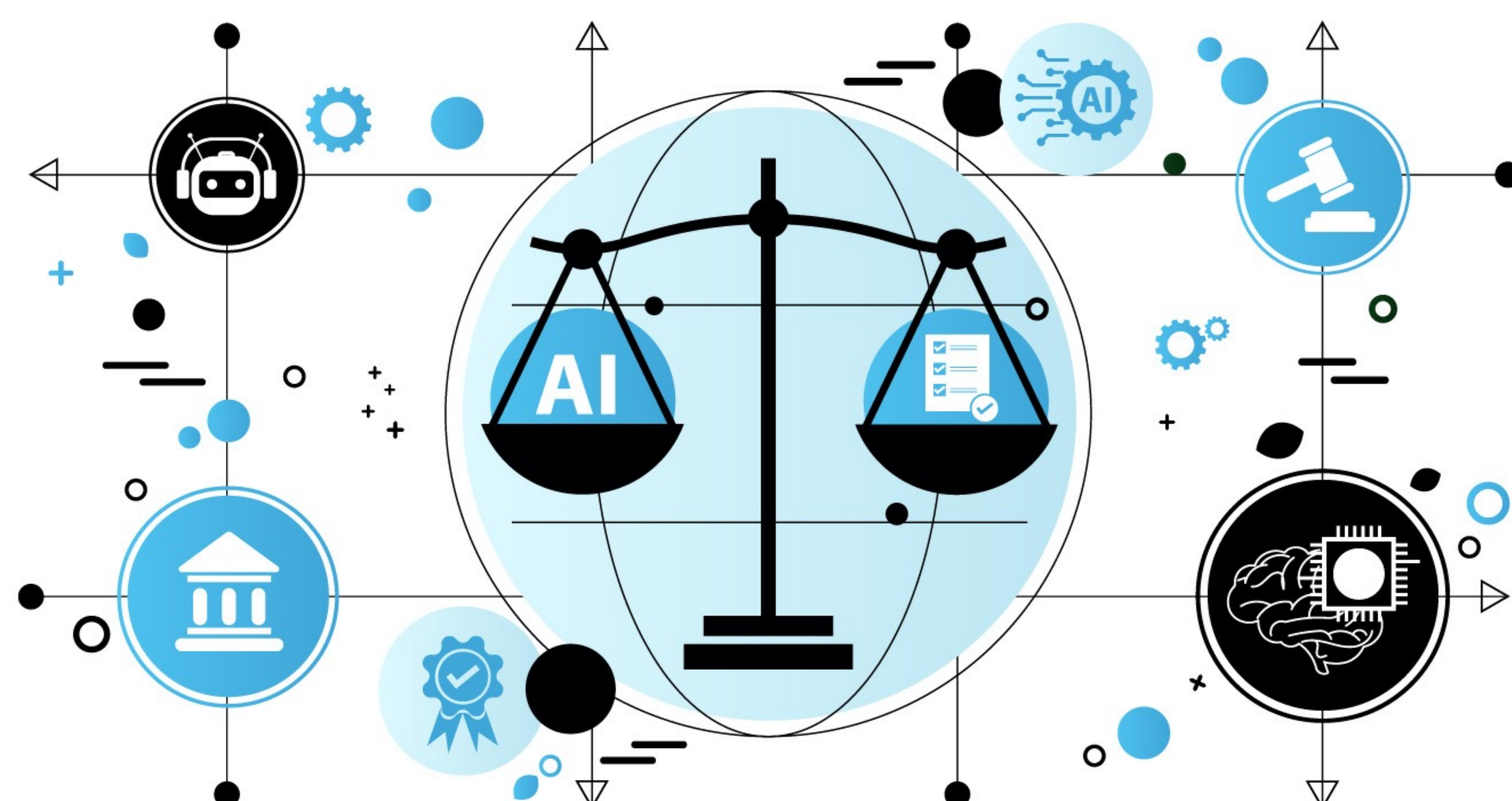
5. 尊重隱私：

涉及個人資訊的蒐集、處理、共享、儲存和刪除，應該要符合個資相關法規的規範。

6. 韌性：

面臨複雜且多變的環境時，AI必須維持可穩定運行的適應能力。在無法正常運作的情況下應提供警告，而非給出不完整的答案。

AI對於資訊安全也是一個重大議題，需要更多專家學者投入進行研究，研究方向除了如何運用AI偵查入侵之外，也應該包含如何防止駭客運用AI作為入侵手段。人類社會已經經歷過多次科學革命帶來的挑戰，社會是一個整體，政府部門、公益組織、各領域專家都有責任共同維護社會的安全，不能只依靠AI公司自己努力。



4 四通八達·小試身手



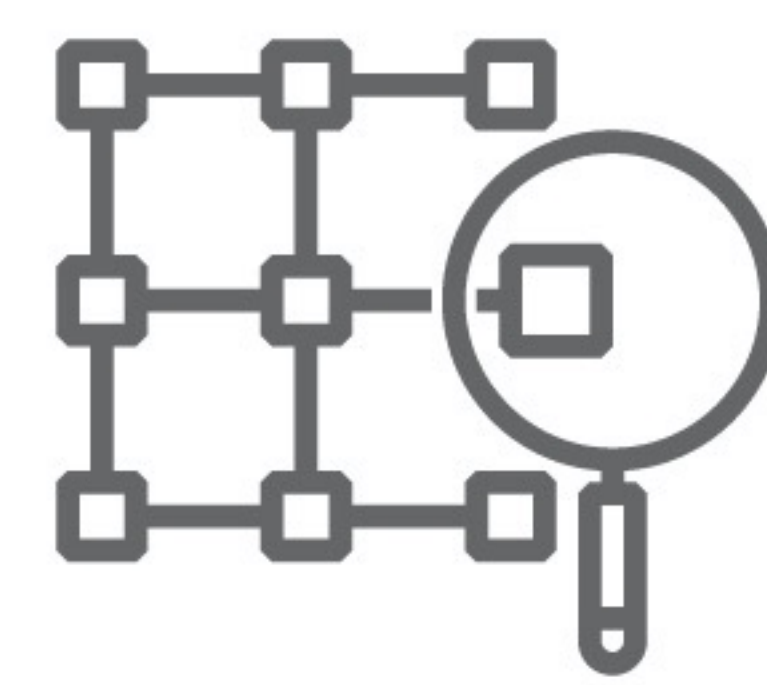
一、請針對以下情境，各寫一句完整的提示詞：

目 標	我的提示詞
寫一篇科幻短篇小說的前情提要	
生成一份健康飲食的每週計畫表	
分析網路購物的優點與缺點，限200字	

二、實際使用AI測試第一大題的提示詞，並將提示詞中使用到的關鍵字進行分類，檢視這些關鍵字是否產生預期的效果。

類 別	使用的關鍵字	有作用嗎？有什麼建議？
文章結構提示		
文章風格提示		
敘事角度與人稱		
回應條件與限制		

5 五彩繽紛·旁徵博引



- 〔1〕深偽技術。維基百科。民114年2月24日，取自：<https://zh.wikipedia.org/zh-tw/深偽技術>
- 〔2〕何謂深偽（Deepfake）趨勢科技。民114年2月24日，取自：https://www.trendmi.cro.com/zh_tw/what-is/ai/deepfakes.html
- 〔3〕顏均萍、甘偵蓉（2023年4月12日）。【AI的多重宇宙】UNIVERSE-02 偏見與歧視。臺灣大學科學教育發展中心。民114年2月24日，取自：<https://case.ntu.edu.tw/blog/p=41761>
- 〔4〕什麼是AI幻覺？Google Cloud。民114年2月24日，取自：<https://cloud.google.com/discover/what-are-ai-hallucinations?hl=zh-TW>
- 〔5〕林以璿（2024年2月20日）。「防止中國AI文化侵略」台灣第一個繁體中文大語言模型TAIDE，能做什麼？。天下雜誌791期。民114年2月24日，取自：<https://www.cw.com.tw/article/5129076>
- 〔6〕行政院及所屬機關(構)使用生成式AI參考指引（民112年8月31日）。行政院教育科學文化處。民114年2月24日，取自：<https://www.ey.gov.tw/Page/448DE008087A1971/40c1a925-121d-4b6b-8f40-7e9e1a5401f2>
- 〔7〕黃維中（2022年12月25日）。Trustworthy AI：邁向可信任的人工智慧治理。工業技術研究院。民114年2月24日，取自：<https://ictjournal.itri.org.tw/xcdoc/cont?xsmsid=0M208578644085020215&sid=0M348549487608256517>
- 〔8〕黃詩淳（2023年11月）。AI可解釋性的法學意義及其實踐。民114年2月24日，取自：https://homepage.ntu.edu.tw/~schhuang/Publication_files/publication_articles/臺大法學論叢/52_S_931AI可解釋性.pdf
- 〔9〕Trust and Artificial Intelligence (NISTIR 8332) (2021, March 2). Available at: <https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8332-draft.pdf>
- 〔10〕AIRisk Management Framework 2nd Draft (2022, August 18). Available at: <https://www.nist.gov/document/ai-risk-management-framework-2nd-draft>
- 〔11〕Dylan Patel and Gerald Wong(2023, July 10)，GPT-4 Architecture, Infrastructure, Training Dataset, Costs, Vision, MoE. Available at: <https://semianalysis.com/2023/07/10/gpt-4architecture-infrastructure/>



語詞解釋：

- 〔1〕肯恰那：是起源於2025年流行的短影音舞蹈歌曲，原曲源自於越南，因歌詞發音似韓語「沒關係」的諧音，在網路竄紅。
- 〔2〕綠幕：是一種去背合成技術，把被拍攝的人物或物體放置於綠色布幕的前面，並進行去背後，將其替換成其他的背景。
- 〔3〕深偽技術（Deepfake）：基於人工智慧的人體圖像、語音合成技術的應用。此技術可將已有的圖像、聲音或影片疊加至目標圖像或影片上，因此被有心人濫用於換臉、偽造身份等非法用途。
- 〔4〕資料偏誤：是指會導致結果與事實之間存在差異的系統性傾向。數據分析的許多過程，包括數據的來源、選擇的估計量和分析數據的方式，都可能存在偏誤。
- 〔5〕ChatGPT：OpenAI開發的人工智慧聊天機器人程式，於2022年12月推出。該程式使用基於GPT-3.5、GPT-4、GPT-4o架構的大型語言模型並以強化學習訓練。ChatGPT一開始只接受單一模態使用文字互動，新的版本已經轉變成多模態 AI，可以使用語音、影像方式互動。
- 〔6〕Suno AI：一款生成式人工智慧音樂創作程序，旨在產生人聲與樂器相結合的逼真歌曲。該程式根據用戶提供的提示詞來產生歌詞和歌曲。
- 〔7〕多模態AI：具備更進階的生成能力，可以處理圖片、影片和文字等多種形式的資訊。多模態就好比賦予AI處理及瞭解不同感官模式的能力。具體來說，輸入內容和輸出內容不再受限於單一類型，您可以給予近乎任何類型的提示，並生成絕大多數的內容類型。
- 〔8〕AGI：從字面翻譯為「通用人工智慧」，是指具備與人類同等智慧或超越人類的人工智慧，能表現出正常人類所具有的所有智慧行為。也有些學者譯為「統合思考人工智慧」。
- 〔9〕提示詞：提示詞工程是您引導生成式AI解決方案以產生所需輸出的程序。儘管生成式AI設法模仿人類，但它仍需要詳細的指示來建立高品質和相關的輸出。在提示工程中，您可以選擇最合適的格式、片語、單詞和符號，以引導AI更有意義地與使用者互動。
- 〔10〕RAG擷取擴增生成：對大型語言模型輸出最佳化的過程，就是在產生回應之前，讓它參考其訓練資料來源以外的權威知識庫。大型語言模型(LLM)在大量資料上訓練，並使用數十億個參、數來生成原始輸出，用於回答問題、翻譯語言和完成句子等任務。RAG將原本就很強大的LLM功能擴展到特定領域或組織的內部知識庫，而無需重新訓練模型。這是改善LLM輸出具成本效益的方法，可讓LLM在各種情況下仍然相關、準確且有用。

<

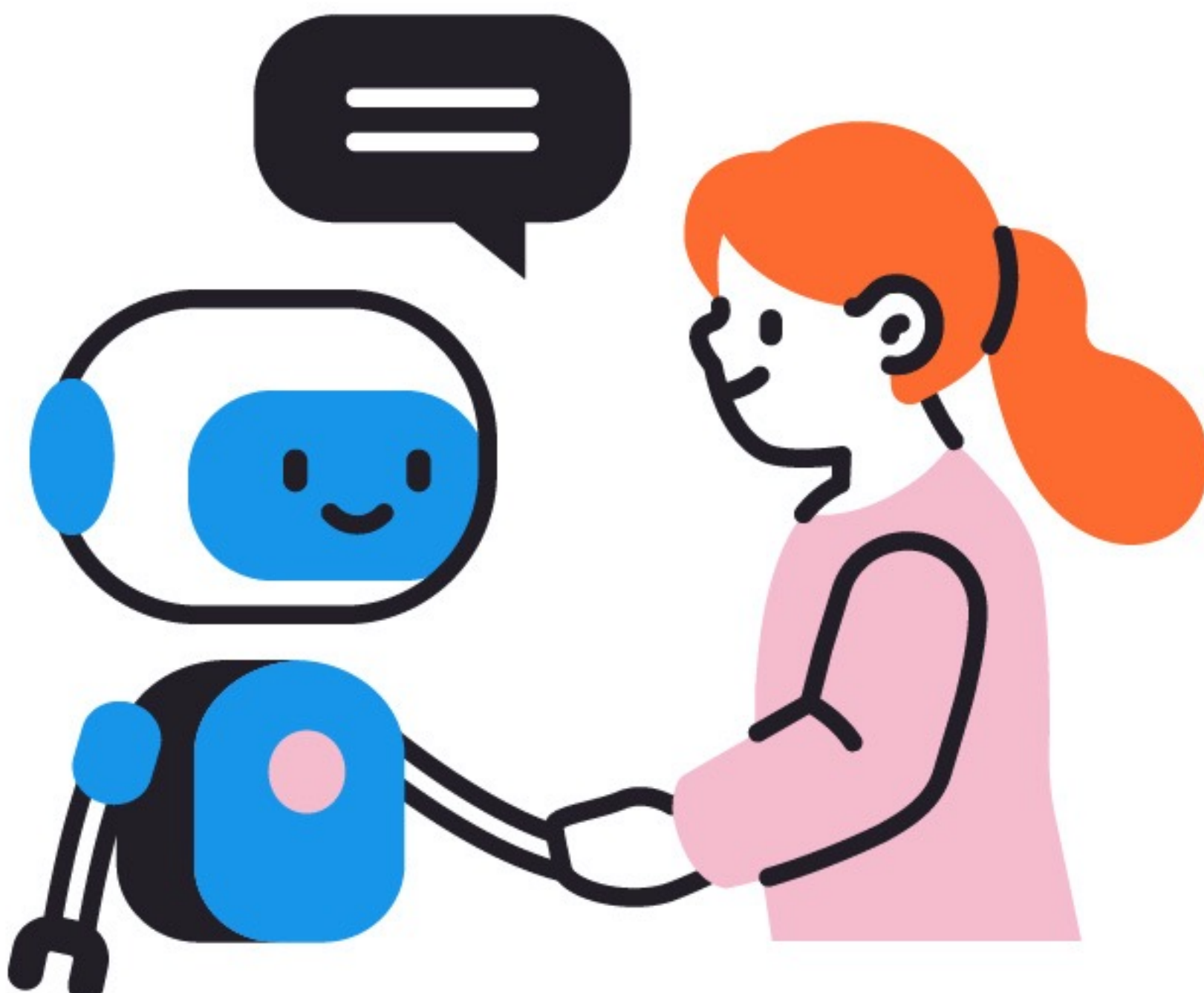
>

學習結語



有圖未必是真相，真假難辨費思量，
生成技巧要磨練，深度交談靈感現。





MEMO

MEMO

資訊素養與倫理 國小 5 版

出版 / 臺北市政府教育局

召集人 / 湯志民 臺北市政府教育局局長

副召集人 / 卓育欣 臺北市政府教育局資訊教育科科长

林裕勝 臺北市文山區永建國民小學校長

指導委員 / 盧東華 臺北市立大學助理教授

楊凱翔 國立臺北教育大學教授

總編輯 / 張仁音 臺北市文山區永建國民小學教務主任

執行編輯 / 蔡佳儒 臺北市文山區永建國民小學專案教師

編審委員 / 江淑娟 臺北市中山區長安國民小學資訊組長

何詩慧 臺北市日新自造教育及科技中心主任

李忠憲 臺北市國語實驗國民小學教師

吳建勳 臺北市大安區新生國民小學教師

花梅真 臺北市北投區明德國民小學教師

姜保煌 臺北市士林區社子國民小學資訊組長

張宇軒 臺北市文山區永建國民小學資訊組長

蔡明佑 臺北市中正區東門國民小學學務主任

蔡孟憲 臺北市松山區健康國民小學學務主任

(按姓氏筆畫排列)

承辦單位 / 臺北市文山區永建國民小學

出版日期 / 中華民國114年3月

AI 未來新世界

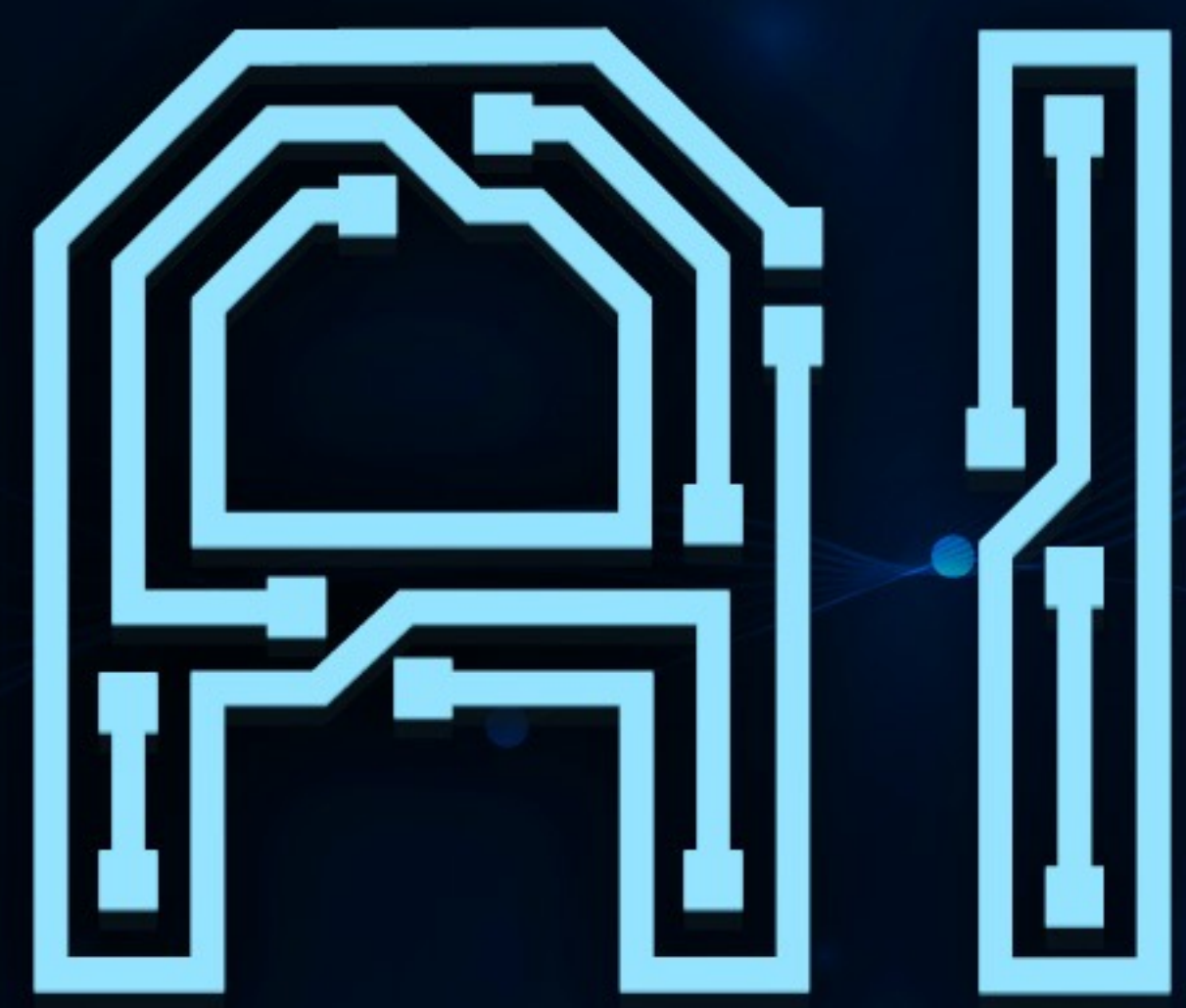
弄假成真-生成式AI



臺北市政府教育局
DEPARTMENT OF EDUCATION
TAIPEI CITY GOVERNMENT

資訊素養與倫理 國小 5 版

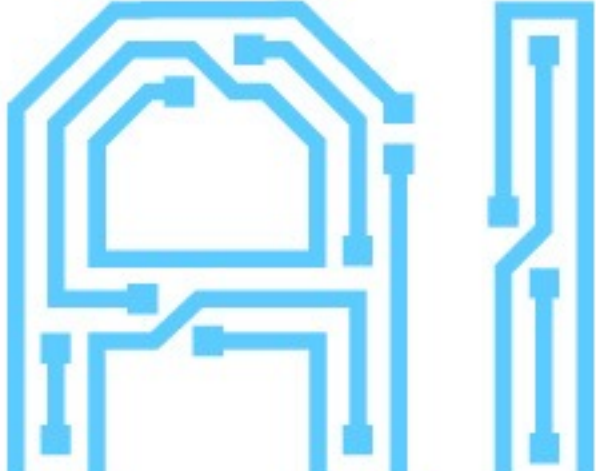
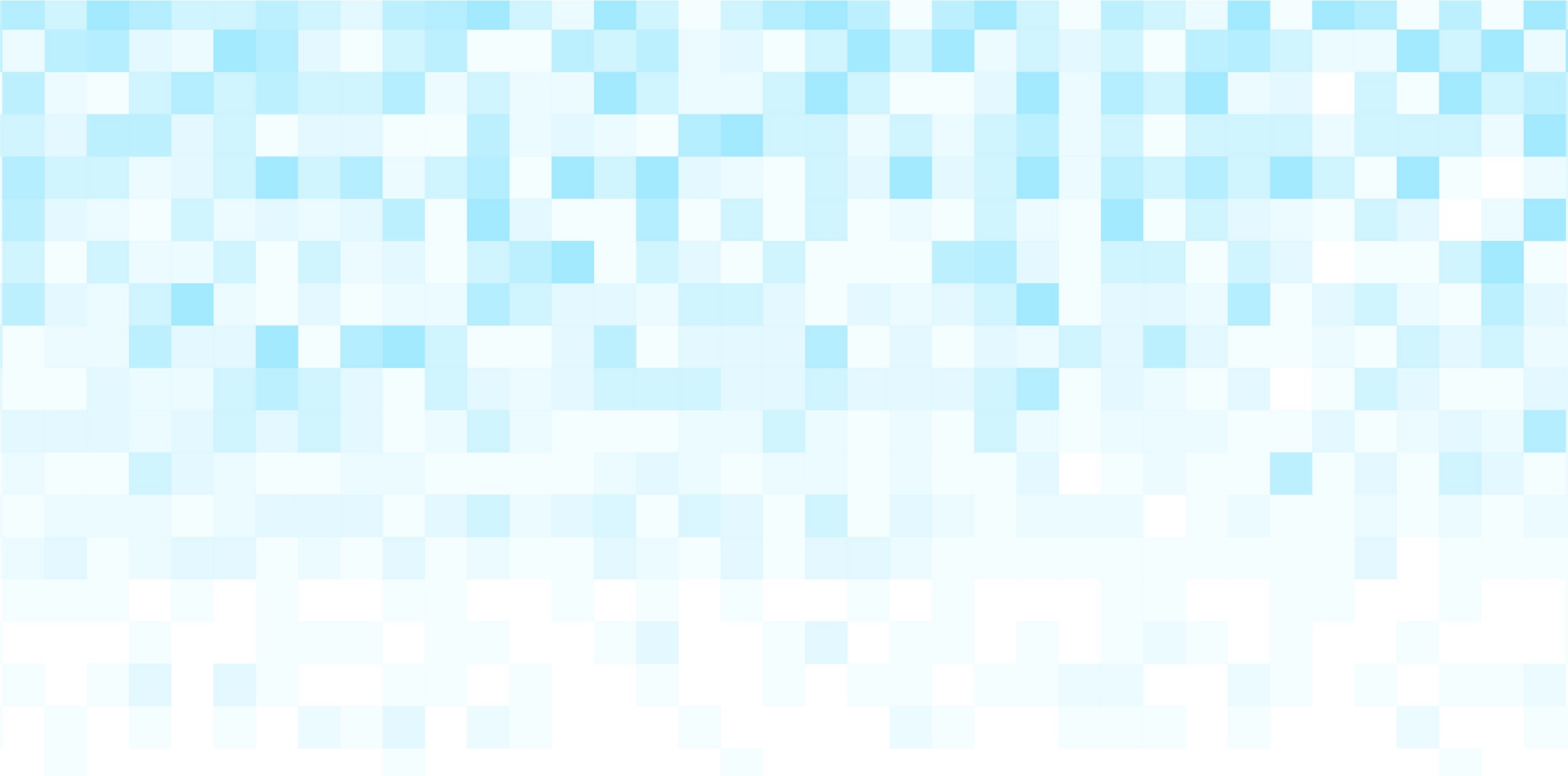
編撰：臺北市松山區健康國民小學 蔡孟憲
審稿：臺北市大同區日新國民小學 何詩慧



未來新世界

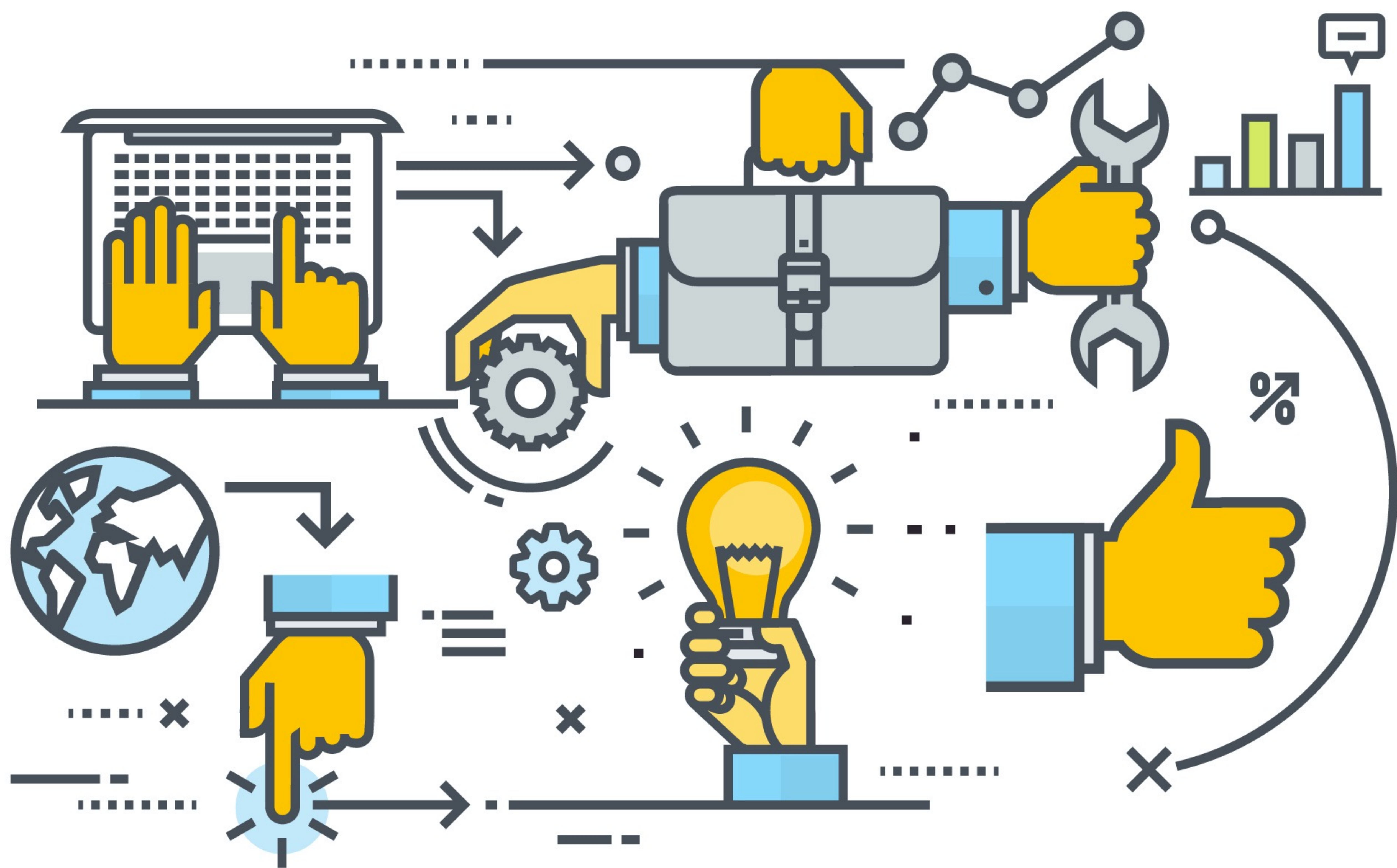
人機對決-AI參加競賽之相關規則及防範





未來新世界

人機對決-AI參加競賽之相關規則及防範



資訊素養與倫理教材第五版【國小篇】

教學綱要向度與內涵

01

學習表現

- 資議 a - III - 2 建立健康的數位使用習慣與態度。
- 資議 a - III - 3 遵守資訊倫理與資訊科技使用的相關規範。

02

學習內容

- 資議 T - III - 2 網路服務工具的應用。
- 資議 H - III - 2 資訊科技合理使用原則的理解與應用。

03

學習重點

- 認識AI在各類型競賽中的影響與應用。
- 理解不同類型的競賽需要不同的AI使用規範。
- 了解使用AI參加競賽的相關規範與倫理界線。
- 學習如何在競賽中正確且負責任地運用AI。

1 一葉知秋・學思並進



電腦課前，沃德和小皮正熱烈討論著即將到來的作文比賽，而偉柏則神情嚴肅地盯著手中的比賽簡章，臉色越來越凝重……

沃德：「這次作文比賽的主題是『科技與生活』，我一定要好好發揮，拿個好名次！」

小皮：「我也想參加，但我只希望我能寫得完整就好了……」

偉柏突然抬起頭，眼神充滿憂慮：「你們有沒有仔細看這次的比賽規則？」

小皮：「規則？就那些啊，字數限制、格式要求……有什麼特別的嗎？」

沃德也湊過來看：「對啊，我也沒發現有什麼奇怪的地方。」

偉柏搖搖頭，指著簡章上的一小段文字：「你們看這裡，關於AI使用的說明……」

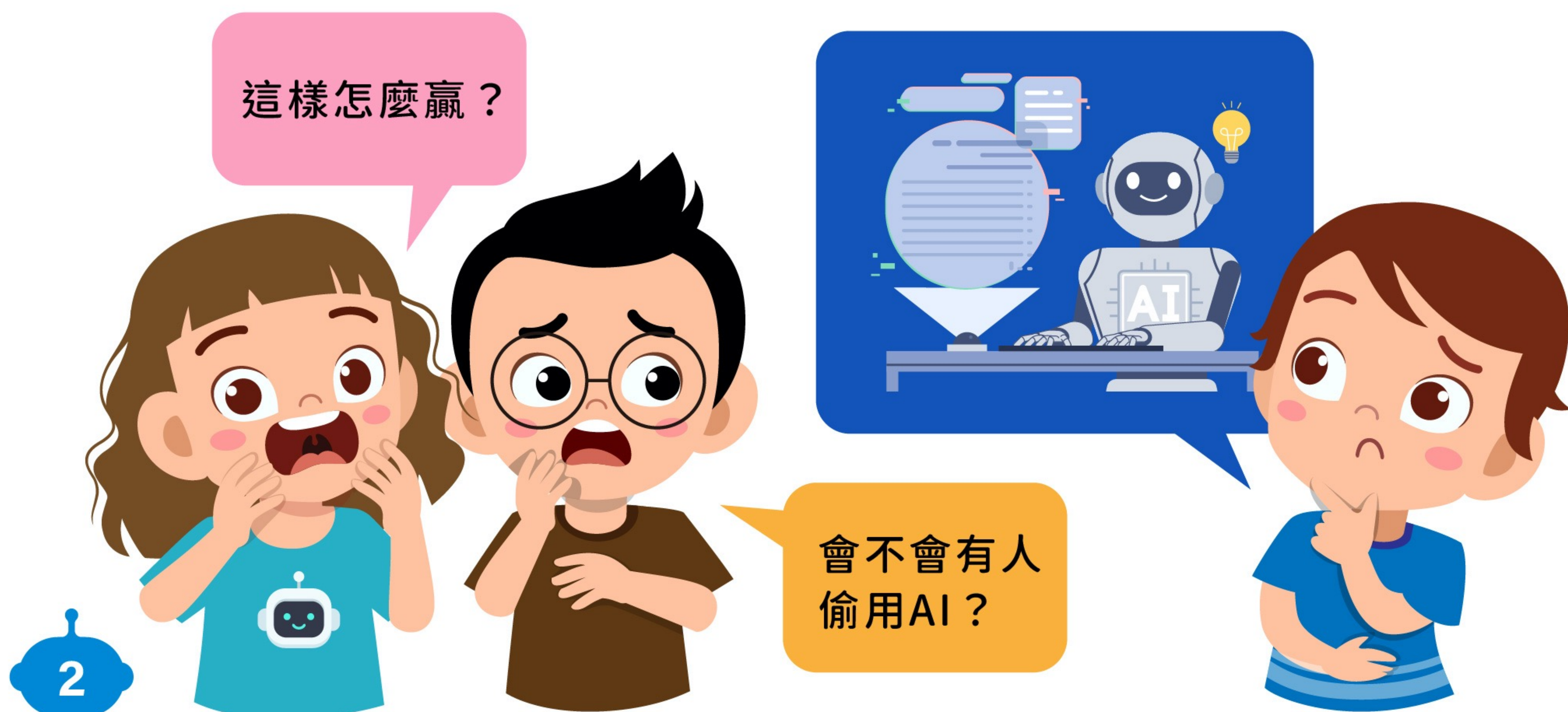
沃德和小皮順著偉柏的手指看去，只見那段文字寫著：「本比賽鼓勵學生發揮原創性，如使用AI輔助，需於作品中明確標示，違者取消參賽資格。」

小皮不以為意地說：「這很正常啊，就是要我們誠實標示嘛。」

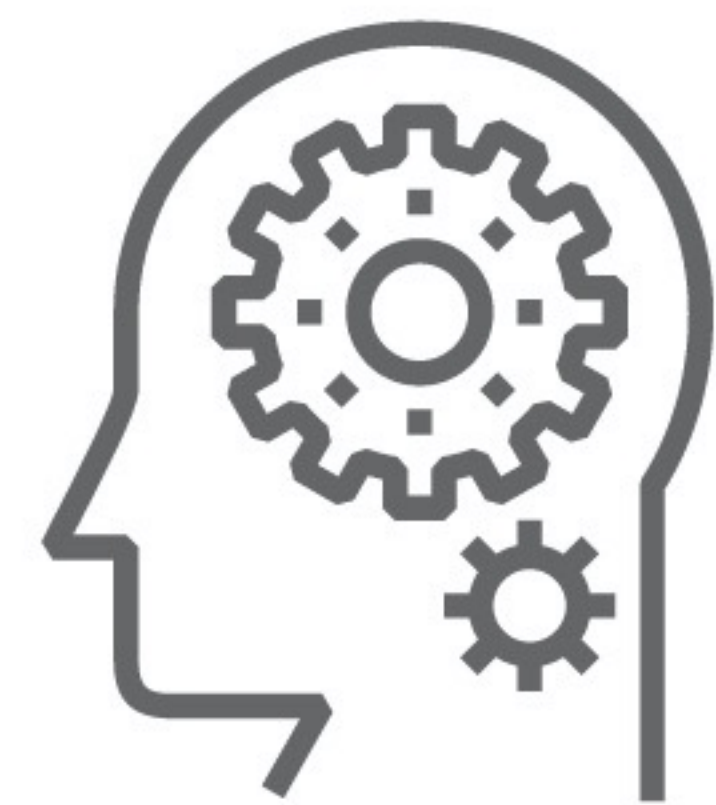
偉柏卻更加不安地說：「事情沒那麼簡單……你們有沒有想過，如果有人偷偷使用AI，卻沒有標示呢？」

沃德深吸一口氣說：「對喔！那……那我贏得過AI嗎？」

小皮聽了，臉色也變得凝重起來。就在這時，上課鐘聲也響起了。



2 兩全其美・觸類旁通



一、世紀大對決—人腦 vs 電腦

李老師走進教室，發現同學們正圍著偉柏，七嘴八舌地討論著什麼。

李老師：「大家在聊什麼呢？這麼熱鬧。」

沃德：「老師，我們在討論這次的作文比賽，偉柏說他擔心有人用AI作弊！」

李老師笑了笑：「哦？怎麼說？」

偉柏：「老師，您知道以前IBM的超級電腦『深藍（Deep Blue）』打敗西洋棋世界冠軍加里·卡斯帕洛夫（Garry Kasparov）的故事嗎？如果有人在作文比賽中偷偷用AI，我們根本沒辦法獲勝！」

小皮：「對啊，老師！這樣很不公平！」

李老師點點頭：「你們的擔心很有道理，AI的確在某些領域超越了人類，尤其是在邏輯和計算方面。就像剛剛偉柏提到的深藍，它是歷史上第一臺在正式比賽中擊敗世界冠軍的電腦，但你們知道嗎？當時的深藍其實並不是真正的『智慧』，它主要是依靠強大的計算能力，每秒分析數百萬步棋局來決定它的下一步棋。」



李老師繼續說道：「西元2000年之後，AI在棋類競賽的技術發展開始突飛猛進。2016年，AlphaGo擊敗了世界頂尖圍棋高手李世乭，這次不是靠暴力計算，而是學習人類棋譜，運用『深度強化學習』來思考最佳棋步。」

「哇！AI真的好厲害！」佩琪驚嘆道。

「更厲害的還在後面！」李老師繼續說：「2017年，AlphaZero出現了，它不再依賴人類棋譜，而是透過自我對弈學習，自己和自己下棋，短短幾個小時就超越所有過去的AI棋手。」

「那現在最強的棋手是哪個AI？」沃德好奇地問。

「目前，MuZero比AlphaZero更進一步，因為它甚至不需要事先知道任何遊戲規則，而是透過觀察來學習規則並進行最佳決策，這不僅在棋類競賽上強大，在其他領域也能應用。」

小皮皺著眉：「可是，這樣人類還有機會贏AI嗎？」

李老師：「這就是我們今天要討論的重點。AI在棋類遊戲上的勝利，其實給我們很多啟示，就像AlphaGo和李世乭的圍棋大戰，李世乭雖然輸了，但他當時有名的一步『神之一手』，展現了人類的獨特智慧。回到作文比賽，AI可以生成流暢的文章，但它很難有真正的情感和獨特的觀點。這就是人類可以勝過AI的地方。」

米迪：「老師，所以我們還是要相信自己，努力寫出最好的作文！」

李老師：「沒錯！而且也要了解AI的限制。AI是工具，它可以幫助我們，但不能取代我們。就像棋類AI，它很擅長計算，但對棋局並非真正理解。我們要學習如何善用AI，同時也要保持自己的原創性和思考能力。」



問題與討論

1. 除了棋類，還有哪些競賽可能會受到AI影響？
2. 人類在面對AI的挑戰時，有哪些獨特的優勢？
3. 你認為未來AI是否應該被允許參加所有類型的比賽？



二、競賽規則面面觀

「老師～」伊妹舉手發問：「所以這次作文比賽，到底能不能用AI啊？規則寫得好模糊喔！」

李老師：「的確，規則的制定是個難題。每個比賽對AI的使用規定都不一樣。有的可以用，但要註明，有的鼓勵參賽者和AI合作，有的則完全禁止AI輔助！像是國際象棋比賽，早就有明確的規定，禁止選手在比賽期間使用電腦輔助。前陣子有選手因為偷偷用AI作弊，被終身禁賽！」

佩琪驚訝地說：「哇！終身禁賽？這麼嚴重？！」

李老師：「是的，因為AI的計算能力太強，如果允許使用，就不再是比人類智慧，而是比誰的AI更厲害。」

米迪：「那如果只是用AI來進行前期準備呢？像是讓AI幫忙分析過去的比賽數據，或是練習對手的風格？」

李老師：「這就涉及到一個重要概念——『AI輔助』和『AI主導』的界線。在許多競賽中，允許參賽者在賽前使用AI工具，但最終的決策還是必須由人類來做。」

沃德若有所思地說：「那作文比賽應該也要訂出明確的規則，像是AI可以當作輔助工具，但不能讓它直接寫完整篇文章？」

李老師點頭：「沒錯，完全禁止AI，可能會扼殺學生探索新科技的機會；完全開放，又可能導致比賽失去公平性。所以，這次比賽採取的是『有條件開放』。」

沃德：「什麼是有條件開放？」

李老師：「簡單來說，就是允許使用AI輔助，但必須誠實標示，並且在作品中清楚標示哪些部分是AI生成的。」

偉柏：「那如果有人故意隱瞞呢？要怎麼發現？」

李老師：「這就是規則制定困難的地方。目前，要完全準確地判斷一篇作品是否由AI生成，仍然非常困難。不過，我們可以從幾個方面來判斷。」



李老師在黑板上寫下幾個要點：

風格一致性：AI生成的文章，有時會出現風格不一致的問題。

例如，前後文的用詞、語氣差異過大。

內容真實性：AI可能會生成一些缺乏事實根據，或與常識不符的內容。

原創性分析：可利用AI檢測工具，分析文章的原創性。

雖然這些工具並非百分之百準確，但可以作為參考。



佩琪：「老師，那如果我用了AI，但是有修改過，這樣也要標示嗎？」

李老師：「是的，只要有使用AI輔助，無論修改程度如何，都應該誠實標示。這不僅是比賽的要求，也是一種負責任的態度。」

小皮：「可是，如果我只是用AI查資料，這樣也要標示嗎？」

李老師：「單純的資料蒐集，不算是AI輔助寫作。但如果你是使用AI工具來整理、分析資料，並將其直接運用到你的文章中，那就需要標示。」

米迪：「我覺得好複雜喔！一不小心就會違反規定。」

佩琪：「這樣的話，未來是不是很多比賽都會加入AI使用規範？」

李老師：「很有可能，因為AI的發展速度很快，我們現在所討論的規則，可能不久後就需要再調整。」

米迪：「那我們參加比賽時，就要特別注意規則，不然可能會不小心違規！」

李老師：「是的，這就是為什麼要仔細閱讀比賽規則，並且誠實面對自己。記住，比賽的重點不只是結果，更重要的是過程。透過這次比賽，我們可以學習如何正確、負責任地使用AI。」



問題與討論



1. 除了上述提到的判斷方法，你認為還有哪些方法可以偵測AI作弊？
2. 你認為哪些競賽項目應該完全禁止使用AI？為什麼？
3. 如果比賽允許使用AI，評分標準應該如何調整？

三、AI輔助的倫理界線

「老師～」偉柏舉手發問：「只要誠實標示用了AI輔助，這樣就完全沒問題了嗎？如果AI幫忙太多，那這樣還算是我們自己的作品嗎？」

李老師：「這是一個非常好的問題！即使誠實標示，我們仍需要思考AI輔助的倫理界線。」

佩琪：「倫理界線？是什麼意思？」

李老師：「簡單來說，就是在使用AI時，應該遵守的道德規範和行為準則。例如，我們不能為了追求好成績，過度依賴AI，甚至讓AI完全取代我們的思考和創作。」

小皮：「可是，如果我用了AI，文章寫得更好、更流暢，這樣不是很好嗎？」

李老師：「當然，AI可以幫助我們提升文章的品質。但是，使用時也要注意，不能讓它完全控制我們的作品。AI應該是我們的助手，而不是主人。」

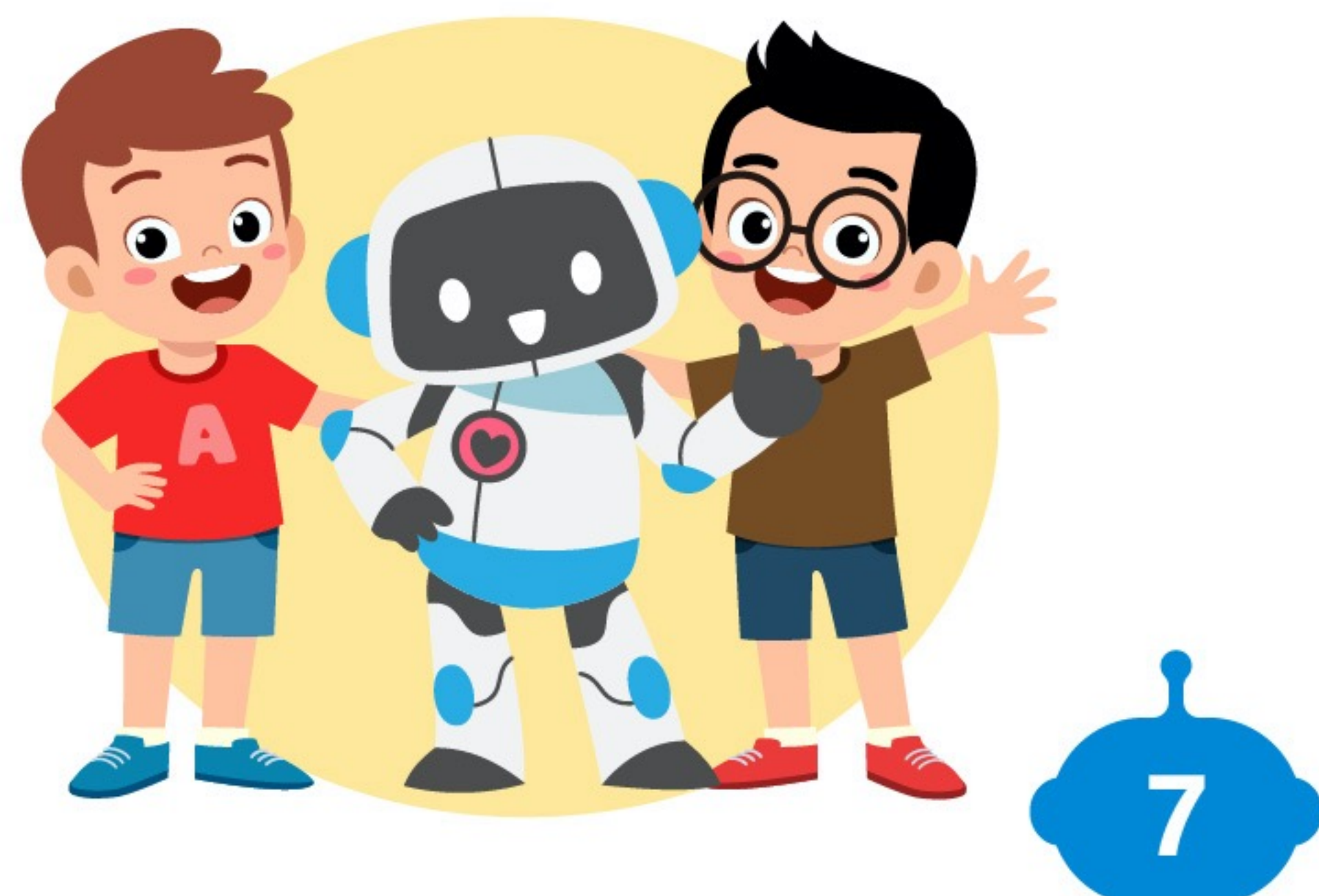
偉柏：「我覺得可以把AI當作一種工具，而不是讓它取代我們。例如，讓AI幫忙整理資料，但最終的分析和寫作還是要靠自己完成。」



李老師：「說得好！這樣既能充分利用AI的優勢，又不會失去自己的創意和努力。」

米迪：「我還是不太懂……那要怎麼判斷是否過度依賴AI？」

李老師：「可以從幾個方面來思考：首先，要清楚了解AI在文章中扮演的角色。AI是提供靈感、修改語法，還是直接生成大部分的內容？其次，要確保文章的核心觀點和思想，仍然是來自我們自己的思考和判斷。最後，要對AI生成的內容進行審核和修改，確保其符合我們的價值觀和道德標準。」



李老師接著說：「除了避免過度依賴，還要尊重AI的智慧財產權。如果使用了AI生成的內容，就應該清楚標示來源，並遵守相關的授權協議。」

沃德：「智慧財產權？AI也有智慧財產權嗎？」

李老師：「這是一個非常複雜的問題，目前法律界還在討論。但我們應該抱持著尊重的態度，不要隨意抄襲或盜用AI生成的內容。」

偉柏：「老師，那如果AI生成的內容，侵犯了別人的權益，責任應該由誰來承擔？」

李老師：「這也是一個很重要的問題。目前法律上對於AI侵權的責任歸屬，還沒有明確的規定。但是，作為AI的使用者，我們有責任確保AI生成的內容，不會侵犯他人的名譽、隱私或其他權益。在使用AI輔助時，必須小心謹慎，遵守倫理規範，才能真正發揮AI的正面作用。」

「我懂了！」佩琪點點頭：「就像使用網路資料一樣要註明出處，避免侵犯到著作權。」

李老師：「是的，AI輔助也是一樣的道理。有些競賽開放AI輔助，但評分標準也會調整，例如，更加強調原創性和創意，而不僅僅是文章的流暢度。這時候，比的不只是誰的AI用得好，還有誰能最有效地結合AI與個人創意。」

偉柏若有所思地點點頭：「這樣我們就要更聰明地運用AI，而不是單純依賴它。」

「沒錯。」李老師微笑著說，「AI是時代的趨勢，我們應該學習如何與它共存，並且以負責任的態度來使用它，尊重他人的權益，並保持自己的獨立思考能力。只有這樣，才能在AI時代，成為一個有素養、有倫理的數位公民。」



問題與討論



1. 如果比賽主辦方沒有明確規定AI的使用規範，參賽者應該如何自律？
2. 你認為在什麼情況下使用AI輔助參賽是合適的？什麼情況下是不合適的？
3. 如果比賽允許使用AI，但AI生成的內容與你的實際經驗不符，你覺得這樣的AI輔助有價值嗎？

3 三絕韋編・鑑往知來



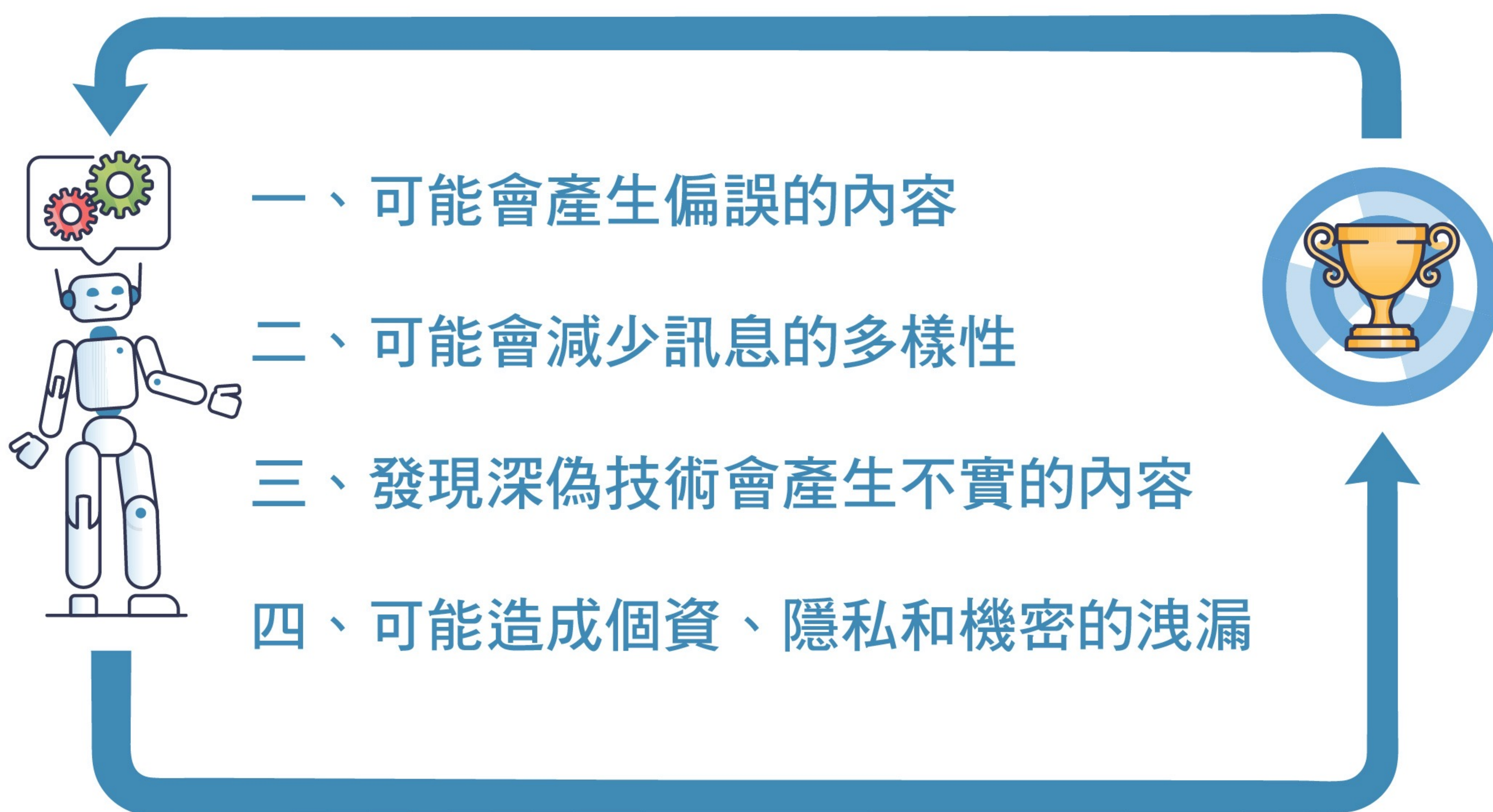
一、棋類競賽AI的重要發展歷程

- ◆ **1997年**：IBM的超級電腦「深藍」（Deep Blue）在西洋棋比賽中擊敗了世界冠軍加里·卡斯帕羅夫（Garry Kasparov）。深藍主要依靠強大的運算能力和龐大的棋譜資料庫，透過快速搜尋和評估大量的棋局來做出決策，還算不上真正的AI。
- ◆ **2016年**：Google DeepMind的AlphaGo使用蒙地卡羅樹搜尋（Monte Carlo tree search，MCTS）和兩種深度神經網路（策略網路和估值網路）成功擊敗韓國圍棋冠軍李世乭。AlphaGo的成功在於其結合了深度學習和強化學習，透過分析大量的圍棋棋譜和自我對弈，學習圍棋的策略和技巧。
- ◆ **2017年**：DeepMind發表了AlphaGo Zero和AlphaZero。AlphaGo Zero不需要任何人類棋譜就能學會圍棋，並且超越了所有之前的AI版本。AlphaZero除了圍棋，還能下西洋棋和日本將棋，展示了在多種棋類遊戲中的通用性和強大能力
- ◆ **2020年**：DeepMind 再度發表新一代AI系統 MuZero，能夠在完全不知規則的情況下，熟悉圍棋、西洋棋和將棋，以及57款Atari遊戲，可望用來解決現實生活中，規則太複雜或完全不知規則的各式難題。



二、使用AI參加競賽之相關規則

使用AI參賽的相關規則，通常依照比賽性質、主辦單位的規範及公平競爭原則制定。AI技術改變了我們原本獲取知識、傳播訊息和創造內容的方式，也迫使越來越多的各類型競賽必須明定使用AI參賽的規則，例如：「第二十六屆臺大文學獎」首次試辦使用AI軟體共作項目、「2024永續智慧創新黑客松競賽」參賽辦法明訂使用人工智慧(AI)科技輔助相關規定…等，都是因應AI技術所訂定之比賽規則，參賽者除了遵守比賽規則以外，教育部亦針對中小學學生訂定「中小學使用生成式人工智慧注意事項（學生版）」，提供以下四個要點作為使用參考：



三、使用AI產生的智慧財產權議題

AI技術日益發達，在使用AI所產生的智慧財產權議題，例如：誰擁有AI生成內容的著作權？用別人的資料訓練AI是否有侵權疑慮？AI生成內容是否涉及模仿或抄襲？生成的內容是否能夠被註冊為商標、專利或作為商業財產？這些議題在各國之間想法不同，尚未有共識，仍待各國政府立法來規範，但在人類對這些議題有共識之前，以下是一些建議的做法：

(1) 提高個人創作的參與程度、避免直接複製或模仿

- 明確參與：在AI生成作品中，應該明確加入個人的創意，例如：提供具體指令、參數設定、選擇素材等，以確保作品具有獨特的創作元素。
- 作品修改與編輯：對AI生成的作品進行進一步的修改、編輯，以展現個人的創意和原創性。

(2) 謹慎使用AI訓練數據

- 檢視合理使用條件：在使用AI學習或生成作品時，應確保符合著作權法的合理使用條件，例如：轉換、批評、新聞報導、教學和研究等。
- 避免侵權：如果AI生成的作品涉及他人的著作權，應避免未經授權的使用，以免構成侵權。

(3) 標示AI生成內容，避免誤導

- 在公開發表或商業使用時，標註「由AI生成」以避免誤導受眾。
- 某些國家或平台可能要求標示AI生成內容，遵守相關規定可降低法律風險。

(4) 了解AI工具的使用條款、注意商業用途

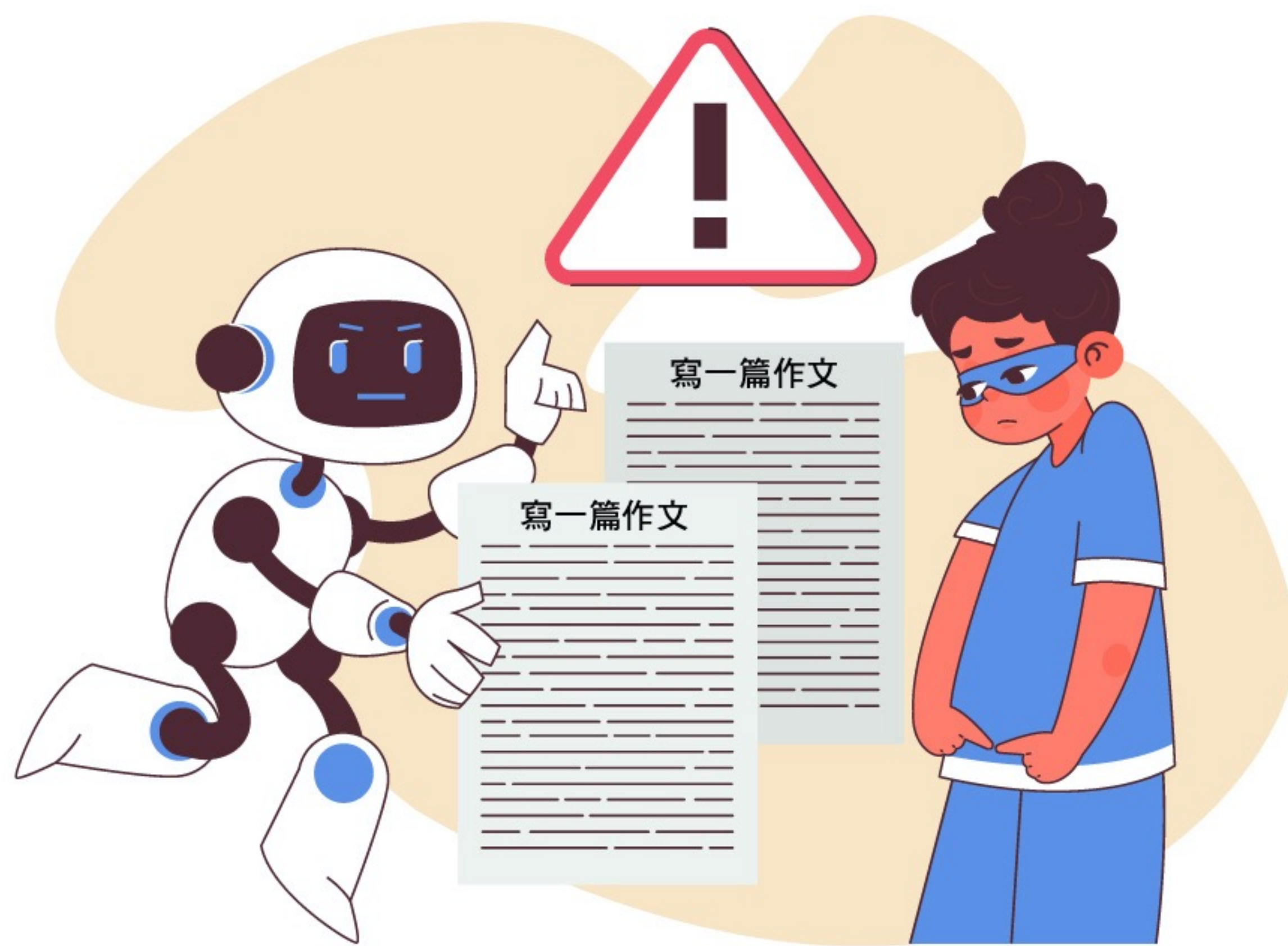
檢視使用條款：在使用AI工具時，應仔細閱讀並理解其使用條款，確保不會違反相關的著作權規定。

(5) 關注法律發展

依循法律規定：隨著AI技術的發展，各國的法律規範也在不斷更新，因此應關注相關的法律變化，以確保創作行為的合法性。

(6) 尋求專業意見

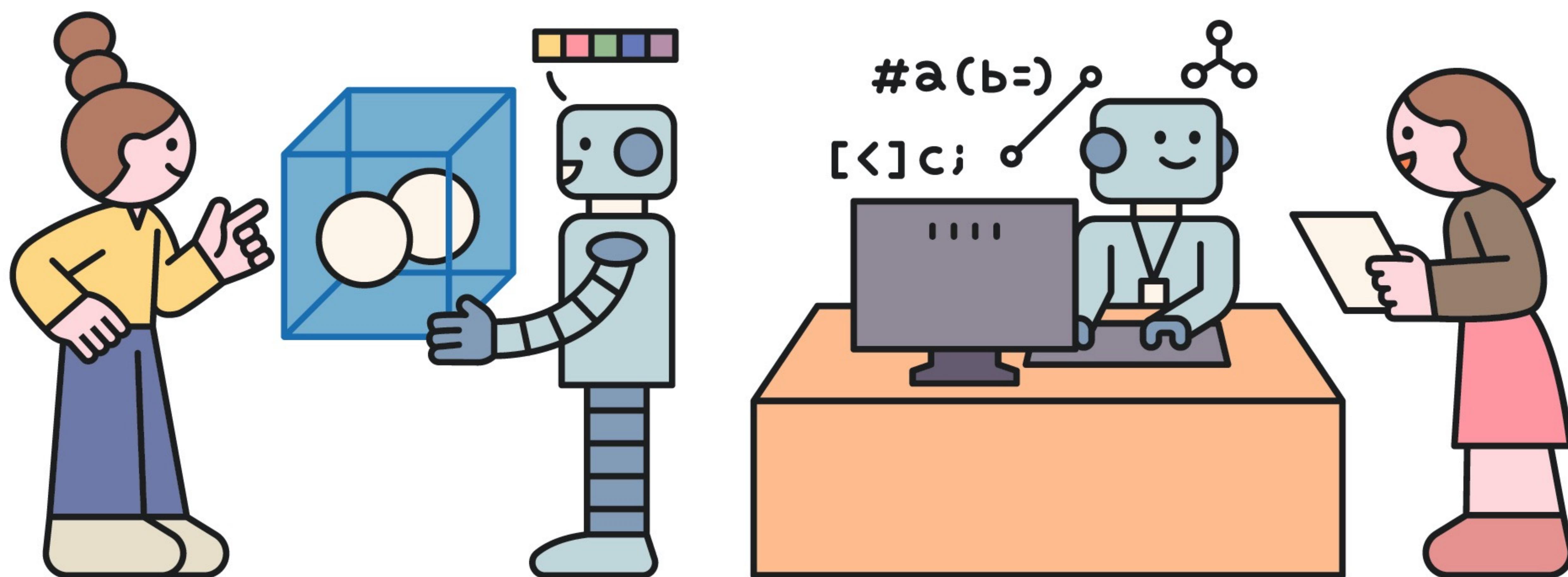
諮詢律師或專家：如果對AI生成作品的著作權問題有疑慮，應尋求律師或智慧財產權專家的意見，以確保創作行為的合法性。



4 四通八達·小試身手



- 一、請分享一個你認為可以使用AI輔助的競賽項目，並說明如何規範參賽者運用AI才能兼顧公平性和創意？

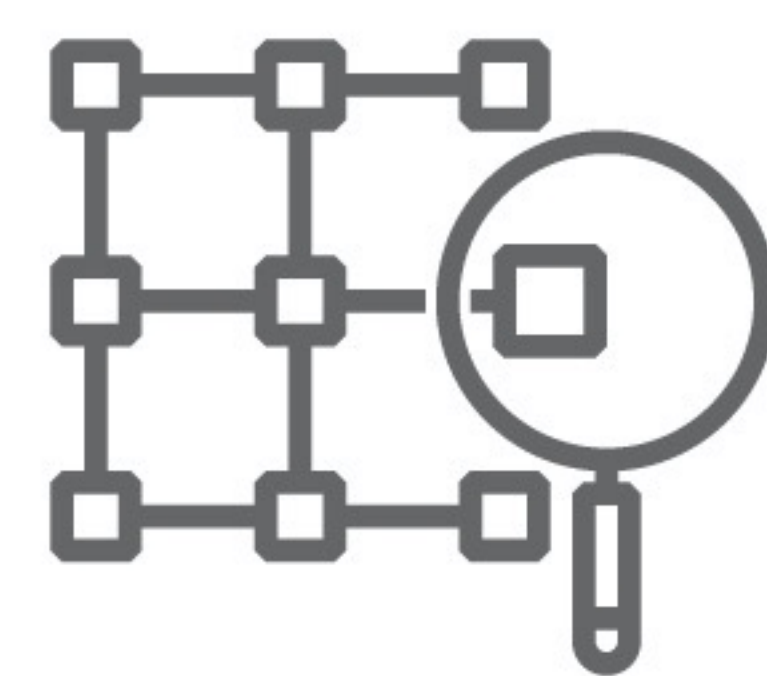


- 二、請分享一個你認為應該完全禁止使用AI的競賽項目，並說明原因。



5

五彩繽紛・旁徵博引

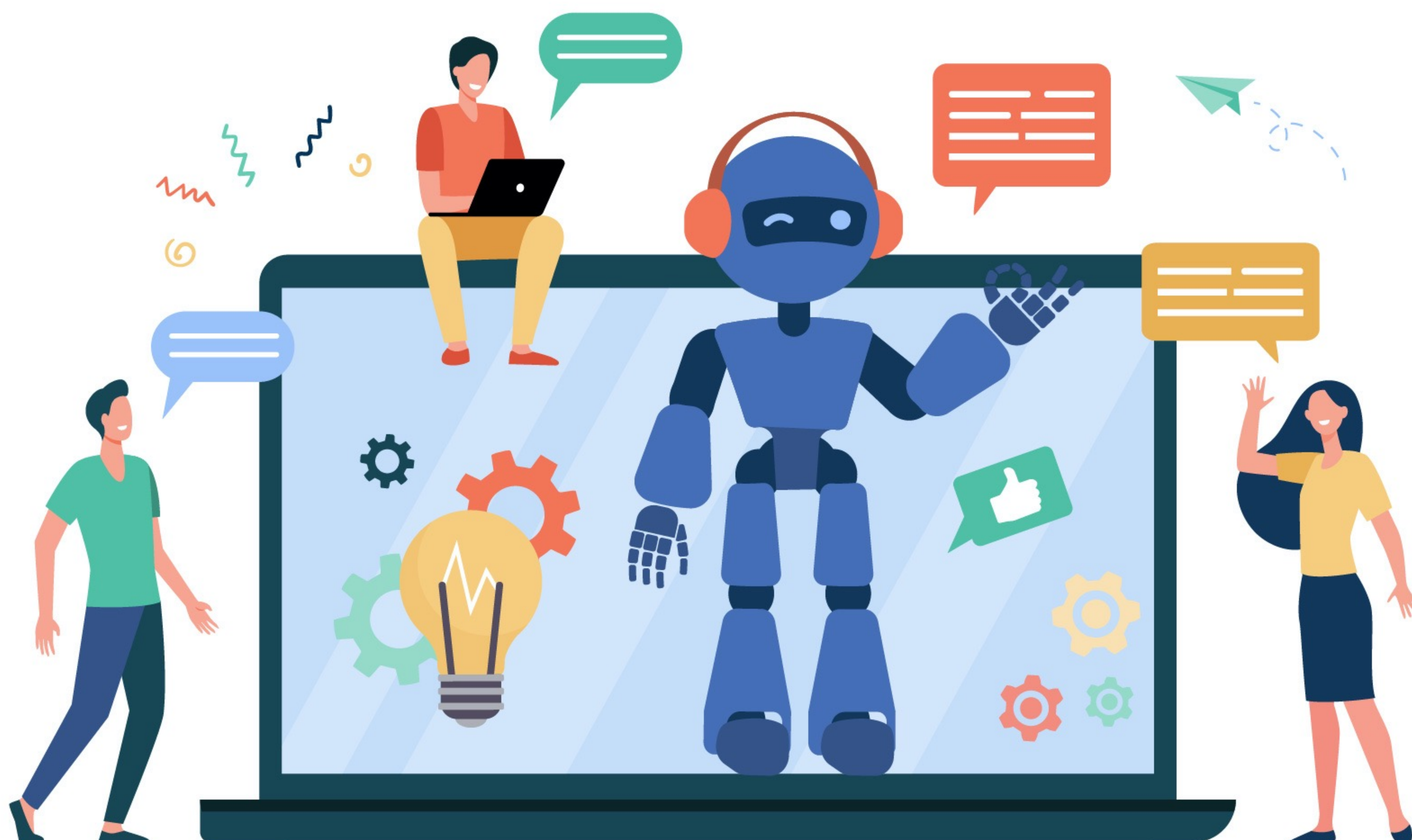


- 〔1〕中小學使用生成式人工智慧注意事項（學生版）（民113年7月1 日核定）。教育部資訊及科技教育司。民114年2月24日，取自：https://pads.moe.edu.tw/pads_front/index.php?action=news_content&uuid=02188ffe-bcf2-4ebd-bc00-f9aa75629d91
- 〔2〕利用AI創作參加比賽？主辦單位如何辨別？逢甲大學圖書館部落格。民114年2月24日，取自：<https://blogcastle.lib.fcu.edu.tw/archives/8640>
- 〔3〕不只是下棋！遊戲AI如何解決現實世界的難題？中央研究院研之有物。民114年2月24日，取自：<https://research.sinica.edu.tw/ai-game-deep-reinforcement-learning-ti-rong-wu/>
- 〔4〕深藍(超級電腦)。維基百科。民114年2月24日，取自：[https://zh.wikipedia.org/zh-tw/深藍_\(超級電腦\)](https://zh.wikipedia.org/zh-tw/深藍_(超級電腦))
- 〔5〕智慧財產權。維基百科。民114年2月24日，取自：<https://zh.wikipedia.org/zh-tw/智慧財產權>
- 〔6〕著作權。維基百科。民114年2月24日，取自：<https://zh.wikipedia.org/zh-tw/著作權>
- 〔7〕「深藍」Deep Blue：第一部在棋盤上擊敗人腦的「偽AI」超級電腦。techapple。民114年2月24日，取自：<https://techapple.com/archives/26438>
- 〔8〕使用AI創作參加比賽蕭雄淋：應揭露使用程度和範圍。文化＋。民114年2月24日，取自：<https://www.cna.com.tw/culture/article/20240714w001>
- 〔9〕DeepMind揭露新一代AI系統MuZero：不知遊戲規則也能有效規畫行動。iThome。民114年2月24日，取自：<https://www.ithome.com.tw/news/141845>
- 〔10〕AlphaGo Zero: Starting from scratch(18 October 2017) . Available at: <https://deepmind.google/discover/blog/alphago-zero-starting-from-scratch/>
- 〔11〕AlphaZero: Shedding new light on chess, shogi, and Go(6 December 2018) . Available at : <https://deepmind.google/discover/blog/alphazero-shedding-new-light-on-chess-shogi-and-go/>
- 〔12〕MuZero: Mastering Go, chess, shogi and Atari without rules(23 December 2020) . Available at: <https://deepmind.google/discover/blog/muzero-mastering-go-chess-shogi-and-atari-without-rules/>



語詞解釋：

- 〔1〕深藍：Deep Blue，是由IBM公司開發，專門用以分析西洋棋的超級電腦。1997年5月曾擊敗西洋棋世界冠軍加里·卡斯帕羅夫（Garry Kasparov）。
- 〔2〕深度強化學習：是機器學習的一個子領域，結合了強化學習和深度學習。強化學習探討如何在嘗試錯誤的過程中，讓AI學習做更好的決策。深度強化學習採用了深度學習的方法，讓AI可以直接基於非結構化資料來做決策，而不需要人為設計的狀態空間。
- 〔3〕神之一手：AlphaGo的其中一項特點在於可以判斷職業棋手下在各個位置的機率，如果機率太低的選擇，就不用特別考慮，如此一來，便不用逐步「窮舉」出所有結果，而能較為深入的分析後續盤面的變化，讓整體判斷更加精準又收斂。但也是因為AI具有這個特性，所以2016年AlphaGo與李世乭對弈的第四局，李世乭在第78手突如其來下出機率較低的選擇，殺得AI措手不及，不僅贏得了該次棋局，更被大家稱為「神之一手」。
- 〔4〕智慧財產權：人類用智慧創造出來的無形財產，主要涉及著作權、專利、商標等領域。智慧財產權可以分為工業產權與著作權兩類，工業產權包括專利、商標、工業外觀設計和地理標誌，著作權則包括文學和藝術作品等。
- 〔5〕著作權：俗稱版權，是著作人對其作品所專有的權利。著作權是有期限的權利，在經過一定期限後，著作財產權即歸於失效，而屬公有領域，任何人皆可自由利用。在著作權的保護期間內，即使未獲作者同意，只要符合「合理使用」的規定，亦可利用。凡此規定皆在平衡著作人與社會對作品進一步使用之利益。



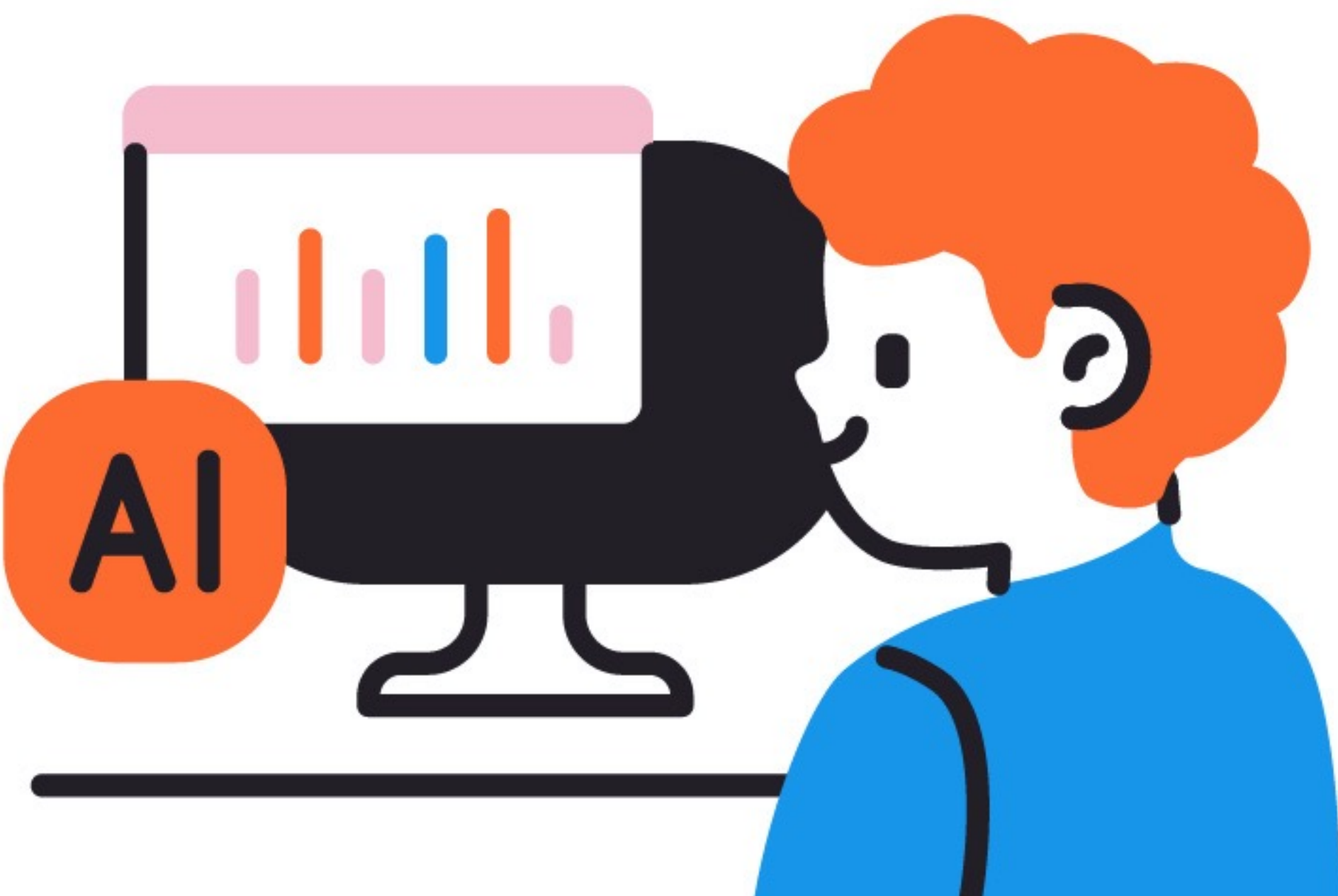
<

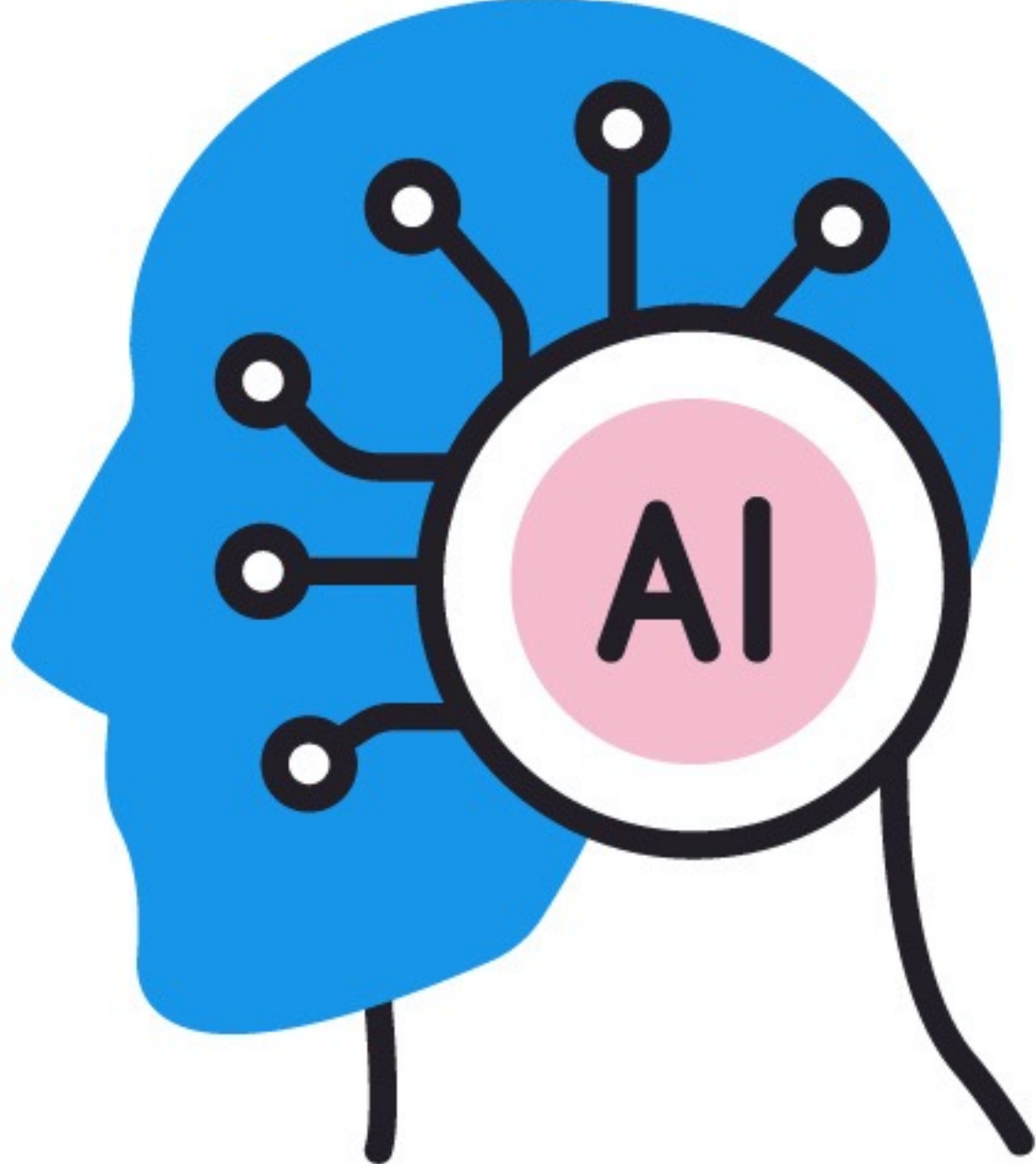
>

學習結語



A I 輔助展身手，競賽規則要遵守，
誠實比賽不強求，智慧倫理齊步走。





MEMO

資訊素養與倫理 國小 5 版

出版 / 臺北市政府教育局

召集人 / 湯志民 臺北市政府教育局局長

副召集人 / 卓育欣 臺北市政府教育局資訊教育科科长

林裕勝 臺北市文山區永建國民小學校長

指導委員 / 盧東華 臺北市立大學助理教授

楊凱翔 國立臺北教育大學教授

總編輯 / 張仁音 臺北市文山區永建國民小學教務主任

執行編輯 / 蔡佳儒 臺北市文山區永建國民小學專案教師

編審委員 / 江淑娟 臺北市中山區長安國民小學資訊組長

何詩慧 臺北市日新自造教育及科技中心主任

李忠憲 臺北市國語實驗國民小學教師

吳建勳 臺北市大安區新生國民小學教師

花梅真 臺北市北投區明德國民小學教師

姜保煌 臺北市士林區社子國民小學資訊組長

張宇軒 臺北市文山區永建國民小學資訊組長

蔡明佑 臺北市中正區東門國民小學學務主任

蔡孟憲 臺北市松山區健康國民小學學務主任

(按姓氏筆畫排列)

承辦單位 / 臺北市文山區永建國民小學

出版日期 / 中華民國114年3月

未來新世界

人機對決-AI參加競賽之相關規則及防範



臺北市政府教育局

DEPARTMENT OF EDUCATION
TAIPEI CITY GOVERNMENT