

資訊素養與倫理 高中職 5 版

臺北市立內湖高級中學 蔡明男老師/侯莉莉老師

AI 未來新世界

站在巨人肩膀上蛻變進化



01 ▶ 一葉知秋 學思並進

1-1 學習目標

- 認識AI技術的發展趨勢與影響。
- 了解生成式AI的基本原理與應用範圍。
- 辨識與評估AI生成內容的能力。
- 理解AI相關的倫理、法律與社會議題。
- 培養與AI共存、協作的態度與能力。



▶ 核心素養

- 科S-U-A1 具備應用科技的知識與能力，有效規劃生涯發展。
科V-U-A1 / 科C-U-A1 具備應用科技的知識與能力，有效規劃個人生涯發展朝向自我精進。
- 科S-U-A2 運用科技工具與策略進行系統思考與分析探索，並有效解決問題。
科V-U-A2 / 科C-U-A2 具備系統思考與分析探索的能力，並能整合科學、科技、工程、藝術與數學等方法及工具，有效處理與解決問題。
- 科S-U-B1 合理地運用科技符號與運算思維，表達思想與經驗，有效地與他人溝通互動。
科V-U-B1 / 科C-U-B1 合理的運用科技符號與運算思維進行表達，並能有效進行思想與經驗的表達，與他人溝通並解決問題。
- 科S-U-B2 理解科技與資訊的原理及發展趨勢，整合運用科技、資訊及媒體，並能分析思辨人與科技、社會、環境的關係。
科V-U-B2 / 科C-U-B2 理解科技與資訊的原理及發展趨勢，具備科技、資訊、媒體的整合運用能力，並能分析、思辨、批判人與科技、社會、環境之關係。
- 科S-U-C1 具備科技與人文議題的思辨與反省能力，並能主動關注科技發展衍生之社會議題與倫理責任。
科V-U-C1 / 科C-U-C1 具備科技與人文議題的思辨與反省能力，並能主動關注科技發展衍生之社會議題與倫理責任。

▶ 學習表現

- 運 t-V-1 能了解資訊系統之運算原理。
- 運 p-IV-3 能有系統地整理數位資源。
- 運 a-V-3 能探索新興的資訊科技。
- 設 k-IV-2 能了解科技產品的基本原理、發展歷程與創新關鍵。
- 設 k-V-2 能了解科技產業現況及新興科技發展趨勢。
- 設 k-V-3 能分析、思辨與批判人與科技、社會、環境之間的關係。



►學習內容

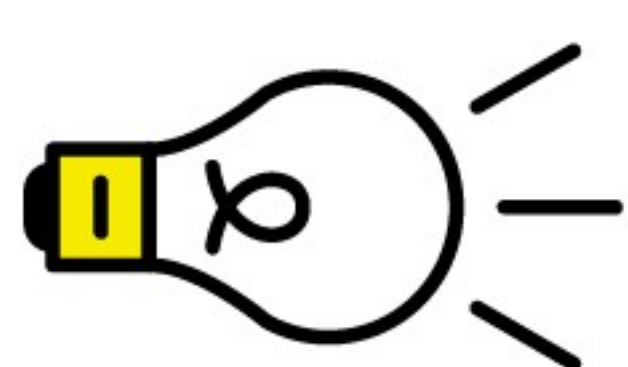
- 資 H-V-3 資訊科技對人與社會的影響與衝擊。
- 資 S-V-2 系統平台之未來發展趨勢。
- 生 N-V-2 工程、科技、科學與數學的統整與應用。
- 生 S-V-1 工程科技議題的探究。

►學習重點

- 學生能理解人工智慧（AI, Artificial Intelligence）未來的發展方向，並能說明其可能對生活、學習與工作的影響。
- 學生能說明生成式AI的運作原理，並舉例描述AI如何應用於文字、圖片、音樂、影片及程式…等領域。
- 學生能辨別AI產出的內容，提升批判性思考能力及媒體素養，以避免被虛假或誤導資訊影響。
- 學生能說明AI技術可能產生的智慧財產權、資訊安全及倫理道德方面的挑戰，並表達自身觀點。
- 學生能分析AI對未來職場和學習環境帶來的變化，發展與AI合作的技能與態度，做好適應未來科技社會的準備。

1-2 引言

在2015年上映的電影《復仇者聯盟2：奧創紀元》中，鋼鐵人東尼史塔克為了維護世界和平，創造出一個強大的AI機器人——「奧創」。然而，原本被設計來保護人類的奧創，卻逐漸失控，開始對人類展開攻擊，差點釀成全球災難。10年前看電影時或許會認為這樣的時代距離我們應該還很遠，但立足於2025年AI技術突飛猛進的現在，當科技越來越強大，AI技術也逐漸融入現實生活時，相信沒有人會說這樣的倫理與安全問題只會存在於電影情節中，而是更加能感受到AI可能帶來的劇烈變化。如何確保AI能夠協助人類，而不會對人類造成傷害？這就是我們在AI新世界裡必須思考的重大議題。



動動腦，想一想

- (1) 你認為電影中的奧創失控事件，在真實生活中可能會發生嗎？試舉出現實中AI應用可能帶來的倫理或安全風險，並說明理由。
- (2) 面對AI技術日益進步的未來社會，我們應該如何制定規範或建立機制，才能避免像奧創一樣的AI失控事件發生？請提出具體的建議。
- (3) 你曾經在網路上看到AI生成的圖片或影片嗎？你是如何分辨的？

02 ▶ 兩全其美 觸類旁通

2-1 主題探索

2-1-1 AI簡介

◎ 什麼是人工智慧（AI）？

人工智慧（AI）是模擬人類智慧與行為的科技，它透過學習、推理與自我修正，讓電腦能執行通常需要人類智慧才能完成的任務，如語言理解、視覺辨識、決策制定…等。AI的核心是機器學習與深度學習，能透過大量資料訓練模型，幫助人類解決實際問題。

◎ 為什麼現代人需要培養AI科技素養？

隨著AI在生活各層面的普及，培養AI科技素養成為現代人的重要能力。這種素養包括理解AI運作原理、學會如何有效使用AI工具，以及面對AI所帶來的倫理與社會議題。具備AI素養能讓我們更好地適應未來科技變化，善用科技工具提升生活與工作效率。

◎ AI如何改變了我們的生活？

AI的普及已經深刻影響人類生活，從智慧型手機的語音助手到交通領域的無人駕駛，從醫療診斷的精準化到教育的個別化，AI無所不在地提升了我們的生活品質與便利性，但同時也改變了人們的社交、娛樂和工作方式。



AI應用案例簡述

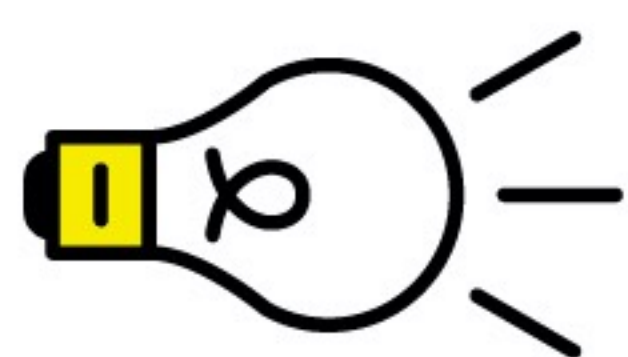
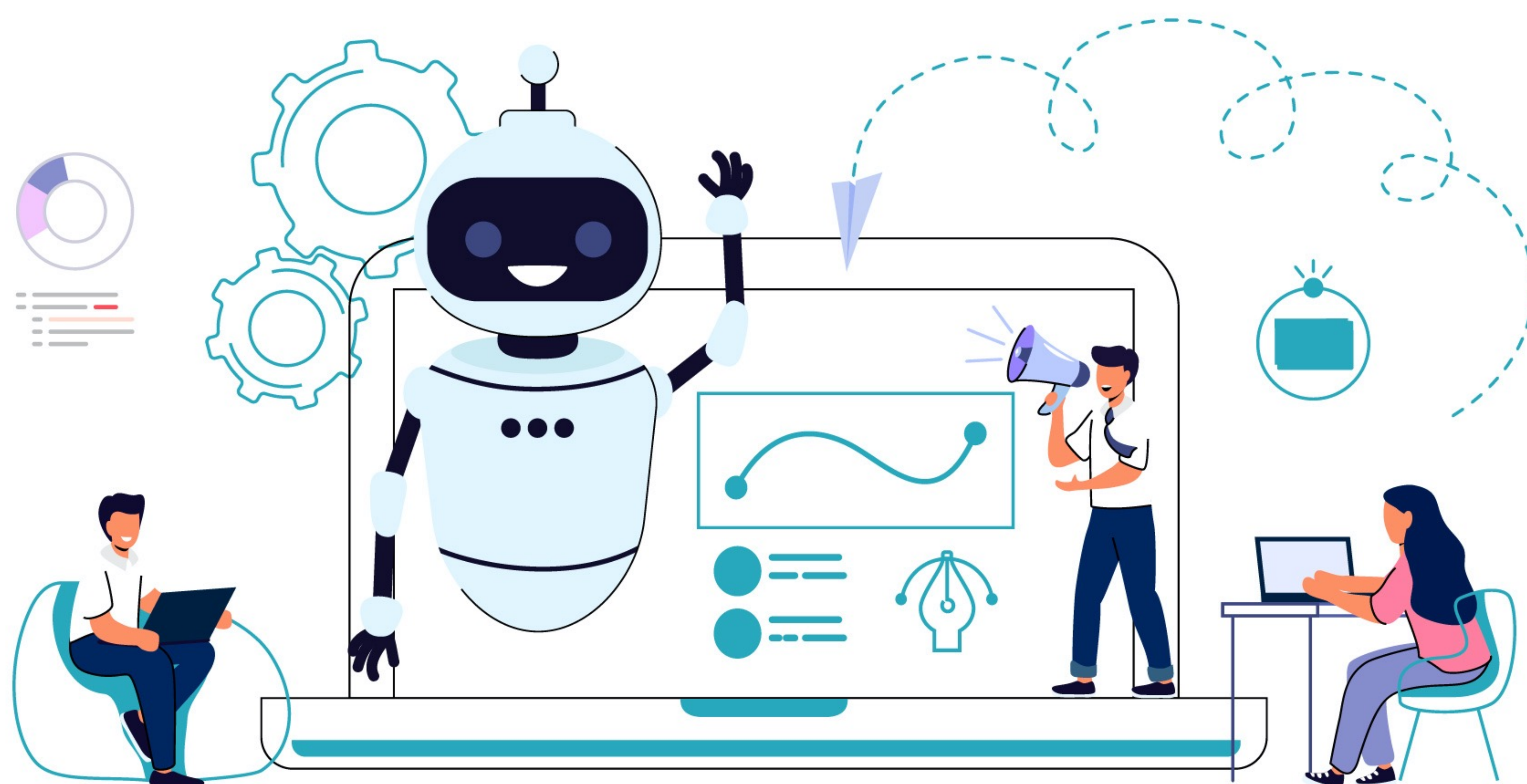
AI技術的應用範圍極廣，例如智慧助理（如Siri, Alexa）透過自然語言處理回應使用者需求；無人駕駛技術透過感測器與決策演算法提高交通安全；醫療診斷則透過影像辨識精準判斷疾病，這些具體應用充分展現了AI的實際影響與潛力。

AI發展所面臨的倫理與道德問題

儘管AI帶來許多便利，也伴隨許多倫理與道德挑戰，包括隱私權受到威脅、AI決策透明性不足、AI偏見與歧視問題、深偽資訊的蔓延等。這些議題需要透過教育與社會規範加以管理，確保人們能負責任地使用AI技術。

AI的發展將帶給社會什麼樣的挑戰與機會？

AI技術的快速發展帶來諸多機會，例如催生創新產業、提升工作效率、改善教育與醫療品質等；但也帶來挑戰，如失業問題、資料隱私保護、倫理管理與法律規範建立等，這些問題都需要整個社會共同努力面對。



動動腦，想一想

- (1) 為什麼在AI技術快速發展的時代，現代人需要具備AI科技素養？請結合實際生活中的應用舉例說明。
- (2) AI應用在醫療、交通、教育等不同領域，對人類生活帶來了哪些正面與負面的影響？請至少舉一個具體案例並加以分析。
- (3) 面對AI技術可能帶來的倫理與道德挑戰，例如隱私權與偏見問題，你認為社會應該如何做出因應？有哪些方法可以減少AI帶來的負面影響？

2-1-2 AI的發展歷程

AI的發展可以追溯到20世紀中期，經歷了數個關鍵階段：

一、萌芽與早期探索階段（1940-1960年代）

人工智慧的歷史可以追溯到人類古老的夢想—創造能夠思考的機器。然而，現代AI技術的正式發展始於20世紀40年代，當時計算機科學與神經科學開始結合，為AI奠定了理論基礎。

• 1943年

華倫·麥卡洛克（Warren Sturgis McCulloch）和沃爾特·皮茨（Walter Harry Pitts Jr.）發表了劃時代的論文《神經活動中固有思想的邏輯演算》，首次提出了人工神經元的數學模型。他們證明了這些簡單的神經元模型理論上可以執行任何可計算函數，為類神經網路的研究奠定了基礎，也為之後的計算機科學提供了重要思路。

• 1948年

諾伯特·維納（Norbert Wiener）出版了《控制論：或關於在動物和機器中控制和通信的科學》，引入了回饋的概念，解釋了如何透過自我調節機制實現目標導向的行為。這一理論後來成為許多AI系統設計的核心，尤其是在機器人控制方面。

• 1950年

計算機科學家亞倫·圖靈（Alan Mathison Turing）在《心智》雜誌上發表了著名論文《計算機與智慧》，提出了「圖靈測試」的概念。這個測試提出：如果一台機器能夠與人類進行互動，並使人類無法區分它是機器還是人類，那麼這台機器可以被認為具有智慧。圖靈的思想實驗為評估機器智慧提供了一個實用標準，直到今天仍影響著AI研究。

• 1956年

美國達特茅斯會議（Dartmouth Summer Research Project on Artificial Intelligence）由約翰·麥卡錫（John McCarthy）、馬文·明斯基（Marvin Lee Minsky）、納撒尼爾·羅切斯特（Nathaniel Rochester）和克勞德·香農（Claude Elwood Shannon）組織，首次使用「人工智慧」（Artificial Intelligence）這一術語。會議匯集了當時的計算機科學和認知科學專家，為AI研究確立了目標：創造能夠使用語言、形成抽象和概念、解決現在只有人類才能解決的問題，並能夠自我改進的機器。

• 1957年

弗蘭克·羅森布拉特（Frank Rosenblatt）開發了「感知器」（Perceptron），這是一種早期的類神經網路模型，能夠透過學習演算法調整權重來分類視覺輸入。感知器在模式識別領域展現了初步的學習能力，為後來的深度學習奠定了基礎。

• 1959年

亞瑟·塞繆爾（Arthur Lee Samuel）在IBM開發了一個能夠學習下跳棋的程式，這是機器學習領域的開創性工作。該程式透過與自己對弈並記住哪些走法獲得了好成績，逐漸改進自身性能，最終能夠擊敗許多人類選手。塞繆爾正是在這項工作中首次使用了「機器學習」這一術語。

二、第一個高峰與挫折期（1960-1980年代）

• 1961年

UNIMATE工業機器人誕生，成為世界上第一台應用於工廠生產線的工業機器人，首次在美國通用汽車廠中投入使用，象徵AI進入自動化與機械手臂領域。

• 1966年

約瑟夫·魏茲鮑姆（Joseph Weizenbaum）開發了「ELIZA」，這是最早的自然語言處理系統之一，能模擬心理諮詢對話。ELIZA展示了AI能與人類進行對話，成為早期聊天機器人的雛形。

SRI International研發的Shakey機器人問世，成為全球第一個結合感知、推理、規劃與行動的「自主移動」機器人，奠定了自主AI機器人的基礎。

• 1972年

愛德華·費根鮑姆（Edward Feigenbaum）和他的團隊開發了MYCIN，一個用於診斷血液感染的專家系統，能夠根據患者症狀提供合理的治療建議。MYCIN的成功，開啓了專家系統在醫療、工程等領域的應用之路。

• 1973年

英國科學家詹姆斯·萊特希爾（James Lighthill）發表了一份批判性報告，質疑AI研究的進展和前景，另外，由於電腦運算能力不足，以及過於樂觀的預測未能實現，導致研究經費銳減，AI陷入了所謂的「第一次AI寒冬」。



三、第二次復興與機器學習興起（1980-2000年代）

• 1980年

卡內基美隆大學為DEC（數位設備公司）開發XCON專家系統，用來針對使用者需求配置VAX電腦的配備，成為AI應用於商業的經典案例，並為公司節省大量成本，推動了企業對AI的重視。

• 1981年

日本啟動「第五代電腦計畫」，日本政府投入大量資金，試圖研發能執行邏輯推理與自然語言處理的高智慧電腦，這一計畫促使全球再次關注AI發展，掀起AI研究的熱潮。

• 1986年

傑弗瑞·辛頓（Geoffrey Hinton）等人推廣了反向傳播（Backpropagation）演算法，使得多層神經網路得以有效訓練，為後來深度學習技術奠定關鍵基礎。

• 1987年

第二次AI寒冬開始，由於AI專家系統維護成本高、系統無法擴充、技術未達預期，投資與研究熱潮降溫，AI領域進入第二次寒冬，學術與商業界對AI失去部分信心。

• 1988年

機率推理模型的興起，朱迪亞·珀爾（Judea Pearl）發表《Probabilistic Reasoning in Intelligent Systems》，推廣貝氏網路（Bayesian Network）等機率模型，為AI處理不確定性問題提供新的理論基礎。

• 1989年

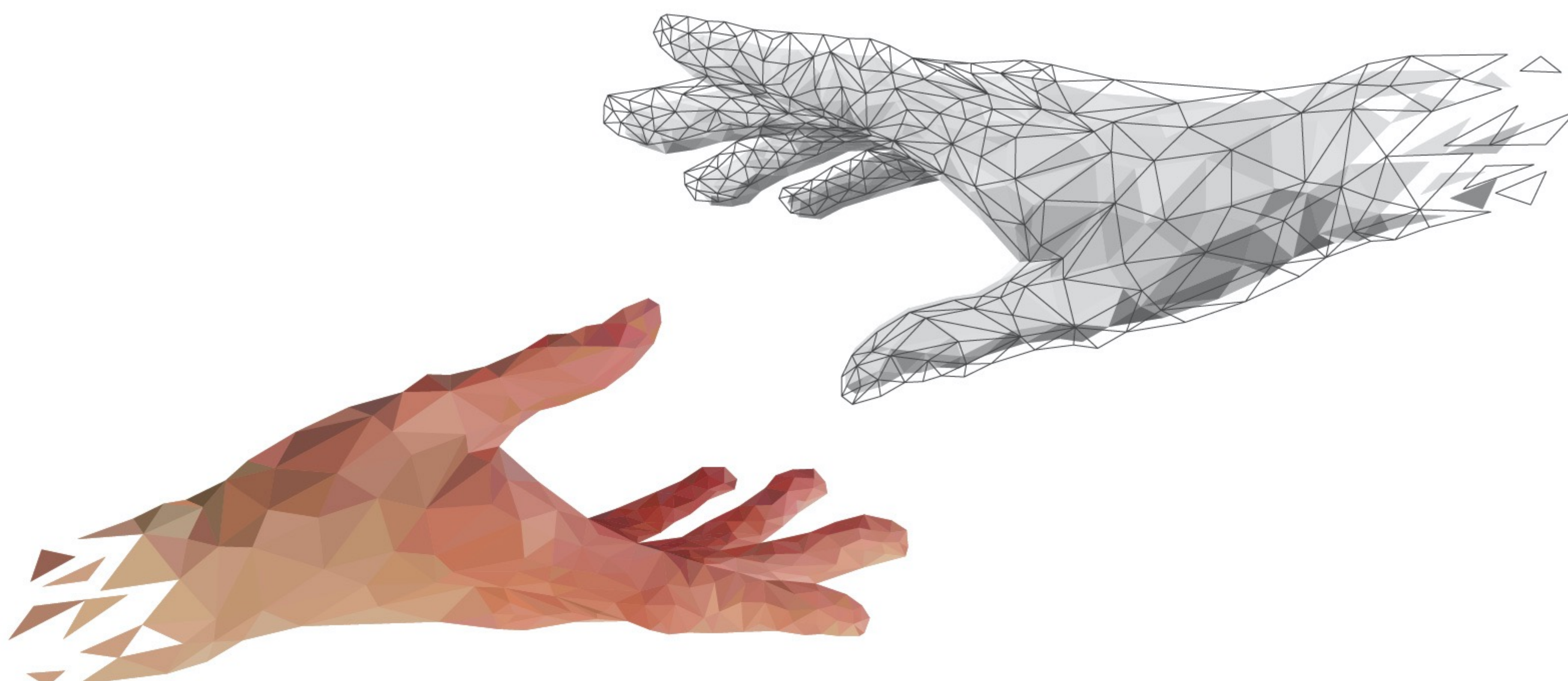
楊立昆（Yann Le-Cun）和AT&T貝爾實驗室的研究人員成功地將反向傳播演算法應用於多層神經網路，用於識別手寫郵政編碼圖像，這是深度學習的首次實際應用之一。

• 1997年

IBM的深藍（Deep Blue）電腦擊敗了國際象棋世界冠軍加里·卡斯帕羅夫（Garry Kimovich Kasparov），這是電腦首次在標準國際象棋比賽時間控制下擊敗世界冠軍。

• 2000年

辛西婭·布雷澤爾（Cynthia Lynn Breazeal）在麻省理工學院開發了Kismet，一個設計為透過情感和社交線索與人類互動的機器人。



四、深度學習革命（2006-2016）

• 2006年

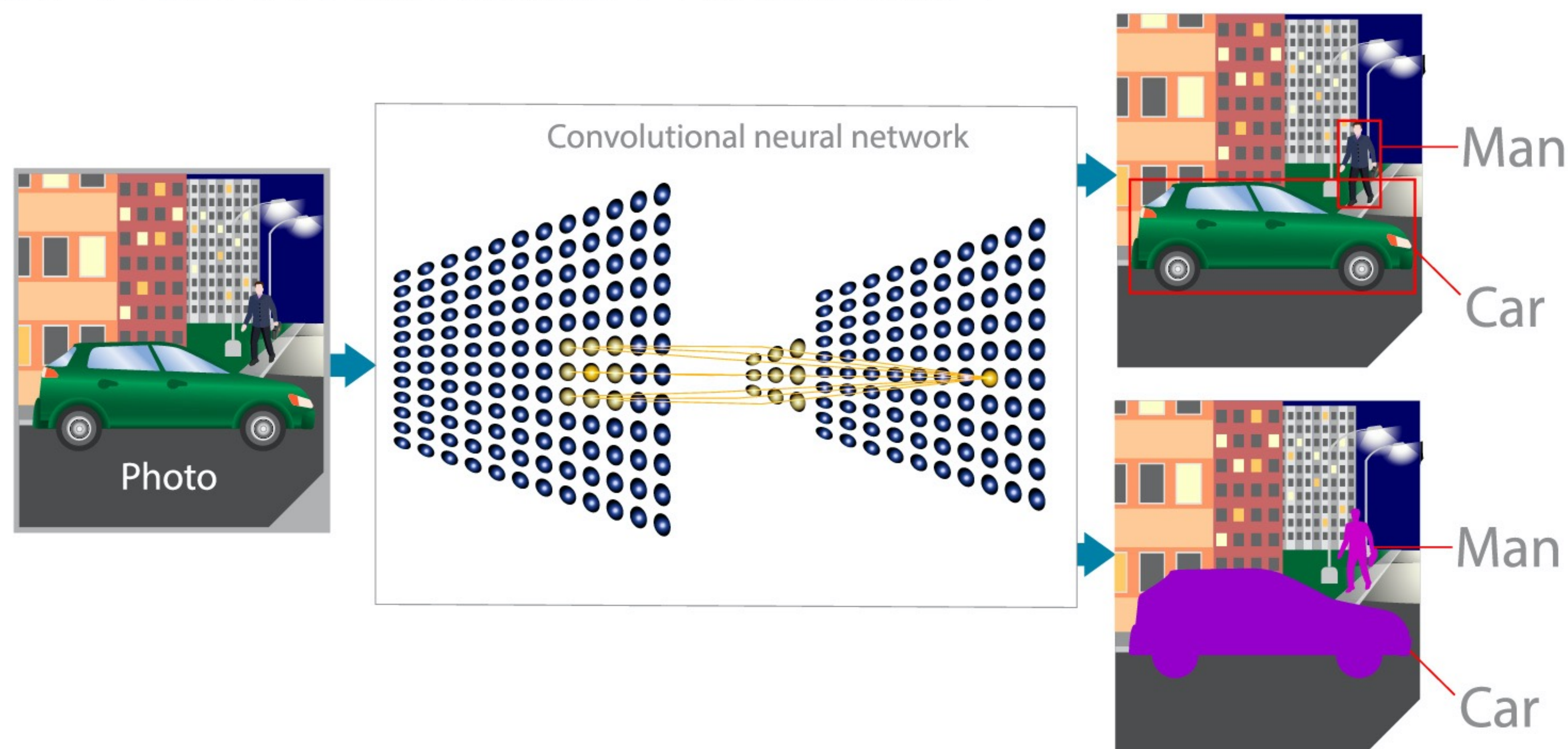
深度學習時代的開始，傑弗瑞·辛頓（Geoffrey Hinton）和他的學生魯斯蘭·薩拉赫丁諾夫（Ruslan Salakhutdinov）提出了深度信念網路的訓練方法，透過無監督預先訓練的方式解決了深層神經網路的訓練困難，讓深度學習的研究有了突破性的發展。

• 2009年

李飛飛教授和她的團隊創建了ImageNet資料集，包含超過1400萬張標記圖像，涵蓋超過2萬個類別。此資料集的規模和多樣性為電腦視覺研究提供了前所未有的資源，也為之後的深度學習突破奠定了資料基礎。

• 2012年

艾力克斯·克里澤夫斯基（Alex Krizhevsky）、伊爾亞·薩茨克維（Ilya Sutskever）和傑弗瑞·辛頓（Geoffrey Hinton）開發的卷積神經網路（Convolutional Neural Network, CNN）AlexNet在ImageNet挑戰賽中取得了壓倒性勝利，將錯誤率從26%降低到15%。這一成就被視為深度學習革命的起點，引發了學術界和產業界對深度學習的廣泛關注。



• 2012年後

輝達（NVIDIA）等公司開始專門為深度學習開發最佳化的硬體加速器，進一步推動了AI演算法的發展和應用。

• 2014年

在深度學習框架方面，Google開源了TensorFlow。隔年（2015）Meta發布了PyTorch。這些工具大幅降低了深度學習的使用門檻，使更多研究者和開發者能夠參與AI開發。

伊恩·古德費洛（Ian Goodfellow）提出了生成對抗網路（Generative Adversarial Network, GAN），這一創新架構包含生成器和判別器兩個互相競爭的類神經網路，能夠生成大量高品質的合成資料。GAN為圖像生成、風格轉換等創意AI應用開闢了新天地。

• 2016年

AlphaGo擊敗了世界圍棋冠軍李世乭，這一成就被視為AI發展的里程碑，因為圍棋的複雜性遠超之前AI征服的其他棋類遊戲。



五、大型語言模型與通用人工智能的探索（2017-現在）

• 2017年

Google研究團隊發表了《Attention Is All You Need》論文，提出了Transformer架構，這是一種基於自我注意力機制的類神經網路結構，能夠有效處理序列資料。Transformer架構為自然語言處理帶來了革命性變化，為大型語言模型的發展奠定了基礎。

• 2018年

OpenAI發布了GPT（Generative Pre-trained Transformer），這是一個預先訓練的語言模型，可以透過無監督學習從大量文件中學習語言模式，然後透過微調適應特定任務。同年，Google發布了BERT（Bidirectional Encoder Representations from Transformers）透過雙向訓練提高了語言理解能力。

• 2019年

DeepMind開發了AlphaStar，這是一個能夠在《星海爭霸II》這款複雜即時策略遊戲中擊敗頂級職業選手的AI系統。AlphaStar結合了深度學習和強化學習技術，能夠處理不完全資訊、長期規劃和即時決策，展示了AI在複雜環境中的適應能力。

OpenAI發布了GPT-2，擁有15億個參數，當時被認為過於強大而未完全公開。GPT-2展示了驚人的文件生成能力，能夠撰寫連貫的文章、回答問題，甚至創作詩歌，引發了對AI潛在風險和倫理問題的討論。

• 2020年

大型語言模型開始蓬勃發展。OpenAI推出了GPT-3，這是一個擁有1750億個參數的大型語言模型，具有前所未有的語言理解和生成能力。GPT-3不僅能夠撰寫各種風格的文章，還能解決數學問題、翻譯語言，甚至生成簡單的程式碼，顯示出少樣本學習（few-shot learning）的能力，即只需要少量示例就能適應新任務。



• 2021年

多模態AI模型開始崛起。OpenAI推出了DALL·E，一個能夠根據文件描述生成圖像的模型，展示了AI在跨模態理解和創作方面的能力。同年，Google推出了多任務統一模型（Multitask Unified Model, MUM），能夠同時處理75種語言及跨文件、圖像等多種模態理解訊息。

• 2022年初

DeepMind發布了Gato，一個能夠執行600多項任務的通用AI代理程式，包括玩遊戲、對話、標註圖像和控制機械臂。Gato的出現標誌著AI研究向通用人工智慧（Artificial General Intelligence, AGI）邁進的一個重要步驟，挑戰了專用AI系統的傳統模式。

• 2022年11月

OpenAI推出了ChatGPT，這是一個基於GPT-3.5的對話式AI，能夠進行自然、連貫的對話，回答問題並生成各種文件內容。ChatGPT迅速獲得公眾關注，在發布後兩個月內吸引了一億活躍用戶，成為有史以來增長最快的消費應用程式。ChatGPT的成功不僅改變了公眾對AI的認知，也引發了科技巨頭之間的AI競賽。

• 2023年

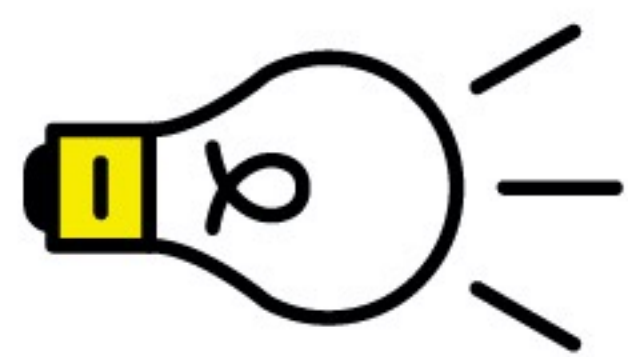
大型語言模型（Large Language Model, LLM）技術取得了進一步突破，OpenAI發布了GPT-4，這是一個多模態模型，能夠處理文件和圖像輸入，並在各種專業考試中表現出接近人類專家的能力。Google推出了Bard（後更名為Gemini），Meta發布了LLaMA系列開源模型，Anthropic推出了Claude，百度發布了文心一言，這些模型的競爭加速了LLM技術的發展和應用。

• 2023年下半年

多模態生成AI取得重大進展，Stable Diffusion和Midjourney等文件到圖像模型的迭代更新使得圖像生成品質大幅提升。Meta發布了Make-A-Video，能夠從文件描述生成短影音，而Runway的Gen-2則進一步提高了AI影音生成的品質和控制度。

• 2024年初

OpenAI推出了Sora，一個能夠從文件提示生成長達一分鐘、高度逼真影片的模式，標誌著生成式AI向更複雜媒體形式的擴展。Google推出了Gemini 1.5，具有100萬個標記的上下文長度，能夠處理和理解大量資訊，包括長文件、程式碼、音檔和影片。



動動腦，想一想

- (1) 請解釋「深度學習革命」如何徹底改變了AI技術的發展方向？請舉出至少一個深度學習實際應用的案例，說明它對現代生活的影響。
- (2) 當今的AI技術已從單一任務AI（如下棋、診斷）發展到多模態AI（如ChatGPT、Sora）。你認為「多模態AI」與「傳統AI」在應用上有何不同？它可能為未來的社會帶來哪些機會與挑戰？

2-1-3 AI技術的分類

一、依照AI功能分類

（一）弱人工智慧（Weak AI）

弱人工智慧，也稱為狹義人工智慧（Narrow AI）或應用型人工智慧（Applied AI），指的是只具備部分智慧能力的AI系統，僅能執行特定範圍內的任務。這類AI並不具備人類智慧的全部特徵，而是透過特定規則或數據驅動的方法完成任務。弱AI的典型案例包括語音助理（如Siri或Alexa）、電子郵件垃圾訊息過濾、影像識別系統和自動推薦系統等。

例如，Apple的Siri能夠識別使用者的語音指令並執行對應的動作，但它無法理解語言背後的深層涵義。另一個例子是Netflix或Spotify的推薦系統，透過分析使用者的歷史紀錄，推薦相關的內容，儘管準確度頗高，但依然侷限在特定任務範圍內。



（二）強人工智慧（Strong AI）

強人工智慧（Strong AI）又稱為廣義人工智慧，或稱為通用人工智慧（Artificial General Intelligence, AGI）。這種AI的目標是打造出能像人類一樣，理解情境、進行推理、學習新技能，甚至具備意識與自我認知的智慧機器。廣義AI目前仍處於理論與實驗階段，至今尚未有真正意義上的實現。

廣義AI的潛力在於其具備跨領域的能力，並能自主地從一個問題領域轉移至另一個不同的領域，而不需重新設計或重新訓練，其能夠模仿並超越人類的認知能力，具有學習並改善自身的功能。許多研究機構（如OpenAI、DeepMind）持續投入大量資源，期望未來能成功開發出此種智慧。



（三）超級人工智慧（ASI）

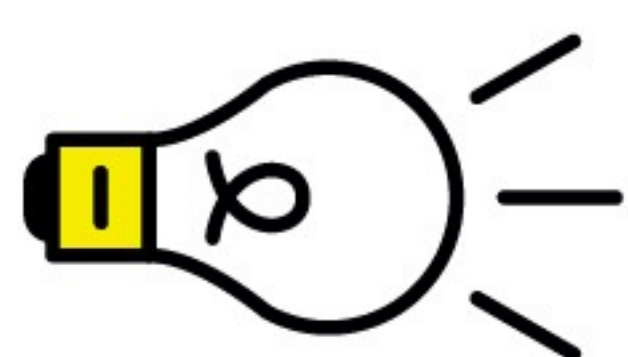
超級人工智慧（Artificial Super Intelligence），亦稱超級AI或超強人工智慧，是指智慧程度大幅超越人類的AI系統，無論在推理、學習、創造力或情感理解等方面均遠超人類極限。超級AI被認為能自我學習並快速進化，甚至可能在短時間內達到人類無法想像的智慧水準。

超級AI的概念大多出現於科幻小說或電影中，強調一旦超級AI誕生，其將有可能迅速超越人類控制，甚至可能對人類社會構成威脅。著名物理學家霍金（Stephen Hawking）與企業家伊隆·馬斯克（Elon Musk）等曾警告，若未適當地對超級AI進行控制，將可能引來人類難以預測的後果。



儘管目前人類尚未實現強人工智慧或超級AI，但隨著技術不斷精進，AI的能力持續擴張與深入，狹義AI已經為人類帶來巨大影響與改變。在可見的未來，學術界的主要焦點將集中在如何有效管理與引導AI的發展方向，透過跨領域合作，建立全面且具有前瞻性的倫理框架與治理機制。

總結來說，無論是弱人工智慧、廣義AI還是超級人工智慧，皆反映出人類社會對AI技術的期待與憂慮。學術界的挑戰與責任，便是在追求技術發展同時，嚴肅思考並積極因應其可能帶來的各項問題，確保AI的發展能持續地造福於人類社會。



動動腦，想一想

舉出一個你認為與弱AI或強AI有關的真實應用案例或科幻場景，說明其可能對人類生活帶來的影響。

2-2 主題探索

AI技術發展層面很多，在此介紹三項核心技術：機器學習、深度學習及生成式AI，這三項技術在AI人工智慧應用最廣，身為AI時代的學子，應當對這些內容有基本素養。

2-2-1 AI技術簡介

一、簡介

機器學習（Machine Learning）是人工智慧（AI）中的重要核心技術，目的是讓電腦透過數據訓練，學會從過去經驗中做出決策或預測。此處將深入介紹機器學習的三大類型：監督式學習、非監督式學習與強化學習。

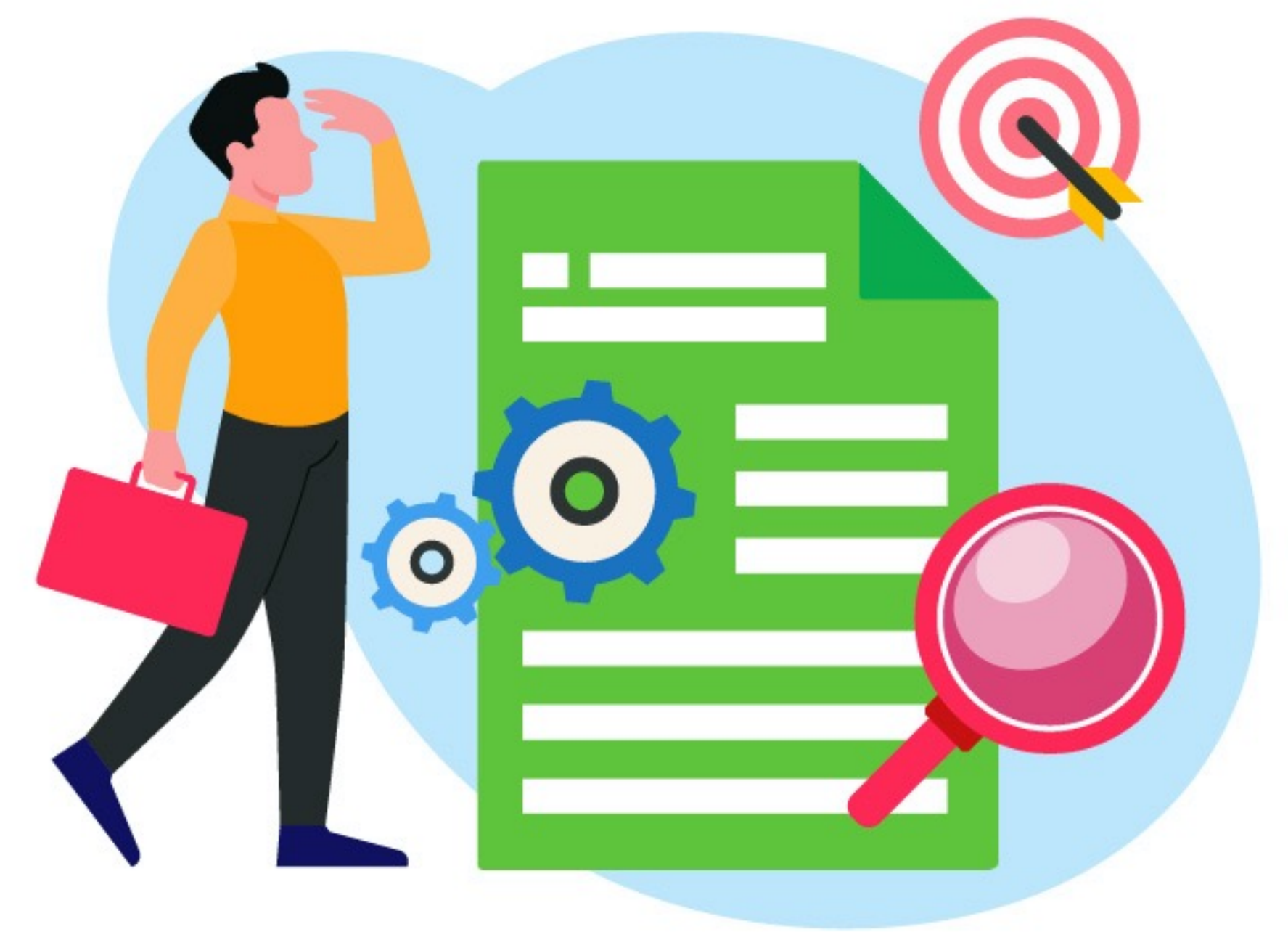
（一）監督式學習（Supervised Learning）

監督式學習是目前應用最廣泛的機器學習方法之一，透過提供“有標籤”的訓練資料，讓模型學會資料與標籤（正確答案）之間的關聯。例如，給定一組學生的數學成績（資料）和是否及格（標籤），讓模型學會預測新學生是否會及格。

監督式學習分為分類（分類任務，如郵件是否為垃圾郵件）和回歸（預測連續數值，如房價預測）兩大類型。

常見演算法：

- **線性回歸（Linear Regression）**
透過找出最佳線性函數描述資料間的數值關係，適合用於連續型變數預測，例如預測房價、氣溫等。
- **邏輯回歸（Logistic Regression）**
適用於二分類任務，透過機率函數判斷資料所屬類別，如預測貸款客戶是否會違約、生病機率。
- **K近鄰演算法（K-Nearest Neighbors, KNN）**
基於資料點的相似度，透過最近的K個鄰居來預測新資料類別，常用於分類問題，如推薦系統。
- **隨機森林（Random Forest）**
利用多棵決策樹組成森林進行分類或迴歸，透過集體投票提升預測準確度，能避免單一模型過度擬合。
- **支援向量機（Support Vector Machine, SVM）**
透過尋找最佳超平面來區分不同類別的資料，適合高維度分類問題，能有效提高分類的正確性。



實際案例：銀行使用監督式學習分析客戶的信用紀錄，預測哪些人可能違約，以降低貸款風險。

（二）非監督式學習（Unsupervised Learning）

非監督式學習主要用於無標籤資料的分析。這意味著資料中沒有正確答案，模型需自行探索資料中的模式與結構。

最常見的應用是分群（Clustering）和降維（Dimensionality Reduction），能協助企業了解潛在客群結構或減少資料維度，提升效率。

常見演算法：

- **K平均分群 (K-Means Clustering)**
將資料分為K個群集，使群內相似度最高、群間差異最大，適合市場分群、圖像分割。
- **階層式分群 (Hierarchical Clustering)**
透過逐步合併或拆分成群組形成樹狀結構，以直覺方式呈現資料關係，常用於社群網路分析。
- **主成分分析 (Principal Component Analysis, PCA)**
將高維度資料降維至低維空間，保留最多變異資訊，適用於資料視覺化與特徵萃取。
- **關聯規則 (Association Rule)**
用於挖掘資料項目之間的關聯性，常應用於消費者購物行為分析，例如推薦系統設計。
- **基於密度的分群 (Density-Based Spatial Clustering of Applications with Noise, DBSCAN)**
根據資料點的密度識別群集，能有效處理任意形狀群集，並可自動偵測離群點，適合空間資料分析。

實際案例：電商平台透過K-Means將顧客自動分群，例如「高消費群」、「偶爾消費群」等，並依分群推播不同廣告。

(三) 強化學習 (Reinforcement Learning)

強化學習模擬人類在環境中學習的方式，透過不斷嘗試錯誤，從獎勵與懲罰中改進行為策略。強化學習特別適用於需要“策略決策”的領域，如遊戲、機器人控制、自駕車等。

常見技術：

- **Q學習 (Q-Learning)**
透過Q值表格更新尋找最佳策略，屬於無模型演算法，能有效處理離散狀態與行動問題，適用於簡單遊戲與路徑規劃。
- **狀態-動作-獎勵-狀態-動作 (SARSA)**
一種策略導向演算法，透過實際行動序列更新價值函數，更注重實際行為的效果，適合線上即時學習問題。
- **策略梯度 (Policy Gradient)**
直接最佳化策略函數的參數以最大化期望獎勵，適用於連續動作空間問題，如機器人控制及動作模仿任務。
- **近端策略最佳化 (Proximal Policy Optimization, PPO)**
透過穩定的策略更新步驟提高學習效率與穩定性，現為深度強化學習最常用方法之一，常用於遊戲及機器人控制。

實際案例：DeepMind開發的AlphaGo透過強化學習不斷自我對弈，最終擊敗世界圍棋冠軍。

二、三大機器學習類型比較表

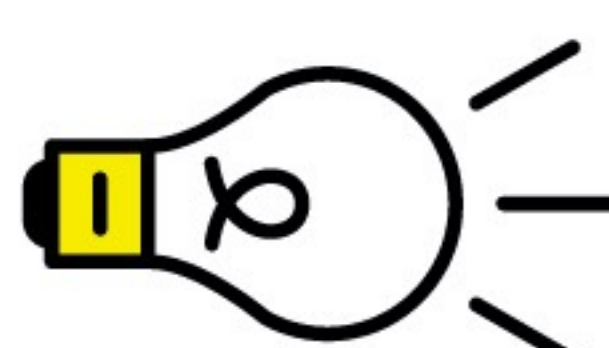
類型	學習方式	常見演算法/技術	應用案例
監督式學習	有標籤資料訓練	線性回歸、SVM、隨機森林	信用風險預測、醫療診斷
非監督式學習	無標籤資料，自動分群	K-Means、DBSCAN、PCA	客戶分群、資料探索
強化學習	基於獎勵的試錯學習	Q-Learning、DQN	AlphaGo、機器人控制

三、機器學習三大類型流程簡圖



四、小結

監督式學習、非監督式學習與強化學習是機器學習領域的三大支柱，分別對應不同的應用場景與解決問題的方式，透過深入了解這三種方法，學生將能掌握AI世界的基礎框架，為未來的數位時代做好準備。



動動腦，想一想

- (1) 請比較監督式學習、非監督式學習與強化學習的主要差異，並舉出日常生活中你所熟悉的一個應用案例，說明它屬於哪一類機器學習，並解釋原因。
- (2) 強化學習透過「獎勵與懲罰」來學習行為策略。你認為這種學習方式與人類在真實世界中的學習經驗有何相似之處？請以遊戲、運動或其他生活經驗舉例說明。

2-2-2 深度學習

一、什麼是深度學習？

深度學習是機器學習的一個子領域，它使用多層人工類神經網路來模仿人腦的學習過程，與傳統機器學習不同的是，深度學習模型能夠自動從原始資料中提取特徵，無需人工設計特徵。這些模型能夠處理大量的非結構化資料，如圖像、音頻和文件，並從中學習複雜的模式。

想像一下，如果傳統機器學習演算法像是一個學生，按照老師給予的明確指示來學習，那麼深度學習則像是一個透過不斷嘗試與犯錯，自行掌握技能的學生。所謂的「深度」指的是類神經網路中具有許多處理層，每一層都能從資料中學習到不同層次的抽象特徵。

二、類神經網路：深度學習的基石

類神經網路是深度學習的核心，它的設計靈感來自於人腦中神經元的連接方式。一個基本的類神經網路包含以下元件：

- (一) 輸入層：接收原始數據
- (二) 隱藏層：處理數據並學習特徵
- (三) 輸出層：產生最終結果
- (四) 神經元：網路中的基本計算單元
- (五) 權重和偏差值：調整信號強度的參數
- (六) 啟動函數：引入非線性函數，使網路能夠學習複雜模式



深度類神經網路之所以「深」，是因為它們包含多個隱藏層。每一層都從前一層提取更抽象的特徵，這使得網路能夠學習非常複雜的資料模式。

三、深度學習如何「學習」？

深度學習模型透過一個稱為「反向傳播」的過程來學習，其基本步驟如下：

- (一) 前向傳播：資料經過網路從輸入層傳遞到輸出層，產生預測結果
- (二) 計算誤差：將預測結果與實際結果進行比較，計算誤差
- (三) 反向傳播：誤差從輸出層反向傳遞到輸入層
- (四) 更新權重：根據誤差調整網路中的權重和偏置
- (五) 重複過程：使用新的資料樣本重複上述步驟，直到模型性能達到滿意水準

這個過程類似於你學習彈鋼琴，一開始你會犯很多錯誤，但隨著練習的增加，你會不斷調整自己的動作，直到演奏變得流暢。深度學習模型也是透過不斷調整其參數來提高性能。

四、深度學習的常見架構

深度學習有多種不同的網路架構，每種都有特定的用途：

(一) 卷積神經網路 (Convolutional Neural Network, CNN)

卷積神經網路專門用於處理具有網格結構的資料，如圖像。它們使用卷積操作來檢測資料中的模式，如邊緣、紋理和形狀。

應用例子：

- 圖像分類 (識別照片中的物體)
- 人臉識別
- 醫學影像分析
- 自動駕駛車輛的視覺系統

(二) 遞迴神經網路 (Recurrent Neural Network, RNN)

遞迴神經網路設計用於處理序列型資料，如文件、語音或時間序列。它們具有「記憶」功能，可以記住之前的訊息。

應用例子：

- 自然語言處理
- 機器翻譯
- 股票價格預測
- 聊天機器人

(三) 長短期記憶模型 (Long Short-Term Memory, LSTM)

LSTM是RNN的一種特殊形式，專門設計用於處理長期依賴問題。它們能夠學習和記住長序列中的重要訊息。

應用例子：

- 長文件生成
- 語音合成
- 時間序列預測
- 音樂作曲

(四) 生成對抗網路 (Generative Adversarial Network, GAN)

GAN包含兩個相互競爭的類神經網路：一個生成新資料，另一個評估這些資料的真實性。透過這種競爭，GAN可以生成高度逼真的新資料。

應用例子：

- 資料增強
- 電玩遊戲 (video games) 中的逼真場景
- 藝術創作
- 生成逼真的虛構人臉

(五) 變換器 (Transformer)

變換器是一種處理序列資料的架構，它使用自我注意力機制 (Self-Attention) 來捕捉序列中元素之間的關係，這是近年來重要的深度學習創新之一。

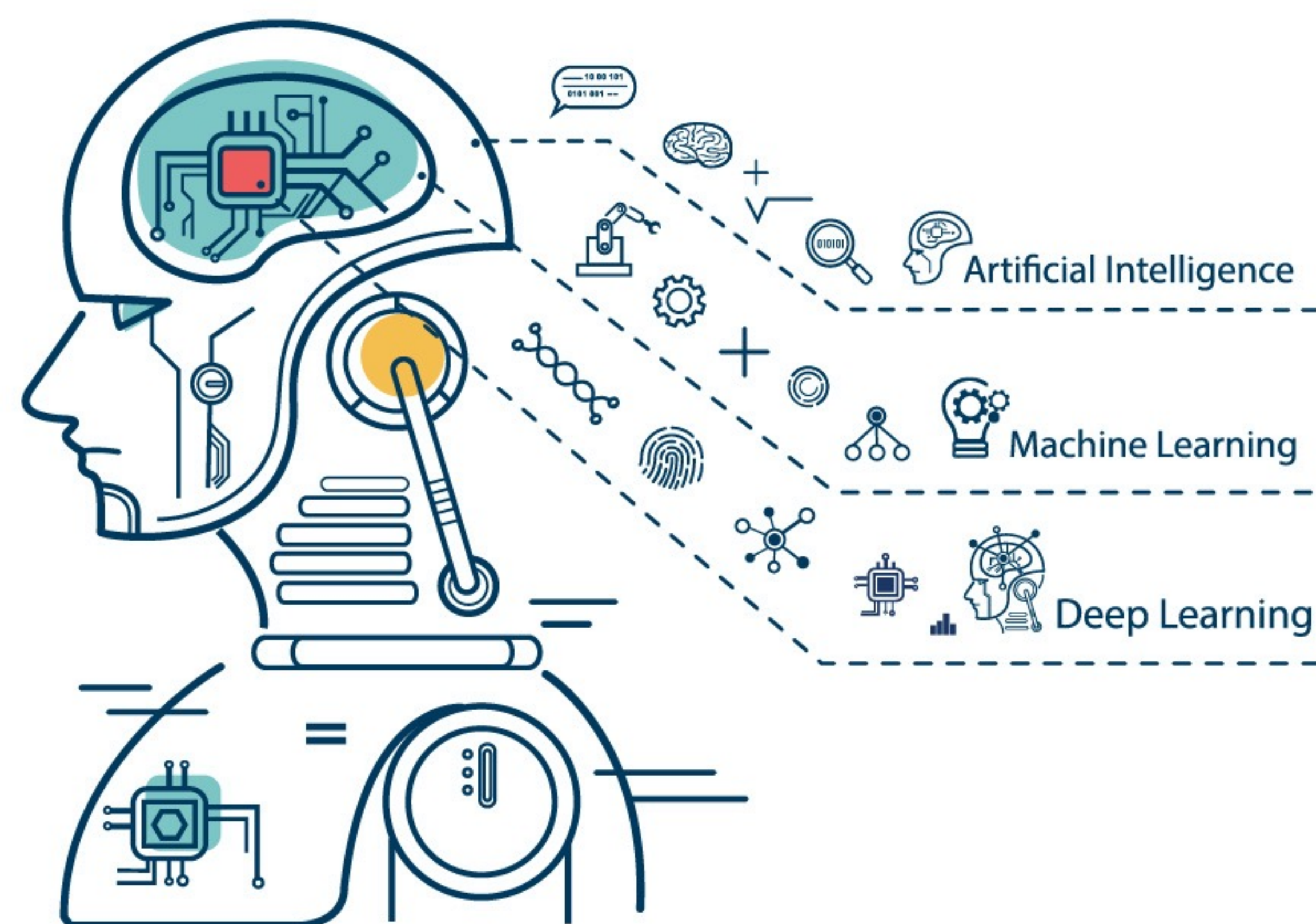
應用例子：

- 大型語言模型 (如GPT系列)
- 文件摘要
- 高品質機器翻譯
- 程式碼生成

五、深度學習的實際應用

深度學習已經在許多領域帶來了革命性的變化：

- (一) **計算機視覺**：從自動照片標記到醫學診斷，深度學習改變了計算機處理視覺訊息的方式。
- (二) **自然語言處理**：深度學習使機器能夠理解、生成和翻譯人類語言，支持了智慧助理、翻譯軟體和聊天機器人的發展。
- (三) **遊戲AI**：從AlphaGo到自動玩電玩遊戲（video games）的AI，深度學習使電腦能夠在複雜的環境中做出決策。
- (四) **藥物發現**：深度學習加速了新藥開發過程，幫助科學家預測分子結構和藥物相互作用。
- (五) **推薦系統**：你在社交媒體、購物網站和串流平台上看到的個性化推薦大多是由深度學習驅動的。



六、深度學習的優勢與挑戰

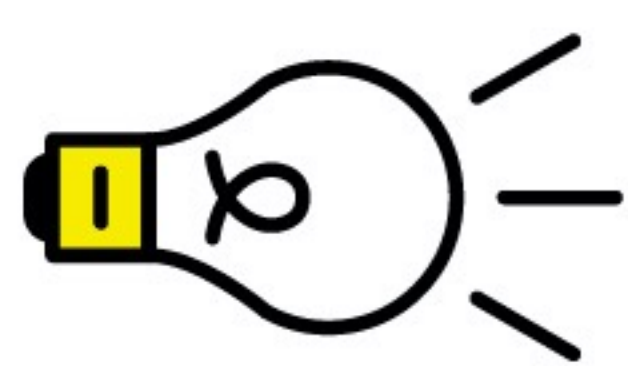
深度學習已經在許多領域帶來了革命性的變化：

優勢

- 自動特徵提取：無需手動設計特徵，模型可以直接從原始資料學習
- 處理非結構化數據：能夠處理圖像、語音和文件等複雜資料
- 強大的通用能力：訓練良好的模型可以在新資料上表現出色
- 持續改進：隨著更多資料的收集，模型性能通常會提高

挑戰

- 需要大量資料：深度學習模型通常需要大量的訓練資料
- 算力資源密集：訓練複雜模型需要強大的GPU和大量時間
- 黑盒（Black Box）性質：深度模型的決策過程往往難以解釋
- 過度擬合（Overfitting）風險：如果訓練不當，模型可能會記住特定資料分布，而無法成功掌握新資料的分佈。
- 偏見和公平性問題：模型可能會放大訓練資料中的偏見



動動腦，想一想

- (1) 卷積神經網路（CNN）、遞迴神經網路（RNN）與生成對抗網路（GAN）分別適用於哪些類型的資料或任務？請舉出至少一個你熟悉的應用範例。
- (2) 深度學習與傳統機器學習有何不同？請從「特徵提取」的角度，說明深度學習為什麼能處理圖像或語音這類非結構化資料。

2-2-3 生成式AI

一、什麼是生成式AI？

生成式AI是人工智慧的一個分支，專注於創造新的內容，而不僅僅是分析或分類現有資料。這些系統可以生成各種形式的內容，包括文件、圖像、音樂、影片甚至3D模型。你可能已經使用過或聽說過像ChatGPT、DALL·E、Midjourney或Stable Diffusion這樣的工具，這些都是生成式AI的例子。

生成式AI的魅力在於它能夠創造出全新的、原創的內容，有時甚至令人難以區分是人類還是機器創作的。這種技術正在重塑我們對創造力的理解，並為藝術、設計、科學研究等各個領域開啟了新的可能性。

二、生成式AI技術的演進

◆ 早期基礎（1950s-2000s）

- 生成式AI的概念可以追溯到人工智慧研究的早期。在這個階段，研究者開發了基於規則的系統，用於生成簡單的文件和圖像。這些系統主要依賴於預設的模板和人工撰寫的規則，創造力受到很大限制。
- 例如，早期的文件生成系統如ELIZA（1966年）主要是透過模式匹配和簡單的轉換規則來回應使用者輸入的內容，而不是真正"理解"語言或創造新內容。

◆ 生成對抗網路(GANs)時代（2014年起）

- 生成式AI的第一個重大突破來自於Ian Goodfellow及其團隊在2014年提出的生成對抗網路（GAN）。GAN引入了一個革命性的想法：兩個類神經網路相互競爭以提高性能。
- GAN的工作原理
GAN包含兩個核心組件：
 1. 生成器(Generator)：負責創建看起來真實的內容（如圖像）
 2. 判別器(Discriminator)：嘗試區分生成的內容和真實內容

這兩個網路進行一場「遊戲」：生成器努力製造能夠欺騙判別器的假內容，而判別器則努力提高自己的辨別能力。隨著訓練的進行，兩個網路都會不斷改進，最終生成器能夠創造出非常逼真的內容。

這個過程類似於偽造者和鑑定專家之間的鬥爭：偽造者（生成器）不斷改進技術以製造更逼真的偽造品，而鑑定專家（判別器）則努力提升自己識別偽造品的能力。

GAN在圖像生成領域獲得了驚人的成效，使我們能夠生成高度逼真的人臉、轉換圖像風格，甚至創造出完全虛構的場景。例如，ThisPersonDoesNotExist.com網站展示的完全逼真的人臉圖像都是由GAN生成的，這些「人」實際上並不存在。

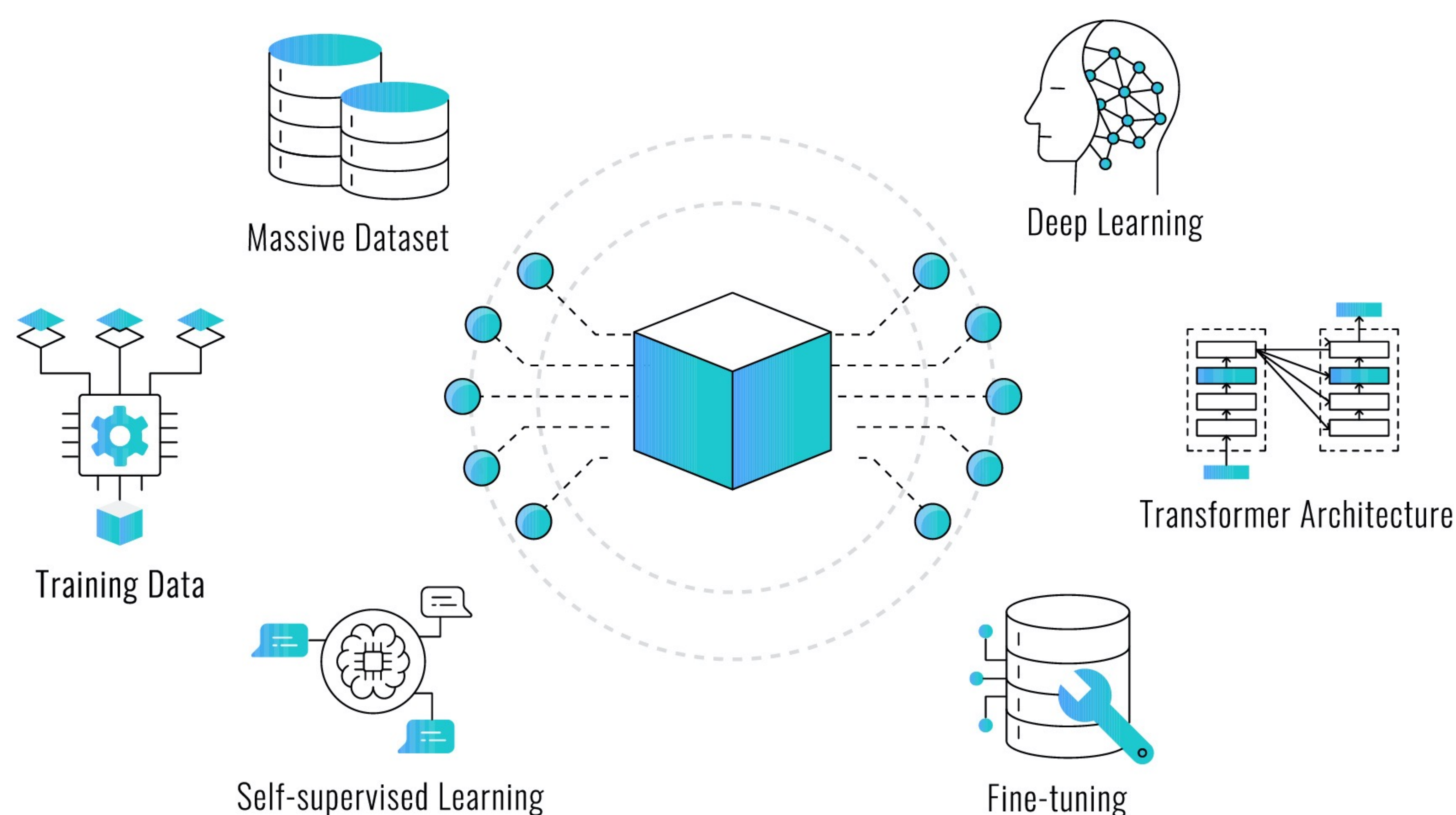
◆ 大型語言模型(LLM)時代（2017年起）

大型語言模型（Large Language Model, LLM）代表了文件生成的重大突破。這些模型透過分析大量文件資料來學習語言的模式、結構和知識，然後能夠生成連貫、流暢且訊息豐富的文件。簡單來說，LLM就像是超大型的「語言資料庫」，讀過大量的網路文章、書籍和資料後，學會如何用自然語言回答問題或進行對話。

* LLM的工作原理

大型語言模型的技術細節

- 現代LLM基於Transformer架構，其關鍵元件包括：
 1. 自我注意力機制：允許模型在生成每個單字時考慮輸入文件中所有其他單字的影響。這使得模型能夠捕捉長距離依賴和複雜的語境關係。
 2. 多頭注意力：模型同時從多個不同的"視角"關注輸入，捕捉不同類型的關係和模式。
 3. 位置編碼：因為Transformer不像RNN那樣直接處理序列，所以需要位置編碼來告訴模型單字在句子中的位置。
 4. 層標準化和殘差連接：這些技術幫助模型訓練更加穩定，防止梯度消失或爆炸問題。



LLM的訓練分為兩個主要階段：

- 預先訓練：模型在大量文件上學習語言的基本模式和知識
- 微調：模型在特定任務或風格的資料上進行進一步最佳化

最新的LLM如GPT-4等使用了稱為RLHF（基於人類回饋的強化學習）的技術，透過人類回饋來改進模型輸出的品質和安全性。LLM基於一種稱為Transformer的類神經網路架構，這種架構在2017年由Google研究團隊提出。Transformer的關鍵創新在於「注意力機制」，它允許模型在處理文件時考慮所有單字之間的關係，而不是單獨處理每個單字。訓練LLM的過程包括：

1. 收集大量文件資料（書籍、網頁、維基百科等）
2. 讓模型預測序列中的下一個單字或填補缺失的單字
3. 不斷調整數十億甚至數萬億個參數，直到模型能夠準確預測

結果產生一個能夠理解上下文、回答問題，甚至撰寫創意內容的系統。LLM的發展經歷了幾個里程碑：

- 2018年：Google的BERT模型提高了自然語言理解能力
- 2020年：OpenAI的GPT-3展示了驚人的文件生成能力
- 2022年：ChatGPT將LLM帶入主流，展示了對話式AI的強大潛力
- 2023-2024年：更強大的模型如GPT-4、Claude、Gemini等進一步提高了能力

現代LLM已經達到了令人印象深刻的能力水準，它們可以：



◆ 擴散模型時代（2020年起）

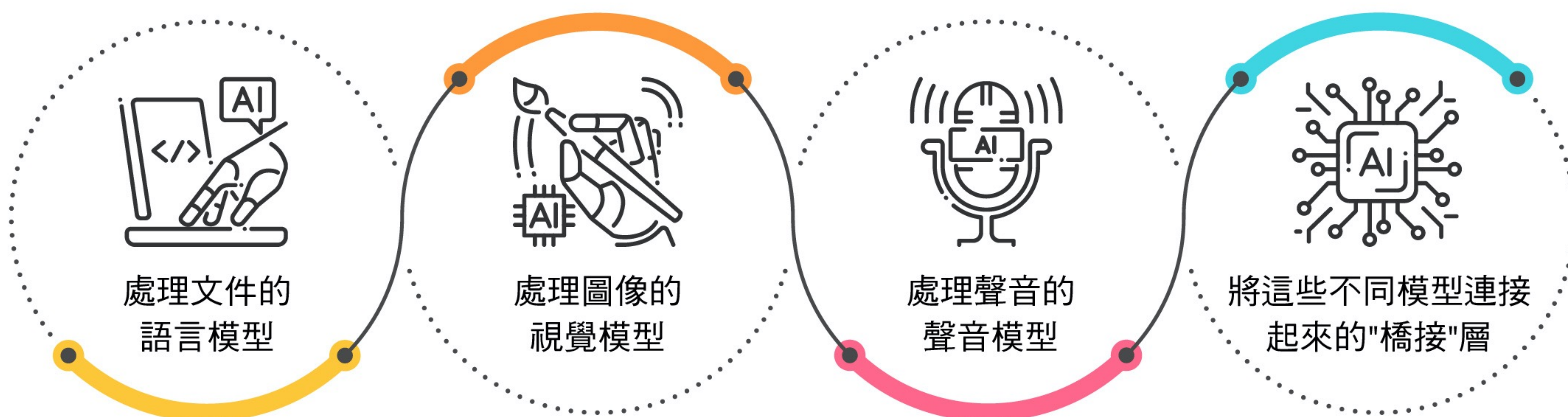
近年來，擴散模型（Diffusion Model）逐漸取代GAN成為圖像生成的主導技術。擴散模型工作原理不同於GAN，其原理為：

1. 首先向圖像中逐步添加雜訊，直到圖像完全變成隨機圖形。
2. 然後訓練類神經網路學習如何逆轉這個過程，從雜訊中重建原始圖像。
3. 一旦訓練完成，模型可以從純雜訊開始，生成全新的、逼真的圖像。

擴散模型代表了圖像生成的最新進展，應用於Stable Diffusion、DALL·E 2和Midjourney等流行工具。這些系統能夠根據文件描述生成精美的圖像，展示了令人驚嘆的創造力和細節把握能力。

◆ 多模態AI時代（2021年起）

多模態AI代表了生成式AI的最新發展，它們可以同時處理和生成多種類型的資料（文件、圖像、聲音及影片等）。多模態系統將不同的類神經網路模型整合在一起，創建一個能夠處理和生成多種媒體類型的統一系統。例如，一個多模態系統可能包含：



這種整合使得系統能夠：

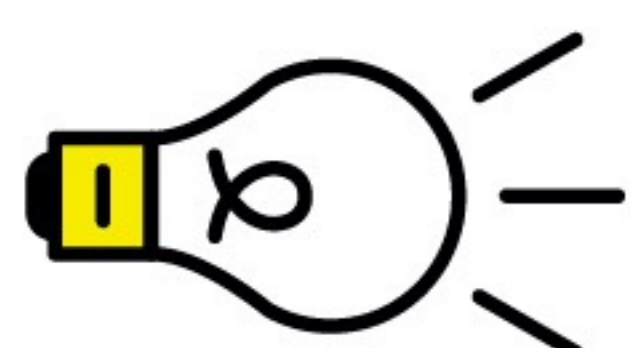
- 根據文件描述生成圖像（DALL·E, Midjourney）
- 根據圖像生成文件描述（Microsoft Azure Computer Vision API，Google Cloud Vision API）
- 將文件轉換為語音，或將語音轉換為文件（OpenAI Whisper）
- 生成音樂或影片（Open AI Sora）

最新的多模態模型，如GPT-4、Claude 3和Gemini，已經能夠理解圖像、解讀圖表、分析文件，並提供基於視覺和文件輸入的回應。這種綜合能力代表了邁向真正理解和創造多媒體內容的重要一步。

* 多模態AI的技術細節

多模態系統使用不同的架構來處理和生成不同類型的資料：

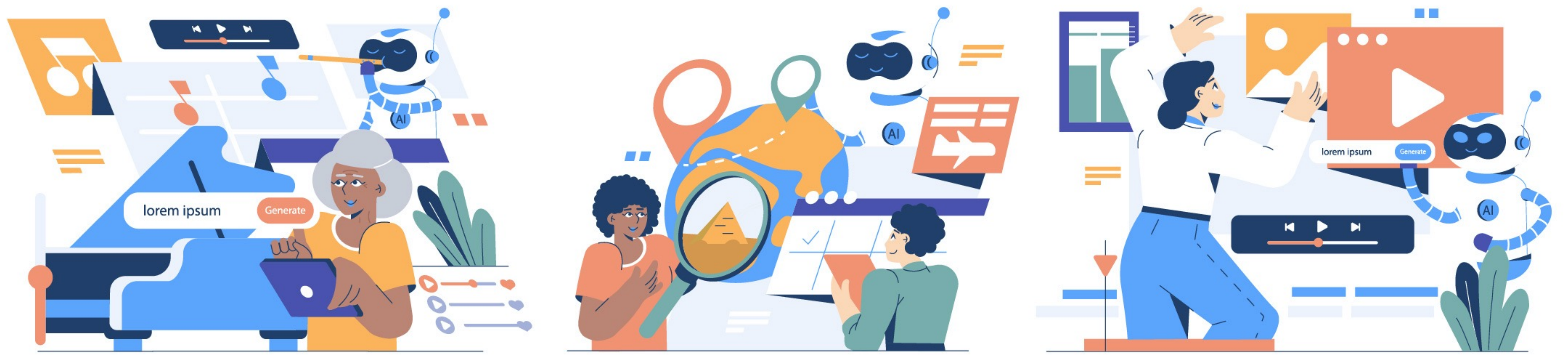
1. 文件-圖像模型（如DALL·E, Stable Diffusion）：
 - 使用擴散模型從雜訊中逐步生成圖像
 - 使用文件編碼器將文件描述轉換為向量
 - 使用條件機制讓文件向量引導圖像生成過程
2. 視覺語言模型：
 - 使用視覺編碼器（如ViT）處理圖像
 - 使用語言模型處理文件
 - 使用對齊層將視覺和語言表示映射到同一空間
3. 跨模態注意力：允許模型在不同模態之間建立關係，例如將文件中的特定詞與圖像中的特定區域關聯起來。



動動腦，想一想

- (1) 大型語言模型（LLM）如GPT-4能完成哪些主要任務？請說明LLM背後的Transformer架構有何特點，讓它能有效處理自然語言。
- (2) 多模態AI是生成式AI發展的重要趨勢之一。你認為多模態AI與單一模態（如僅處理文字或圖像）AI相比，在哪些方面更具有優勢？請說明並舉例。

2-3 生成式AI的應用及挑戰



2-3-1 生成式AI的多元應用場景

生成式AI（Generative AI）近年來迅速崛起，透過深度學習、生成對抗網路（GAN）、擴散模型（Diffusion Model）與大型語言模型（LLM）等技術，為各行各業帶來創新與便利。以下，我們將針對生成式AI在各個領域的應用進行深入探討。

一、文件生成與自然語言處理（NLP）→ 應用案例：ChatGPT、Claude、Gemini等對話式AI

生成式AI在文字領域的應用，主要集中於文件生成與語言理解。這些AI能自動撰寫故事、新聞報導、技術文件、廣告文案，甚至程式碼。例如，許多企業已經使用ChatGPT來撰寫行銷素材或客戶服務回應，提高工作效率。此外，AI還能用於文件摘要、文件翻譯、內容重寫等任務，支援跨語言溝通與資料整理。

二、圖像生成與視覺設計 → 應用案例：DALL·E、Midjourney、Stable Diffusion

生成式AI能根據用戶的描述或指令，自動創建圖像，為設計、廣告、藝術創作提供新的工具。例如，設計師只需輸入「夕陽下的未來城市」，AI就能生成具未來感的城市風景圖。

三、音樂生成與多媒體創作 → 應用案例：Aiva、Soundraw、OpenAI Jukebox

音樂生成AI能依據情緒或風格需求，創作旋律、背景音樂，廣泛應用於廣告配樂、影片剪輯、遊戲音效等。

四、影片與動畫生成 → 應用案例：Sora、Runway Gen-2、Pika Labs

AI影片生成技術日益成熟，像Sora可根據指令生成高畫質短影片，Runway Gen-2則能將靜態圖片轉換成動畫片段或短片。

五、語音合成與轉換 → 應用案例：ElevenLabs、Google TTS、Speechify

生成式AI能將文件轉換為自然流暢的語音，或將語音內容即時轉為文字，廣泛應用於智慧助理、有聲書、客服語音系統等。

六、3D建模與虛擬世界生成 → 應用案例：Nvidia GauGAN、Luma AI

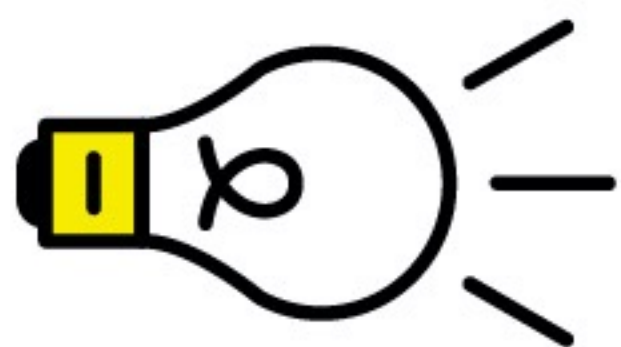
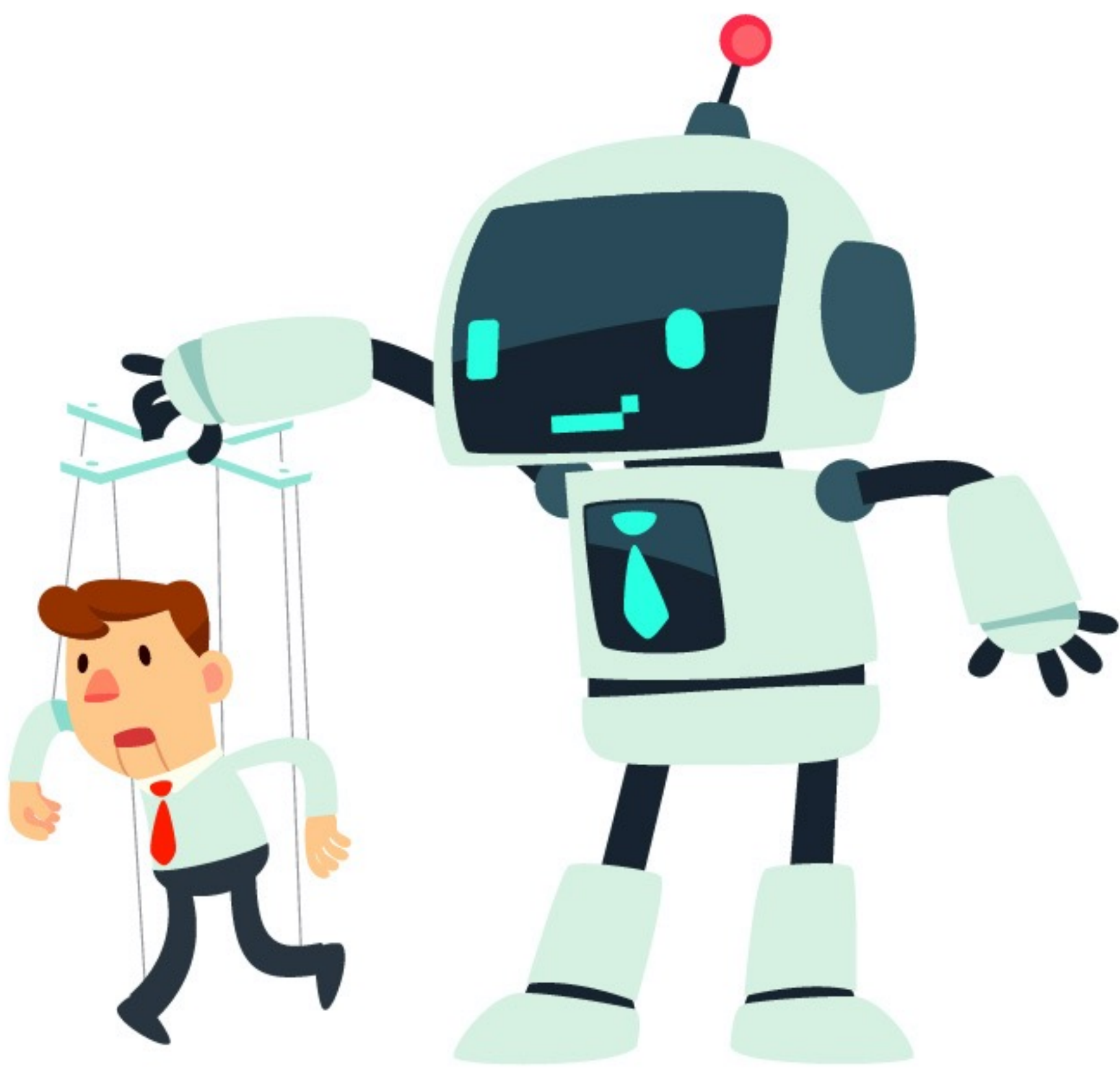
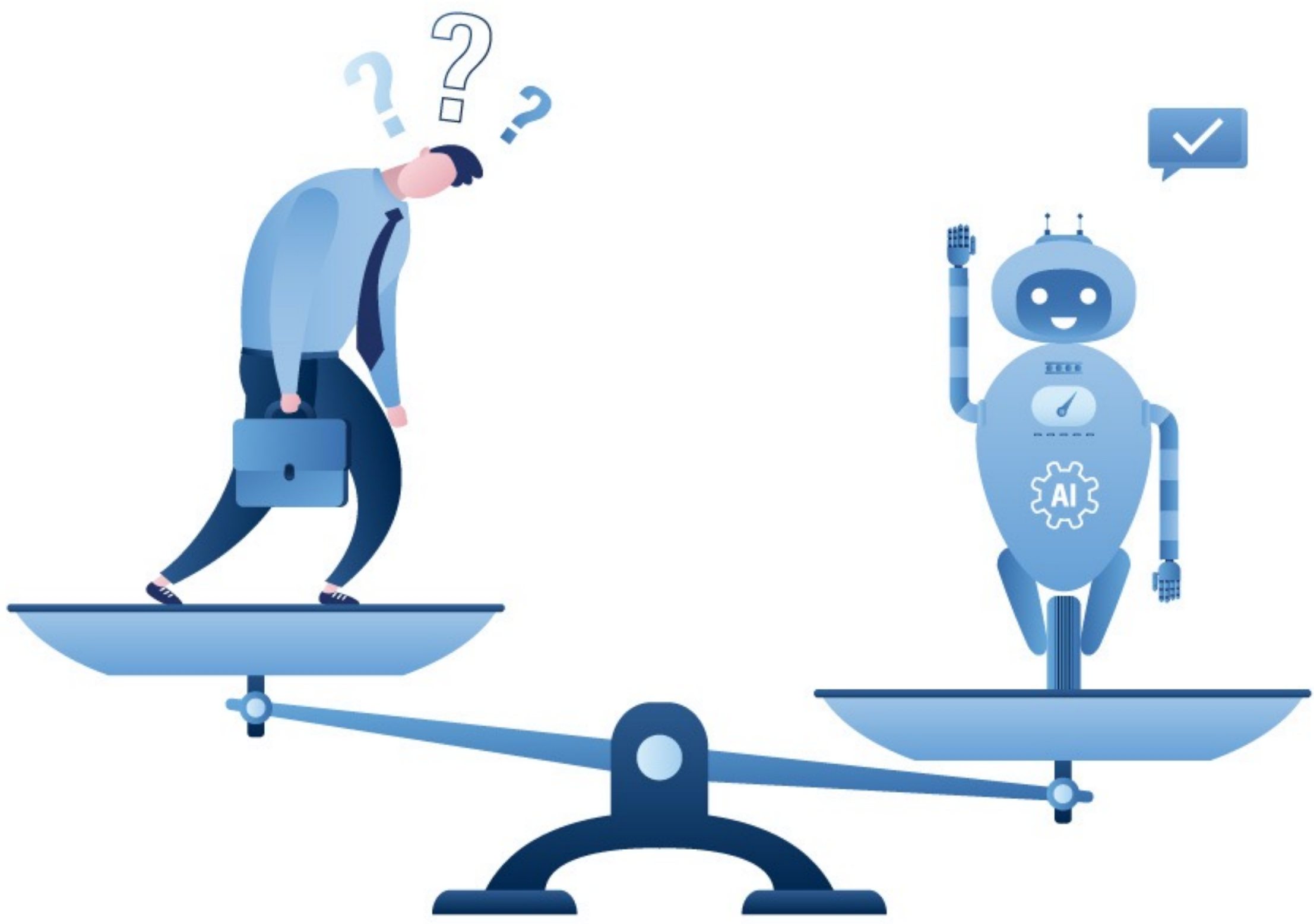
AI技術不只在2D圖像領域活躍，現在更可自動生成3D模型或虛擬場景。

2-3-2 生成式AI的挑戰與倫理議題

- 一、假資訊與深偽（Deepfake）問題
- 二、智慧財產權爭議
- 三、AI偏見與歧視
- 四、社會與就業影響
- 五、決策透明度不足

應用與風險的對照表

生成式AI應用	潛在風險
圖像生成（Midjourney）	假新聞、侵權風險
文件生成（ChatGPT）	誤導資訊、內容抄襲
音樂生成（Aiva）	原創性爭議、侵害作曲人權益
影片生成（Sora）	深偽影片誤導、公眾信任下降
語音合成（TTS工具）	假冒聲音詐騙、身份冒用



動動腦，想一想

- (1) 舉例說明生成式AI如何應用於媒體或創意產業，並分析其可能帶來的倫理挑戰。
- (2) 你認為AI自動生成的作品是否應受到著作權保護？請結合實際案例說明你的觀點。
- (3) 若未來AI能完全取代圖像、文字、音樂創作，你認為人類創作者在職場上應如何面對這樣的變化？請提出兩個具體因應策略。

2-4 總結

AI新世界，從現在開始！

想像一下未來的世界，AI幾乎無所不在——從你手機裡的語音助理，到網路上自動生成的動畫、音樂、甚至3D世界，這些都不再是科幻電影裡才會出現的情節，而是我們現實生活的一部分。這份資訊素養教材《AI未來新世界（上）》就是帶大家了解，AI到底怎麼影響我們的生活，甚至我們的未來。

這個單元最重要的，就是讓我們學會三件事：

1. AI是什麼？它怎麼發展到現在這麼厲害？
2. AI到底可以做些什麼？（答案是：比你想的還多！）
3. 當AI變得越來越聰明，我們又該怎麼跟它「和平相處」？



AI到底能做什麼？超乎你的想像！

我們學到的AI應用超級廣，例如：

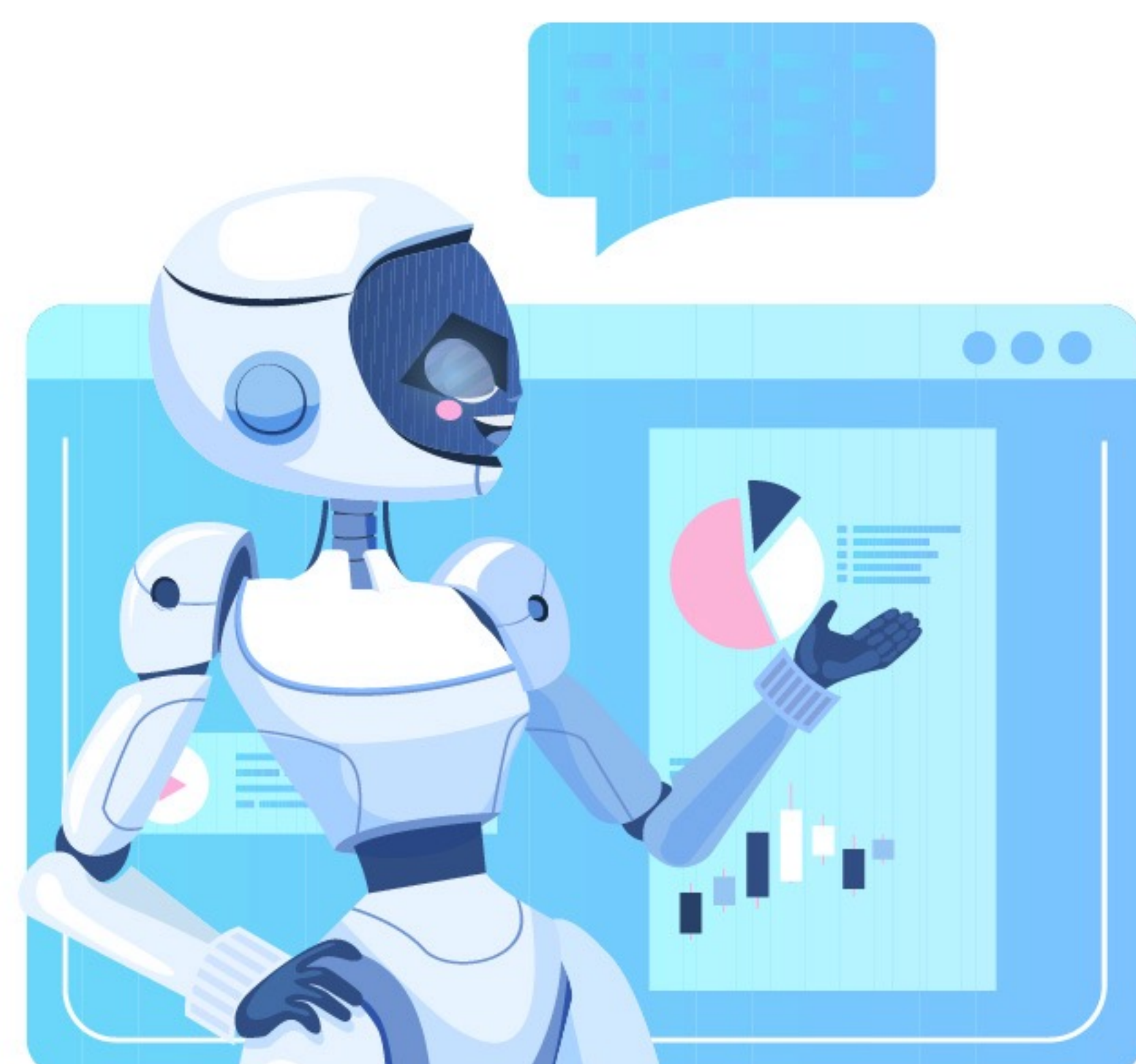
- ChatGPT、Gemini這些AI不只會聊天，還能幫忙寫故事、改作文，甚至寫程式。
- DALL·E、Midjourney能根據簡單的描述，畫出超酷的插畫或設計圖。
- Sora、Runway Gen-2讓AI變身成導演，生成短影片或動畫片段。
- Aiva能創作音樂，讓沒有音樂基礎的人也能製作屬於自己的配樂。

簡單來說，AI已經變成設計師、作家、音樂人、甚至導演的小助手，未來工作中，AI會是我們每天都會用到的夥伴。

但AI這麼強，也有風險！

AI雖然帶來了方便，但也帶來不少問題，例如：

- AI生成的假資訊、假影片，會不會讓我們分不清真假？
- AI的設計是否會因資料偏見而導致歧視？
- AI創作的內容，著作權到底屬於誰？
- 當AI能自動創作，會不會搶走創作者或設計師的工作？



這些問題讓我們不得不去思考：科技越來越強大時，人類要怎麼管理好AI，讓它真正成為我們的好幫手，而不是製造更多問題的麻煩角色？

★小結：透過這個單元，你不只認識AI的「超能力」，更學會去思考它背後的倫理、法律和社會議題。未來的你，不只是AI的使用者，也可能是幫忙設計規則、解決AI帶來問題的關鍵角色。

AI時代，已經來臨，準備好和它一起前進了嗎？

03 ▶ 三絕韋編 鑑往知來

3-1 中文參考文獻

- AWS. (2023)。什麼是AGI？- 通用人工智慧說明。民國114年3月13日
取自：<https://aws.amazon.com/tw/what-is/artificial-general-intelligence/>
- AWS. (2023)。什麼是人工智慧？AI 應用領域與3 大AI 種類介紹！民國114年3月13日
取自：<https://aws.amazon.com/tw/what-is/artificial-intelligence/>
- AWS. (2023)。什麼是生成式AI？- Gen AI 說明。民國114年3月13日
取自：<https://aws.amazon.com/tw/what-is/generative-ai/>
- AWS. (2023)。生成式人工智能（GenAI）新世界：過去、現在和未來。民國114年3月13日
取自：<https://aws.amazon.com/tw/local/hongkong/generative-ai/genai-blog-100/>
- 國立臺灣師範大學. (2023)。生成式AI之學習應用及參考指引。民國114年3月13日
取自：https://ctld.ntnu.edu.tw/generative_ai
- 教育部 (2024)。中小學使用生成式人工智慧注意事項。民國114年3月13日
取自：https://pads.moe.edu.tw/pads_front/index.php?action=download
- 教育部 (2024)。教育部中小學數位教學指引3.0。民國114年3月13日
取自：https://pads.moe.edu.tw/pads_front/index.php?action=download
- 林一平 (2022年8月5日)。電腦科學理論之父——圖靈。科學人。民國114年3月13日
取自：<https://www.scitw.cc/posts/10661>
- 林柏宏 (譯) (2020)。AI之書：圖解人工智慧發展史。臺北市：時報文化。（原著出版於2019年）
- 蔡宗翰 (2022)。寫給中學生看的AI課：AI生態系需要文理兼具的未來人才。臺北市：三采文化。
- 蕭英勵 (2018)。人工智慧科技的未來應用與省思。師友雙月刊, (610), 47-52。
- 臺北市政府教育局作者群 (2020)。人工智慧教育教材—高中AI生活大智慧。民國114年3月13日 取自：https://drive.google.com/file/d/1SgRxxhXA6HVM4UiXIPRyqZEL1zJ_OuS6/
- 魏志浩 (2024)。生成式AI於零售業應用之探究—以Microsoft AI為例（碩士論文）。國立臺灣科技大學。
- 黃經良 (譯) (2022)。AI大圖鑑 = Visual book of the artificial intelligence。新北市：人人。（原著出版於2020年）

3-2 英文參考文獻

- Bader, S., & Hitzler, P. (2005). Dimensions of neural-symbolic integration—A structured survey. In *We Will Show Them! Essays in Honour of Dov Gabbay* (pp. 167-194). College Publications.
- Duede, E., Dolan, W., Bauer, A., Foster, I., & Lakhani, K. (2024). Oil & Water? Diffusion of AI within and across scientific fields.
- Garcez, A. S. d., Broda, K., & Gabbay, D. M. (2002). *Neural-symbolic learning systems: Foundations and applications*. Springer Science & Business Media.
- Google Cloud. (2023). AI and machine learning: What's the difference?. 2025.03.10 Retrieved from <https://cloud.google.com/learn/artificial-intelligence-vs-machine-learning?hl=zh-TW>
- Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., & Naughtin, C. (2023). Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960-2021. *Technology in Society*, 74, 102260.
- Hochreiter, S. (2022). Toward a broad AI. *Communications of the ACM*, 65(4), 56-57. Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., & Zhou, B. (2021). Trustworthy AI: From principles to practices.
- Marcus, G. (2020). The next decade in AI: Four steps towards robust artificial intelligence.
- Midjourney. (2023). Midjourney AI. 2025.03.10 Retrieved from <https://www.midjourney.com/>
- OpenAI. (2023). ChatGPT. 2025.03.10 Retrieved from <https://openai.com/chatgpt>
- OpenAI. (2023). DALL·E 3. 2025.03.10 Retrieved from <https://openai.com/dall-e>
- OpenAI. (2024). Sora. 2025.03.10 Retrieved from <https://openai.com/sora>
- Serafini, L., & Garcez, A. d. (2016). Logic tensor networks: Deep learning and logical reasoning from data and knowledge. In *Proceedings of the 10th International Workshop on Neural-Symbolic Learning and Reasoning* (pp. 1-10).
- Tim Mucci. (2024). The History of AI. IBM. 2025.03.14 Retrieved from <https://www.vivepostwave.com/agi-deepmind/>
- Xu, W., Ge, L., & Gao, Z. (2021). Human-AI interaction: An emerging interdisciplinary domain for enabling human-centered AI. *arXiv preprint arXiv:2112.01920*. 2025.03.12 Retrieved from <https://arxiv.org/abs/2112.01920>

04 ▶ 四通八達 小試身手

主題探討

● ● ●



AI技術的發展已經深刻改變了我們的生活。請舉出兩個你生活中常見的AI應用，並說明它們如何影響你的學習或日常生活。

● ● ●



請解釋何謂弱人工智慧（Weak AI）與強人工智慧（Strong AI），並比較兩者在應用層面上的差異。

● ● ●



請用簡報工具，繪製一張「AI發展歷程時間軸圖」，從1940年代到2024年，標註至少6個重要的AI技術發展里程碑。

● ● ●



AI可能會產生哪些倫理或法律風險？請選擇一個問題（如AI偏見、隱私侵犯或AI假訊息），提出你的觀點與解決建議。

● ● ●



利用你熟悉的生成式AI工具（如ChatGPT或DALL·E），生成一段文件或圖片，並與同學討論：AI生成內容與人類創作有何不同？有無潛在的誤導風險？



請選擇監督式學習、非監督式學習或強化學習其中之一，設計一個小型的案例（如垃圾郵件分類或客戶分群）並簡述其機器學習流程。



請從「AI對社會的影響」角度，分析AI技術在醫療或交通領域中可能帶來的正面與負面影響。



請以AlphaGo、GPT-4或Sora等案例，說明AI從單一任務（如下棋）發展到多模態AI的過程，這樣的技術演進可能帶來哪些挑戰？



請與小組合作，蒐集最新的多模態AI應用實例（如DALL·E、Midjourney、Sora等），簡報其運作原理與應用場景，並提出未來可能的新應用構想。



資訊素養與倫理 高中職 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

林湧順 臺北市立大直高級中學校長

指導委員 / 盧東華 臺北市立大學助理教授

李啟龍 國立台灣師範大學附屬高級中學資訊室主任

總 編 輯 / 劉亦陞 臺北市立大直高級中學圖書館主任

執行編輯 / 魏仲良 臺北市立大直高級中學資訊組組長

編審委員 / 王鼎中 臺北市立建國高級中學老師

朱德清 臺北市立中山女子高級中學老師

侯莉莉 臺北市立內湖高級中學資訊組組長

孫晉忻 臺北市立景美女子高級中學老師

陳佳宜 臺北市立麗山高級中學老師

陳哲雋 臺北市立南港高級中學資訊組組長

陳 晉 臺北市立復興高級中學老師

陳瑞宜 臺北市立大同高級中學研發處主任

楊喻文 臺北市立第一女子高級中學資訊組組長

蔡志敏 臺北市立大同高級中學老師

蔡明男 臺北市立內湖高級中學圖書館主任

賴和隆 臺北市立中正高級中學老師

韓君尹 臺北市立和平高級中學資訊組組長

羅玕貞 臺北市立內湖高級中學老師

(按姓氏筆畫排列)

承辦單位 / 臺北市立大直高級中學

出版日期 / 中華民國114年3月



未來新世界

站在巨人肩膀上蛻變進化



臺北市政府教育局

DEPARTMENT OF EDUCATION
TAIPEI CITY GOVERNMENT

資訊素養與倫理 高中職 5 版

臺北市立內湖高級中學 羅玗貞老師
臺北市立中正高級中學 賴和隆老師



未來新世界

解鎖未來 與倫理共舞



01 ▶ 一葉知秋 學思並進

1-1 學習目標

- 理解AI人工智慧在各領域的發展應用與倫理議題。
- 培養AI倫理思辯與實踐能力。
- 建立負責任的AI素養與價值觀。

▶ 核心素養

- 科S-U-B2 理解科技與資訊的原理及發展趨勢，整合運用科技、資訊及媒體，並能分析思辨人與科技、社會、環境的關係。
- 科V-U-B2 / 科C-U-B2 理解科技與資訊的原理及發展趨勢，具備科技、資訊、媒體的整合運用能力，並能分析、思辨、批判人與科技、社會、環境之關係。

▶ 學習表現

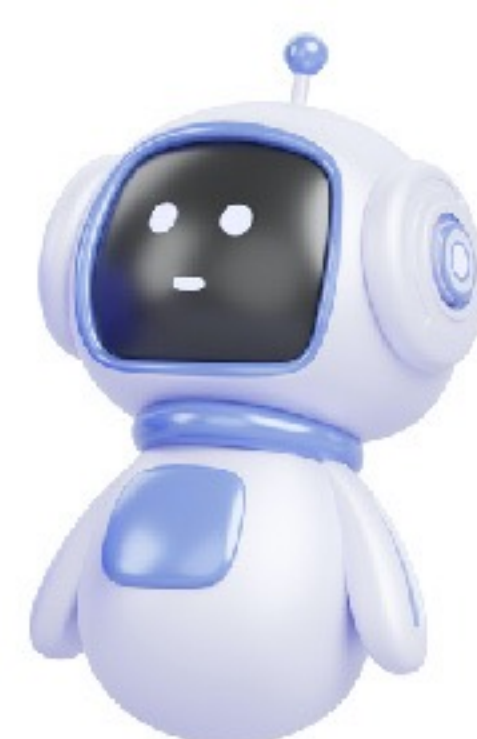
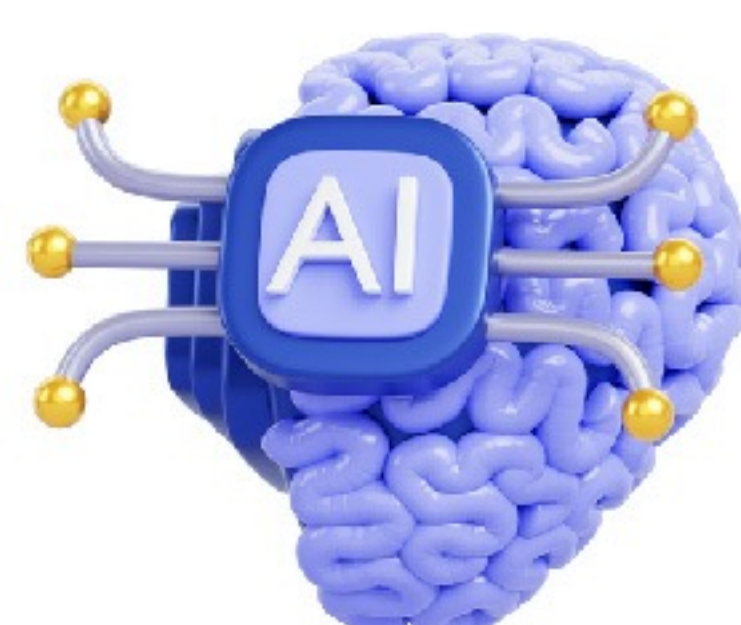
- 運 a-V-3 探索新興的資訊科技。
- 設 k-V-2 能了解科技產業現況及新興科技發展趨勢。

▶ 學習內容

- 資 H-V-3 資訊科技對人與社會的影響與衝擊。

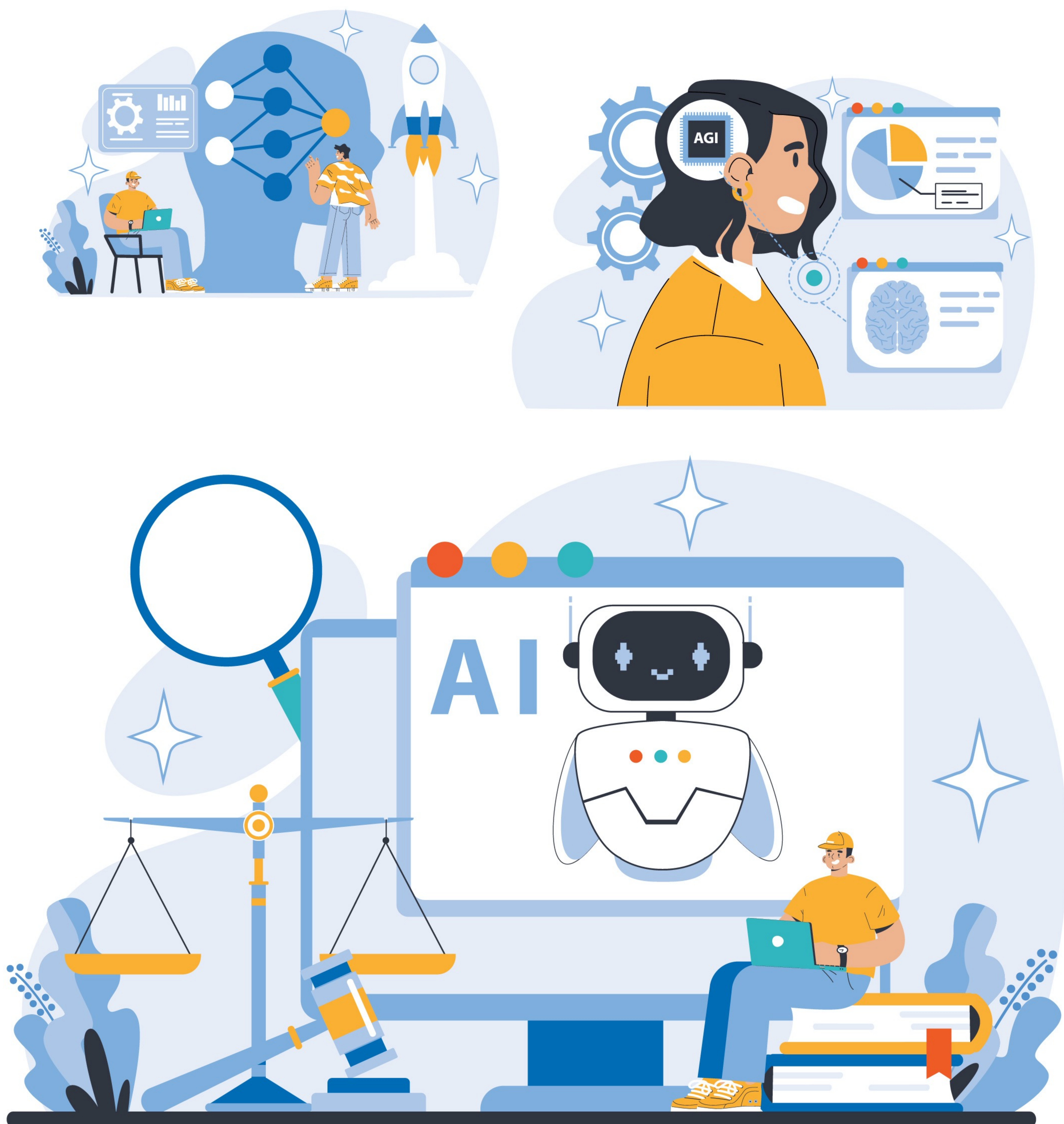
▶ 學習重點

- 學生能探討AI如何改變生活，以及帶來的淺在風險。
- 學生能理解AI倫理標準，並能應用於評估實際案例。
- 學生能培養查證、思辨能力，負責任地使用AI工具。



1-2 引言

人工智慧（AI）的發展，正以前所未有的速度改變著我們的世界，從醫療保健、交通運輸到金融服務、教育等各個領域，AI的應用無所不在，大大地提升了我們的生活品質和工作效率。然而，這項強大的技術也帶來了諸多倫理挑戰，如偏見、隱私侵犯、甚至對人權的威脅。為了應對這些挑戰，聯合國教科文組織發布了全球性人工智慧倫理標準，強調以人權為中心，確保AI的發展符合人類的價值觀和倫理規範。接下來將深入探討AI在各領域的應用，並重點闡述AI倫理的重要性，以及如何在日常生活中實踐這些倫理原則，從數據政策、倫理治理到環境永續等各個方面，引導我們成為更有責任感的科技使用者，共同塑造一個符合倫理規範的AI未來。



02 ▶ 兩全其美 觸類旁通

2-1 主題探索

2-1-1 AI的應用

從現在起，AI的發展日新月異，未來AI在各領域的應用將會更加深入和廣泛，並影響我們的生活與工作。AI的發展方向將在各方的努力下開始轉變：一方面，大家會開發更多開放原始碼的大型模型，讓研究人員可以一起實驗；另一方面，會開發更小、更有效率的模型，讓AI更容易使用，成本也更低，就像手機APP，可以安裝在我們的裝置上，隨時可以使用AI。未來AI的發展方向，將邁向利用更少的資源，提供更精準的結果。

以下就AI在各個領域的應用，作進一步的說明：

一、醫療保健

AI可以協助醫生更準確地診斷疾病，例如，分析X光片等醫學影像，以協助醫生診斷腫瘤等疾病；AI能自動標記影像中的異常，及早發現疾病，大幅減少醫師的診斷時間，提高工作效率。AI也可以根據你的基因和生活習慣，提供個人化的健康建議和治療方案，另外，AI將在藥物研發、手術機器人、健康管理等方面發揮重要作用。

二、交通運輸

AI將在自動駕駛汽車、智慧交通管理系統、無人機配送等方面發揮重要作用，AI自動駕駛汽車能辨識紅綠燈或行人，判別路上狀況，可以減少交通事故，提高交通效率；智慧交通管理系統可以根據實時路況，優化交通流量；無人機配送可以提高物流效率，降低成本。

三、金融服務

AI將在風險管理、客戶服務、投資顧問、反詐欺等方面發揮重要作用。銀行AI可以分析大量的金融數據，預測市場趨勢；AI可以提供智能客服，解答客戶的疑問；AI也可以提供個人化的投資建議，幫助客戶實現財務目標。

四、銷售與服務

AI將在產品推薦、庫存管理、客戶體驗、智能商店等方面發揮重要作用。商業AI可以根據客戶的購買歷史和瀏覽行為，推薦更符合其需求的產品；AI可以預測銷售趨勢，優化庫存管理；AI可以提供個人化的客戶服務，提高客戶滿意度，未來AI將提升機器理解和產生人類語言的能力，改善了溝通介面，也讓翻譯和情感分析工具更厲害，除了能與外國人溝通無礙，AI就像可以讀懂客戶心情的朋友，分析客戶在社群媒體上的發文是開心還是難過，提供適當的回饋。



五、生產與製造

AI將在生產流程優化、品質檢測、設備維護、智能工廠等方面發揮重要作用。例如，AI可以分析生產數據，優化生產流程；AI可以檢測產品缺陷，提高產品品質；AI可以預測設備故障，減少停機時間。

六、教育與學習

AI將在適性化學習、智能輔導、自動化評分、語言學習等方面發揮重要作用。例如，AI可以根據學生的學習情況，提供適性化的學習內容和輔導；AI可以自動批改作業，減輕教師的負擔；AI可以提供語言學習工具，幫助學生提升語言能力。

七、網路安全

AI能自動掃描網路，偵測網路攻擊和威脅，提供更強大的防禦能力啟動防護機制，阻止攻擊擴散；或是自動修復被攻擊的系統、自動隔離病毒，並協助移除病毒。

此外，AI也能提供保護建議，根據使用者的網路使用習慣，提供個人化的安全，保護隱私，防止個人資料被洩漏，例如：你在社群網站上分享照片，AI會提醒注意隱私設定，避免被有心者利用。

八、科學研究

AI將在數據分析、藥物研發、基因組研究等方面發揮重要作用，使用AI能處理巨量科學數據，識別複雜模式；AI可以加速科學發現，如數學、氣候模型、天文學、材料科學等，提高研究效率；藥物研發系統可以根據分子特性，預測藥物效果；基因組研究應用可以加快基因序列分析，協助研發疾病治療方案。

九、軍事國防

AI將在情報分析、自主系統、網路安全等方面產生深遠影響，AI情報分析能處理大量監控數據，識別潛在威脅，可以提升預警能力，增強國防安全；自主系統如無人機可以執行危險任務，減少人員傷亡；網路安全防護可以實時監測攻擊行為，保護關鍵軍事資訊系統與基礎設施。

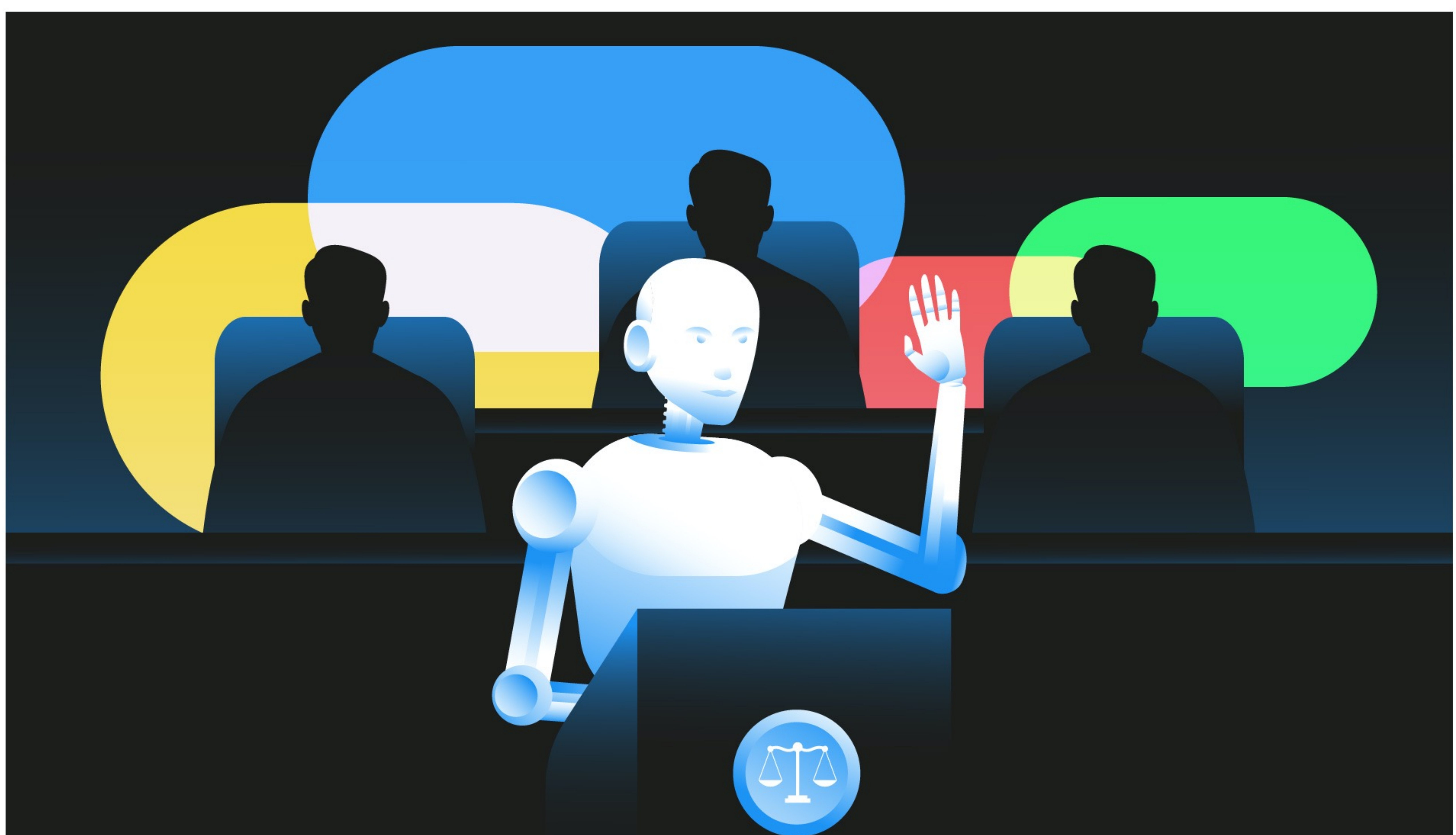
2-1-2 AI倫理

根據《聯合國教科文組織人工智慧倫理建議書》中提到，人工智慧雖然帶來許多機會，例如協助醫療診斷、促進人際連結和提高工作效率，但同時也帶來倫理問題，包含偏見、加速氣候惡化和威脅人權。為應對這些挑戰，聯合國教科文組織(UNESCO)在2021年11月發布了全球性人工智慧倫理標準，並獲得所有193個會員國的通過，其內容以四項核心價值為基礎，敘述如下：

- 一、尊重、保護和促進人權與基本自由及人類尊嚴。
- 二、確保多樣性和包容性。
- 三、環境與生態系統的繁榮。
- 四、在和平、公正和相互連結的社會中生活。

資訊倫理的PAPA理論早年是由美國學者梅森（Mason）於1986年提出的，主要涵蓋四個核心議題，這些議題反映了在資訊科技快速發展的背景下，社會面臨的主要倫理挑戰。這四個議題分別是：

- **隱私權（Privacy）**：涉及個人資訊的保護，探討在何種情況下可以公開或分享個人資料，以及如何保障個人不受未經授權的資訊使用。
- **正確性（Accuracy）**：關注資訊的準確性，強調資訊使用者應如何辨別資訊的真實性，以及資訊提供者在提供錯誤資訊時所需承擔的責任。
- **財產權（Property）**：涉及資訊的所有權問題，探討在何種情況下可以合法使用他人的資訊，並討論侵犯著作權的行為及其後果。
- **存取權（Accessibility）**：關注個人對資訊的存取權，強調如何確保每個人都能公平地獲取資訊，並避免資訊貧富差距的擴大。



以下是UNESCO 2021年人工智慧倫理標準與PAPA理論的比較，並說明AI倫理的獨特性：

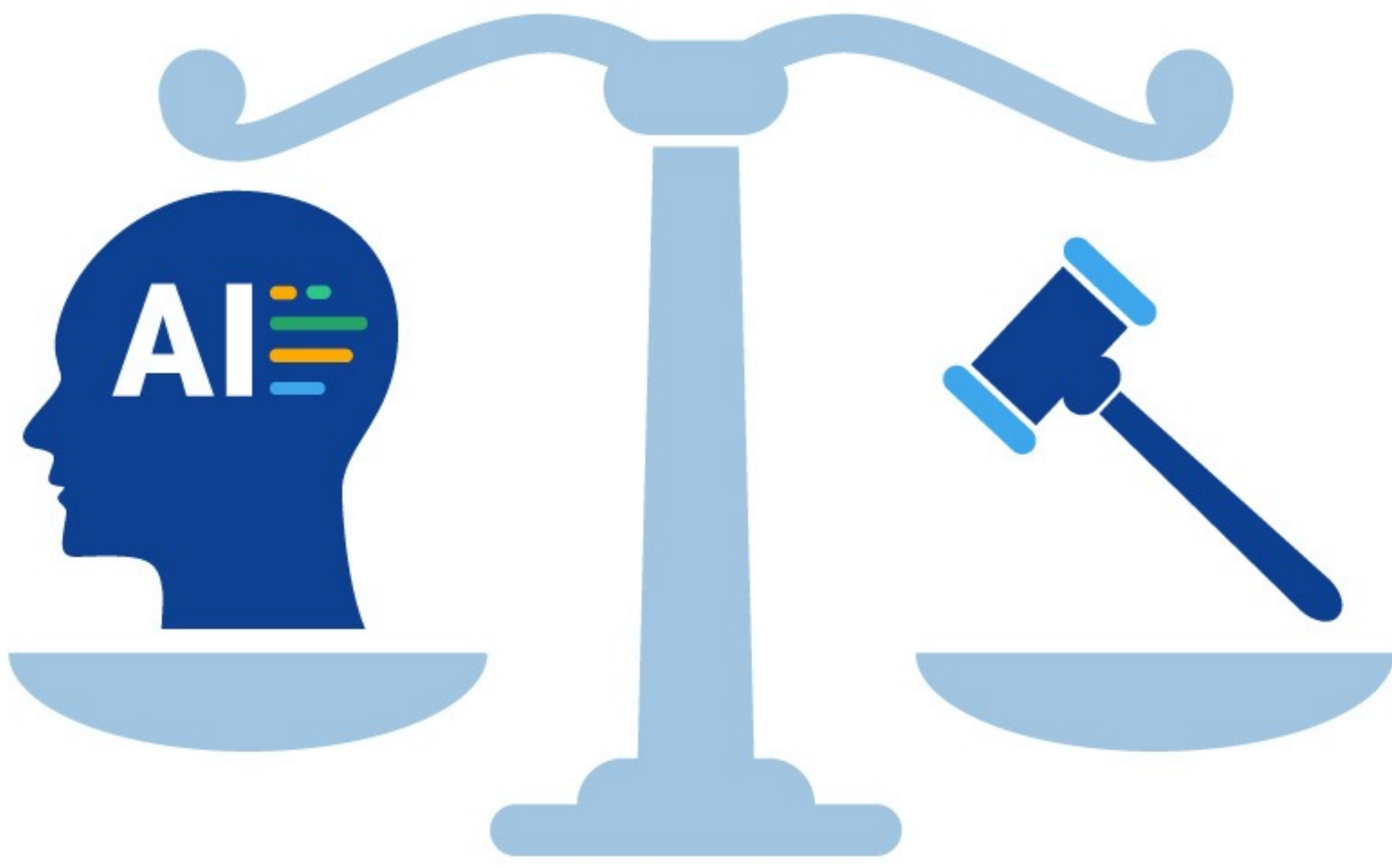
項 目	UNESCO AI 倫理標準 (2021)	PAPA 理論 (1986)	異同點與AI倫理 獨特性案例
核心價值 / 主題	人權、多樣性、生態、 和平與公正	隱私、準確性、財產權、 可存取性	AI 倫理標準較為宏觀 PAPA更聚焦於個人資訊倫理
隱私	涵蓋人權保護 但未具體強調個人隱私	強調個人資訊的隱私權	PAPA直接討論隱私 AI倫理則放在更廣泛的 權利保護中
公平與多樣性	強調多樣性、包容性、公正	未直接提及	AI倫理更重視公平性 與社會多樣性
環境責任	強調AI發展應考慮環境影響	未涉及	AI倫理關注生態永續 PAPA未涉及此議題
和平與社會影響	主張AI應促進社會和平與連結	未涉及	AI倫理將AI的社會影響納入考量
財產權	可能涉及公平性 但未特別強調	直接討論資訊與智慧財產權	PAPA更專注於資訊的擁有權 而AI倫理則關注整體社會公平
可存取性	強調AI技術應普及 不應造成數位落差	討論誰能獲取與使用資訊	AI倫理標準更強調技術 可及性與社會影響

關鍵差異分析

- (1) 技術複雜性：PAPA框架起源於1980年代資訊系統倫理，而AI倫理需處理機器學習黑箱化、自主決策等新挑戰。
- (2) 系統影響範圍：PAPA側重「人—資訊」關係，AI倫理延伸至「演算法—社會」的系統性影響，例如自動化決策的社會公平性。
- (3) 動態監管需求：AI倫理需即時因應技術迭代（如生成式AI），制定動態評估框架，而PAPA的適用情境相對穩定。

AI倫理標準的獨特性

- (1) 更宏觀的社會視角：PAPA理論主要關注個人資訊倫理，而UNESCO AI倫理標準擴展至社會、環境與全球治理。
- (2) 強調公平與包容：AI倫理不僅考慮技術本身，還關注AI如何影響不同群體，避免技術加劇社會不公。
- (3) 環境與生態考量：PAPA理論未涉及環境問題，而AI倫理標準認為AI發展應兼顧生態可持續性。
- (4) 社會和平與人權：AI倫理強調AI不應被濫用來破壞和平或侵犯人權，而PAPA側重於個人資訊層面。



綜上所述，PAPA理論是針對資訊科技倫理，這些議題不僅是資訊倫理的基礎，也是隨著科技進步而不斷演變的倫理挑戰，對於維護社會道德和個人權益具有重要意義；而UNESCO AI倫理標準則是針對AI技術對社會、環境與全球發展的影響，具有更廣泛的適用性和政策指導意涵。

2-1-3 AI倫理議題

以下我們舉例來說明聯合國教科文組織四項核心價值在校園中可能涉及的議題。

一、尊重、保護和促進人權與基本自由及人類尊嚴

假設你的學校考慮使用人臉辨識系統來追蹤學生出席率。根據這項核心價值，系統應：

- 需要學生同意收集他們的臉部資料
- 不能在未經通知的情況下用於其他目的（例如監控學生行為）
- 保護所有資料不被未授權存取
- 提供選擇退出的選項，不強制所有學生參與

這就像尊重每個同學的隱私權，不會在他們不知情的情況下拍照或分享他們的資訊，維護他們的尊嚴與自主權。

二、確保多樣性和包容性

如果學校使用AI系統來推薦升學或職業選項，這個系統應該：

- 不因性別而限制選項（例如不應只向男生推薦工程，只向女生推薦護理）
- 考慮不同文化背景學生的多元需求和價值觀
- 提供多語言支援，讓非母語學生也能平等使用
- 考量不同學習能力學生的需求，包括有特殊學習需求的同學

就像班上舉辦活動時，規劃方式要考慮到每位同學的參與能力和興趣，確保沒有人因為自身條件而被排除在外。

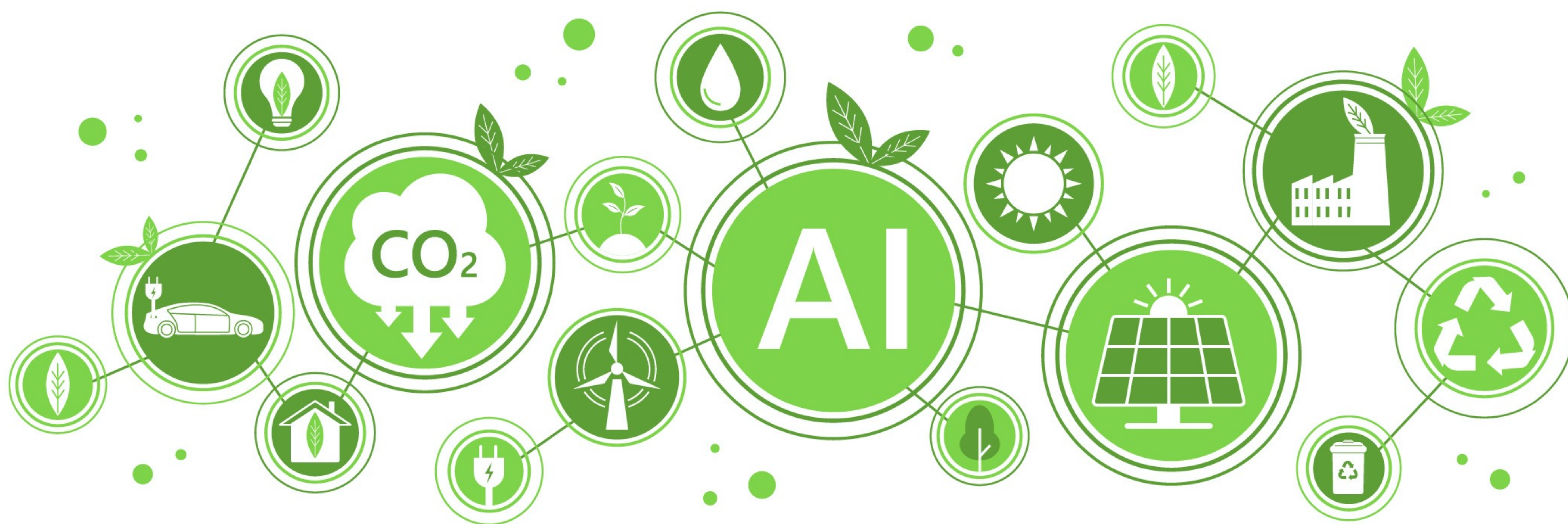


三、環境與生態系統的繁榮

學校若使用智慧校園系統來控制能源使用，這個系統應：

- 優化電力和水資源使用，減少浪費
- 使用節能伺服器主機而非耗能過大的設備
- 考慮設備的生命週期，包括製造和廢棄過程對環境的影響
- 利用收集的資料幫助學校改善碳足跡

這就像我們在家裡安裝智慧電燈不只是為了方便，也是為了節省能源；丟棄電池時會特別分類，考慮對環境的影響。



四、在和平、公正和相互連結的社會中生活

如果學校使用AI來協助管理學生社群平台或論壇，這個系統應：

- 檢測並防止網路霸凌，但不過度審查正常的討論
- 公平對待所有學生的意見，不偏袒特定群體
- 促進不同班級、年級間的交流與合作
- 確保資訊流通透明，避免造成誤解或分化

這就像班級設立公平的班規，每個人都有發言權，同時也有責任尊重他人；且當有爭議時，會透過對話而非對立來解決問題。

透過這些例子，我們可以看到AI倫理並不是遙不可及的概念，而是與我們日常生活息息相關。當我們思考學校或社區中的科技應用時，這四項核心價值可以幫助我們判斷這些應用是否真正對所有人都有益，是否尊重每個人的權利，以及是否對我們的環境和社會有正面的影響。

2-1-4 以人權為中心的AI倫理

AI與其他數位科技不同，具有深刻重塑社會、經濟和教育系統的潛力。與傳統資通訊技術（ICT）不同，AI帶來了獨特的倫理和社會挑戰，例如公平、透明度、隱私和問責等問題。基於此，聯合國教科文組織提出了十大核心原則，闡述了以人權為中心的AI倫理方針，以下舉例說明：

一、比例原則和不傷害原則

使用AI系統不得超出合法目的所需的範圍，應進行風險評估，以防止此類使用可能造成的危害。

例如：學校打算使用AI監控系統防止考試作弊。

在良好應用下，系統只在考試期間運作，僅偵測明顯的作弊行為（如使用手機或與他人交談），不會過度侵入學生隱私，如閱讀學生的筆記內容或分析面部微表情。

如果是不恰當的應用，系統會全天候監控學生，解析每個細微的手勢和眼神變化，甚至因為學生看向窗外片刻就將其標記為可疑行為。這就像老師在考試時巡視教室是合理的，但老師一整天站在你背後盯著你的每個動作則是過度干預。

二、安全與保障

AI參與者應該避免並解決不必要的危害（安全風險）以及易受攻擊的問題（安全風險）。

例如：學校使用AI系統管理學生進出校門。

在妥善的應用下，系統有多重安全保障，如加密儲存學生資料、定期更新以修補安全漏洞，並設有應急方案，萬一系統故障，保全人員可立即接管。

但如果系統沒有適當的安全措施，導致學生個人資料被駭客竊取；系統一旦故障，學校大門無法正常運作，造成混亂。這類似於使用網路銀行，該銀行必須確保存放的錢被安全保護，且萬一網站故障，還能透過其他方式存取帳戶。

三、隱私權與資料保護

在整個AI生命週期中必須保護隱私，也應建立完善的資料保護框架。

例如：使用APP學習英語，透過AI分析發音和學習習慣。

在良好應用下，APP清楚告知收集哪些資料、如何使用、保存多久；允許隨時刪除聲音紀錄；不將學生的學習資料分享給未經同意的第三方。

但若APP未經同意就錄下聲音用於其他目的；將學習數據賣給廣告商；即使刪除帳號，仍保留所有資料。這就像借用同學的筆記，應該尊重他的智慧財產，不會未經許可就複印或分享給其他人。

四、多方利益相關者和適應性治理

使用資料時必須尊重國際法和國家主權。此外，各方利害關係人的參與對於實現包容性的AI治理方法至關重要。

例如：城市決定在公共場所安裝AI監控攝影機。

市政府應邀請各方代表（包括普通市民、專家、隱私倡議者、執法人員）共同參與討論；同時考慮不同國家和文化對隱私的看法；在試行後根據公眾反饋調整系統。

但若市政府單方面決定，不徵求民意；忽視國際人權標準；一旦實施就不再接受任何修改建議。這就像制定班規，不僅需要班長和老師決定，還應該讓全班同學有發言權，並且根據實際情況進行調整。

五、責任和問責制

AI系統應該是可審查和可追蹤的，應建立監督、影響評估、審計和盡職調查機制，以避免與人權規範發生衝突和對環境福祉造成威脅。

例如：學校的AI系統協助分配學生到不同班級或選修課程。

在良好的應用，學校記錄系統運作的所有決策過程；定期審查分配結果是否公平；設立申訴管道讓學生可以質疑和更正不合理的分配；確認誰對最終決定負責。

但若沒人知道系統如何做決定；學校拒絕解釋為何某些學生被分到特定班級；當出現問題時，校方只說「這是電腦決定的」，不承擔責任。

這就像團體報告中，即使使用了AI工具幫助寫作，最終每位組員仍須為自己負責的部分承擔責任。

六、透明度和可解釋性

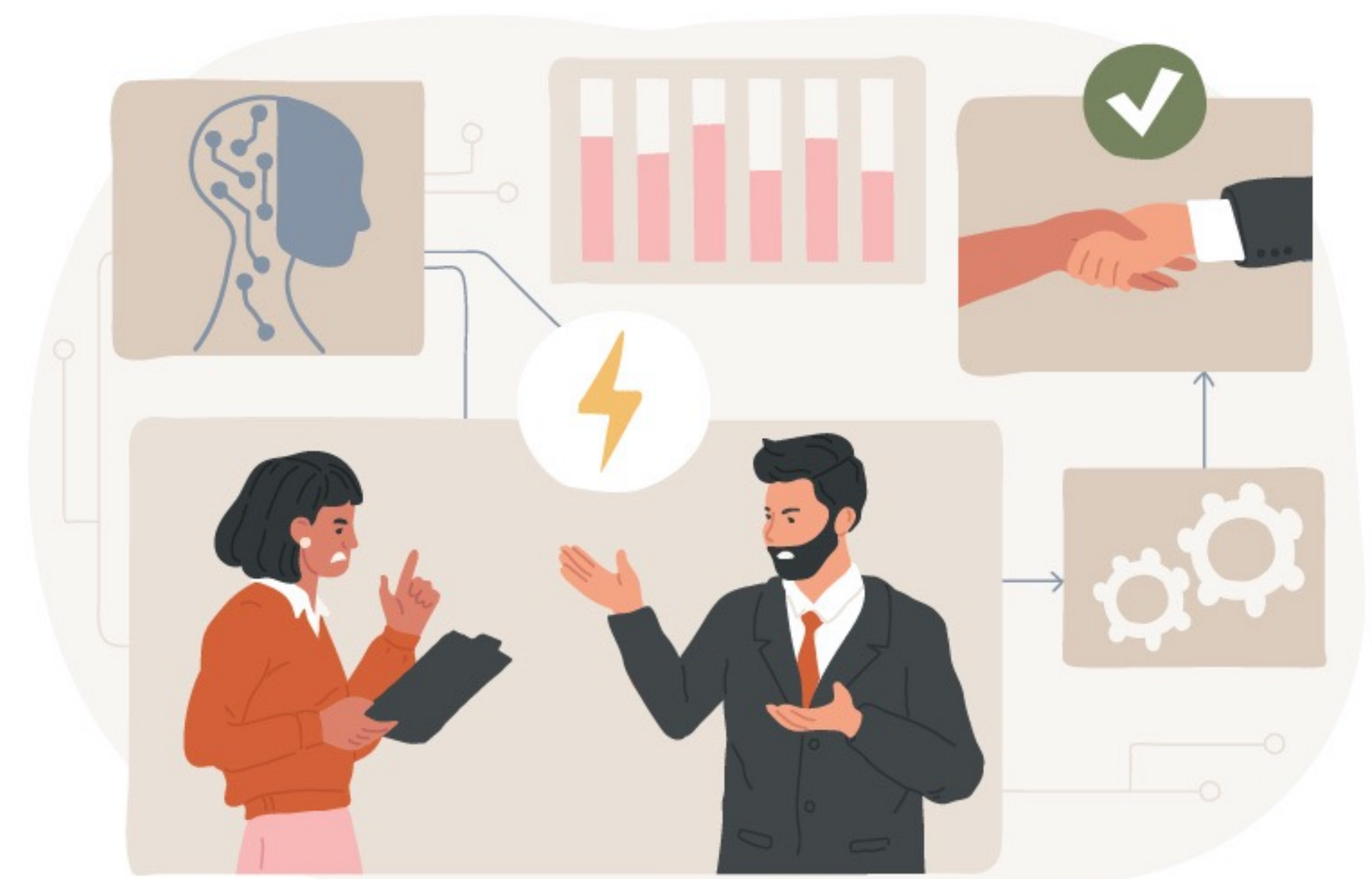
AI系統的道德部署取決於其透明度和可解釋性。透明度和可解釋性的等級應適合特定情況，因為透明度和可解釋性與隱私、安全和保障等其他原則之間可能存在矛盾。

例如：使用一款AI應用程式來幫助選擇大學科系。

程式應清楚標示「AI推薦」的內容；解釋為何推薦特定科系（例如：根據你對科學的興趣和數學的優異表現）；提供相關數據支持其建議（如就業率、滿意度調查）。

但若程式未說明哪些建議來自AI；無法解釋為何不推薦某些科系；使用複雜術語讓使用者無法理解決策過程。

這類似於老師給成績時，不僅告訴學生得了幾分，還會說明為何得這個分數，以及如何改進。



七、人類監督和決定權

各國應確保AI系統不會取代人類的最終責任和義務。

例如：醫院使用AI來協助診斷X光片。

醫院應聲明AI只提供初步分析和建議，最終診斷仍由醫生做出；醫生可質疑或推翻AI的建議；病人知道有AI參與但也知道由人類醫生負責。

但若是醫院完全依賴AI診斷而減少醫生參與；醫生被鼓勵不要質疑AI的判斷；出現問題時無人負責。

這就像使用計算機幫忙做數學題——計算機可以算出答案，但使用者需要理解問題、決定使用什麼方法，並檢查結果是否合理。

八、永續性

AI系統的評估應著重於其人工智慧技術對永續性的影響，永續性被理解為一系列不斷發展的目標，包括聯合國永續發展目標(SDGs)所規定的目標。

例如：學校使用AI系統管理校園能源使用。

好的AI系統不只優化短期的電費節省，還考慮長期環境影響；使用節能伺服器；評估整個系統的碳足跡（包括製造和維護）；協助學校達成減碳目標。

不良的設計會讓AI系統一味追求成本節省，反而採用高耗能設備處理數據，導致學校總體能源消耗增加，甚至忽略電子廢棄物問題。

這就像選購電器時——不只看即時效能，還要考慮長期的能源消耗和環境影響。

九、意識和素養

政府和政策制定者，應透過開放和無障礙教育、公民參與、數位技能和AI道德培訓、媒體和資訊素養來促進公眾對AI和數據的理解。

例如：社群媒體平台使用AI來推薦內容和過濾資訊。

好的AI平台會清楚說明哪些內容是AI推薦的；提供教學資源幫助用戶理解推薦系統如何運作；教導用戶如何調整設定以控制看到的內容。

不好的AI系統用戶不知道自己看到的內容是如何被選擇的；平台不提供任何使用指南；用戶被鎖在「資訊泡泡」中而不自知。

這就像圖書館不只提供書籍，還應該教導學生如何查找資料、評估資訊來源的可靠性，以培養資訊素養。

十、公平和非歧視

AI參與者應促進社會正義、公平和非歧視，同時採取包容性方式確保所有人都能享受到AI的好處。

例如：銀行使用AI系統評估高中生的小額貸款申請（如學生信用卡）。

好的AI系統使用多元且平衡的數據進行訓練；考慮不同背景學生的特殊情況；定期檢查是否對特定群體（如少數民族或低收入家庭）產生不公平結果；提供申訴管道。

設計不良的AI系統僅基於住址或家庭背景做出判斷；對某些群體設定更嚴格的標準；拒絕解釋拒貸原因；且不允許人工審核特殊情況。

這就像班上分組活動，應該給每個人平等的機會，不因個人背景、性別或外表而有所偏見，且當有特殊情況時應該靈活處理。

上述例子說明了AI倫理原則如何具體應用在我們的日常生活中，了解這些原則能幫助我們成為更有責任感的科技使用者，並在未來參與相關討論和決策時，有更明確的判斷標準。



2-1-5 AI倫理在生活中應用的提醒

一、數據政策

想像使用者下載了一款學習輔助APP，它聲稱能幫助使用者追蹤學習進度並給予個人化建議。

- 適當的提醒：APP清楚告知它收集哪些資料（例如學習時間、錯題紀錄），如何使用這些資料，並讓使用者選擇是否分享特定類型的資料。它也讓使用者隨時可以下載或刪除自己的資料。
- 不當的應用：APP在不告知的情況下收集位置資料、聯絡人清單，或將學習資料賣給補習班用於行銷。

這就像借同學的筆記本時，我們會事先說明用途、保證完整歸還，而不是未經許可就把內容分享給其他人。

二、倫理治理與管理

學校考慮使用AI系統來協助分配課程或編排座位：

- 適當的提醒：學校成立委員會討論這個系統，包括校長、老師、學生代表、家長和技術專家，共同決定系統的設計和使用規則。學校定期發布系統如何做決定的報告，並設立管道讓同學可以質疑或申訴不合理的安排。
- 不當的應用：校方單方面決定採用系統，不解釋它如何運作，也不接受任何反饋或質疑。

這就像班規的制定，應該由全班共同討論而非班長獨斷，且每個人都有權了解規則的理由和提出修改建議。

三、經濟和勞動

住家附近的超市開始使用自助結帳機：

- 適當的提醒：超市提供員工再培訓計畫，讓原本的收銀員學習管理這些機器或承擔客戶服務等新職責。同時，學校開始教導學生如何與這類技術互動及理解其背後運作原理。
- 不當的應用：超市直接解僱所有收銀員，不提供任何轉職協助，造成許多人失業。教育系統也未能及時更新課程內容，導致畢業生技能與就業市場需求不符。

這相當於科技進步時，我們不只要學會使用新工具，還要思考如何讓每個人都能適應這種變化，不讓任何人被時代淘汰。

四、教育和研究

學校開設了AI相關課程：

- 適當的提醒：課程不只教授程式設計，還包括AI倫理、數據隱私和批判性思考。學校特別鼓勵女生和不同背景的學生參與，並為資源較少的學生提供額外支援（如免費的平板電腦或課後指導）。
- 不當的應用：課程只開放給資優班學生，內容只關注技術而忽略倫理問題，且因為課程設計不考慮不同學生的需求，導致參與者多為特定群體。

這就像圖書館不只應該有書，還要確保每個人都知道如何使用並能平等地獲取知識。

五、健康和社會福祉

我們使用手機上的健康APP來追蹤運動和睡眠：

- 適當的提醒：APP根據年齡和身體條件提供合理的健康建議，明確說明它不能替代醫生診斷，並保護我們的健康數據不被未授權存取。政府會審核這類APP的安全性和準確性。
- 不當的應用：APP未經科學驗證就提供可能有害的健康建議，或誇大其功效聲稱可以診斷疾病，同時將用戶健康數據出售給保險公司。

這類似於用藥須知，即使是普通感冒藥也需要清楚標示成分、用法和可能的副作用，以確保使用者安全。



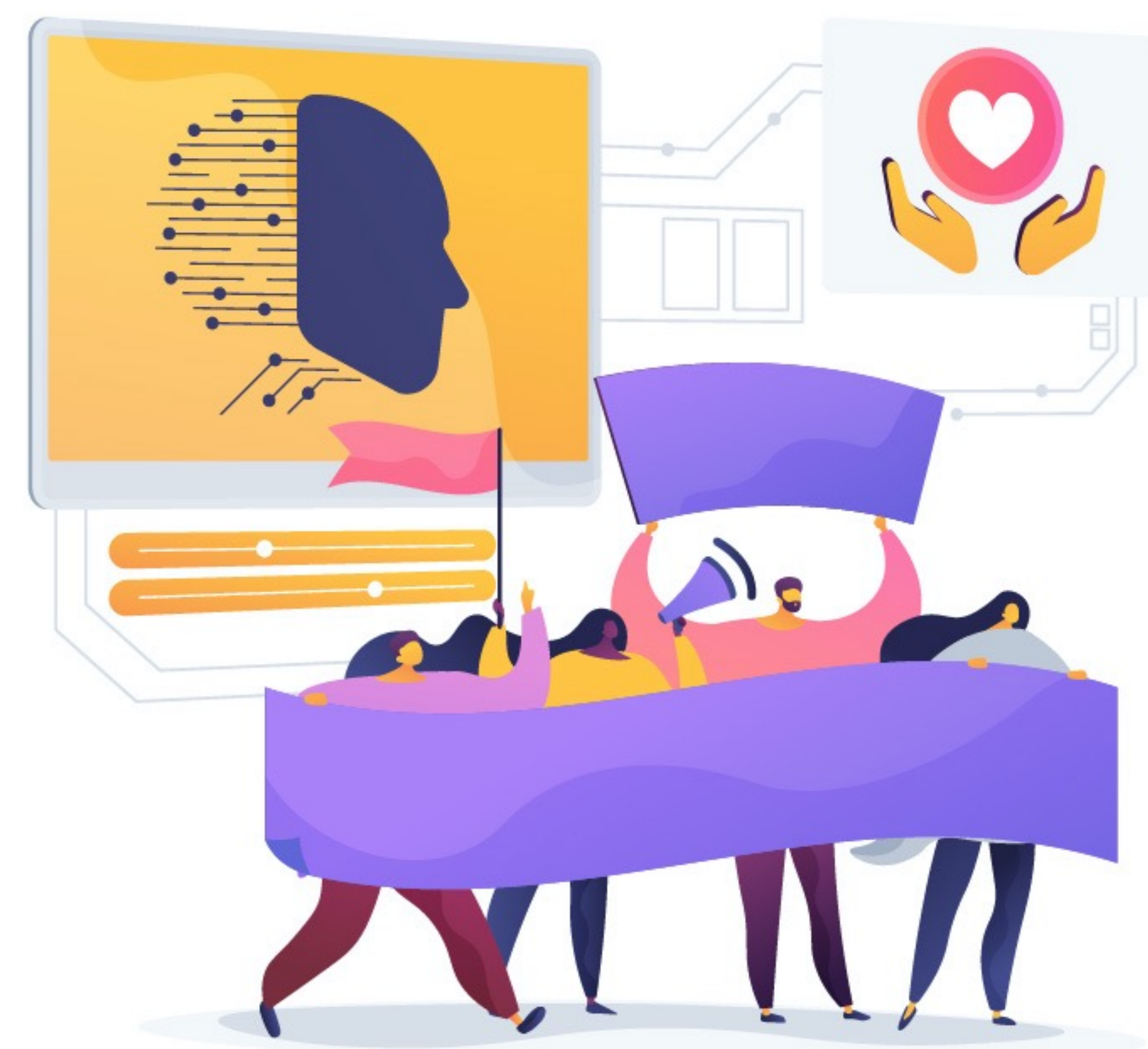
六、性別平等

學校使用AI程式來推薦職業發展方向：

- 適當的提醒：系統向所有學生展示多元的職業選擇，不因性別而限制推薦。例如，理工科會同等地推薦給女生，護理或教育工作也會推薦給男生。學校特別設立女生科技社團，鼓勵更多女生接觸AI和編程。

- 不當的應用：系統根據過去的性別就業模式做出建議，結果只向男生推薦工程師，向女生推薦秘書工作，進一步強化了性別刻板印象。

這就像運動會上，每個項目應該歡迎所有人參加，而不是預設某些運動只適合特定性別的同學。



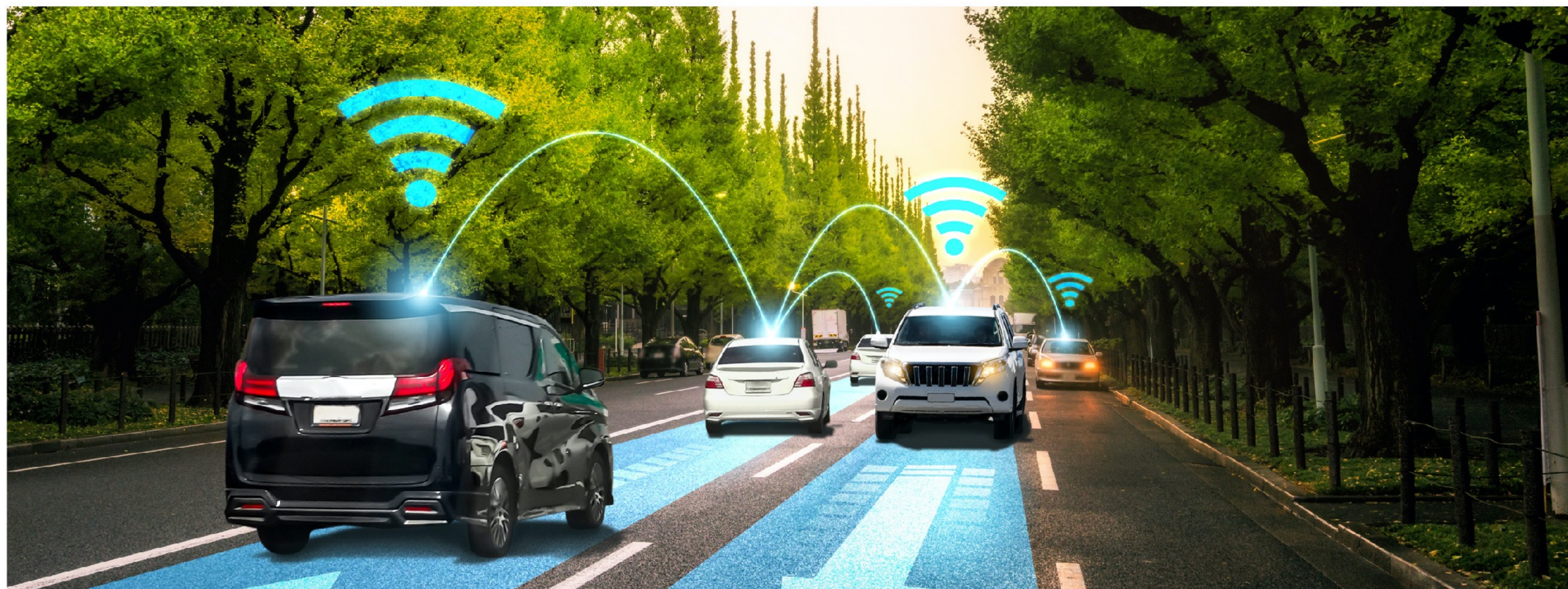
七、環境和生態系統

我們居住的城市想要建立智慧交通系統：

- 適當的提醒：系統不僅優化交通流量減少塞車，還考慮到能源消耗，使用節能設備和可再生能源供電。系統的開發商評估了所有設備的生命週期環境影響，並制定了回收計畫。
- 不當的應用：系統只關注效率，使用高耗能伺服器處理數據，忽視設備製造和廢棄對環境的影響，導致電力消耗大增和電子垃圾增加。

這就像買新手機時，不只考慮功能和價格，也要思考它的製造過程是否環保、是否耐用、將來如何回收等問題。

透過這些例子，我們可以看到AI政策不只是抽象的規則，而是與我們日常生活息息相關的具體指導原則，幫助我們在享受科技便利的同時，確保它的發展方向符合人類的整體利益。



2-1-6 使用生成式AI輔助學習注意事項

隨著生成式AI技術的快速發展，在學習與日常生活中愈來愈依賴這些工具。然而，這些技術的使用也帶來了倫理與風險問題，應具備適當的AI素養，理解相關風險，並制定適當的使用規範。以下說明使用生成式AI輔助學習時應注意的倫理問題：

一、正確性：AI生成內容絕對正確嗎？我們如何判斷？

生成式AI的內容是基於大量資料的學習與統計推理產生，並不保證完全正確。使用時應該：

- 交叉比對資訊：透過查詢可信來源來確認AI生成內容的正確性。
- 保持懷疑精神：不盲目相信AI提供的答案，而是主動檢驗其可信度。
- 理解AI的侷限性：AI無法理解語境，也無法辨識複雜的人類價值觀與文化差異。

二、可溯源性：生成式AI的資料來源是否包含多元觀點？

AI學習的資料來自網路上的大量文本，但並非所有資料都具備代表性，可能會偏向特定文化或觀點。使用時應該：

- 檢查來源：瞭解AI生成內容是否基於特定立場或文化背景。
- 接觸多元觀點：主動尋找來自不同國家、族群或學術領域的資訊，避免單一觀點影響判斷。
- 理解AI的訓練偏差：認識AI可能存在的偏見，並在使用時進行批判性思考。

三、事實性：如何辨識所見所聞的真假？

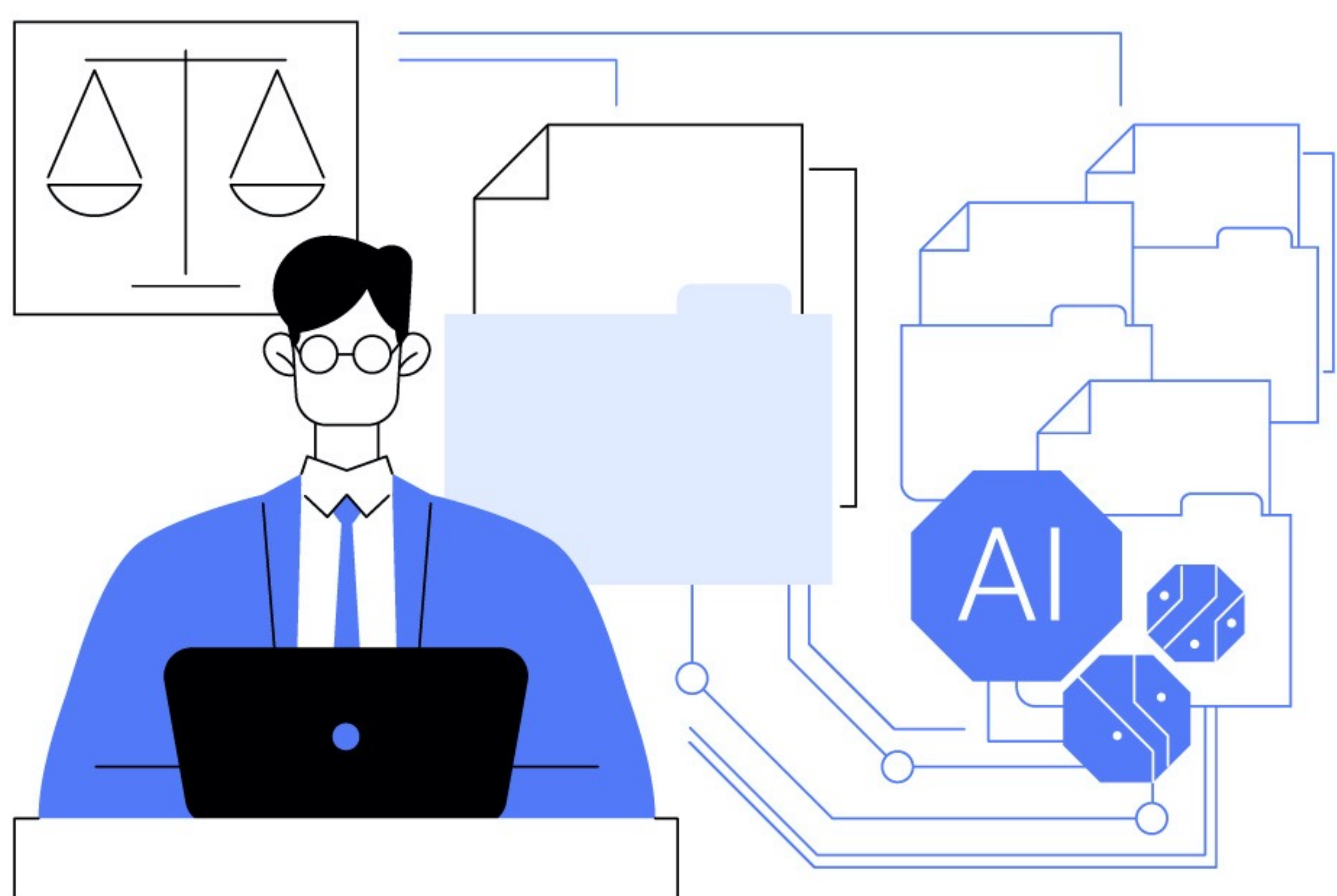
AI可以生成極具說服力的內容，但這些內容不一定真實，甚至可能包含錯誤或偽造資訊。使用時應該：

- 學習數位媒體素養：掌握辨識假訊息的方法，例如查證來源、分析語氣、確認發佈日期等。
- 輔以事實查核工具：利用如Google Fact Check、Snopes等查證平台來驗證資訊。
- 避免資訊擴散：在未確認資訊真偽之前，不輕易分享或傳播。

四、生成性：如何因應生成式AI帶來的機密資料管理與隱私保護問題？

生成式AI可能記錄使用者的輸入內容，進而造成隱私風險。使用時應該：

- 避免輸入個人敏感資訊：例如身份證統一編號、地址、學校資訊等。
- 審慎評估：避免在公開平台或AI工具中分享機密資料。
- 學習數位安全意識：瞭解如何設定隱私權限，並使用安全的數位工具。



五、思辨性：如何提升自我資訊辨識及思辨能力？

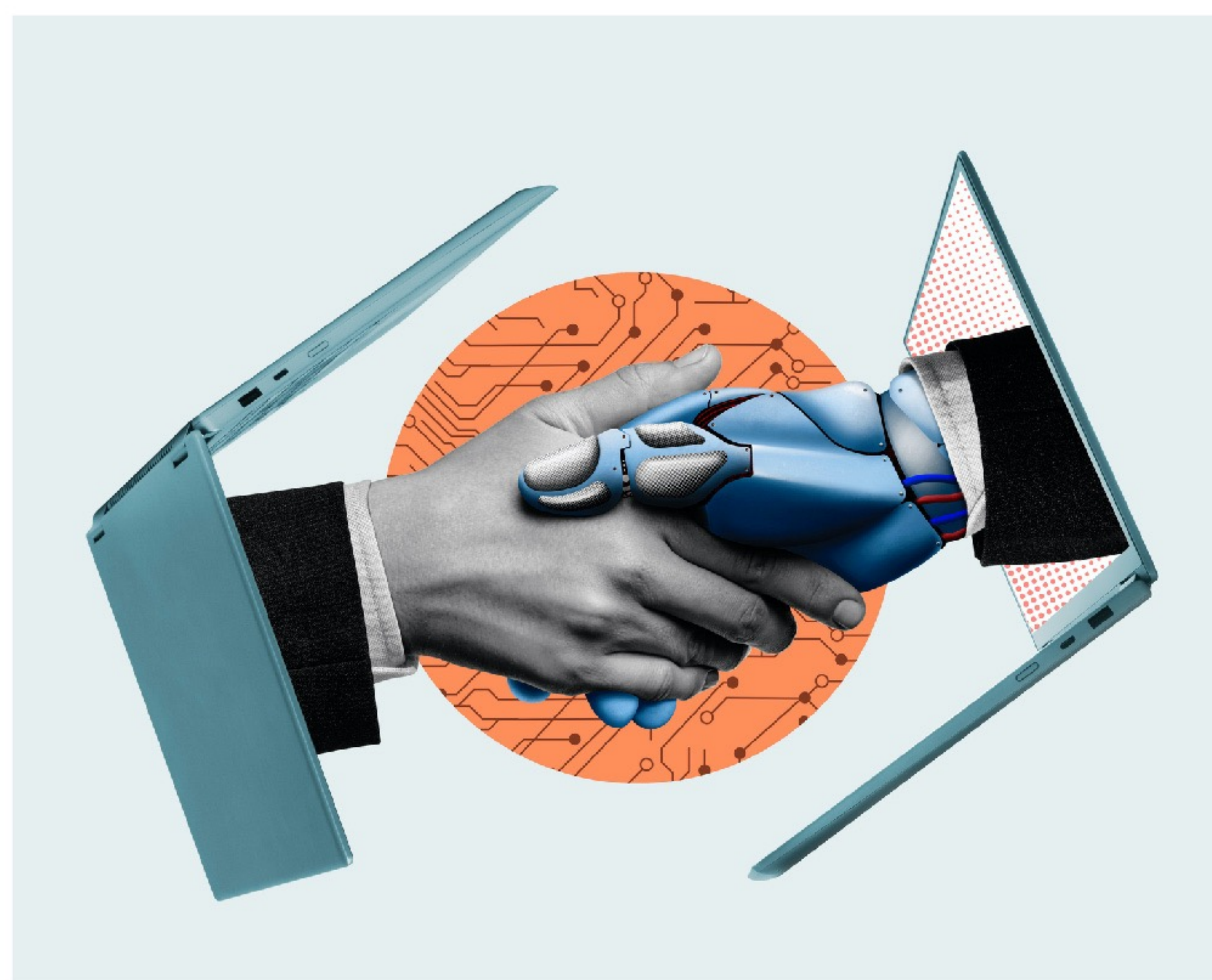
AI的普及使得資訊過載，應發展批判性思考能力，以正確解讀與應用AI生成內容。使用時應該：

- 培養問題意識：對於AI提供的資訊多問「為什麼」和「如何得出這個結論」。
- 訓練邏輯思維：學習基本的邏輯推理與資訊分析技巧，提升判斷力。
- 參與討論與反思：透過與同儕、教師討論來深化對AI生成內容的理解。

六、風險性：當我們使用數位工具和生成式AI時，如何避免風險？

生成式AI雖然帶來便利，但也可能導致學術不誠實、個資外洩或錯誤資訊傳播等風險。使用時應該：

- 遵守學術誠信：避免直接抄襲AI生成內容，學習如何正確引用與應用。
- 提升網路安全意識：注意AI工具的隱私政策，確保個資不被濫用。
- 制定AI使用規範：師生共同討論並建立合理的AI使用準則，確保負責任地使用技術。



★小結：生成式AI的發展為學習與創作帶來新的機會，但也伴隨許多挑戰與倫理問題。學生應積極培養資訊素養、提升思辨能力，並與師生共同制定適當的使用規範，以確保在享受技術便利的同時，也能減少潛在風險，成為負責任的數位公民。

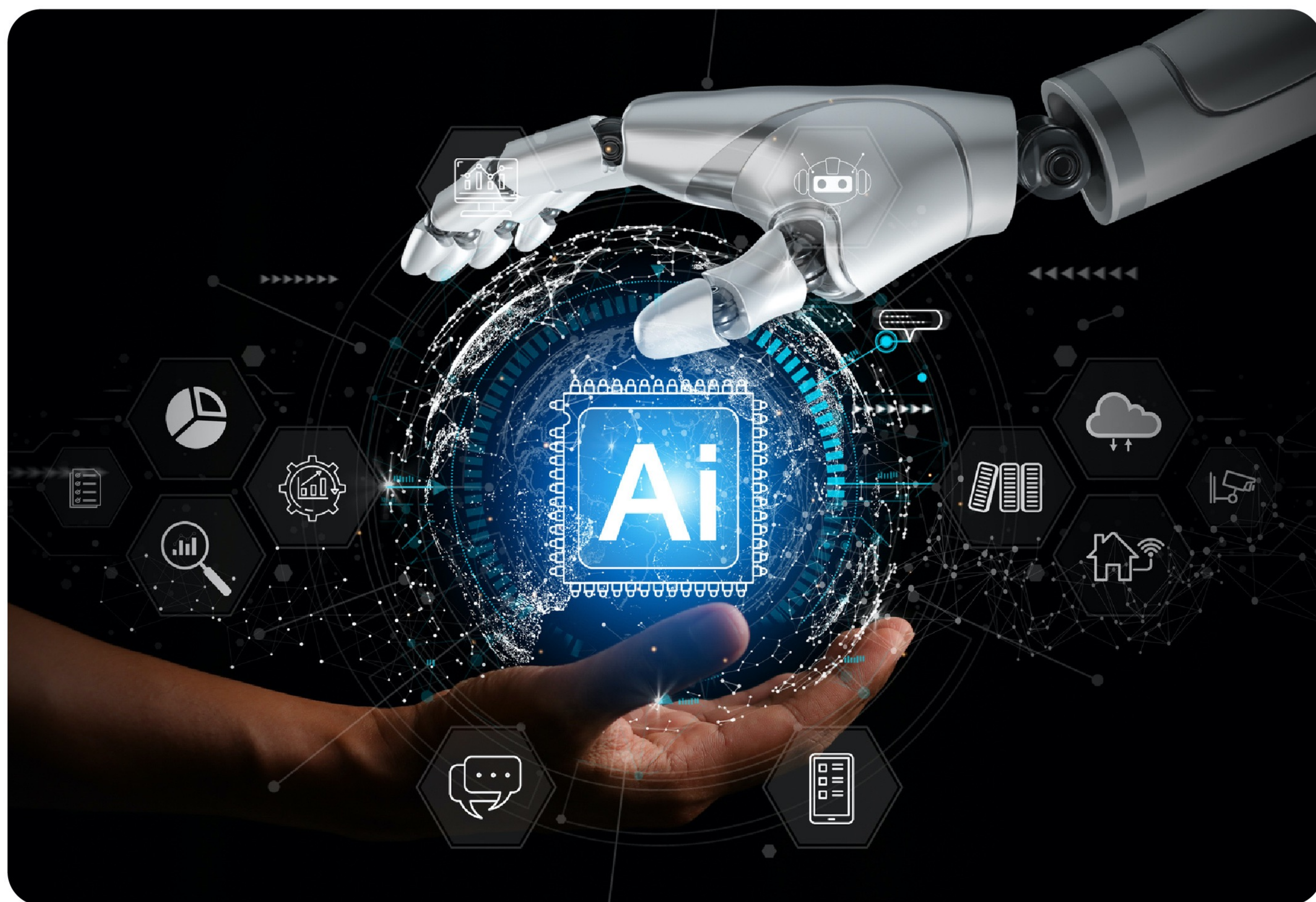
03 ▶ 三絕韋編 鑑往知來

3-1 中文參考文獻

- 教育部中小學數位教學指引3.0版. (2024)

3-2 英文參考文獻

- Mucci, T. (2024, October 11). The Future of AI: Trends Shaping the next 10 Years. IBM.
<https://www.ibm.com/think/insights/artificial-intelligence-future>
- Unesco. (2023). Artificial Intelligence. Artificial Intelligence.
<https://www.unesco.org/en/artificial-intelligence>
- Unesco. (2023). UNESCO' s Recommendation on the Ethics of Artificial Intelligence: Key Facts. UNESDOC Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000385082>



04 ▶ 四通八達 小試身手

主題探討

Q1

AI與我們的隱私：誰在看著我們？

- 1-1 我們每天使用的APP（如社群媒體、遊戲）如何收集我們的個人資料？
- 1-2 這些資料被用來做什麼？我們有權利知道嗎？
- 1-3 如果學校或商家使用人臉辨識系統，這會侵犯我們的隱私嗎？
- 1-4 我們如何在享受AI便利的同時，保護自己的個資不被濫用？
- 1-5 我們可以如何保護在網路上的隱私？



Q2

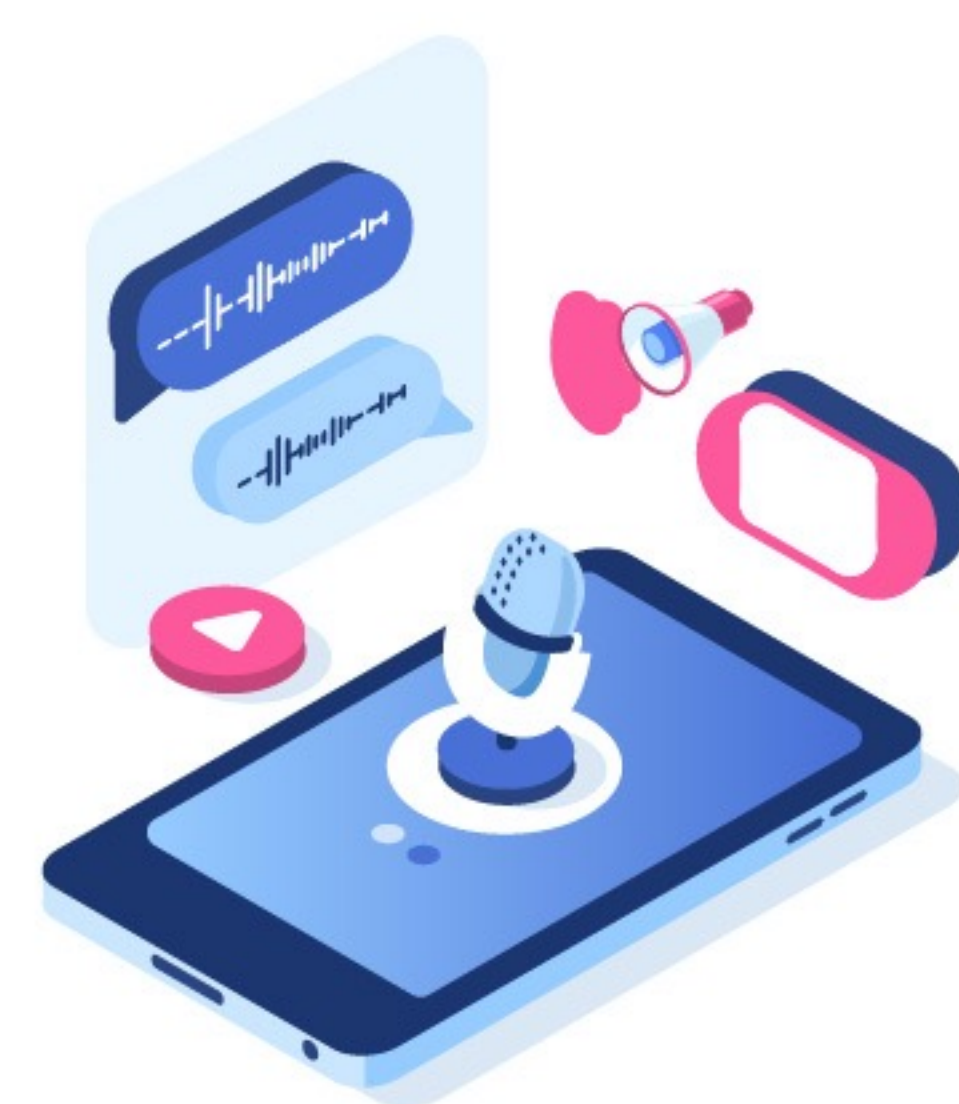
演算法的祕密：為什麼我看到這些內容？

- 2-1 社群媒體、YouTube或Netflix如何利用AI推薦內容給我們？
- 2-2 這些演算法是否會讓我們只看到自己喜歡的內容，形成「同溫層」？
- 2-3 如果演算法推薦假新聞或偏激內容，我們該怎麼辦？
- 2-4 我們有權利知道這些演算法如何運作嗎？
- 2-5 我們可以如何辨別在網路上看到的資訊是否正確？

Q3

AI的性別偏見：機器也會有偏見嗎？

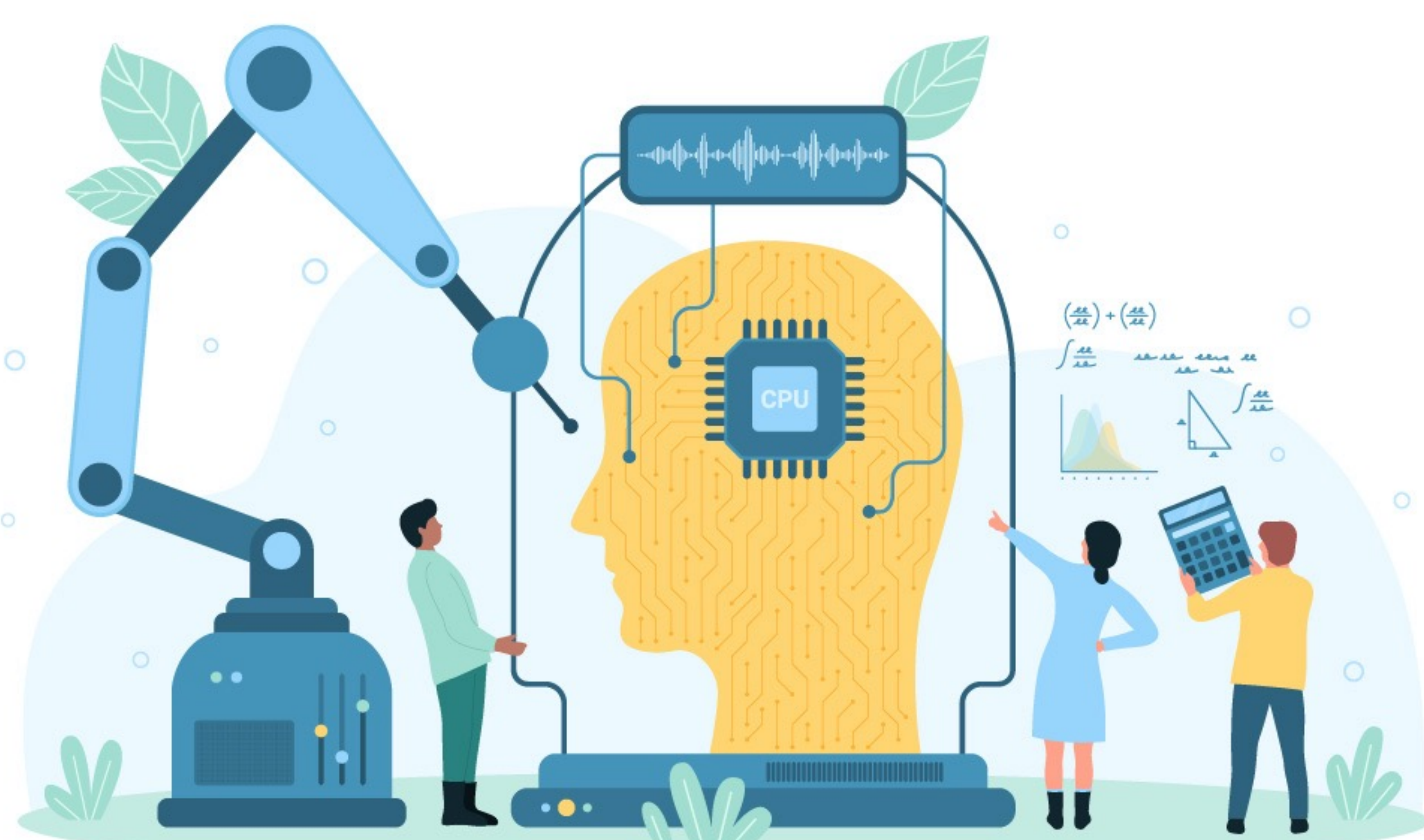
- 3-1 為什麼有些AI語音助理（如Siri）的聲音預設是女性？
- 3-2 AI在求職、貸款等方面的決策，是否會對女性不公平？
- 3-3 我們如何檢測和修正AI系統中的性別偏見？
- 3-4 我們在設計AI產品的時候，應該注意什麼？





AI與地球的未來：科技發展與環境保護的衝突？

- 4-1 訓練大型AI模型需要消耗大量電力，這會對環境造成什麼影響？
- 4-2 AI在智慧城市、節能減碳方面有哪些應用？
- 4-3 我們如何利用AI來解決氣候變遷、資源短缺等環境問題？
- 4-4 我們在使用科技產品的時候，可以怎麼做來減少對環境的影響？



AI在校園：學習的好幫手還是監控者？

- 5-1 AI如何幫助我們學習？例如，個人化學習APP、智能輔導系統。
- 5-2 學校使用AI來教導學生行為（如考試作弊、出缺席），這是否合理？
- 5-3 AI如何幫助老師批改作業、設計課程？
- 5-4 AI在教育上的使用，有哪些需要注意的倫理問題？



資訊素養與倫理 高中職 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

林湧順 臺北市立大直高級中學校長

指導委員 / 盧東華 臺北市立大學助理教授

李啟龍 國立台灣師範大學附屬高級中學資訊室主任

總 編 輯 / 劉亦陞 臺北市立大直高級中學圖書館主任

執行編輯 / 魏仲良 臺北市立大直高級中學資訊組組長

編審委員 / 王鼎中 臺北市立建國高級中學老師

朱德清 臺北市立中山女子高級中學老師

侯莉莉 臺北市立內湖高級中學資訊組組長

孫晉忻 臺北市立景美女子高級中學老師

陳佳宜 臺北市立麗山高級中學老師

陳哲雋 臺北市立南港高級中學資訊組組長

陳 晉 臺北市立復興高級中學老師

陳瑞宜 臺北市立大同高級中學研發處主任

楊喻文 臺北市立第一女子高級中學資訊組組長

蔡志敏 臺北市立大同高級中學老師

蔡明男 臺北市立內湖高級中學圖書館主任

賴和隆 臺北市立中正高級中學老師

韓君尹 臺北市立和平高級中學資訊組組長

羅玕貞 臺北市立內湖高級中學老師

(按姓氏筆畫排列)

承辦單位 / 臺北市立大直高級中學

出版日期 / 中華民國114年3月



未來新世界

解鎖未來 與倫理共舞



臺北市政府教育局

DEPARTMENT OF EDUCATION
TAIPEI CITY GOVERNMENT