

AI的道德與社會責任的規範

GROUP 1



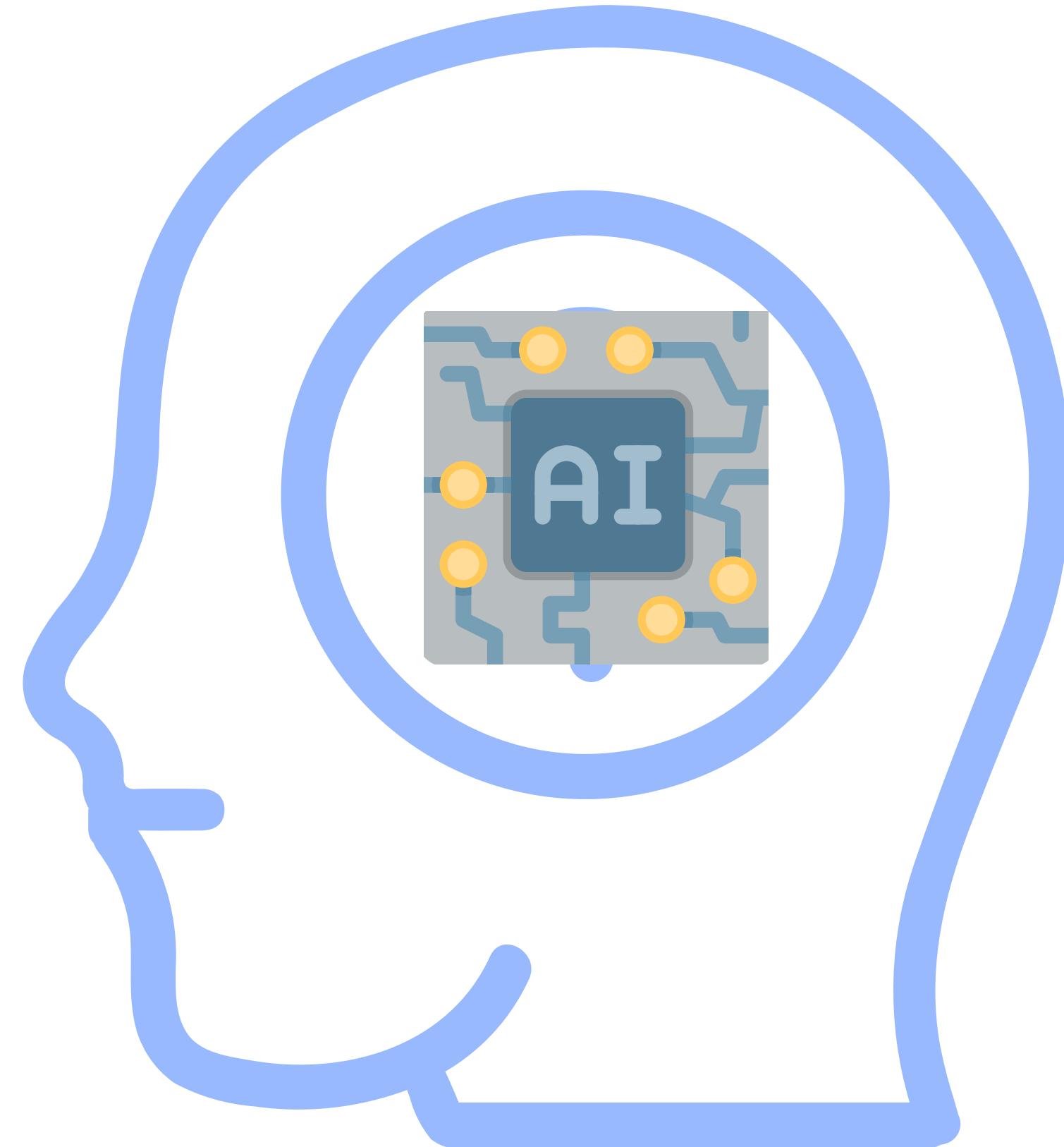
目錄

- 1 目的和動機 引言
- 2 定義基本概念
- 3 現行狀況與趨勢
- 4 步驟與規範
- 5 案例分析
- 6 主題探討

1. 目的和動機

目前世界上的 AI 都是這種只具備特定功能的「弱 AI」 (artificial narrow intelligence, ANI)
我們熟知的各種AI就是其中

但也正是如此，讓人們這幾年在使用 AI 時，逐漸發現它可能造成的問題或傷害



AI的快速發展和其對社會的影響

經濟影響

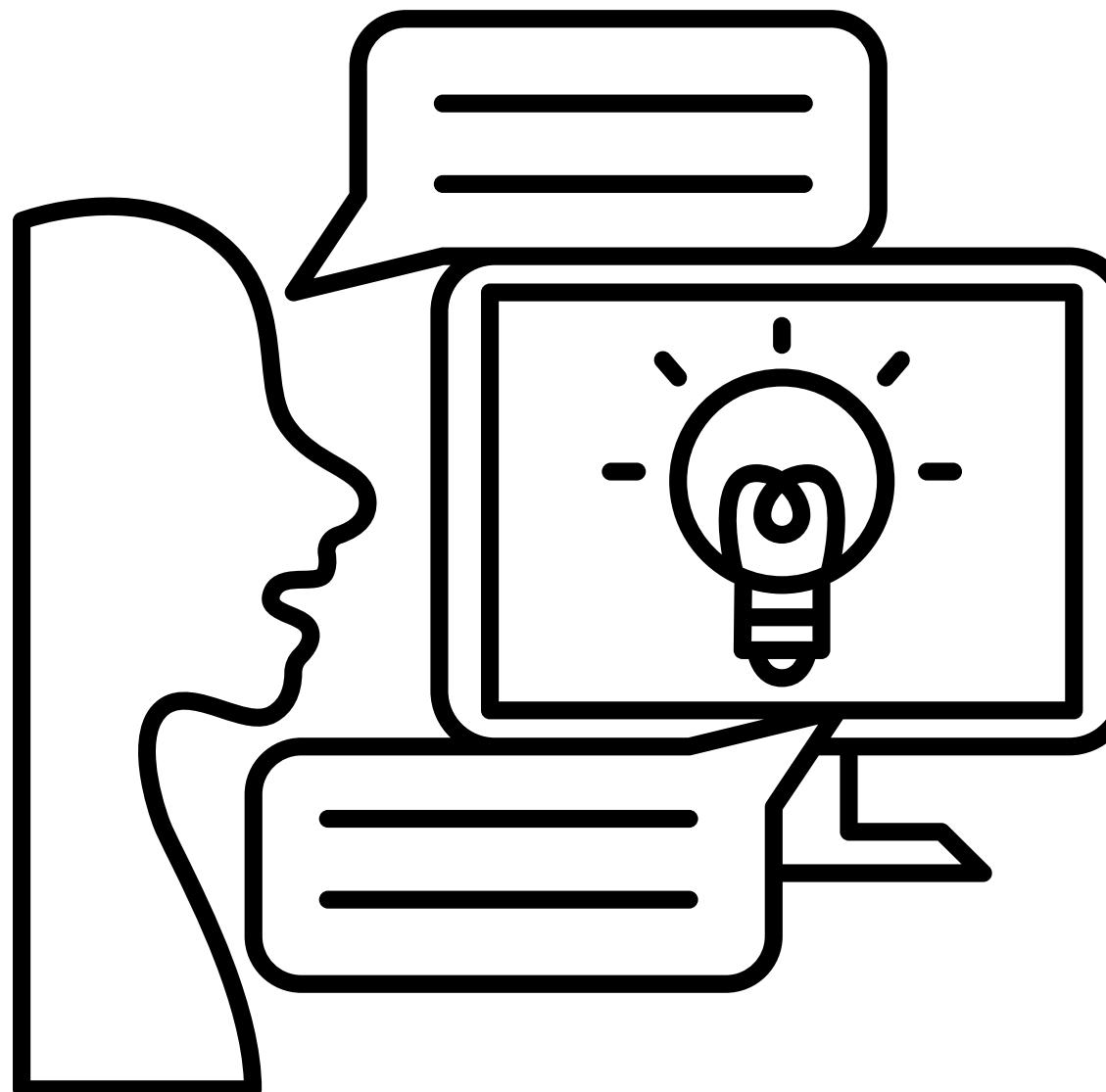
社會影響

倫理和法律問題

教育和培訓

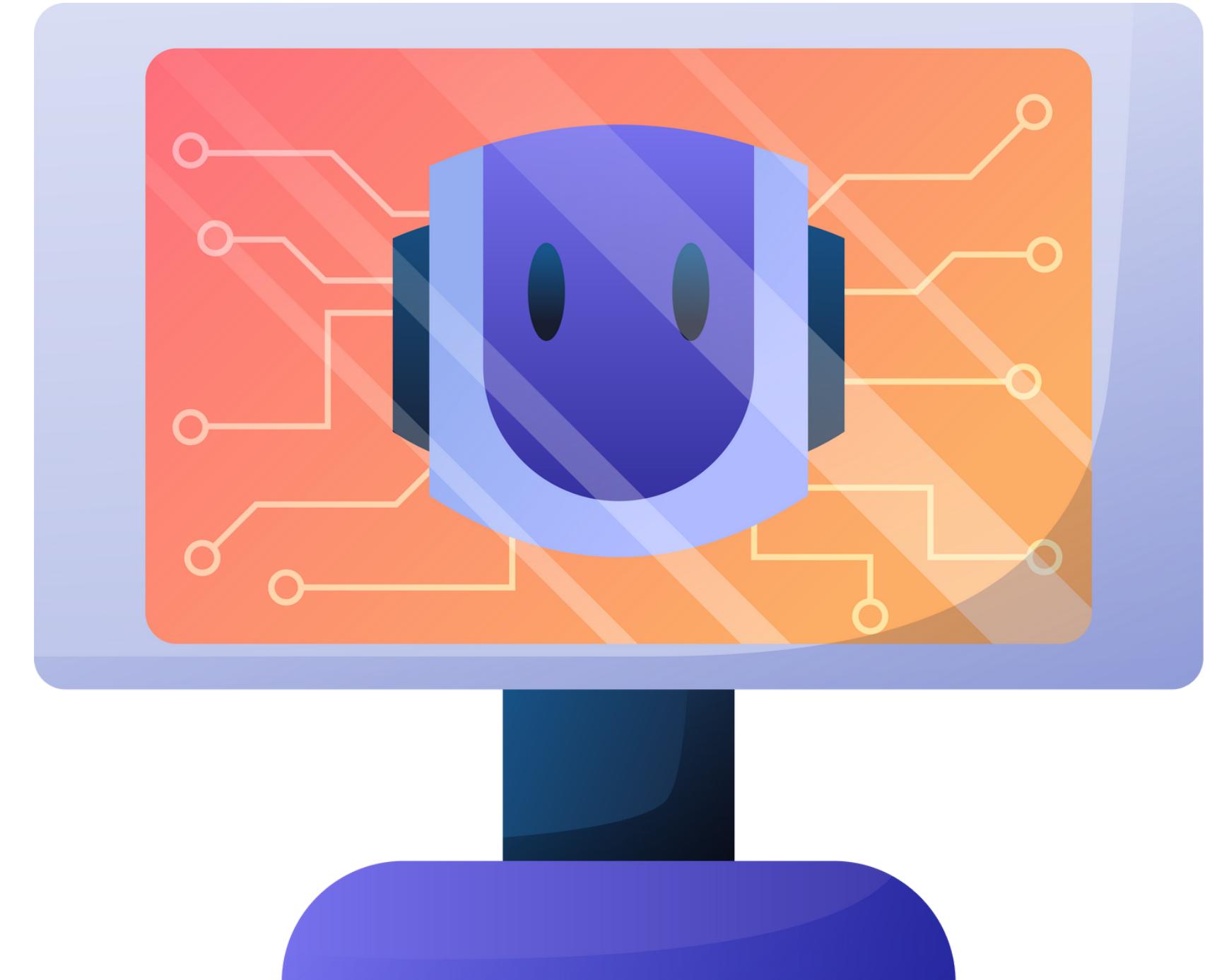
政治與治理

全球合作和競爭



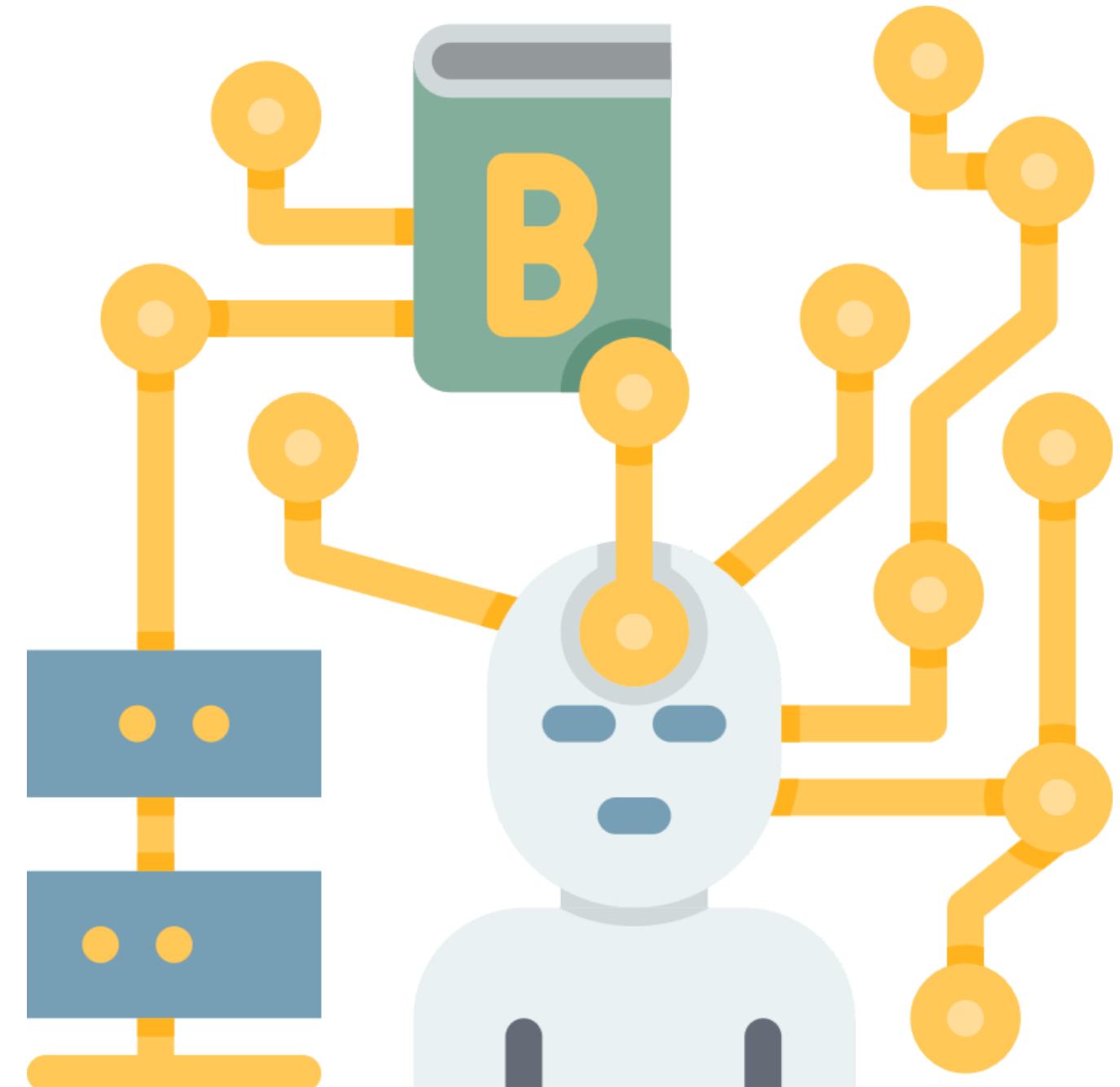
2. AI定義基本概念

人工智能 (AI) 是電腦科學的一個領域，致力於解決與人類智慧相關的常見認知問題，例如學習、創造和影像辨識。



3. 當前的使用狀況

- 自然語言處理 (NLP)
- 自動駕駛技術
- 醫療保健
- 金融領域



• 自然語言處理 (NLP)

NLP

正面影響

1. 改進溝通
2. 自動化客戶服務
3. 信息擷取與分析
4. 多語言溝通

潛在風險

1. 語言偏見與不公平
2. 隱私問題
3. 安全性問題

文
A

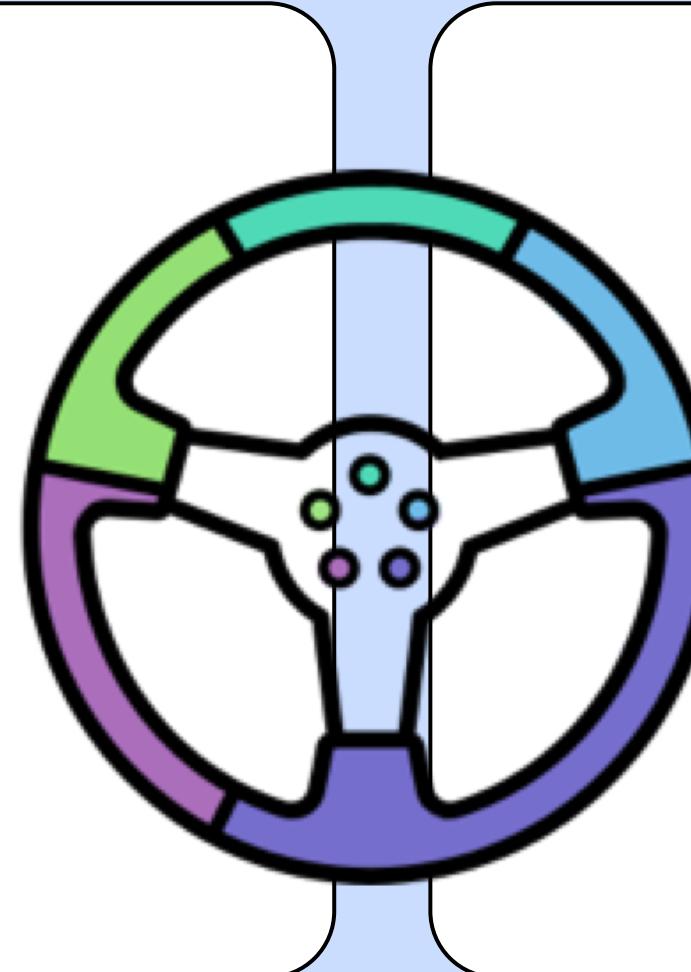
• 自動駕駛技術

正面影響

1. 改進交通流量和擁堵
2. 能源效率
3. 城市規劃

潛在風險

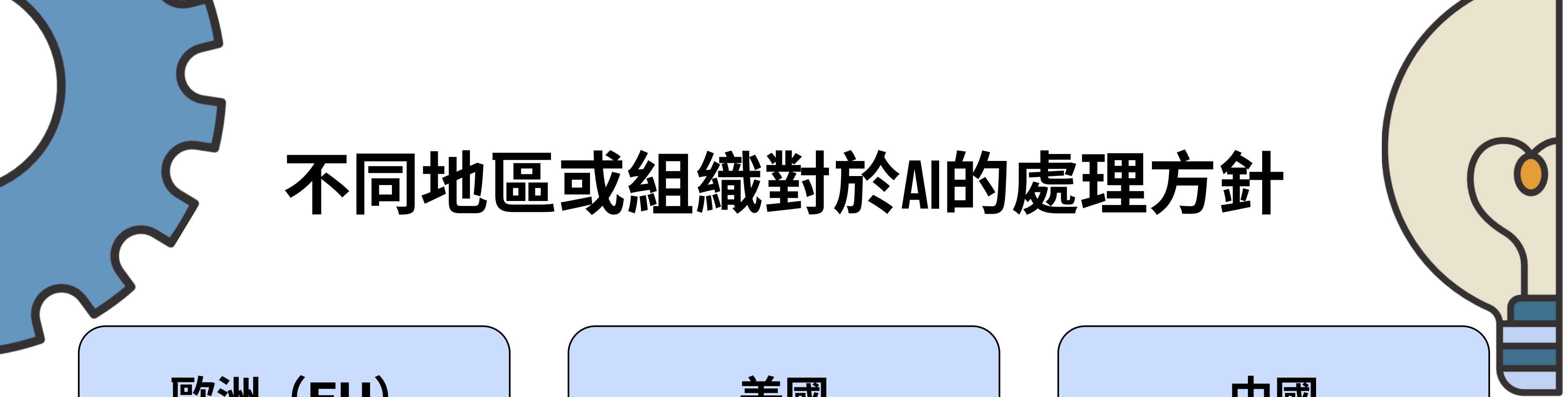
1. 技術可靠性
2. 軟體安全性
3. 肇事責任歸屬
4. 技術僵化



4. 制定對於道德及法律的規範



- 多方參與和跨學科合作
- 透明和公開的流程
- 倫理原則的確立
- 風險評估和影響評估
- 解釋性
- 公平性和反歧視



不同地區或組織對於AI的處理方針

歐洲 (EU)

- GDPR(歐盟通用數據保護規例):對個人數據處理的法規
- AI法規框架:確應用符合倫理原則

美國

- 分散性監管:不同的機構和州進行監管
- 倡導創新:減少規範，以促進技術發展

中國

- AI發展計劃:《新一代人工智能發展規劃》等
- 數據監管:數據的收集和使用相對開放



成功案例

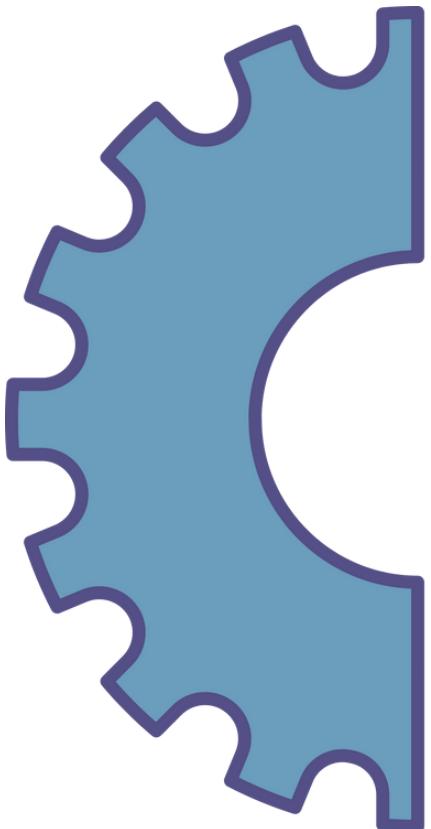
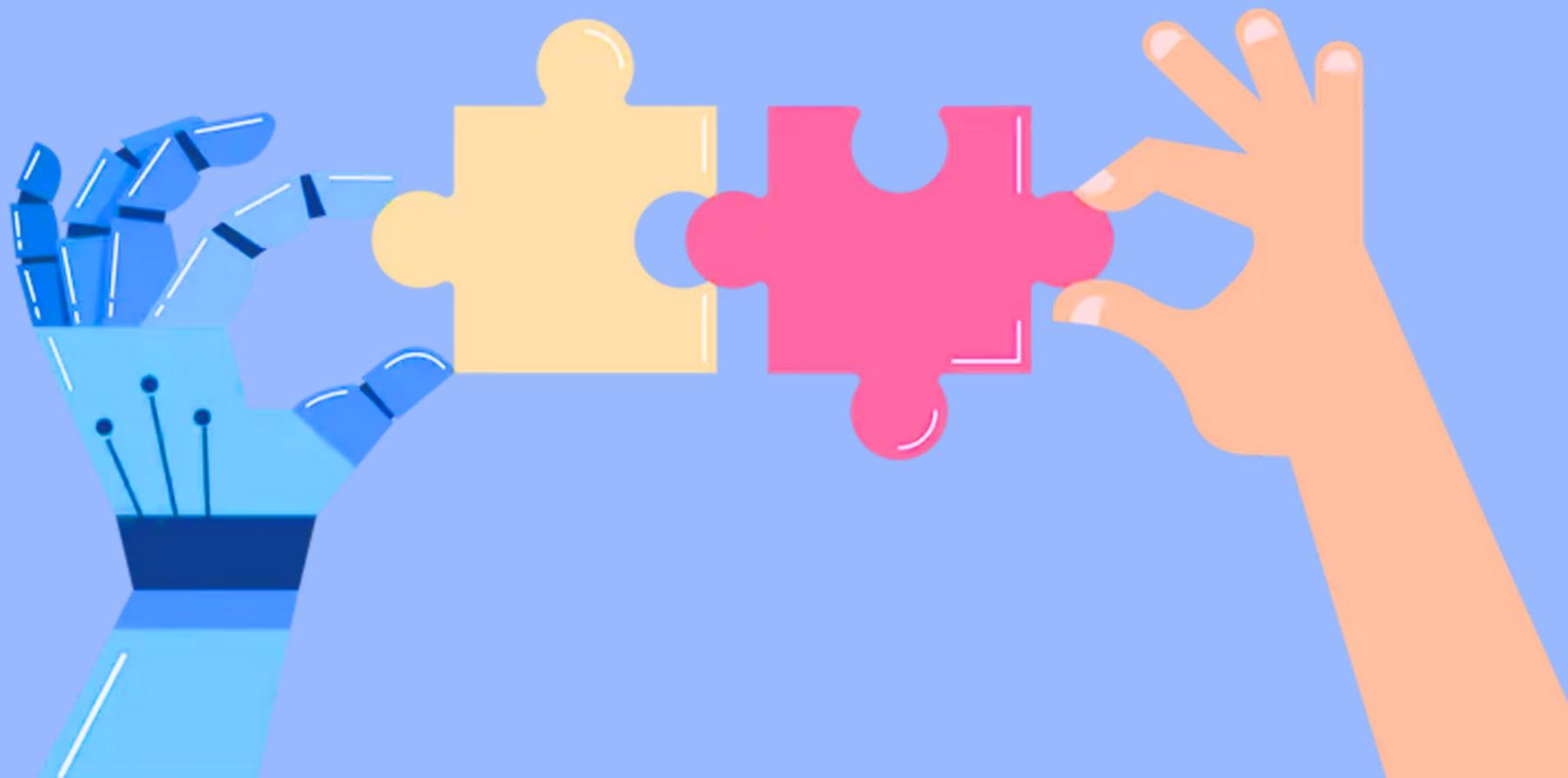
1. OpenAI 的 GPT 模型
2. AI 醫療診斷
3. 自動駕駛技術

失敗案例

1. 臉部辨識技術的濫用
2. 性別歧視的 AI 演算法
3. 社交媒體的信息過濾

6. 探討的五項主題

- a. 透明度與解釋性**
- b. 公平與非歧視**
- c. 隱私與資料安全**
- d. 社會影響與就業**
- e. 國際合作與標準**



A. 透明度與解釋性

- 告知與解釋權利
- 通知與透明度
- 深入分析與保障措施
- 對於嚴重威脅的揭露要求



B. 公平與非歧視

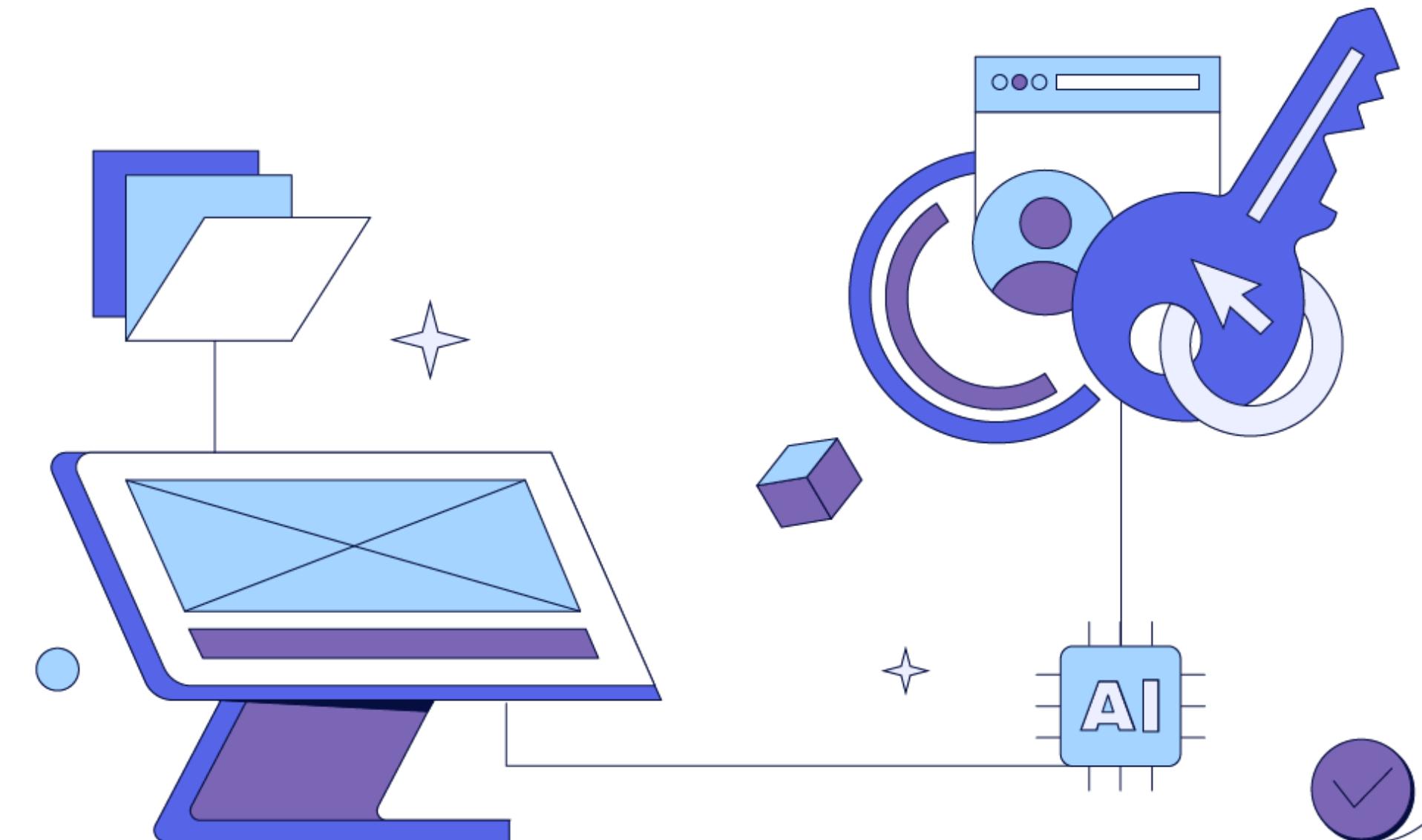
- 社會正義與公平
- 避免歧視性或偏見
- 救濟措施
- 考慮不同需求
- 國際法遵循



EQUITY

C. AI與個人隱私

探討AI對個人隱私的影響，包括數據收集、使用，以及應對措施。強調道德規範的必要性，在科技進步與個人隱私的保護之間取得平衡



AI如何處理個人資料？

網頁抓取與網頁 爬蟲

從第三方網站、維基百科、數位圖書館等公開互聯網來源中提取資料

用戶查詢存儲在 AI模型中

AI模型儲存用戶輸入的問題或其他數據，有些工具不僅收集而且保存數據以供未來訓練使用

生物特徵技術與監 視設備

利用監視設備，包括安全攝像機、臉部和指紋掃描器，以及能檢測人聲的麥克風，收集生物特徵數據

物聯網 (IoT) 傳感器和設備

IoT傳感器和邊緣運算系統收集大量數據，AI軟體透過學習、數據吸收、安全的IoT協議和API等利用這些數據

SECURITY



COMPLIANCE



PRIVACY



確保資料安全的應對措施

設計AI算法以最
小化個人數據的
收集和處理

監控和測試AI系
統，識別和減輕
可能存在的偏見
和歧視

透明和負責任地開
發和部署AI技術，
確保個人數據的透
明和道德使用

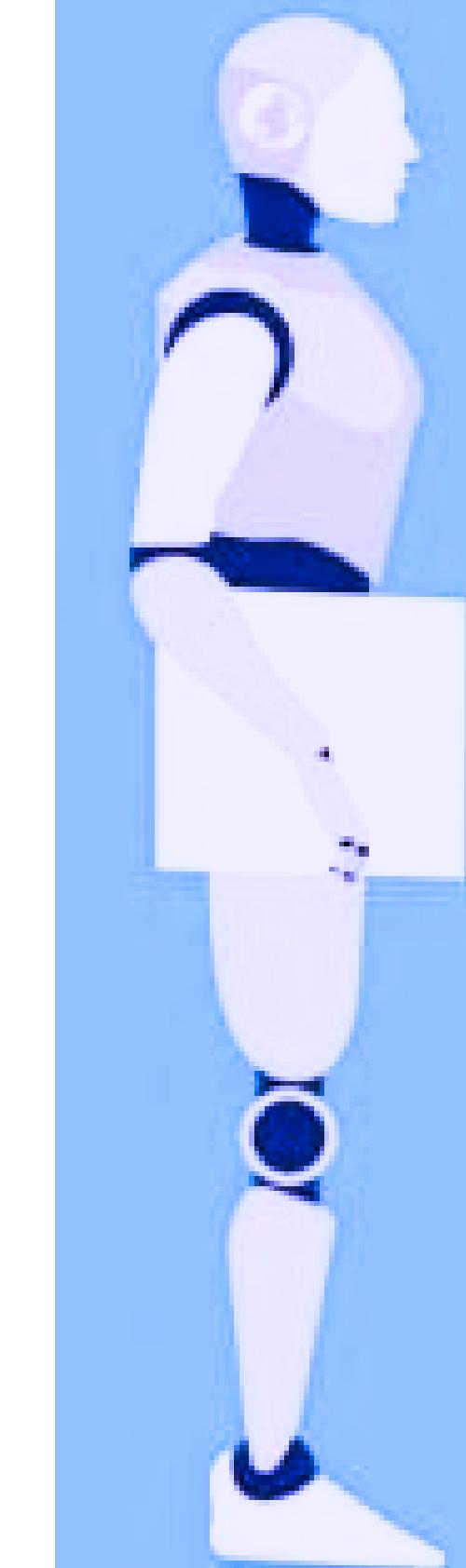
發展機制，讓個
人能夠控制其數
據的收集和使
用，防止濫用AI
技術



PURPOSE

D.AI 對社會與 就業影響

在AI時代的浪潮中，見證了技術的巨大飛躍，對企業生產力和就業結構產生深遠影響。欲深入剖析了AI技術的興起，探討對台灣電子產業的影響。然而隨著挑戰與機會共存，正面臨著塑造未來職場和社會結構的關鍵時刻。



可能的挑戰

失業風險

- 隨著人工智慧 (AI) 技術的進步，對於大多數工人失業的擔憂日益增加。這主要源於AI技術的應用可能取代一些例行工作，特別是可以被自動化執行的任務

技能轉型需求

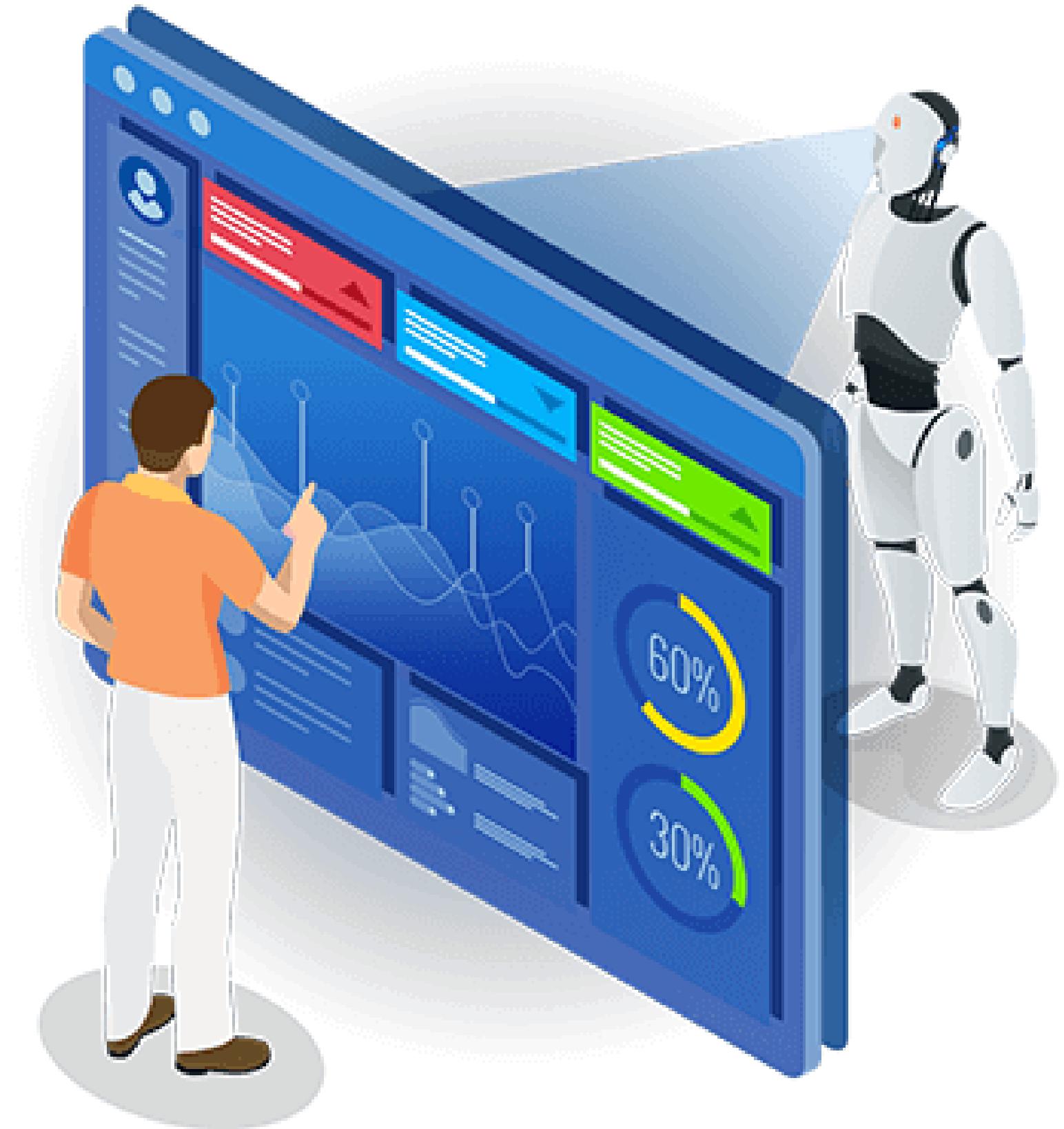
- 工作者可能需要學習新的技能，以適應新的工作環境。需要具備AI技術相關的知識和技能，以有效地與機器協同工作。

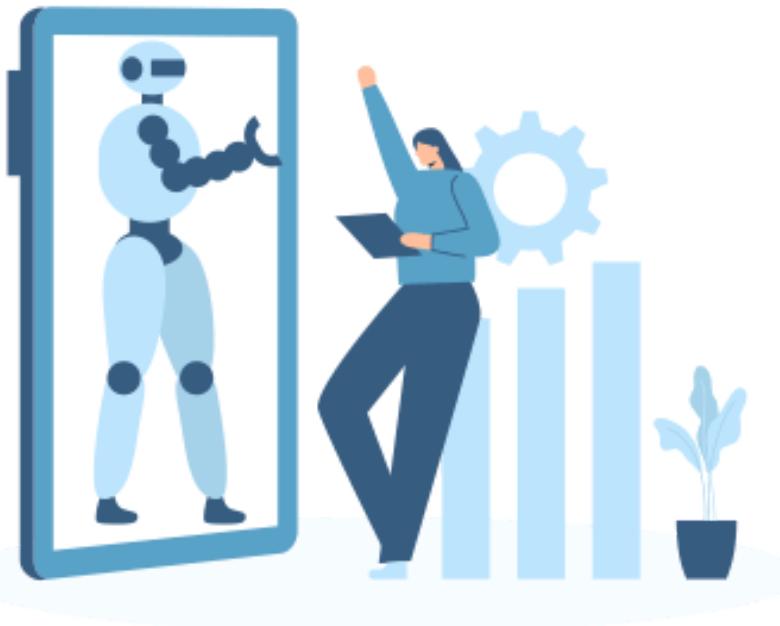
就業結構變化

- 低技能和高技能領域增加就業機會，而在中等技能領域可能面臨挑戰。可能導致勞動市場的不平等現象。

E.AI與國際合作和相關標準

許多國家紛紛制定AI政策，全球AI標準制定工作在國際組織中取得重要進展。國際合作有助於最大化AI發展規模優勢、構建負責任AI並解決全球挑戰。以實現更開放、全球化的AI市場。





AI 和國際之間之變化

全球AI政策趨勢

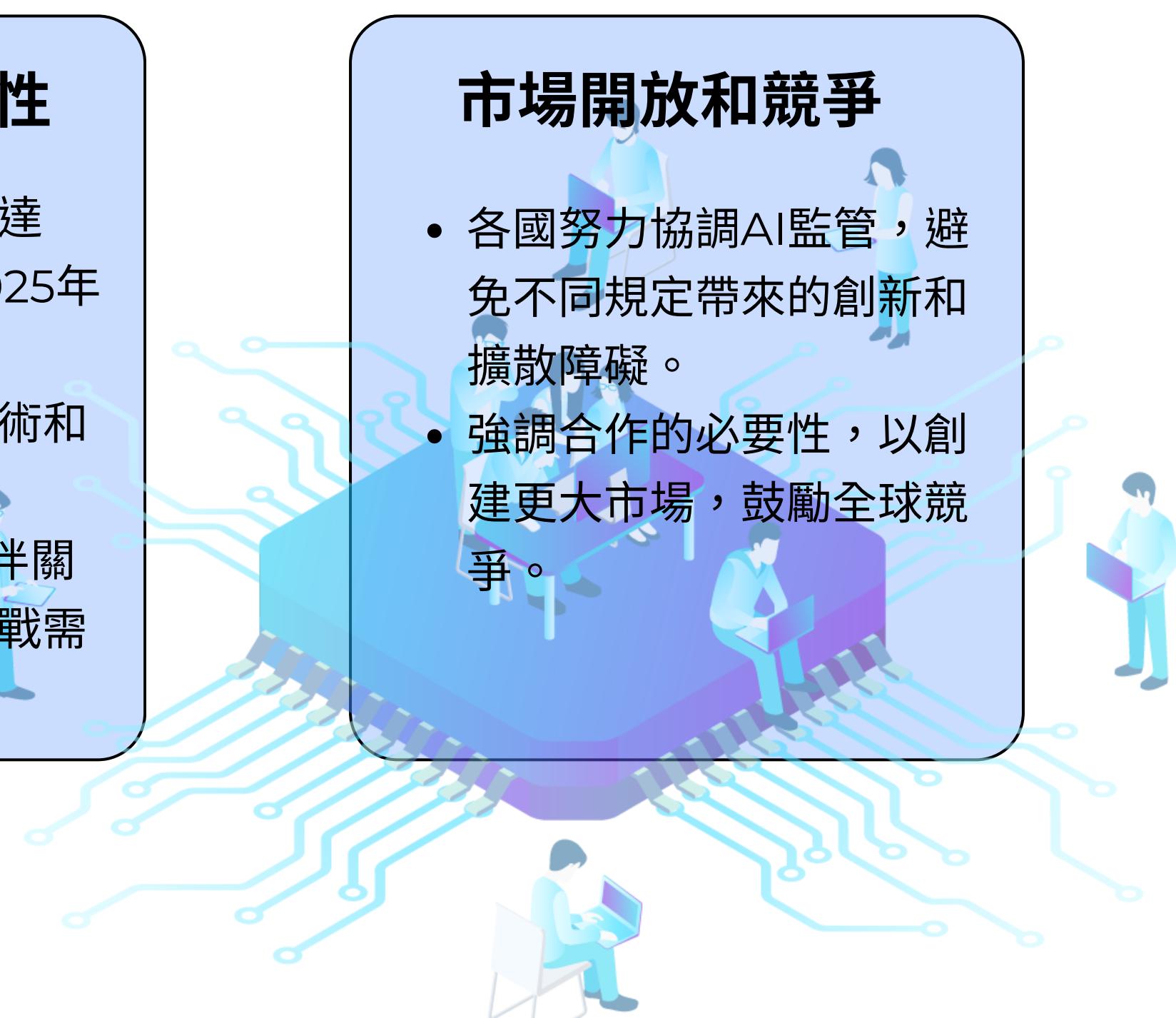
- 加拿大制定全國AI戰略，啟動全球AI政策浪潮。
- 多國制定政策，在工業、消費市場和政府服務中推動AI應用。
- 預估到2030年，全球產出將增長16%

國際合作的必要性

- 全球企業對AI的投資達600億美元，預計2025年將翻一以上。
- 國際標組織推動AI技術和倫理標準。
- G-7建立全球合作夥伴關係，強調解決全球挑戰需要合作。

市場開放和競爭

- 各國努力協調AI監管，避免不同規定帶來的創新和擴散障礙。
- 強調合作的必要性，以創建更大市場，鼓勵全球競爭。





**THANK YOU FOR
LISTENING!**