



SWANKY

史旺基工作室



NCC 媒體素養公民培力計畫

媒體人的 **AIGC** 工具箱： 提高生產力和創意的技術

<https://swanky.github.io/>

主講：史旺基 (Swanky Hsiao)

大綱

- 前言
- 生成式 AI 的基礎與背景
- 虛擬人物與 AI：新時代的名人與代言人
- 負責任的 AI
- 文字生成：ChatGPT
- 圖像生成：Midjourney / Stable Diffusion



蕭宇程 (史旺基)

<https://swanky.github.io/>

swanky.hsiao@gmail.com

國立臺灣師範大學資訊工程博士

「史旺基加密貨幣交易實戰班」講師

《制服·女孩×史旺基》系列與
《高校制服戀物論》攝影作者

朱學恒「阿宅反抗軍」及
「制服地圖」特約攝影師

曾任職 資策會數位所、前瞻中研院
ITM 國際信任機器股份有限公司 技術研發總監

現為 資策會數轉院 正工程師
東吳大學資料科學系 兼任助理教授





前言



初音未來，暱稱「殿下」、「公主殿下」，是克理普敦未來媒體（Crypton Future Media）以 Yamaha 的 VOCALOID 2 語音合成引擎為基礎，開發販售的**虛擬女性歌手**，角色主唱系列的第一作、VSTi 規格的電子樂器；或此軟體的印象角色（為軟體的象徵，不會在實際使用時出現）。2007年8月31日發售，發售後大受歡迎，並出現大量使用者製作的翻唱歌曲、原創曲等。

初音未來除了吸引新一代和從不認識電子音樂的人接觸電子音樂創作，令更多人認識電子音樂製作，亦令約20年前創作熱潮時的老手「回歸」。電子音樂創作再度捲起熱潮，也刺激到音樂創作以外的繪畫、動畫創作。雖然初音未來並非第一個可以模仿人類歌唱的軟體，但像真度比以往同類軟體高，因而引起的熱潮帶來業餘音樂製作的革命，促進日本的消費者自組媒體發展。

初音未來的角色設定年齡為16歲、生日8月31日，身高與體重則分別是158cm與42kg，有著蒼綠色雙馬尾，腳穿黑色過膝靴，擅長流行歌曲，搖滾樂和舞蹈。其代表色為蒼綠色（#39C5BB）。



專有名詞

- 人工智慧 (**Artificial Intelligence, AI**)
- 通用人工智慧 (**Artificial General Intelligence, AGI**)
- 生成式 AI (**Generative AI**)
- AI 生成內容 (**AI Generated Content, AIGC**)
- 大型語言模型 (**Large Language Model, LLM**)



人工智慧（Artificial Intelligence, AI）

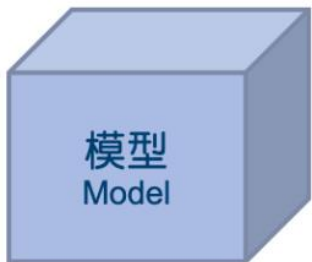
亦稱機器智慧，指讓機器展現類似人類的智能行為，包括學習、推理、規劃、創造和解決問題等能力。

人工智慧的四個主要組成部分：

- 機器學習：使電腦能夠從數據中學習和改進
- 自然語言處理：使電腦能理解和生成人類語言
- 專家系統：模仿人類專家的決策能力
- 機器感知：使電腦能夠感知世界，如視覺和聽覺

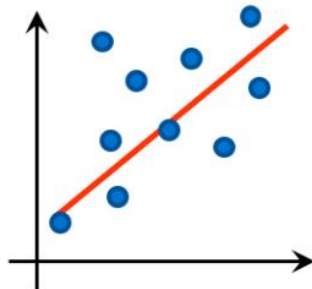


輸入
Input

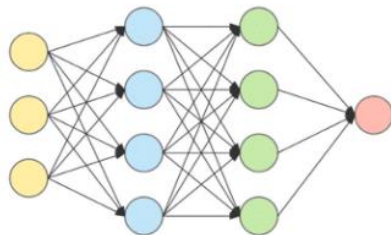


輸出
Output

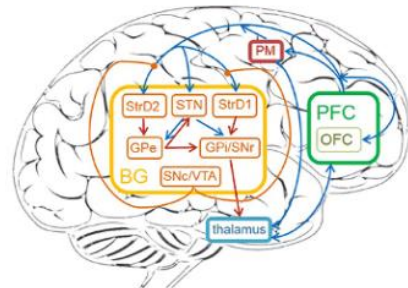
統計模型



人工神經網路



大腦神經迴路

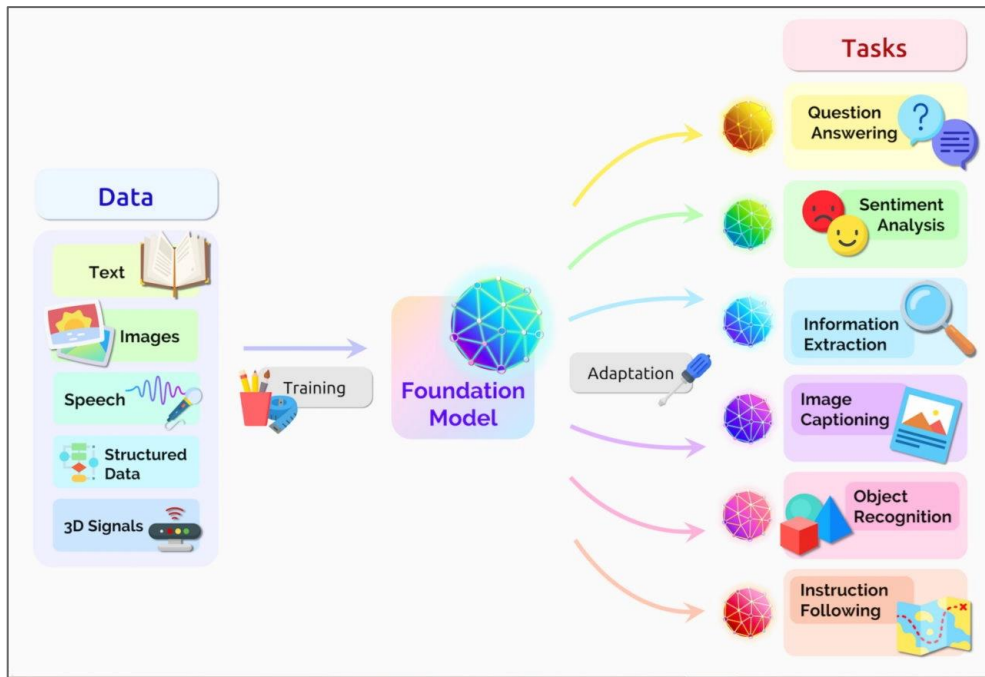


複雜程度



生成式 AI (Generative AI)

生成式 AI 是 AI 技術的一種分支，透過讓機器學習模型研究歷史資料，能進一步創造全新內容，相關技術包含生成對抗網路 (General Adversarial Network, GAN) 與 Transformer 模型等



Transformer 模型，偶爾稱為基礎模型，是許多應用程式使用之許多資料的來源



ChatGPT

ChatGPT 是基於 Transformer 模型所訓練出的大型語言模型（Large Language Model, LLM），OpenAI 使用大量文本資料預先訓練模型，使模型能夠學習自然語言的結構和規律，從而生成高品質、流暢的自然語言文本。再加上微調後，可進一步提高 ChatGPT 生成文本的準確度和流暢度。

大語言模型 (large language model, LLM) 是一種語言模型，由具有許多參數（通常數十億個權重或更多）的人工神經網絡組成，使用自監督學習或半監督學習對大量未標記文本進行訓練。大型語言模型在 2018 年左右出現，並在各種任務中表現出色。

AIGC 應用



- **文案 (Copywriting)**：越來越多人需要個性化的網頁和電子郵件內容來推動銷售和行銷策略以及客戶支持，這是語言模型的完美應用。這些文案往往形式簡單，並且都有固定的模版，加上這些團隊的時間和成本壓力，大大推動對自動化和增強解決方案的需求。
- **垂直產業的寫作助手 (Vertical specific writing assistants)**：現在大多數寫作助手都是通用型的，未來為特定的終端市場建立更好的生成式應用程式有著巨大機會，例如從法律合約編寫到劇本編寫。這裡的產品差異化體現在針對特定工作流程的模型和 UX 交互作用的微調。
- **程式碼產生 (Code generation)**：目前的應用推動了開發人員的發展，使他們的工作效率大大提高。在安裝了 Copilot 的專案中，它產生了近 40% 的程式碼。但更大的機會可能是為 C 端消費者賦能程式設計開發能力，學習提示 (learning to prompt) 可能會成為最終的高階程式語言。
- **藝術生成 (Art generation)**：整個藝術史和流行文化的世界現在都被編碼進了這些大型模型中，這將允許任何人隨意探索在以前可能需要花人一輩子的時間才能掌握的主題和風格。

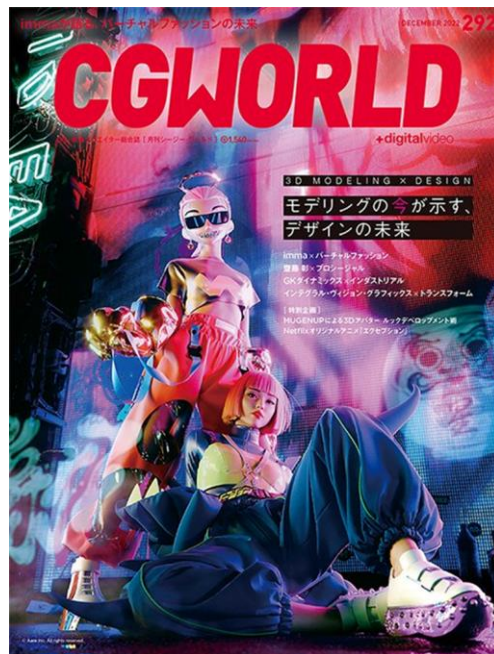
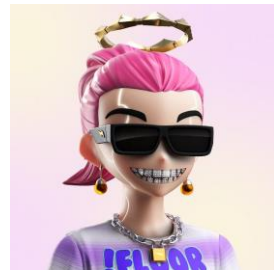


虛擬人物與 **AI**： 新時代的名人與代言人

imma

[Aww inc](#) 從 2018 年創建和製作的第一個虛擬人類 imma 開始，到其他幾個虛擬人：

從我們的原始 IP（如 imma）到與世界各地的其他公司合作創建虛擬人，包含了 Ayayi 等不同的虛擬人。我們希望專注創造虛擬人並且結合元宇宙的探索，而從公司的角度上則更多地在 IP 創造以及探索不同的可能性。例如，創建非虛擬人的 IPs。







負責任的 AI





負責任人工智慧（**Responsible AI**）

- 開發、部署和管理人工智慧系統時，遵循一套原則和規範聲明，以確保符合倫理和法律
- 確保 AI 技術的正面影響，減少潛在的負面效果
- 注重道德與倫理，保護使用者權益，確保技術公正與不歧視
- 提高 AI 的透明度與可解釋性，清楚展示 AI 系統的運作和決策過程，使非專業人士能理解 AI 的決策



文字生成：ChatGPT



2018 GPT-1 (1.17 億參數)

GPT-1 有一定的泛化能力，能用於和監督任務無關的 NLP 任務中

2020 GPT-3 (1750 億參數)

GPT-3 作為一個自監督模型，可以完成自然語言處理的絕大部分任務：將網頁描述轉換成對應程式語言、模仿人類敘事、創作詩歌、生成遊戲劇本，甚至模仿已故的哲學家預測生命的真諦

2022-11 ChatGPT (GPT-3.5)

ChatGPT 全稱聊天生成預訓練轉換器 (Chat Generative Pre-trained Transformer)，在 GPT-3.5 之上用基於人類回饋的監督學習和強化學習微調。這兩種方法都用人類教練來提高模型效能，以人類干預增強機器學習效果，獲得更逼真的結果

2019 GPT-2 (15 億參數)

除了理解能力外，GPT-2 的生成能力表現出強大的天賦：閱讀摘要、聊天、續寫、編故事，甚至可以生成假新聞、釣魚信件或進行角色扮演

2022-01 InstructGPT

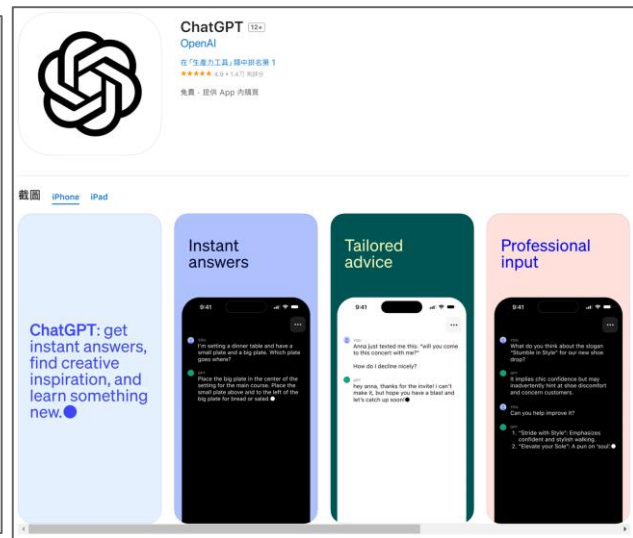
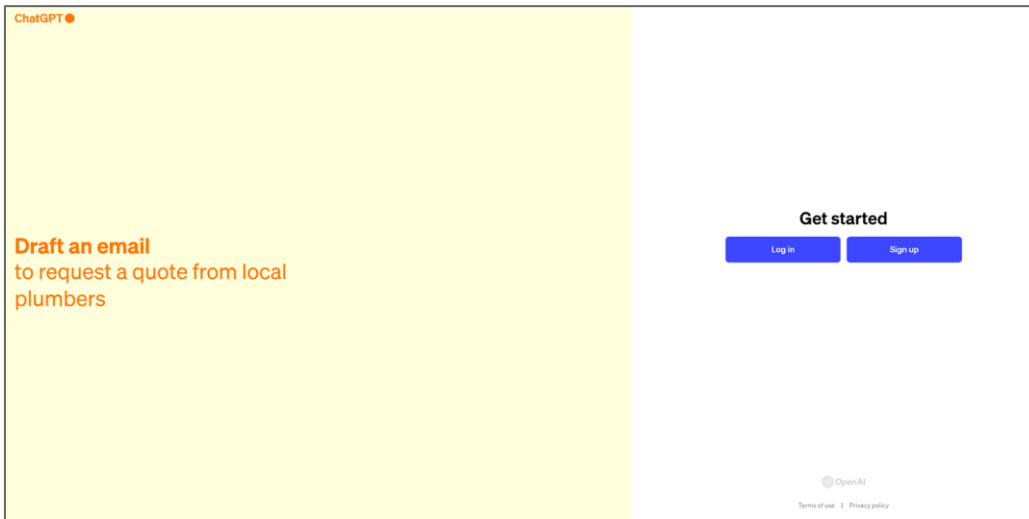
InstructGPT 是一個經過微調的新版 GPT-3，可以將有害的、不真實的和有偏差的輸出最小化

2023-03 ChatGPT (GPT-4)

比 GPT-3.5 更強的推理力，GPT-4 在律師資格考試獲得 PR 90 的佳績，表示 GPT-4 的表現贏過 9 成考生成為頂尖的前 10%；在 SAT 的閱讀測驗達到 PR 93、SAT 的數學考試得到 PR 89 的佳績。能支援圖片輸入功能，讓 GPT-4 用文字來解釋圖中內容，甚至給予建議、回答問題



實際體驗



<https://chat.openai.com/>



圖像生成：

Midjourney / Stable Diffusion



Midjourney

Midjourney 是一個由位於美國加州舊金山的同名研究實驗室開發之人工智慧程式，可根據文字生成圖像，於 2022 年 7 月 12 日進入公開測試階段，使用者可透過 Discord 的機器人指令進行操作。該研究實驗室由 Leap Motion 的創辦人大衛·霍爾茲（David Holz）負責領導。

AI 法規注意事項、生成心法



1. AI 版權

- a. 無版權，但避免使用：別人的IP/插畫家作品/設計師作品
- b. 補充知識! 著作權保護期間

著作權保護期間指的是著作權持有者可以享有版權保護的時間範圍。在台灣，著作權保護期間是計算至作者過世後的 50 年，過了這個期間，作品進入公有領域，可供大眾使用。一些國家也可能有特定的規定或協定，因此建議在具體情況下查詢當地法律或相關法規以確認著作權保護期間。

2. AI 圖像生成心法

- a. 清楚且具體的需求與目標
- b. 模型這麼多，prompt 通用嗎？

AI 不知道自己什麼，所以要清楚地指令中詳細描述你的需求與目標
繼續學習和探索AI技術的發展，以保持競爭力並不斷提升自己。

Prompt 的結構：(主體) + (細節描述) + (風格) + (參數)

參數可能因為不同的模型和不同的 UI 工具，自行輸入文字或是使用介面調整

Thanks!

Any questions?

You can find me at FB: [史旺基](#) / IG: [@swanbear](#)



活動測驗及滿意度問卷

