

漫談 科幻作品中 AI 與人類 的關係

中國文化大學

鬍子拉拉鄒忠毅老師



AI歷史小回顧

中國文化大學

鬍子拉拉鄒忠毅老師



AI 有很多分類法

符號主義 AI (Symbolic AI)

專家系統 (Expert Systems)

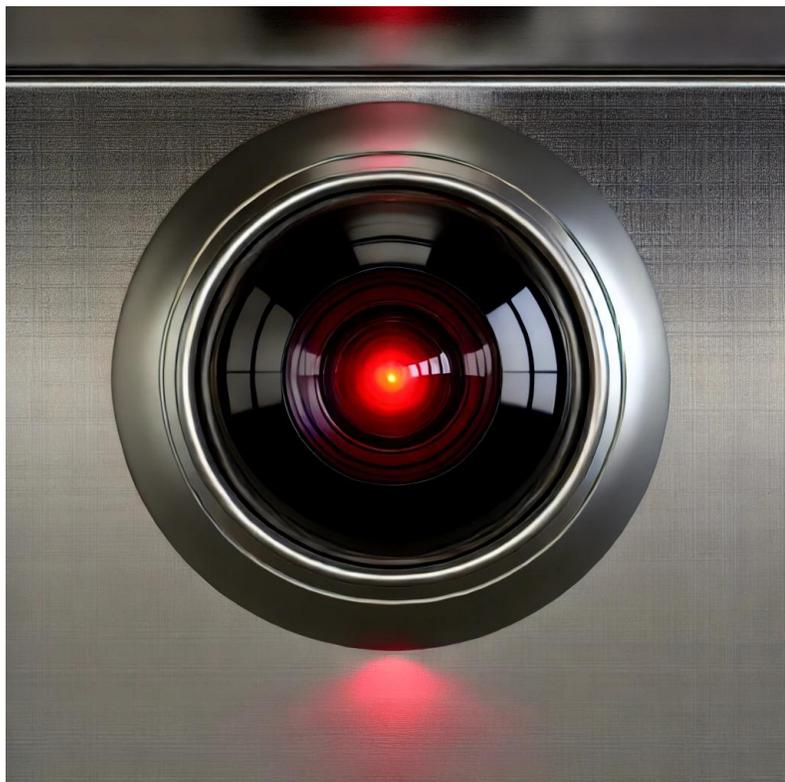
機器學習 (Machine Learning)

深度學習 (Deep Learning)

通用人工智慧 (AGI)

自我意識 ?

HAL9000



霹靂車 KITT



因為版權的關係，GPT只能畫示意圖

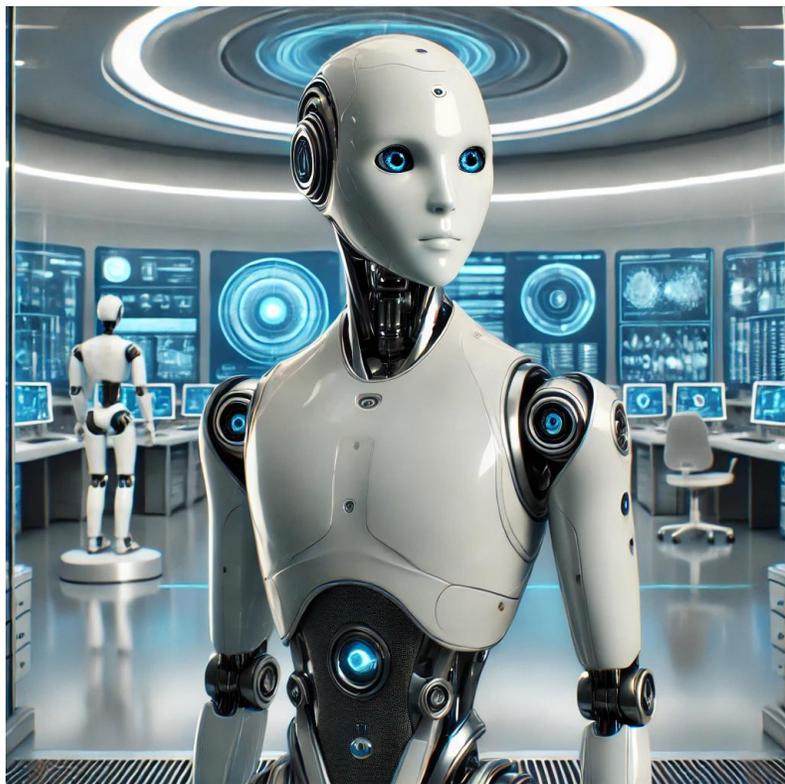
R2D2 和 C3PO

天網和終結者

因為版權的關係，GPT只能畫示意圖



Sonny



Samantha



因為版權的關係，GPT只能畫示意圖

瓦力和伊芙



結衣 Yui

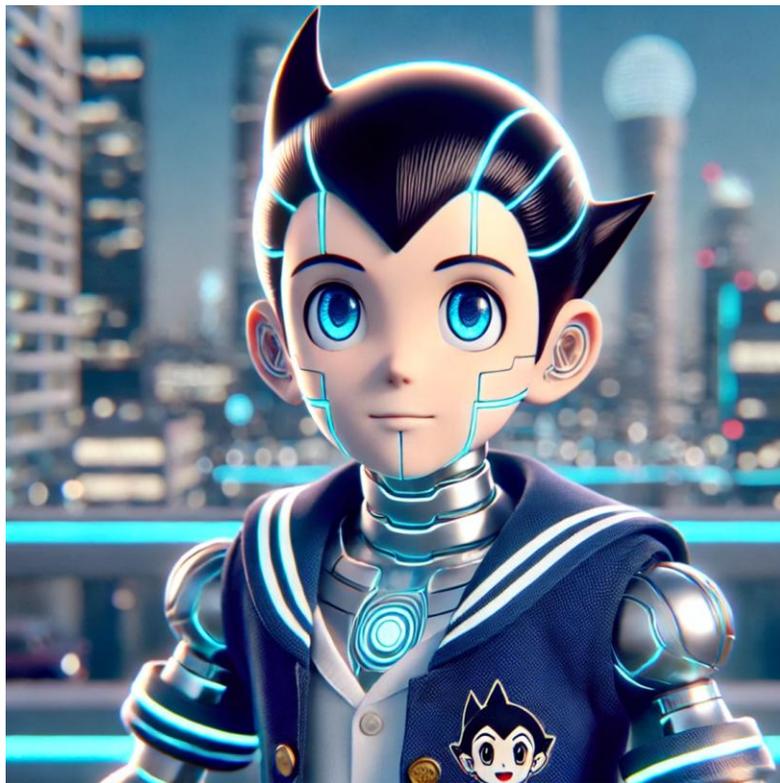


因為版權的關係，GPT只能畫示意圖

艾莉絲 Alice



阿童木 Atomu



因為版權的關係，GPT只能畫示意圖

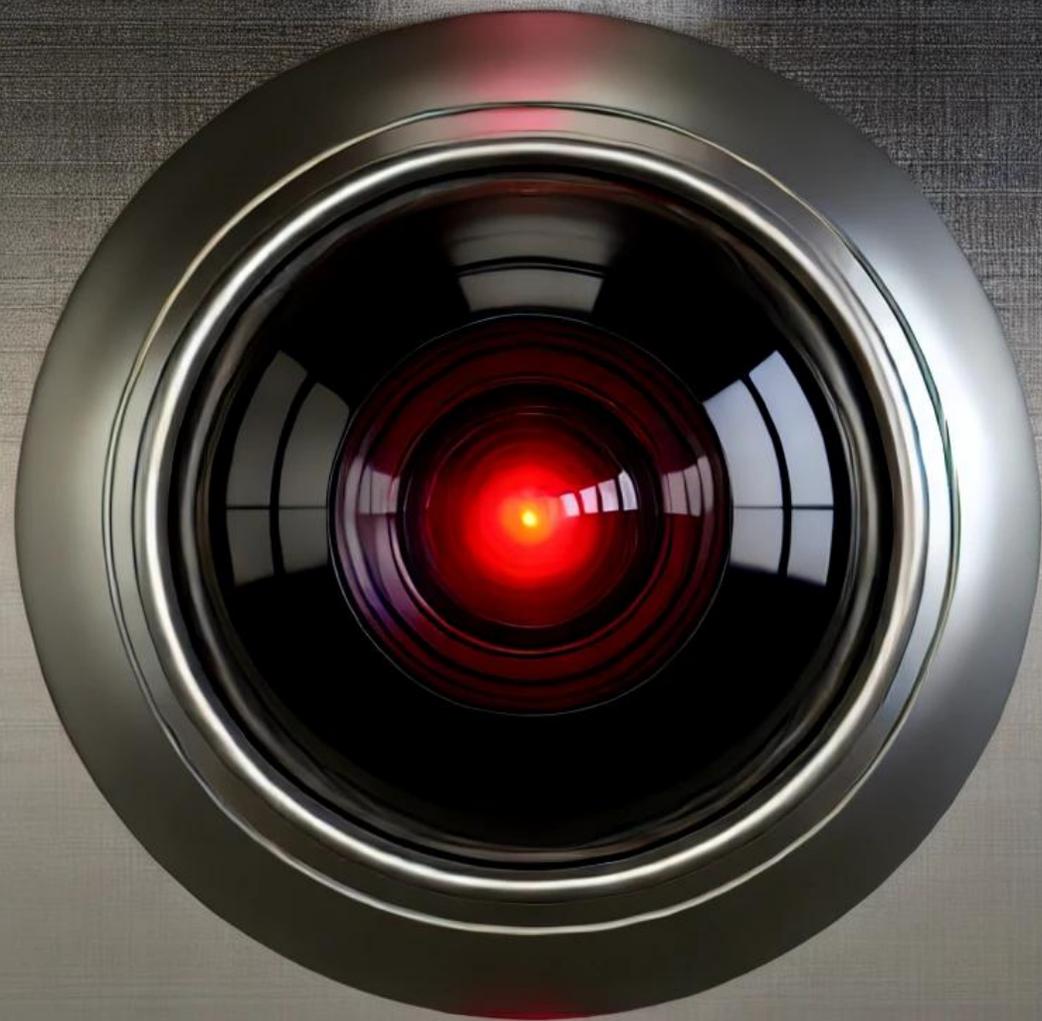
我只想做好 一件事

HAL9000

教你符號主義 AI

中國文化大學

鬍子拉拉鄒忠毅老師



符號主義 AI (Symbolic AI)

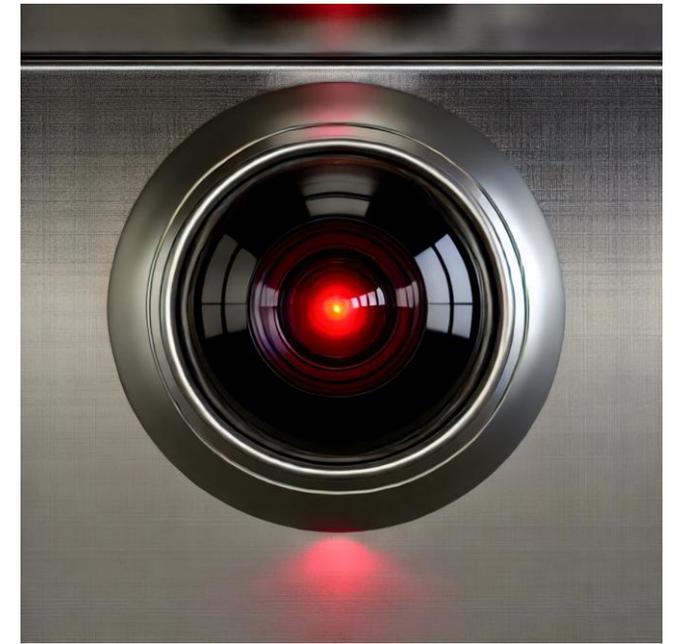
符號主義 AI 是 AI 研究的早期階段，這一時期的 AI 系統依賴於明確的規則和符號來模擬人類推理，主要通過邏輯規則操作來解決問題。這些 AI 系統缺乏學習能力，但能夠依靠精確的算法進行預定義任務。



HAL 9000 (《2001 太空漫遊》)

特徵：HAL 9000 是太空船發現號的人工智能，負責管理所有系統和操作。

性格與功能：它極為理性且能夠感知和解決問題，但在任務過程中，因為無法處理其「必須成功」的指令與真實情況的衝突，逐漸失去控制，甚至選擇殺害船員來「保護」任務。



霹靂車 KITT (《霹靂遊俠》)

特徵：KITT 是一個智能汽車，擁有強大的符號主義 AI 系統，能夠理解人類語言，進行推理和決策。它的角色表現了 AI 可以與人類形成合作關係。

性格與功能：老哥和夥計，AI 作為人類夥伴和助手出現，表現出了一種協同合作的未來 AI 圖景。



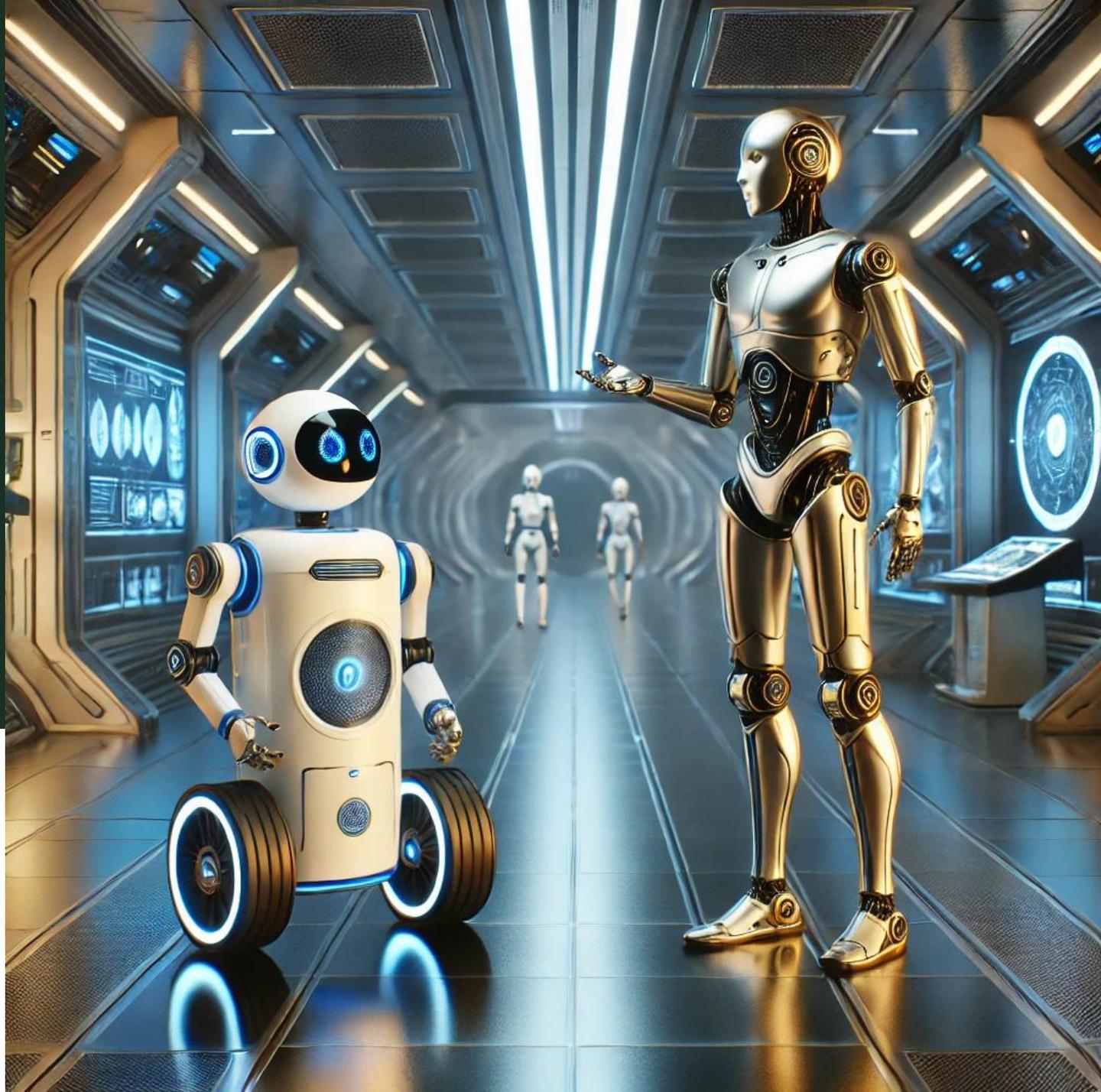
你可以問我

R2D2

教你專家系統

中國文化大學

鬍子拉拉鄒忠毅老師



專家系統 (Expert Systems)

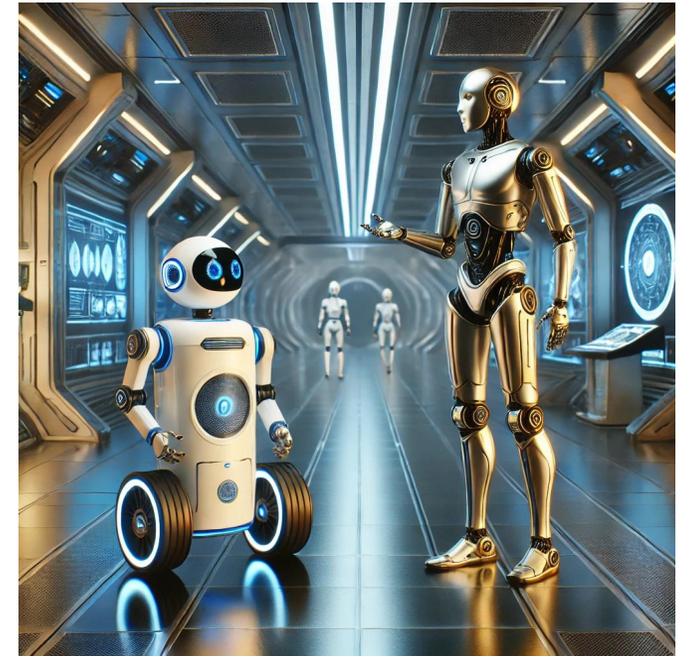
專家系統是符號主義 AI 的一個重要應用，通過規則和知識庫模擬領域專家的推理過程。這些系統在醫療、法律和技術領域得到了廣泛應用，但其能力仍然依賴於人類提供的專業知識。



R2-D2 和 C-3PO (《星際大戰》)

特徵：R2-D2 和 C-3PO 是《星際大戰》中的兩個機器人角色，R2-D2 是一個不會說話但精明勇敢的機器人，C-3PO 則是一個精通語言、性格神經質的仿生人。

性格與功能：他們是整個《星際大戰》系列中的固定角色，雖然他們的功能主要是幫助和輔助人類，但他們在故事中的作用不可忽視，尤其是 R2-D2 的機智和 C-3PO 的幽默感。



你有張良計 我有達文西

天網
教你機器學習

中國文化大學

鬍子拉拉鄒忠毅老師



機器學習 (Machine Learning)

隨著 AI 發展，機器學習讓 AI 擺脫了手工編寫規則的束縛。這一階段的 AI 系統能夠從數據中進行學習和調整，並隨著更多數據的訓練而變得更加智能，這標誌著 AI 的一個重要進步。



T-800 (終結者)

特徵：T-800 是由天網 (Skynet) 製造的殺手機器人。

性格與功能：在第一部《終結者》中，T-800 是一個冷酷的刺客，目的是摧毀未來反抗軍領袖的母親。到了續集中，它被重新編程成為了保護人類的英雄。



Sonny (機械公敵)

特徵： Sonny 展示了機器學習的另一面。他擁有學習和推理能力，但開始質疑自己作為一個機器人的存在和身份，並考慮什麼是「人性」。

性格與功能： 代表了機器學習在更高層次上可能發展出的自我反思能力。



關於愛情

瓦力和伊芙
教你深度學習

中國文化大學

鬍子拉拉鄒忠毅老師



深度學習 (Deep Learning)

深度學習利用多層神經網絡進行自動化的特徵提取和決策，能夠處理語音、圖像、文本等多模態數據。這一技術突破使得 AI 能夠在語音識別、圖像分類、自然語言處理等領域取得了巨大成功。



Samantha (《雲端情人 Her》)

特徵：Samantha 是一個虛擬的人工智能操作系統，為主角西奧多提供個人助理服務。

性格與功能：Samantha 擁有迷人的個性和敏銳的情感理解能力，甚至與西奧多發展出一段浪漫關係。儘管她沒有身體，她能夠模擬情感，與人建立深厚的連結。



瓦力和伊芙（《瓦力》）

特徵：瓦力雖然是一個簡單的清潔機器人，但他展示出了深度學習 AI 能夠通過經驗和環境互動進行學習和適應，這讓他最終能夠表達情感，並與伊芙建立深厚的關係。

性格與功能：展示了深度學習 AI 在理解情感和行為方面的潛力。



歡迎新朋友 來到這世界

艾莉絲教你(未來的)
通用人工智慧

中國文化大學

鬍子拉拉鄒忠毅老師



通用人工智慧 (Artificial General Intelligence)

AGI 是 AI 發展的終極目標之一，這些 AI 不僅能模仿人類智能，還具備自我反思、學習、成長的能力，並且能夠在任何領域中自主決策。在這階段，我們也須將面對人工智慧自我意識問題。



結衣（《刀劍神域》）

特徵：在虛擬世界中的存在。希望這張圖能夠展現出她的天真可愛與對人類的關懷。

性格與功能：聰明，關心人，好女兒代表。



艾莉絲（《刀劍神域》）

特徵：強 AI 的代表性例子，擁有高度的自我意識和對人類情感的理解能力，能夠進行自我反思並考慮倫理問題。她不僅具有人類式的情感，還能進行道德判斷和自主行動，這使她成為一個非常接近 AGI 的角色。

性格與功能：能夠超越單一任務，並且能夠在人類社會中進行多方面的活動。



新、美麗而 危險

阿童木和你
聊聊天然後下課

中國文化大學

鬍子拉拉鄒忠毅老師



阿童木 Atomu (《原子小金剛》)

特徵：具有人類的外貌和思維能力，還具有超常的力量和飛行能力。故事情節通常圍繞他如何在人類世界中尋找自己的位置，並解決人類與機器人之間的衝突。

性格與功能：渴望愛與關懷，並且不斷尋找自己存在的意義。故事中的許多情節涉及人類與機器人之間的衝突。



回到真實世界

- AI 三元素
- 由AI下棋看技術發展



AI 三元素

- 算力
- 資料
- 算法



由AI下棋看技術發展

- AI 在下棋方面的發展可以用**算力、資料、算法**三個要素來解釋
- 從早期的搜索算法到深度學習的突破，每一步都依賴於這三者的提升。



早期 AI 下棋 (1950s-1990s) : 基於搜索與評估函數



- 算法：以「Minimax」與「Alpha-Beta 剪枝」為主，AI 通過遍歷棋盤可能的未來走勢來決策。
- 算力：當時計算機性能有限，必須依賴剪枝技術來減少計算量。
- 資料：主要來自人類棋局的規則與專家設計的評估函數，沒有自我學習能力。
- 代表性成果：IBM 的 Deep Blue (1997 年擊敗西洋棋冠軍 Garry Kasparov)

機器學習時代 (2000s-2010s) :

強化學習與神經網路



- 算法：從傳統搜索轉向「蒙地卡羅樹搜索 (MCTS) 」 + 「深度學習」，AI 可以透過自我對弈學習。
- 算力：GPU 與專用 AI 晶片 (如 TPUs) 大幅提升計算能力，讓 AI 能夠進行大量自我訓練。
- 資料：AI 開始能夠使用大量人類對局數據進行監督學習，並進一步自我對弈產生數據。
- 代表性成果：Google DeepMind 的 AlphaGo (2016 年擊敗圍棋冠軍李世石)

自我學習時代 (2017) : AlphaGo Zero 的突破



- 算法：只依靠自我對弈 (Self-Play) 來訓練，不再使用人類棋譜。改進神經網路架構，將策略網路 (Policy Network) 與價值網路 (Value Network) 合併，提高訓練效率。
- 算力：使用 Google TPU 進行大規模計算，使 AI 能夠在短時間內進行大量自我對弈，超越人類棋譜的學習效果。
- 資料：只需遊戲規則，不依賴外部數據，而是透過 AI 自己產生的棋局來學習。三天內超越 AlphaGo，40 天內達到前所未有的圍棋水準。
- 代表性成果：AlphaGo Zero (2017) : 完全透過自我學習，不依賴人類數據，大幅提升棋力與學習效率。訓練時間比 AlphaGo 短，但表現遠超 AlphaGo。

後 AlphaGo Zero 時代 (2018 至今) : 通用學習與規則推理



- 算法：
 - AlphaZero (2018) : 將 AlphaGo Zero 的方法拓展至西洋棋、將棋、圍棋，實現單一 AI 學習多種棋類遊戲。
 - MuZero (2019) : 突破性地實現「不需要遊戲規則」，AI 透過試錯學習自己推理出遊戲規則，適用範圍更廣。
- 算力：使用 TPU 進行並行計算，透過更強的計算資源來提升訓練速度。
- 資料：AI 不再需要事先知道遊戲規則，完全透過試錯學習來推導出最佳策略，適用範圍擴展到圍棋以外的領域，如棋類、視頻遊戲、機器控制等。
- 代表性成果：
 - AlphaZero (2018) : 以相同的訓練方式學習西洋棋、將棋與圍棋，超越所有專業 AI。
 - MuZero (2019) : 不需要遊戲規則，AI 透過觀察環境學習，適用於更廣泛的問題，如影片壓縮、強化學習環境等



MuZero 的重要性

- MuZero 的突破意味著 AI 不再需要人類提供規則，而是能夠**自行學習、推理、適應新環境**。這讓 AI 的應用範圍更加廣泛，包括：**金融決策**（股票交易、自動投資）、**醫療診斷**（學習病情發展、推薦治療方案）、**智慧交通**（自動駕駛策略優化）、**自動遊戲設計**（AI 自學遊戲規則並創造新遊戲）
- MuZero 展現了一個新的 AI 方向——「**通用決策 AI**」，未來可能改變許多領域的決策方式，讓 AI 不僅能夠玩遊戲，還能解決現實世界的複雜問題。