

Marcel Hedman 「領航 AI 新領域：進化、人力動態和倫理」

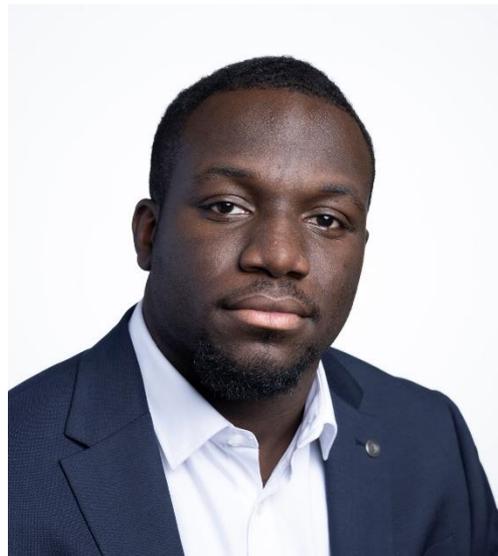


自我介紹

- 創辦人
Nural Research
- 產品經理
Enterprise AI manufacturing start-up

經歷

- 顧問
麥肯錫公司 (**McKinsey & Co.**)
- Choate Memorial 研究員
哈佛大學
- 自然科學 (物理)
劍橋大學





今日目標

1.

2023 年的人工智慧 (AI) 是什麼，誰是關鍵行動者？



2.

人工智慧 (AI) 如何影響產業和人力？



3.

倫理考量和新興法規





迷思破解

迷思

人工智慧 (AI) 和機器學習相同

人工智慧 (AI) 會讓人類被職場淘汰\

人工智慧 (AI) 將超越人類智慧

人工智慧 (AI) 可以做出比人類公平的決定

資料科學家和機器學習 (ML) 工程師正在塑造人工智慧 (AI) 的未來

現實

機器學習是人工智慧 (AI) 的分支

提升人類工作表現時人工智慧 (AI) 的表現最優異

穩健且一般的人工智慧 (AI) 是遙不可及的未來

人工智慧 (AI) 會導致資料集中深植的偏差惡化：垃圾進，垃圾出

人工智慧 (AI) 由所有關鍵決策者塑造。有效並符合倫理的實行需要跨領域思維才得以完成



今日目標

1.

2023 年的人工智慧 (AI)
是什麼，誰是關鍵行動
者？



2.

人工智慧 (AI) 如何影響
產業和人力？



3.

倫理考量和新興法規





瞭解生成式人工智慧 (AI)



傳統人工智慧 (AI)

對比



生成式人工智慧 (AI)



生成式人工智慧 (AI) 已解鎖大眾消費者和商業採用

全球支出提高：
生成式人工智慧 (AI) 將於 2032 年成為 1.3 兆元的市場

達到 1 億使用者的時間
(Visual Capitalist)

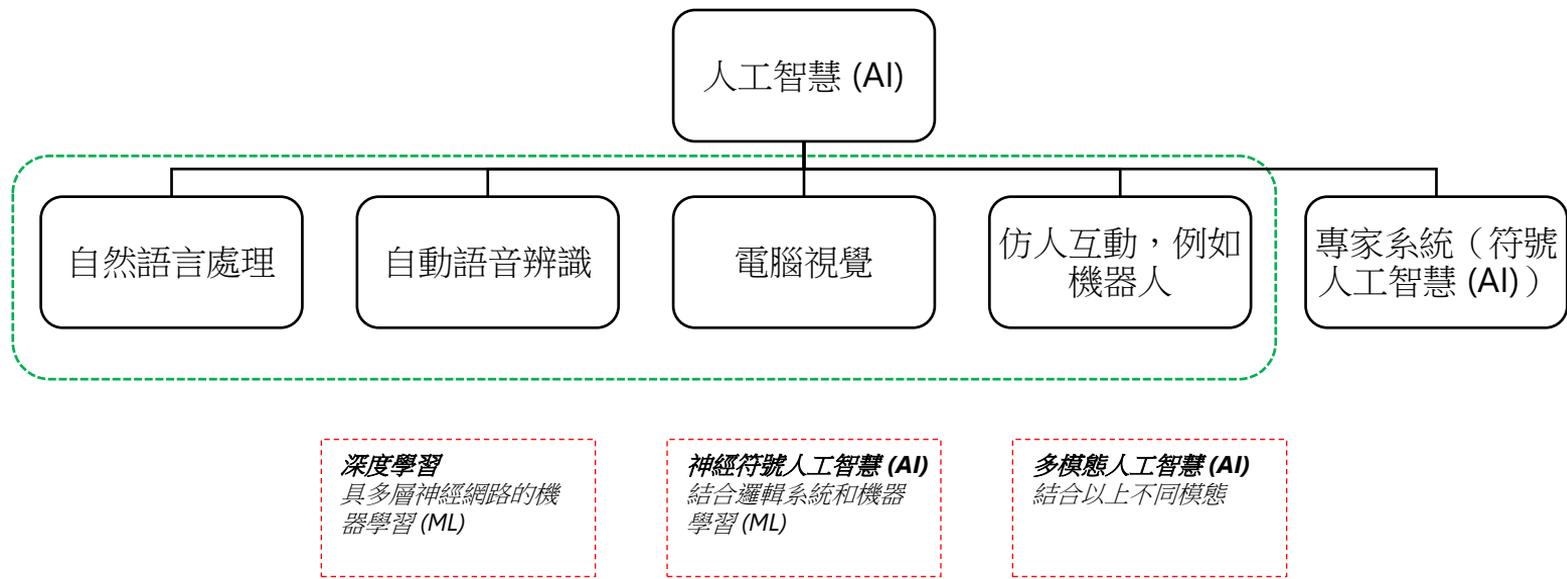


討論題目

生成式人工智慧 (AI) 如何連結人工智慧 (AI) 產業的各個環節？



人工智慧 (AI) 劃分：機器學習是近年人工智慧 (AI) 進步背後勞苦功高的推手



 = 在機器學習 (ML)/強化學習傘下



更深入瞭解機器學習 (ML)

非監督式

- 沒有標記的資料
- 資料中發現先前未知的模式

監督式

- 標記的訓練資料
- 將參數調整為訓練集

強化

- 即時從試驗和錯誤中學習
- 巧妙運用神經網路作為政策近似者
- 獎勵和處罰



可於業界的資料結構中窺見

結構化

- 標準的資料結構描述
- 易於分析和查詢疑問
- 降低可能的儲存量
- 提升整合潛力
- 可能難以修改



半結構化

- 無預先定義的資料結構描述
- 包含更多結構屬性或階層
- 比非結構化資料更方便儲存

例如 JSON、CSV

非結構化

- 無清楚定義的架構（例如自由文字、圖片）
- 快速收集，處理速度慢
- 彈性的結構描述讓查詢問題更快速（沒有接合點）



mongoDB 資料湖

The Foundation Model Stack is Growing More Powerful

基礎模型堆疊更強大

APPLICATION

應用程式

New 新的

Existing 現有的

Jasper HyperWrite runway Adopt tome
 GitHub Copilot coqum Cabline ChatGPT
 code Notion yoodli lexion Microsoft Adobe

Explosion of new applications and reimagining of existing applications based on foundation models.

OPPORTUNITY: In the long run, almost every kind of application will be recreated or reimaged.

新應用程式百花齊放，並根據基礎模型重新發現現有應用程式。

機會：長遠可見，幾乎所有種類的應用程式都會被重新創造或重新發現。

工具層興起，讓開發人員可以更快地打造檔案管理 (FM) 應用程式，並連線至外部端點。

TOOLING

工具層

Orchestration 協調

Evaluation 評估

LangChain Fxfile DUST
 GPTIndex Cognosis
 Humanloop WHYLABS
 HoneyHive

DATA SOURCES

資料來源

External DBs and APIs
 外部 DB 和 API

ACTIONS

行動

External DBs and APIs
 外部 DB 和 API

A tooling layer has emerged to let developers build FM applications faster and connect to external endpoints.

OPPORTUNITY: New ways to differentiate apps, apart from the underlying FM model capabilities.

機會：除了潛在的檔案管理 (FM) 模式容量外，區分應用程式的另一種新方法。

FOUNDATION MODEL

基礎模型

Proprietary FMs 專屬 FM

Model Hub

Open-source FMs 開放原始碼 FM

Open-source Data Sources 開放原始碼資料來源

OpenAI codex
 Anthropic
 Google AI
 Microsoft Turing
 HuggingFace GitHub
 Stable Diffusion EleutherAI BLIP OPT GLM-130B
 LAION The Pile Common Crawl

An iPhone vs. Android style battle is emerging between proprietary and open-source models.

OPPORTUNITY: Different teams with different missions will choose sides based on priorities of development speed versus customization.

專屬模型和開放原始碼模型間興起 iPhone 與 Android 式的對決。

機會：進行不同任務的不同團隊，會根據開發速度或最佳化為優先而抉擇。

FM Ops 讓開發人員更有效最佳化、訓練並執行模型。

FM OPS

FM OPS

Deployment Optimization: 部署最佳化: OctoML

Inference: 推論: Modal beam SanumCloud BANANA

Training: 訓練: mosaic carebras Lightning

Data Tooling: 資料工具: Snorkel xetdata fastdup

FM Ops lets developers optimize, train, and run their models more efficiently.

OPPORTUNITY: Sophisticated teams can use FM Ops to create differentiated capabilities and cost structures.

機會：經驗豐富的團隊可以利用 FM Ops，建立區隔的容量以及成本結構。

策略上，FM 應用程式是如此重要，會影響開發人員的雲端平台選擇。

CLOUD

雲端

Azure aws Google Cloud

FM apps are so strategic that they will influence developers' broader cloud platform choices.

SILICON

矽

NVIDIA AMD amazon SambaNova Google cerebras

Access to a supply of GPUs and AI-specialized silicon is critical for FM training and inference.

得以取得 GPU 和 AI 專用矽的供應，對 FM 訓練和推論十分關鍵。





人工智慧 (AI) 革命的 5 大推手



1. 資料



2. 儲存



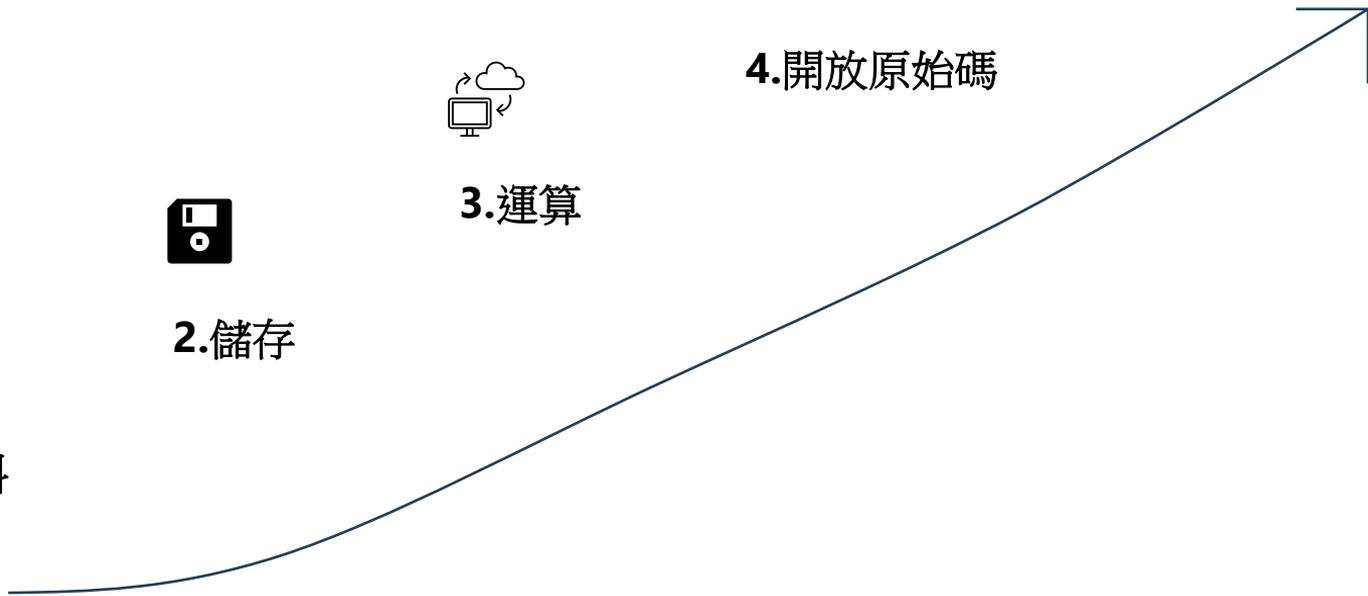
3. 運算



4. 開放原始碼



5. 才華





→ 今日目標

1.

2023 年的人工智慧 (AI) 是什麼，誰是關鍵行動者？



2.

人工智慧 (AI) 如何影響產業和人力？



3.

倫理考量和新興法規





產業焦點：需要創意或重複性任務的產業受到最 多衝擊

各地區、產業和年齡層的受訪者表示，他們已經在使用生成式人工智慧 (AI) 工具。

Respondents across regions, industries, and seniority levels say they are already using generative AI tools.

報告使用生成式人工智慧 (AI) 工具的人數，佔受訪者的 %

Reported exposure to generative AI tools, % of respondents

選擇人口統計

Select demographic

By industry

依據產業

在工作上經常使用

在工作上和工作之外經常使用

在工作之外經常使用

■ Regularly use for work

■ Regularly use for work and outside of work

■ Regularly use outside of work

■ Have tried at least once

■ No exposure

■ Don't know

使用至少一次

沒接觸過

不知道

先進產業
Advanced industries



商業、法律和專業服務
Professional services



消費性商品/零售
Consumer goods/retail



能源與材料

Energy and materials
金融服務
Financial services



醫療保健、藥品和醫學產品
Healthcare, pharma,
and medical products



科技、媒體和電信
Technology, media, and
telecom



14



「人工智慧 (AI) 已經從科技業員工的低階話題，崛起成為公司領袖的目光焦點」 - 麥肯錫

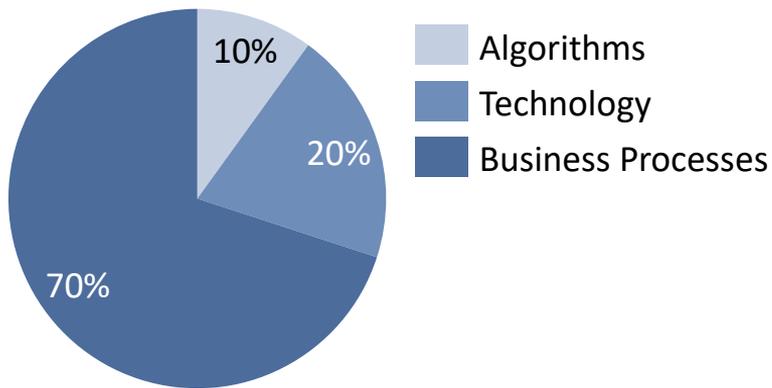


業務功能：採用人工智慧 (AI) 帶動收益提高，並降低成本

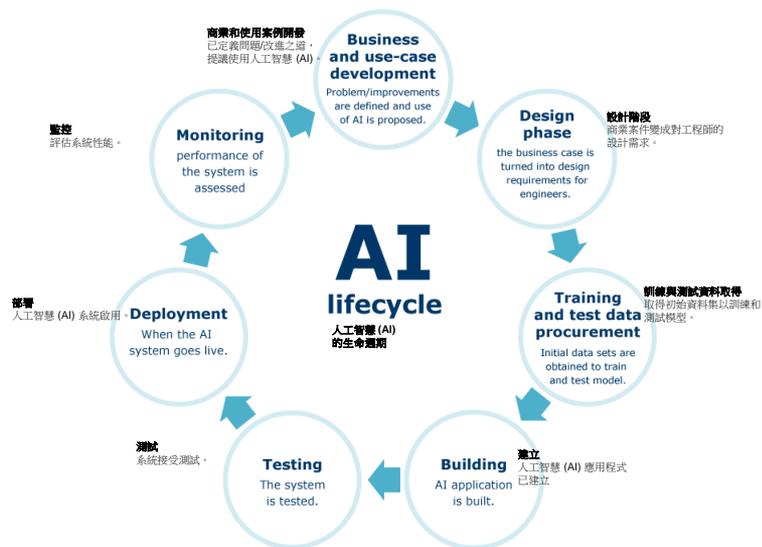
範圍	範例	收益影響	成本影響
風險	影響和可能性 建議系統、保護安全的智慧型裝置	●	●
策略和財務	機器人流程自動化，以自然語言處理 (NLP)/ 電腦視覺進行稽核	◐	◐
製造	按部就班分析以預測需求、機器視覺 – 品質控管	●	●
人力資源	人工智慧 (AI) 面試、公司內部聊天機器人	◐	◐
R&D	試驗自動化、新穎資料洞察	◐	◐
行銷和銷售	量身打造的客戶體驗	◐	◐

→ 在企業內實行人工智慧 (AI)

典型人工智慧 (AI) 投資分配



典型的人工智慧 (AI) 工作流程





人力：人工智慧 (AI) 對知識工作的衝擊最大

當今人工智慧 (AI) 技術的 4 大人力衝擊

1. 截至 2045 年，知識工作者目前花費 60% 的時間用來進行自動化
2. 現有的人工智慧 (AI) 相關職務需求將增加
3. 需要理解語言，但專業程度低，被撤換風險最高的職務
4. 創意職務將受到重大衝擊，並且可能面臨「自我競爭」



新興的人工智慧 (AI) 相關職務將興起，解鎖新的企業產能



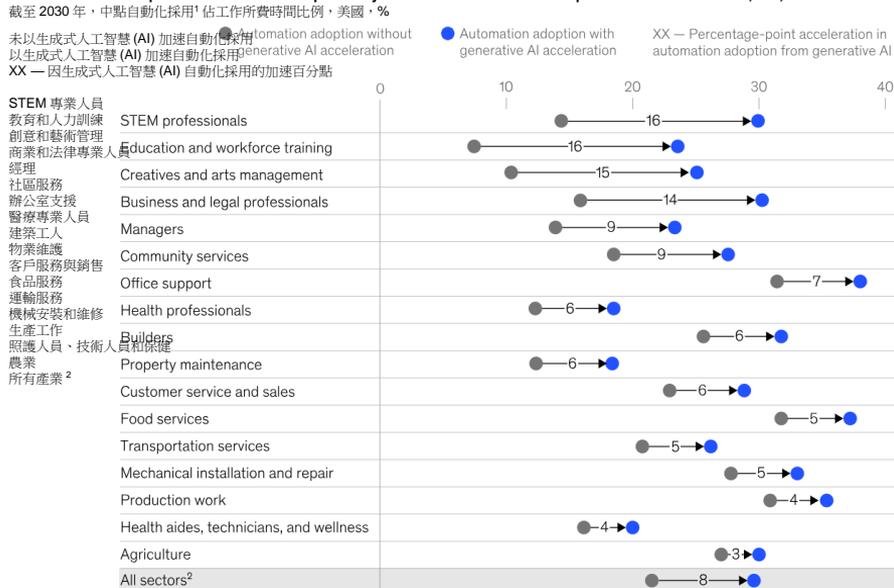
AI 工程師



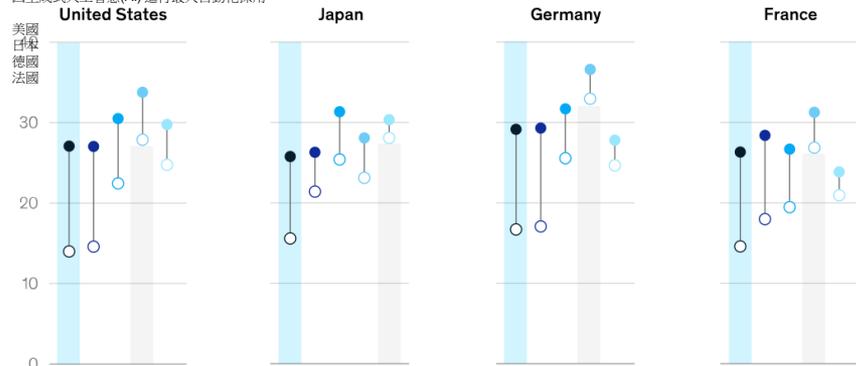
提示工程師

人力：雖然生成式人工智慧 (AI) 對高薪工作的自動化衝擊最大，但對低薪工作的衝擊更嚴重

Midpoint automation adoption¹ by 2030 as a share of time spent on work activities, US, %



Automation adoption per wage quintile, % in 2030, midpoint scenario



麥肯錫



5 項人工智慧(AI) 和產業預測：未來三年

1. 業界罕見採用人工智慧(AI) 成長率將上升至高點並停滯
2. 服務作業的人力規模將大幅縮減
3. 多數企業將配置人工智慧(AI) 和提示工程師
4. 主要的人工智慧(AI) 實驗室將因訓練資料而面臨法律訴訟
5. 穩健、可說明的機器學習 (ML) 將取代目前的生成式方法成為主流



展望：人工智慧(AI) 產業

槓桿

推動者

- ✓ 大數據
- ✓ 提升的處理力 - 摩爾定律
- ✓ 摘要的機器學習 (ML) 平台
- ✓ 資金
- ✓ 要求個人化經驗
- ✓ 人類等級的理解

挑戰

- ✓ 法規行程和領航
- ✓ 可解釋性
- ✓ 責任
- ✓ 邊緣案例
- ✓ 資料和概念漂移
- ✓ 資料所有權

機會

- ✓ 處理不平等問題
- ✓ 非英文人工智慧(AI) 的成長
- ✓ 零樣本學習

→ 今日目標

1.

2023 年的人工智慧 (AI) 是什麼，誰是關鍵行動者？



2.

人工智慧 (AI) 如何影響產業和人力？



3.

倫理考量和新興法規





人工智慧

Artificial intelligence

+ 新增至 myFT

+ Add to myFT

How actors are losing their voices to

演員為何被人工智慧(AI) 奪走他們的工作機會

AI

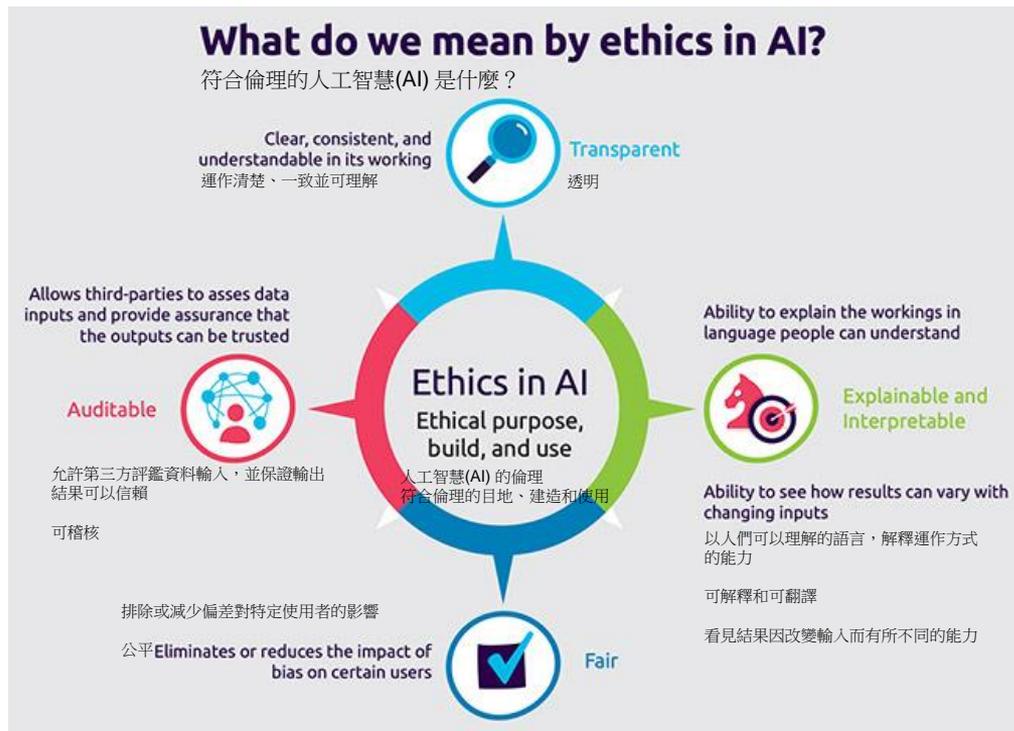
企業使用複製科技激起更新著作權法聲浪的同時，演員們被迫與自己競爭

Performers forced to compete with themselves as companies' use of technology for cloning prompts calls to update copyright law

→ 什麼是符合倫理的人工智慧(AI)

符合倫理的人工智慧(AI) 會
考量：

1. 人工智慧(AI) 使用更廣大的外部空間
2. 巧妙運用人工智慧(AI) 解決重大挑戰

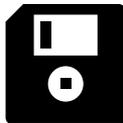




架構：偏差發生之處

1.

資料來源



2.

建立模型和訓練



3.

最終使用者實施狀況不良



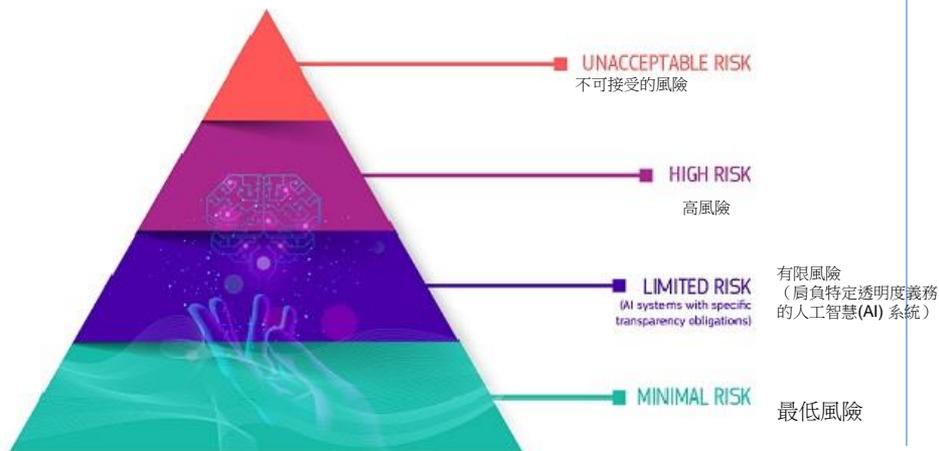
→ 風險

類別	詳細資訊	風險等級
安全和性能	高風險產業性能不佳 - 資料漂移、產生幻覺	●
	社會武器 - 錯誤資訊、操弄和深偽	●
	自主武器	◐
公平和隱私	偏差 - 因資料集的統計分佈而強化	●
	資料隱私	●
外部性	失業	◐
	自由行動人工智慧(AI)	◑



法規：監管機關正在開始縮小與科技的鴻溝

歐盟人工智慧法案：依據風險的方法



注意：提出初始法案時未預期 LLM

中國：生成式人工智慧服務管理辦法過渡期措施

- 取得創新和法規的平衡



基礎模型供應商遵循歐盟人工智慧法案草案程度的評分

Grading Foundation Model Providers' Compliance with the Draft EU AI Act

資料來源：史丹佛基礎模型研究 (CRFM) · Human Centered Artificial Intelligence (HAI)

Source: Stanford Research on Foundation Models (CRFM), Institute for Human-Centered Artificial Intelligence (HAI)

	OpenAI	cohere	stability.ai	ANTHROPIC	Google	BigScience	Meta	AI21labs	ALEPH ALPHA	EleutherAI		
人工智慧 (AI) 法案草案規定	GPT-4	Cohere Command	Stable Diffusion v2	Claude	PaLM2	BLOOM	LLaMA	Jurassic-2	Luminous	GPT-NeoX	總計	Totals
Draft AI Act Requirements	GPT-4	Cohere Command	Stable Diffusion v2	Claude	PaLM 2	BLOOM	LLaMA	Jurassic-2	Luminous	GPT-NeoX		
資料來源 Data sources	● ○ ○ ○	● ● ● ○	● ● ● ●	○ ○ ○ ○	● ● ○ ○	● ● ● ●	● ● ● ●	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●		22
資料治理 Data governance	● ● ○ ○	● ● ● ○	● ● ○ ○	○ ○ ○ ○	● ● ● ○	● ● ● ●	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ● ○		19
受著作權保護的資料 Copyrighted data	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●		7
運算 Compute	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●	● ● ● ●	○ ○ ○ ○	● ○ ○ ○	● ● ● ●		17
能源 Energy	○ ○ ○ ○	● ● ○ ○	● ● ● ○	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●	● ● ● ●	○ ○ ○ ○	○ ○ ○ ○	● ● ● ●		16
產能和限制 Capabilities & limitations	● ● ● ●	● ● ● ●	● ● ● ●	● ○ ○ ○	● ● ● ●	● ● ○ ○	● ● ○ ○	● ● ○ ○	● ○ ○ ○	● ● ● ○		27
風險和緩解 Risks & mitigations	● ● ● ●	● ● ● ●	● ● ○ ○	● ○ ○ ○	● ● ● ○	● ● ○ ○	● ● ○ ○	● ● ○ ○	○ ○ ○ ○	● ○ ○ ○		16
評估 Evaluations	● ● ● ●	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ○ ○	● ● ○ ○	● ● ○ ○	○ ○ ○ ○	● ○ ○ ○	● ○ ○ ○		15
測試 Testing	● ● ● ○	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ○ ○	● ● ○ ○	○ ○ ○ ○	● ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○		10
機器產生的內容 Machine-generated content	● ● ● ○	● ● ○ ○	○ ○ ○ ○	● ● ● ○	● ● ● ○	● ● ○ ○	○ ○ ○ ○	● ● ● ○	● ○ ○ ○	● ● ○ ○		21
會員國 Member states	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ○ ○	● ● ● ●	○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ○ ○ ○	○ ○ ○ ○		9
下游文件 Downstream documentation	● ● ● ○	● ● ● ●	● ● ● ●	○ ○ ○ ○	● ● ● ●	● ● ● ●	● ● ○ ○	○ ○ ○ ○	○ ○ ○ ○	● ● ● ○		24
Totals	25 / 48	23 / 48	22 / 48	7 / 48	27 / 48	36 / 48	21 / 48	8 / 48	5 / 48	29 / 48		



人工智慧 (AI) 4 優良



符合倫理和法律的實行



深具影響力的應用程式



關於 Nural Research



探索人工智慧 (AI) 的使用情形，以迎戰全球大挑戰



縮小使用人工智慧 (AI) 和正在開發人工智慧 (AI) 間的鴻溝



人工智慧 (AI) 和資料科學諮詢



訂閱 – 週報和文章

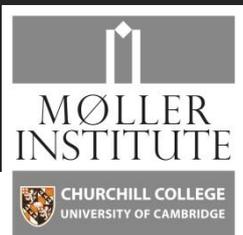
自我介紹

- Nural Research 創辦人
- 前哈佛大學 Choate Memorial 研究員
- 熱中普及知識和天馬行空打造新東西

聯絡方式

網站：www.nural.cc

LinkedIn：Marcel Hedman



感謝您的聆聽

Marcel Hedman
Nural Research

Møller Institute Limited
Churchill College, Storey's Way, Cambridge CB3 0DE



The Generative AI Application Landscape

生成式人工智慧 (AI) 應用程式概況



應用程式層 APPLICATION LAYER	Marketing (內容)						
	銷售 (電子郵件)	程式碼產生 Code generation	圖片產生 Image generation				
支援 (聊天/電子郵件)	程式碼文件 Code documentation	消費者/社會 Consumer / Social					RRPA Music
General writing 一般寫作	Text to SQL	媒體/廣告 Advertising					音樂 Music
Note taking 筆記	網頁應用程式 Web app builders	設計 Design	聲音合成 Voice Synthesis	影片編輯/產生 Video generation	3D 模型/場景 3D models/scenes		生物科學 & 化學 Biology & chemistry
Other							生物和化學 Biology and chemistry
	TEXT 其他	CODE	IMAGE	SPEECH	VIDEO	3D	OTHER
	OpenAI GPT-3 文字	OpenAI GPT-3 程式碼	OpenAI Dall-E 2 圖片	OpenAI 語言	Microsoft X-CLIP 影片	Dream Fusion 3D	其他 Other
	DeepMind Gopher	Tabnine	Stable Diffusion		Meta Make-A-Video	NVIDIA GET3D	
	Facebook OPT	Stability.ai	Craiyon			MDM	
	Hugging Face Bloom						
	Cohere	模型層 待確認					
	Anthropic						
	AI2						
	Alibaba, Yandex, etc.						

→ Ignitarium



背景：

Ignitarium 是位於印度的產品工程設計公司，打造即時噪音抑制系統等產品

問題：

傳統的噪音抑制方法不完全有效（包括濾波器和傳統的 DSP 演算法）

解決方案：

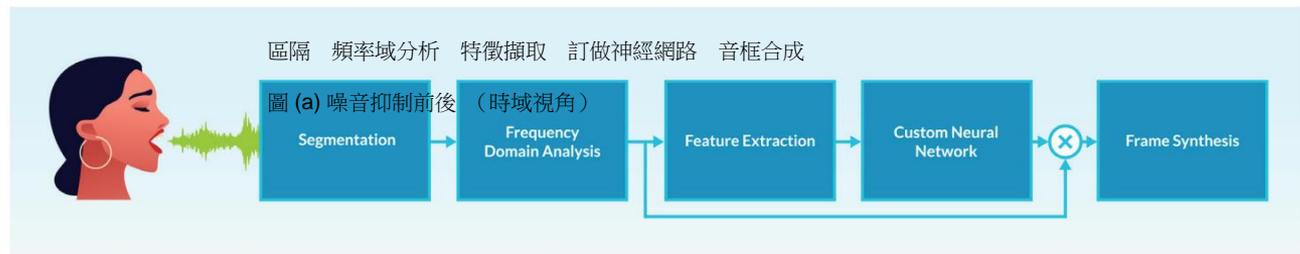
該公司開發深度學習的 (GRU) 小型記憶噪音抑制系統，可在 RAM 低的裝置上運作

成果：

穩態和非穩態噪音環境下，抑制噪音



Fig.(a) Before and after noise suppression (time domain view)





Disney



背景：

Disney 於 1923 年創辦，始終保留內容資料庫，供日後創作者汲取靈感。先前的實體資料庫現已數位化。

問題：

查詢超過 100 年的內容非常耗時，如果沒有相應的大數據標籤，工作會窒礙難行

解決方案：

Disney 開發一套深度學習標籤系統，強化標籤流程（「內容基因」）

訓練在動畫中區分類似的特徵

分類法必須穩健

成果：

特定且量身訂做的搜尋（例如爆炸）

縮短手動觀看時間

可能改變影片搜尋的未來