何彥南 Yen-Nan (Aaron) Ho

新北市，台灣 | 個人網站：https://aaron1aaron2.github.io/
Email：hoyennan@gmail.com

AI Engineer

專業摘要

現任皓博 Data Scientist，專注工業 AI 產品與客戶交付。主導 Harbor AI 2.0 監控平台的產品 / 領域設計，以及設備出廠品質檢測從需求到現場交付的端到端專案；結合資料判斷、系統整合與 AI 協作，把分析方法做成可操作、可驗證的產品。

研究以 MCIENet 碩論為最完整案例，涵蓋模型、資料流程、消融與跨資料集評估；另累積客流、金融文本與人物再辨識研究 / 專案經驗。

技術能力

- **資料科學與工業診斷：** Python、NumPy、pandas、SciPy、scikit-learn、PyTorch；振動 / PRPD、FFT / WAV、RUL、統計檢定與效果量、CNN、消融、XAI 及資料可用性評估。
- **AI 與報告系統：** LangChain / LangGraph、vLLM、OpenAI-compatible APIs、確定性分析與 LLM 敘事分離、多設備隔離、失敗備援與人工驗證。
- **後端與產品：** Django / Django Ninja、Pydantic、PostgreSQL、Redis、RabbitMQ、REST、非同步工作流、React / TypeScript。
- **工程與交付：** Docker / Compose、離線與跨架構安裝、更新 / 回復、資料遷移、Git、自動化測試、AI agents 與人工發布審查。

工作經歷

Data Scientist，Harbor Tech | 台北，台灣 | 2025.10-目前

參與 Harbor AI 2.0 兩項定位不同、共用基礎技術的產品：設備常態監控與診斷，以及逐件出廠品質檢測。本人主導產品 / 領域與客戶交付工作；AI agents、工程夥伴依功能參與實作與驗證，具體責任如下。

工業智慧監控平台：從領域流程到獨立運作的產品

- **產品與操作邏輯：** 主導產品及領域設計，與團隊從 0 到 1 建構多數功能。將振動、PRPD、RUL、趨勢與資料品質轉為分層設備 / 感測器狀態、事件檢視及操作流程；以設備為報告入口，串接多語、圖表、行動版與真實 API。
- **新舊系統演進：** 在共存期間釐清設備資料與使用者設定的權責，避免舊側同步覆寫新側設定；推進設備控制權轉移、事件換源與可稽查遷移。感測器對應不明時停止匯入，補正映射再驗證。2026-10-23 本人確認 2.0 已不依賴 1.0，這不等於所有客戶 / 資料均已切換。
- **分析工作流：** 與平台工程夥伴協作 Django 非同步編排，定義 API、資料存取、檔案引用、任務生命週期及失敗行為；整合以資料庫為準的逐感測器排程、手動觸發、批次與歷史，按資料可用性選擇彙總或 WAV 輸入。
- **邊緣資料與控制：** 主責 Gateway 資料接收規格與方案驗證，對齊批次上傳、心跳、重複資料、完整性及容量邊界，移交對應工程師實作。另參與伺服器端軟體套件分發，以及 RUL 建議到平台閾值、歷史與 Gateway 更新的控制回饋，分開線上接收、遷移、OTA 與閾值策略。

智慧設備出廠品質檢測平台：需求、操作與現場交付

- **端到端責任：** 主導客戶需求、規格、系統實作、UI / UX 優化與現場交付，與 AI agents 及工程夥伴建立逐件檢測流程。釐清受測件、量測批次、良品基準、多點同步、插單、序號 / 工單、人工複核與報告；基本功能驗證與現場判定校準分開推進。
- **操作重整：** AI 初稿符合規格，卻需 5-6 層頁面跳轉並出現循環導航。本人將日常操作收斂為計畫清單、工單、看板三類核心介面，將報告、複查、基準管理整併或改為對話框；集中主要動作，加入方向鍵 / Enter、狀態與感測器脫落提醒。DQA 回饋為定性改善。

- **量測防錯：** 將感測器安裝確認與正式量測拆成兩階段，以 sensor ID 與安裝示意圖作硬性檢核，避免錯誤配對後直接開始量測。
- **現場判斷與修正：** 2026-10-12 到場架設、示範、教育訓練並回答客戶問題。連線中斷時親自定位主機時間偏移，輸入排查命令並完成現場及後續修正；Claude 參與現象與假設討論。將回饋理清後帶回團隊，並說明操作與設計邏輯。
- **交付界線：** 協調 DQA、測試版本、操作手冊與驗收紀錄；專案持續推進，正式簽核與最終驗收結論待確認。本人曾估整體貢獻約 80%，為全流程自估，非手寫程式比例或量測指標。

模型、診斷與報告：讓分析結論對得上資料

- **假設審查：** 審查 RUL 分析時追查來源，發現測試常數被當成觀測資料，作廢原結論並要求多參數、多曲線及正反向消融。讓研究 notebook 直接呼叫產品程式，檢查快速劣化漏報、健康序列誤報與實作漂移。
- **診斷與輸出規則：** 串起前處理、有效運轉區段、異常分類、ISO 嚴重度、RUL 與報告，分開故障型態、嚴重度及剩餘時間。檢查取樣、覆蓋、選擇偏差與分母；資料不足時顯式降級，保留引擎輸出。現場壽命真值與信心校準未完成，不主張 RUL 準確率。
- **報告服務產品化：** 從人工驗證的案場報告整理方法及品質規則；七月用 LangChain / LangGraph POC 驗證節點契約、多設備隔離、局部重產與備援。確定性節點負責數值與分類，LLM 組織敘事；八月推進服務接入平台，九月 DOCX 與前端下載合入主線。本人主導方法與品質條件，團隊及 AI 參與實作 / 整合，各路徑部署與客戶驗收分別追蹤。
- **模型服務介面：** 完成 vLLM 真實端點的 chat、embedding 與 tool calling 驗證，分開連線、功能及語意健康檢查；不以端點可呼叫代表所有報表路徑均已切換。

AI 協作工程與版本交付：讓並行工作可接續、可審查

- 建立串起需求、分析、決策、issue、測試與版本的共用紀錄。本人拆需求、提供持久上下文與工具，訂驗證及停止條件；AI agents 在隔離分支研究、實作、修復與複驗。
- 開發機每日執行各功能回歸與監控，異常進入 issue 及 Bitbucket PR；本人審查資料假設與產品行為，與 reviewer 決定合併及發布。將反覆問題回饋為可重用規則、測試及檢查，未量測整體效率倍數或自動修復成功率。
- 整合 Docker / 離線套件、跨架構安裝、版本資訊及平台自更新 / 回復；將更新執行與平台停機風險隔離，加入前置檢查、進度、失敗訊號及回復程序。真機問題回饋至停寫、資料結構、約束、儲存與重啟檢查。
- 收斂交付包內容、排除客戶原始資料，完成全新安裝驗證，主導出貨版本整合與切版判斷。遷移另核對欄位、時區、可讀性與連續性；以 VitePress 串接程式、issue、版本及手冊。

案場分析、技術交付與協作

- 在半導體與大型多感測器場域的局部放電分析中，與現場 / 品質團隊對齊 DB、事件與 WAV 範圍。以完整資料庫定位事件，再取對應波形，串接批次分析、累積 PRPD 圖、AI 草稿與人工逐頁複核。
- 在馬達 / 泵浦分析中，以 max / P99 區分持續嚴重度與單秒尖峰，檢查運轉區段、RUL 依據及 VRMS 波動；後續以分鐘中位數、非參數檢定、多重比較校正與效果量複核客戶觀察。分析與客戶說明已交付，根因仍需現場轉速、流量等資訊。
- 逆向 legacy code、notebooks 及歷史報告，重建可追溯的 ADC-to-g、FFT、SNR、切割點與頻域流程，以可執行程式及手冊交付，保護公司演算法 IP。
- 為研究實習設計離線題目、教材與 starter notebooks，按週報檢視資料前提及方向；實習生保留分析作者歸屬。另主責公司側智慧城市邊緣平台提案的技術文件、版本與跨公司介面 / 查核點對齊，成果屬未送件草案。

資料科學實習生，中央研究院社會學研究所 | 台北，台灣 | 2019.05-2020.12

- 整合正式抽樣地址、點間距離及 fallback distance，建立 TSCS 訪區分組、OR-Tools TSP 與 Folium 流程，支援實地訪查路線、離散加給與預算估算。
- 使用 Selenium、multiprocessing、Tor multi-port 與 Folium 進行距離蒐集與 GIS 輸出，為當時的專案技術方案。

代表研究與專案

MCIENet | 碩士論文，2024

- 以 DNA 序列預測 3D chromatin interactions / CTCF loops，設計多尺度 CNN / Inception 架構，檢查不同 receptive fields 可提供的訊號。
- 獨立主導研究問題拆解、模型與資料流程整合、實驗設計與執行、消融、跨資料集評估、XAI、論文與可重現 repository 工程化。
- 透過 BaseCNN 比較、不同序列長度及消融設定檢查模型設計；保留 DeepLIFT、ChINN / DNABERT2 比較入口與共用切分配置，讓比較方法可追溯。評估結果以已完成實驗為準，比較入口及共用設定另留作後續重現。
- 將資料集、位置與解析度差異列為泛化檢查條件，保留跨細胞類型泛化限制；部分正負樣本前處理改寫自 ChINN，與本人模型、實驗及工程貢獻分開歸屬。

金融文本情緒分析 | TCSE 研究，2020

- 研究股票文章的情緒分析，結合 Hierarchical Attention Network 與情緒詞典；前期獲科技部大專學生研究計畫核定，後續發表於 TCSE。
- 本人完成此研究的全部研究工作；論文為共同發表，不以研究工作範圍宣稱唯一作者。
- 初期反覆調整模型但效果有限，逐步補上資料與參數比較，將比較條件與實驗結果留作判斷依據。

動態嵌入客流估計 | IIAI AAI 研究，2021

- 探討以動態嵌入處理客流估計；本人負責方法發想、資料處理、模型建構與實驗執行。
- 透過資料與參數比較檢查模型選擇，研究獲 IIAI AAI Competitive Paper Award；共同發表與本人研究工作分開歸屬。

逐步建立的研究實驗方法 | 2020-2024

- 從早期反覆調模型但改善有限，逐漸轉向資料與參數比較；到 MCIENet 碩論，進一步建立較完整的消融設計及參數調整邏輯。
- 在 GPU 資源有限且需執行多組實驗時，以 Shell 安排並行；曾在同一 GPU 同時跑兩組不同參數，分開保存 logs 與 checkpoints。
- 先將前處理結果存為 H5，搭配 early stopping；調參時先選三個點觀察參數曲線，再決定後續調整策略。
- 上述方法在 2020 情緒研究、2021 客流、PropGMAN 與 MCIENet 中逐步累積，各案使用程度不同；不主張每案皆採相同完整流程，也沒有量測整體提速倍數。

動態交通網路表示 | KBS 合作研究，2024

- 共同發表以動態交通網路表示改善客流預測的研究；本人實際工作偏向程式碼與技術支援的輔助角色。

Horus Human Identification System | 大學畢業專題，2020

- 團隊整合 multi-object tracking、person ReID 與 face recognition，建立多鏡頭人物追蹤及辨識系統。
- 本人主要負責 ReID 特徵擷取整合、資料夾取樣 / 匹配、餘弦相似度、雙門檻判斷及追蹤紀錄佇列流程。
- ReID 使用第三方 FastReID 基礎模型，與團隊的 MOT / 臉部辨識模組整合；團隊獲東吳大學畢業專題第一名。

PropGMAN Asset Activation | 團隊專案，2022

- 三人團隊以實價登錄資料進行土地區域指標的時空建模，預測區域土地交易價格與交易量，支援資產活化分析。
- 本人發想專案方向、帶領團隊分工，負責模型實作、實驗執行及決賽報告；保留資料前處理、模型設定、實驗結果與重現方式。基礎 GMAN 方法採第三方既有研究實作，團隊獲資產活化 AI 應用創新競賽創意獎。

其他獨立專案

動態公車路線與班次規劃 | 專案期間待補

- 整合花東觀光人流、Google Trends、Popular Times、景點座標與行車距離，建立需求分攤及距離矩陣，以 OR-Tools TSP 探索路線，並以 Folium、Gantt 與 Sankey 呈現路線、班次與人流方向。
- 此為探索性獨立專案，與中研院社會調查路線規劃分開；專案期間待確認，不宣稱交通營運採用或實際營運成效。

其他團隊專案

- **Distributed Spark Crawler / Recovery，2022：** 政大分散式系統課程團隊串接資料擷取、缺漏偵測、Spark 預測補值、資料庫與 dashboard，展示資料擷取失敗後的補救流程；本人負責 Spark cluster、GitHub 與 demo 影片，參與系統整合。
- **金融資訊推薦與自動推播（E.SUN Fugle），2020：** 本人發想題目方向並帶領團隊分工，主要負責 Telegram chatbot 架構、實作與部署，將夥伴模組整合為統一入口。團隊以 item-based、user-based 及共現關係設計股票推薦 / 通知方案；本人提供推薦算法的想法與建議，不將所有算法歸為個人完成。團隊獲金融程式設計課程競賽優勝。
- **雙北防災資源分析系統，2025：** 負責題目發想與資料蒐集，與團隊探討旅遊景點周邊防災資源配置，作品晉級雙北城市儀表板黑客松決賽。公開儀表板以朋友提供的範本 fork 為基礎，保留團隊與範本來源歸屬。
- **託你辦 | 授權協作申辦與預約平台（2026）：** 主責提案題目、應用場景與 SPEC，並擔任決賽報告者；團隊入圍城市通微服務大黑客松決賽。 規格涵蓋授權範圍、操作流程、資料權限、API 與驗收條件；決賽展示口徑為單次案件、本人確認後模擬送出。

研究與發表

- Dynamic traffic network representation model for improving the prediction performance of passenger flow. Knowledge-Based Systems, 2024；本人主要提供程式碼與技術支援，屬輔助研究角色。
- A Dynamic Embedding Method for Passenger Flow Estimation. IIAI AAI, 2021；本人負責方法發想、資料處理、模型建構與實驗。Competitive Paper Award。
- Hierarchical Attention Network with Sentiment Lexicons for Stock Article Emotion Analysis. TCSE, 2020；本人完成全部研究工作，論文為共同發表。研究前期獲科技部大專學生研究計畫核定。
- MCIENet: Multi-scale CNN-based Information Extraction from DNA Sequences for 3D Chromatin Interactions Prediction. 碩士論文，2024。

學歷

- 國立政治大學，資訊科學系碩士 | 2021.09-2024.09。
- 東吳大學，巨量資料管理學系學士 | 2017.09-2021.07。

獎項

- IIAI AAI Competitive Paper Award，2021。
- 資產活化 AI 應用創新競賽創意獎，2022。
- 東吳大學畢業專題第一名，Horus，2020。
- 金融程式設計課程競賽優勝，2020。
- 雙北城市儀表板大黑客松決賽晉級，2025。