

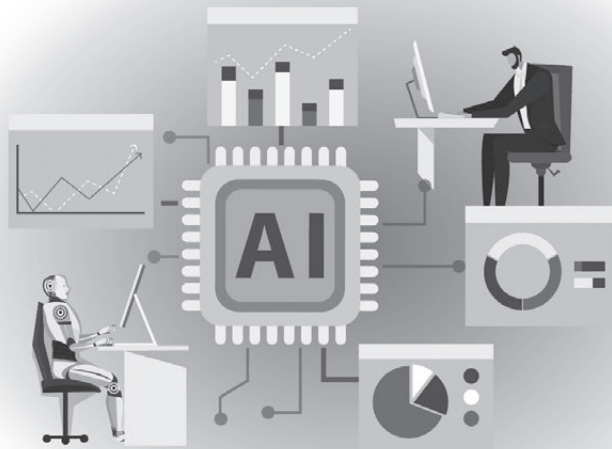
AI 工具確效策略 – AI 品質

AI Tool Validation Strategy - AI Quality

蕭君航 Andrew Hsiao / 作者現職於艾斯伯特資訊有限公司

序言

AI 已經來到第五階段，『生成式 AI / 大型語言模型』AI 大爆炸的時代，硬體算力的大型突破，訓練速度大幅提升，已經能夠理解自然語言，可以自行產生新內容並進行推理，突破各種知識領域的限制，讓 USER 能從茫然無知跨越到通材的程度，知識的啟發讓 USER 能夠快速進入所需的狀況，同時 AI 已經突破單純的系統程式，上升到助理的程度，開始能夠拆解人類的工作，維護錯誤內容、跑流程，蒐集資料，最近串上 AI Agent 已經能夠開始代理部分工作。



改程式碼。

(2) Lovable：可生成具備前後端跟完整數據庫的完整網頁。

(3) Google AI Studio：開發者能直接調適參數，測試 Prompt。

4. AI 工作流：透過 AI 瀏覽器與 AI Agent 檢視畫面、執行操作

(1) manus AI：AI 代理人能依模糊指令進行自主思考，跨工具規劃並執行複雜的任務。

(2) Comet：AI 瀏覽器將結合搜尋並行動，可以執行購物跟訂位之類的操作。

(3) Wispr Flow：可以將即時口說轉化為排版正確、語法流暢的文字。

四大領域 AI 工具清單

1. 通用型 AI：

(1) ChatGPT：通用型 AI

(2) Google Gemini：整合 Google 生態系。

(3) Claude：擅長長文寫作處理複雜分析。

(4) NotebookLM：知識管理神器、根據有限資訊進行精準對話；會議記錄最佳夥伴。

2. 創意、影音生成

(1) Gamma：自動生成有設計感的簡報或網頁。

(2) Lovart：可以用自然語言輕鬆完成廣告跟海報設計。

(3) Suno：只要給她歌詞、曲風就能完成音樂創作。

(4) ElevenLabs：音訊生成工具，可以完成多語、多聲道配音。

3. 程式開發與技術 Vibe Coding

(1) Cursor：深度理解整個專案邏輯、自動生成修

AI 品質

企業在如此的 AI 大爆炸的衝擊下，該何去何從？

最基本的就是導入 AI Agent，但該如何確認成效呢？

企業高階管理層目前最有感的應該就是螢幕滿片都是導入AI的廣告，似乎沒導入AI就是跟不上潮流，似乎就會被同業拋得遠遠的或者可能企業組織本身的產品／服務就會被其他行業定義為附屬品，不再存在價值，這是一件令人恐懼的事情，但是導入AI就能解決問題嗎？是否有考量生成式AI的三大問題：AI的幻覺：自信地亂說、成為只是聊天抒發感情的工具以及在沒有流程整合、資料治理不足的情況下，AI的回答能追溯責任來源，符合公司制度和流程？

- 企業／組織該如何因應？
- 導入AI Agent該如何確效，該如何確定AI回覆的品質？
- 如何得到能理解情境場景、持續學習，能夠協助決策的智能夥伴？

『企業組織需要轉型，讓AI Agent落地且確效，需要確保AI品質。』

- AI品質，強化自身價值
- AI工具已成為43.3%企業的標配
- 具備自主性的AI Agent實際落地15.9%

『AI可推倒公司各部門間無形的藩籬，打破與客戶隔閡的壁壘，藉由讓數字開口說話，將隱性需求和資料轉化為顯性實際需求資訊，創造有形價值以及無形的競爭力，但需要一定基礎的建置和維護，才能確保AI品質。』

在現今資訊發達的時代，各行各業對於資料和資訊e化已經是一種趨勢，是一種常態，即使一般民衆在日常生活也與之息息相關：手機APP上的支付系統、網際網路上的網站、大眾捷運的悠遊卡、上班族的打卡出勤系統、訂單管理系統、工廠MES現場管理系統…銀行金融體系的流程梳理…等記錄了各式各樣的資料，這些資料經由串聯介接到各個伺服器儲存、存放構成大數據的一部分。若單純串接到AI模型就能得到昇華？該如何確保單

個大型語言模型(LLM)，甚至多個語言模型能夠協作企業？

『未經處理的資料只是資料，記錄該資料曾經的過往，但不存在產值、判斷依據或趨勢，所以企業組織使用AI需要確保「AI品質」。

「AI品質」目標依據執行順序先後和時間段，區分為：短期、中期以及長期目標。

■ **短期目標：統一標準、提升效率、止血停止內卷、宣告目標，降低公司不再因為資料而內耗；EX：各測試站點品質不良標示。**

>>> 3~6個月，獲益／良率／正向影響(年)：

1. 降低每週產銷會議討論時間30min，每年節省130千元
 $(4000 \text{ 元/日} \div 8 \times 0.5 \text{ 時}) \times 10 \text{ 人} \times 52 \text{ (週)} = 130 \text{ 千元}$
2. 降低品管檢驗員每天不良判定時間60min，每年節省26千元
 $(2000 \text{ 元/日} \div 8 \times 1 \text{ 時}) \times 2 \text{ 人} \times 52 \text{ (週)} = 26 \text{ 千元}$
3. 綜合上述，每年共節省130千元+26千元=156千元
4. 實際生產成本計算：
 成本準確度從84%→92%
5. 報價基礎資訊更新：
 報價精準度從76%→89%
6. 接單策略調整：
 機台閒置率從28%→17%

■ **中期目標：提升管理經營品質，提高資料品質準確度、預留未來擴充性、降低錯誤判斷風險**

>>> 6~18個月，獲益／良率／正向影響(年)：

1. 批量數值FTR標準化，準確度100%
2. 降低管理上造成的Fail Log產生
3. 針對SF擋下LOG另存資料夾，確保異工產生原因，降低發生機率
4. 新世代產品WIFI產品EVM調整完成，降低約4%不良

5. 加快新產品開發速度約2週
6. 降低合約大批量採購單物料交付灰色地帶，確保各批准時交付滿足生產。

7. 平衡各供應商準時交付率基準

8. 提升整體供應鏈強度，降低供應商異常交付率

■ 長期目標：決策數據優化、品質風險降低、接單策略調控、供應商能力提高、產能配置

藉由以上資料的正確性，能讓公司開始因為資料而變強。做出更正確的判斷，AI Agent才能學習到正確資料並給予準確有效的答覆。

主要面向有以下：

1. 文件撰寫：會議記錄、簽到表、客戶提問、白皮書撰寫、整理簡報內容
2. 需求分析：商業流程分析、開發需求分析、系統需求分析
3. 測試案例：系統自動模擬測試
4. 程式輔助：重複性函式、語法套用，元件使用以及套件配套，加快開發速度
5. 內部溝通：定時產出報告、同步資訊到Notion/Sheets
6. 風險盤點：季度比較、月份比較、週期比較、凸波點提醒、8D管控
7. 行政支援：開發票、寄報價單、雲端檔案整理
8. 客服回覆：可開放權限回覆自動寄出提醒信件，必要時轉人工客服
9. 知識管理：整理筆記、彙整學習資源與教材
10. 內部教育：降低新手培訓時間，快速提供正確知識
11. 表單處理：摘要內容並分發給相關負責人
12. LT縮短：生產週期縮短、簽核流程縮短等原因分析
13. 產線成效：CT提升，產量提升
14. 資訊蒐集：爬取網站資料、追蹤競品或即時股價

15.數據整合：串接API，分析行銷成效、製作報表

16.AI模型效益：藉由AI模型針對輸入數據給予最適合的建議、假說假想判定

什麼是AI品質？

AI品質在運用上，是基於大數據內的數據進一步發掘其規律和運作模式，比較各資料個體彼此間關係，比較資料群體和資料群體間關係，比較同一類型資料群體在不同時間點上的關係，比較不同類型資料群體在同時間上的關係，更進一步比較不同系統間資料個體上的關係，以淬鍊具備價值的資訊並應用於企業發展、商業策略、組織結構升級以及加值客戶服務。

AI品質並不局限於以業務端為主的AI品質，而在於能夠開源節流、提升工作效能和產出、拉高品質水準、指出隱藏在作業中的TONE點作為改善切入點的作法皆泛稱為【AI品質】，但並不是將數據資料以圖表、顏色顯示就能夠稱為AI品質，而是以時間為X軸並承接上述指標於Y軸，作為項目衡量，方可稱作AI品質；後續將於本文依照過往工作實例以及認知，歸納並分析AI品質可行步驟，區分短期、中期和長期三種階段，或許部分有些天馬行空，希望供各位讀者參考並期勉能得到些許反饋，作為精進自己的動力。

為什麼一定要做AI品質？

- (1) 企業 / 組織已累積大量系統與資料，但資料未被有效整合，無法支持決策
- (2) 多數管理仍仰賴經驗與感覺，沒有具體實證，跨部門爭議成本高、決策速度慢
- (3) 市場競爭已由「成本與規模」轉為「反應速度與數據決策能力」

AI品質能帶來甚麼效益？

- (1) 建立事實基礎決策(Fact-based decision)，以數據說話。
- (2) 預測、優化跨部門共識，不內捲不拋球，共同

解決問題

- (3)管理決策「有依據、可追溯」，同時間點資訊區分權責
- (4)減少跨部門內耗與無效會議
- (5)為未來 AI、智慧製造和智慧流程預留基礎
- (6)不一次性大投資，分批投入成果可期
- (7)以既有系統為基礎逐步優化，不需要大批投入金額
- (8)每階段皆有明確成果指標(KPI)

AI 品質『短期目標』

《基礎資料架構建立和重整》

一筆完整的資料，具備創建時間(更新時間)、來源索引、數值以及資料標籤；其中最為重要的為資料標識的建立和重整。資料標識代表此筆資料的型態，諸如：姓名、性別、年齡、產品編號、機台名稱、機台型號、起始時間、結束時間、狀態、作業員、訂單/工單號、花費時間等等都屬於資料標識的範疇。一筆有效的資料，需要有效且統一的資料標識：資料標識需要考量現存於系統內的標識是否能滿足現今和未來之需求，且考量資料來源與介接系統的一致性和通用性，統合各部門和各系統間命名方式並將其宣告，此為 AI 品質『短期目標』：基礎的資料架構建立和重整。

提供以下案例供參考：

案例 A1

品管系統-不良品管理統之不良名稱定義：

- 目的：
 - 1.即時確認不良狀況以維持品質水準
 - 2.產品維護修繕
 - 3.生產和品管策略調整
 - 4.供後續研發單位使用
- 資料來源：品檢員檢驗後，由組長將不良品名

稱和數量輸入品質系統

- 狀況：不良品名詞定義不一
- 改善對策：重新定義統一不良品名詞
- 投入成本： $2000 \text{ 元/日} * 1 \text{ 人} * 90 \text{ 天} = 180 \text{ 千元}$

獲益 / 良率 / 正向影響(年)：

- 1.降低每週產銷會議討論時間 30min，每年節省 130 千元
 $(4000 \text{ 元/日} \div 8 * 0.5 \text{ 時}) * 10 \text{ 人} * 52 \text{ (週)} = 130 \text{ 千元}$
- 2.降低品管檢驗員每天不良判定時間 60min，每年節省 26 千元
 $(2000 \text{ 元/日} \div 8 * 1 \text{ 時}) * 2 \text{ 人} * 52 \text{ (週)} = 26 \text{ 千元}$
- 3.綜合上述，共節省 130 千元+26 千元

此案例為 3C 電子產品品檢站；品管系統資料來源為品檢員檢驗，經由組長確認後，於當日下午前填入品管不良管理系統。於此例中，主要問題在於各個檢驗員對於同樣不良狀況具有不同的名詞說法，其問題主要在於個人經歷、文化或者認知不同導致，為使系統能夠正常運作並建立長久追蹤管理機制，重新定義《不良品處理規範作業》以及不良品名詞及定義並納入文件管控。

成品檢驗(FQC)流程為流水線生產完成後，將未檢驗之成品裝箱堆放置待檢驗區，由品檢員使用電源供應器為成品上電開機，依照產品特性以及定義的檢驗項目，調整測試環境並進行環境驗證後，依序為各個成品檢驗測試項目、機構功能以及目視確認外觀；若檢驗確認不良則使用小標籤手寫不良現象貼附於不良處上並放入不良品區域，待現場維修工程師回收處理。

主要問題在於各個品檢員對於不良狀態認知有所差異、測試手法上存在機台設置單位不一的狀態；故以上述兩點作為執行要點。系統面上，建立品管訓練手冊、整理歷年內不良紀錄並由組長整合並與各個品檢員共同討論，重新定義各個不良名詞

和定義，並定義每年更新一次，同時請生產現場工程師一同確認並進行告知宣告。流程面上，則由組長在當日蒐集不良品時，同時確認不良標籤是否出現名詞定義不同的狀況，若有則與各個品檢員溝通並進行確認；若有出現需要調整的部分，除了當下通知生產現場外，則在每個季度品質會議提出並記錄在案。

案例 A2

MES 系統報工狀態：

- 目的：

1. 實際生產成本計算
2. 報價基礎資訊更新
3. 接單策略調整

- 資料來源：作業員於生產各個階段手動輸入工單狀態進入 MES 報工系統

- 狀況：報工狀態定義不同

- 改善對策：重新定義、宣導和訓練並以報工數據每週 double check

- 投入成本： $(2000 \text{ 元} / \text{日} \div 8 \times 0.5 \text{ 時}) \times 2 \text{ 人} \times 52 \text{ (週)} = 13 \text{ 千元}$

獲益 / 良率 / 正向影響：

1. 實際生產成本計算：

成本準確度從 84% → 92%

2. 報價基礎資訊更新：

報價精準度從 76% → 89%

3. 接單策略調整：

機台閒置率從 28% → 17%

此案例為機械加工業；MES 報工系統資料來源為作業員於生產各階段由員工手動輸入 MES 報工系統，藉各個不同狀態按鈕組成生產時間流，狀態按鈕如：開工、生產、完工、停工、復工以及結案等等。於此案例中，主要問題在於各個作業員在操作上對各個狀態按鈕的定義不同；雖然在觀察整

體報工時間流時並無多大問題，但向下一階細看到換線架模時間和作業時間，就出現了不同作業員之間的差異；部分員工出現無換線時間但生產時間較長的狀態。

換線架模時間屬於非正式生產時間，換線時間可能因應工藝、生產機台型號、產品、產品族群而有所不同，其屬於生產前置作業的一個環節；換線架模時間並無實際產出，而是在每一次投產前需要花費時間前置且沒有產出但不得不做的前置作業；對於機械加工業而言，重達上百公斤的模治具，每一次的換線架模都是一種負擔，也是一種成本的支出。

於此案例中為了能夠清楚知道換線架模的成本支出，做為現場改善依據、成本和報價衡量以及接單的考量，以 MES 報工狀態按鈕作為資料標識並重新定義報工方式，確保報工方式和狀態的一致性，另外於 MES 系統外掛增加偵測機制，以每週 MES 報工資料庫分析後核對，再由現場主管於每週週會宣導鞭策。

★短期目標小結

藉由以上案例，3~6 個月的調整各案例有明顯調整

- **案例 A1**：重新定義《不良品處理規範作業》以及不良品名詞及定義並納入文件管控，降低個人主觀意見，以統一且第三方客觀立場定立標準。後續可發展不良品原因分析以及改善對策，作為 AI Agent 可抓取資料。

- **案例 A2**：確保報工方式和狀態的一致性，另外於 MES 系統外掛增加偵測機制，以每週 MES 報工資料庫分析後核對，再由現場主管於每週週會宣導鞭策。

後續可發展生產成本準確度，進一步藉由實際生產成本確保報價精準度，同時可以推估閒置機台率，讓業務單位可以接回閒置機台的訂單，

降低機台閒置率，提高機台稼動率。同時，作為 AI Agent 可抓取資料。

AI 品質短期目標可達到統一標準、提升效率、止血停止內卷、宣告目標，讓公司不再因為資料不一而花費巨量時間進行內耗。

- 對待傳產製造業，可以提升機台整體效率 (OEE)，降低異常工時，提升品質良率。
- 對待科技業，可以提升 FTR (Firstratio)，提升直通率；強化產線生產效率。
- 對待電商，可以提升出貨效率，確保出貨流程正確；進一步分析漏斗分析，提取出。
- 對待軟體資訊服務商，可以提升專案期程，確保里程碑；確保「交付」「驗收」順暢進行

AI 品質『中期目標』

《資料品質優化》&《資料架構優化》

AI 品質『中期目標』為《資料品質優化》和《資料架構優化》，此階段目的為建立涵蓋企業願景的大數據資料庫；資料品質優化除了資料是否有效，更加注重的是蒐集到資料的代表性和涵蓋範圍，畢竟優良的資料才能轉化為有效的資訊，才能轉化出對企業公司組織有效的決策參考數據。資料架構優化則是將目光放長遠，剔除多餘的欄位和無意義的資料，同時為未來資料擴充留下欄位。

《資料品質優化》承接 AI 品質短期目標，針對蒐集得到的資料進行衡量，衡量資料的品質和優缺點、是否滿足企業願景並針對缺失的部分制定改善方針；對於資料品質未達到要求、出現資料健康度不良的情況，則需要訂立改善目標和對策，以改善資料品質、避免出現資料缺口進而導致偏差結果的出現或者重新考量資料來源是否需要進行調優。

《資料架構優化》為依照 AI 品質目標和範圍進行策略性的調整，確保核心價值是否出現偏移，其中需衡量所蒐集資料是否擁有穩定來源，能夠持續

取得資料以及資料健康度和可用價值；若在中期發現當下資料架構不能滿足的情狀下，則需進行新增、修改以及刪除的動作。

案例 B1

某廠電測數值品質衡量系統：

• 目的：

1. 確認產品良品關鍵能力品質水準
2. 排除異常 Log、降低系統運算耗損

• 資料來源：電測站測試 LOG

• 狀況：

1. 不良品不能開機導致 Log 關鍵數據缺少
2. 作業員未依序過站，導致測試 Log 暴增

• 改善對策：

1. 調整測試流程架構，先刷入 BARCODE 並建立 LOG，確保分母正確
2. 刪除 SF 擋下不能測試的 Log，降低防呆措施的 LOG 影響分子

• 投入成本：4000 元/日 * 1 人 * 2 天 = 8 千元

獲益 / 良率 / 正向影響：

1. 批量數值 FTR 標準化，準確度 100%
2. 降低管理上造成的 FAIL LOG 產生
3. 針對 SF 擋下 LOG 另存資料夾，確保異工產生原因，降低發生機率

一般電子產品流水測試線會經過許多測試站點，各個產品都需要經過每一站點的測試，遵循一定的順序和節拍，完成所有測試站點的測試方可算良品計入產出，作為可出貨數量；各測試站點都有技術團隊負責維護各測試站點的 Test Program，以測試站點而言需要記錄每一筆測試紀錄，包含架站試機測試 Log、正常測試 Pass Log、測試 Fail Log 以及非正常 Log。

以數據分析角度而言，若要分析同一世代產品族群其下三個產品的關鍵能力高低優劣，則只

需要分析正常測試 Pass Log，故在篩選資料時會剔除測試 Fail Log 和非正常 Log；在需要比較測試良率、直通率時，才會納入測試 Fail Log 和非正常 Log。

故在本案例中，在分析同一世代產品族群三個產品關鍵能力，在能確保資料有效性，避免被不良品錯誤數值影響前提下，剔除測試 FailLog 和非正常 Log。

案例 B2

電子產品測試站點電測數值系統：

• 目的：

1. 建立全產品品質監控系統
2. 各世代產品測試項目數據資料庫
3. 比較同產品相關測試站點數據以改善品質和降低非必要作業

• **資料來源：**各測試站點測試項目以及特殊產品額外測試項目

• **狀況：**各個測試站點格式不一無法於同一頁面陣列展開

• **改善對策：**以全流程測試站十一種站點展開取最大共通格式

• **投入成本：** $(4000 \text{ 元/日} * 1 \text{ 人} * 6 \text{ 天} = 24 \text{ 千元}) + (4000 \text{ 元/日} * 2 \text{ 人} * 30 \text{ 天} = 240 \text{ 千元}) = 264 \text{ 千元}$

獲益 / 良率 / 正向影響：

1. 新世代產品 WIFI 產品 EVM 調整完成，降低約 4% 不良
2. 加快新產品開發速度約 2 週

此案例較為複雜，原因為全流程 11 種測試站並非為一定機台數目和測試點數。其測試站點的機台數和測試點數，乃參照各測試站架構和能力、Cycle Time 進行規劃各站點機台數目和測試點數。

※ 各測試站點下有大於等於一個機台

※ 各測試機台下有大於等於一個可同時測試產品數目

舉例來說某產品 X 需要經過測試站點 A、測試站點 B 以及測試站點 C，其測試站點測試時間和可同時測試產品數目如下：

測試站點 A 測試時間為 16 秒，A 測試站點擁有 2 個 PORT 可同時測試產品數目為 2 個

測試站點 B 測試時間為 24 秒，B 測試站點擁有 1 個 PORT 可同時測試產品數目為 1 個

測試站點 C 測試時間為 12 秒，C 測試站點擁有 3 個 PORT 可同時測試產品數目為 3 個

以流水生產線線平衡的觀點來看各測試站點需架設機台數量

- 測試站點 A 架設 4 站
- 測試站點 B 需架設 12 站
- 測試站點 C 需架設 2 站

呈上敘述，測試站點 B 可同時測試數目為 1 個產品，故在過往架設測試站點編寫 Test Program 時，Log 並無定義可同時測試產品數量；為建立此電子產品測試站點電測數值系統，新增可同時測試產品數量於 Test Program Log 內，供系統撈取，確保全測試站點資料架構一致性。

案例 B3

某電子零件廠採購系統採購單 & 倉儲系統入庫單

• 目的：

1. 完善物料資訊流架構
2. 物料到貨準確率評估
3. 供應商能力評鑑

• 資料來源：

1. 請採購系統-採購單
2. 倉儲系統-入庫單

• **狀況：**入庫單存在實際入庫日，但不包含採購單（分批）預計到貨日以及供應商欄位

• **改善對策：**以入庫單庫存批增加採購單(分批)預計到貨日

• **投入成本：** $4000 \text{ 元/日} * 3 \text{ 人} * 20 \text{ 天} = 240 \text{ 千元}$

獲益 / 良率 / 正向影響：

採購單準時交付率 - $[\text{準時交付(分批)採購單數量} \div \text{總(分批)採購單數量}] \times 100\%$ 。

1. 降低合約大批量採購單物料交付灰色地帶，確保各批准時交付滿足生產。
2. 平衡各供應商準時交付率基準
3. 提升整體供應鏈強度，降低供應商異常交付率

近來因宅經濟和5G興起，電子產品市場需求持續擴大，又因疫情緣故造成供應商產能起伏不定、貨運也經常出現運輸不及和卡海關的狀況，為避免物料未到導致生產順延造成延遲出貨以及降低企業組織對於原物料庫存的積壓，故對於物料是否能準時抵達產線這件事情，物料實際到料日就越發重要了。本案例以結合請採購系統採購單和倉儲系統入庫單作為衡量指標依據。

採購單歸屬於請採購系統，採購單資料源頭來自需求單位開立請購單，經採購人員詢價確認供應商、數量和交期，主管簽核完成後轉為採購單；採購單由請購單轉化而來，可能合併多張請購單或者多筆項次請購單，屬於多對一或者多對多的資料架構。採購單欄位主要有以下：採購單號、項次、料號、品名、規格、供應商以及預計到貨日；完成簽核的採購單已經具備相當多可進行衡量的資料。

入庫單則隸屬於庫存管理系統。入庫單資料源頭來自物料到貨時，由物管人員接收交貨單位的貨物，經查詢採購系統預到貨資料核對、檢驗、判定後填寫完成；入庫單在原有資料狀況對應上，存在多對一的情況，也就是多張採購單項次物料同一時間段抵達倉庫；為了後續分析將入庫單以各資料

階層展開，若同一項次存在不同供應商的資訊，則添加供應商ID作為展開基礎；同一供應商不同採購單的資訊則再向下展開。確保各筆入庫單資訊能夠追尋到正確的採購單，以利後續分析；這部分由於會與過往習慣導致需要磨合的時間，故在系統上添加防呆機制，避免錯誤資訊產生。

★中期目標小結

藉由以上案例，6~18個月的調整各案例有明顯調整

- **案例 B1：**重新定義測試流程架構並擋下錯誤 LOG，降低防呆措施影響整體 LOG 分析。
後續以新流程進行流通產品，提升產品生產流暢度；另一方面藉由分析 Fail Log 分析會議，降低現場 TE 管理負荷。
- **案例 B2：**整體測試流程表格擴建較為繁瑣，需要將近一個月的時間投入並考量後續產品未來發展趨勢和測試項目，但優點是顯而易見，各產品線能夠橫向比較，啟發優化方向。
後續於新產品開發時，MAT 產品測試部門藉由 AI 模型分析，降低約 4% 不良。
- **案例 B3：**增加欄位簡單，但分析數據較為複雜，部分供應商採取年度合約採購量，一次性採購年需求量，但分批交期卻未準時準量，藉由此變更，能夠提供準確的資料，提升管理效益。
後續可藉由準時交付率達成狀況與供應商討論協商，確保物料準時準量，提高整體供應鏈強度。

AI 品質『長期目標』

《決策與洞察：讓數據開口說話，讓管理不是感覺而是實槌》

AI 品質第三階段需要將隱性資料轉為顯性資訊，也就是讓數據開口說話，揪出真因和趨勢預估，建立完整的事實真相，供判斷和決策。

AI 品質第三階段承接上述短期目標《基礎資料架構建立和重整》和中期目標《資料品質優化》及《資料架構優化》建立的大數據資料庫，基本上已經可以將各個系統內的『資料』轉為有效的『資訊』；筆者在執行上述專案時，經常於訪談過程中遇到各部門因立場不同而有迥異的觀點和意見，更有甚者互相推卸指責、千奇百怪的理由紛紛出籠；一場場偏移主題無意義無結論的會議不斷消耗企業組織及涉及部門其下成員的資源，涉及範圍和方向持續發散，但對於癥結處卻無相對應的解決之道。

AI 品質則藉由分析統計手法比較各資料、資料群、資料類別，在時間點上的關係，收斂各個原因，用基礎理論正確的思考概念，結合靈感不同的角度，淬鍊出正確的指引並提供跨部門且立場不偏頗的真實觀點。

把複雜問題→可控流程→可交付成果

AI 品質在企業的價值，主要本質在於提升文件的撰寫速度、降低重複性工作的繁瑣、減少程式改 A 錯 B、加快開發速度、快速進行風險匯總、24hr 即時回覆，包含內部及外部（客戶）提問。

當然對於商業流程而言，USER 提問時，AI Agent 需要能依據可公開數據，拆解問題，對應不同流程進行檢索、分析對應企業面向、調整成適合語句、重組成自然語言和回應方案、控制輸出給用戶。讓 AI 不是取代人，而是人用 AI 管理並放大能力密度。

當然 AI Agent 輸出的內容不能盡信，則建議用以下進行評估

- Ragas：評估檢索的相關性 (Context Relevance)、答案的忠實度 (Faithfulness)。
- Trulens：評估 RAG 三元組 (Query、Context、Response)

案例 C1

某廠電測數值品質衡量系統：

- 目的：
 1. 確認產品良品關鍵能力品質水準 CPK
- 資料來源：電測站測試 LOG
- 狀況：
 1. 作業員架設站點，未依 SOP 架設
 2. Antenna 耗材耗損狀況不明
- 改善對策：
 1. 重新購買耗材，依據 SOP 架設測試站
 2. 建立 QFD 品質防禦系統
- 投入成本： $(4000 \text{ 元/日} * 1 \text{ 人} * 2 \text{ 天} = 8 \text{ 千元}) + (100 \text{ 千元 CABLE 線}) = 108 \text{ 千元}$
- 獲益 / 良率 / 正向影響：
 1. CPK=0.5 → 1.5
 2. QFD 依據測試數據變化，即時通知線場產線更換耗材
 3. 直通率：80% → 99%
 4. 提升產值： $40\text{USD} * 1\% * 400\text{K} (55,555 * 3 * 24) * (99\% - 80\%) = 760 \text{ 千元}$

案例 C2

某廠金屬加工衡量系統：

- 目的：
 1. 刀具損壞率、反修後再次加工狀態
- 資料來源：MES 加工資料
- 狀況：
 1. 耗材刀具耗損嚴重
 2. 品質良率變化不一
- 改善對策：
 1. 依據 AI 模型得到啓發，降低每次委外研磨後度
 2. 維持良品率

- 投入成本： $(1000 \text{ 元/日} * 1 \text{ 人} * 30 \text{ 天} = 30 \text{ 千元}) = 30 \text{ 千元}$

獲益 / 良率 / 正向影響：

1. 68% 使用次數提升
2. 年省 120 萬，降低 23.8%

案例 C3

某廠金屬加工衡量系統：

- 目的：
 1. 確認產品生產 CT
 2. 確認產品生產 LT
- 資料來源：MES 報工系統 LOG
- 狀況：
 1. 中低管理層對於 CT 不熟悉，不能反饋產線進步幅度
 2. 高階管理層對於 LT 不熟悉，不能反饋產線進步幅度
- 改善對策：
 1. 藉由 AI 工具，讀取 CT 進步幅度，確認是否進步幅度是否符合目標
 2. 藉由 AI 工具，讀取 LT 進步幅度，確認是否進步幅度是否符合目標
- 投入成本： $4000 \text{ 元/日} * 1 \text{ 人} * 2 \text{ 天} = 8 \text{ 千元}$

獲益 / 良率 / 正向影響：

1. CT 進步幅度，達目標 5%，提升比率為 6.57%，超過 5% 目標，達標
2. LT 縮短範疇，達目標 3 天，取得 9 天成績；超過 3 天目標，達標

★長期目標小結

藉由以上案例，6~18 個月的調整各案例有明顯調整

- 案例 C1：提升 WLAN 站架設 SOP，標準化作業流程
後續以 QFD 持續監測 WLAN test station，確保 CPK 符合 1.33 上。
- 案例 C2：以 AI 工具為啓發，降低委外加工鍍鈦需求，讓單一刀具能夠提升更多使用次數。
後續於降低金屬加工刀具，主要耗材的購買，年省 120 萬，降低 23.8%
- 案例 C3：具體列出產線生產工效，提高 CT 進步幅度，達目標 5%，提升比率為 6.57%，超過 5% 目標，達標。
具體列出產線工效，取得降低 9 天成績；超過 3 天目標，達標
後續可藉由準時交付率達成狀況與供應商討論協商，確保物料準時準量，提高整體供應鏈強度。

問題一：AI 品質推行十大問題？

1. 「這跟我們現在的 BI 有什麼不同？」
ANS：BI 是結果呈現，AI 品質是確保數據本身可信，能夠推演更多情境和場景。
2. 「如果資料本來就不準，做這個有用嗎？」
ANS：正因為不準，才需要先朝『短期目標』《基礎資料架構建立和重整》以及『中期目標』《資料品質優化》&《資料架構優化》進行，先做標準與品質優化；在『長期目標』才能避免 AI 幻覺給予錯誤答案的風險，讓 AI Agent 確效，讓企業 / 組織高層能放心使用。
3. 「一定要花這麼多時間嗎？」
ANS：不做，錯誤決策的成本更高，而且是長期隱性成本；故在本文中區分各期目標，短期目標和中期目標能夠更快得到結果，讓企業 / 組織能夠更知道是否需要調整策略或方向。
4. 「這是不是只要 IT 的事情？」
ANS：不是，IT 只是其中一環的工具，AI 品質是經營管理問題。
短期目標：以基層員工為主，依據工作內容提出；由低中級主管匯總
中期目標：以低中層主管提出數據和流程，由資訊管理人員進行匯總架構
長期目標：以中高層主管提出需求和面相，由高層主管定義裁決

主要面向有以下：

- (1) 文件撰寫：會議記錄、簽到表、客戶提問、白皮書撰寫、整理簡報內容
- (2) 需求分析：商業流程分析、開發需求分析、系統需求分析
- (3) 測試案例：系統自動模擬測試
- (4) 程式輔助：重複性函式、語法套用，元件使用以及套件配套，加快開發速度
- (5) 內部溝通：定時產出報告、同步資訊到 Notion/Sheets
- (6) 風險盤點：季度比較、月份比較、週期比較、凸波點提醒、8D 管控
- (7) 行政支援：開發票、寄報價單、雲端檔案整理
- (8) 客服回覆：可開放權限回覆自動寄出提醒信件，必要時轉人工客服
- (9) 知識管理：整理筆記、彙整學習資源與教材
- (10) 內部教育：降低新手培訓時間，快速提供正確知識
- (11) 表單處理：摘要內容並分發給相關負責人
- (12) LT 縮短：生產週期縮短、簽核流程縮短等原因分析
- (13) 產線成效：CT 提升，產量提升
- (14) 資訊蒐集：爬取網站資料、追蹤競品或即時股價
- (15) 數據整合：串接 API，分析行銷成效、製作報表
- (16) AI 模型效益：藉由 AI 模型針對輸入數據給予最適合的建議、假說假想判定

- 5.「現場會不會反彈？」
ANS：會，但透過標準化反而能減少現場爭議與重工，釐清問題的嚴重性、影響範圍和發生頻率；未來會更好進行溝通。
- 6.「能不能一年內看到成果？」
ANS：短期目標階段《基礎資料架構建立和重整》3-6 個月即可看到具體改善
- 7.「如果中途停掉，會不會浪費？」
ANS：每一階段都有獨立價值，不是一次性投資。
- 8.「我們規模適合做這件事嗎？」
ANS：規模越小，越不能承擔錯誤決策成本；規模越大，能夠橫向產生的價值更高。
- 9.「未來能不能接 AI？」
ANS：不能沒有這一層就談 AI，這是必要基礎。可以大幅避免 AI Agent 回答幻覺答案，讓人得到錯誤的輸出
- 10.「在這個專案中的主要推動角色職責是甚麼？」
ANS：PM 為主；橋接經營、現場與系統，確保數據能真正用於決策。

問題二：AI 工具發展概況

AI 工具發展歷史悠久，主要從 1950-1990，以『規則式 AI』為主，以人為定義流程並歸納為程式語言邏輯，以 if-then、if-else 等規則為主，由軟體工程師建立專家系統，專門為單一流程、固定場景服務。

- **系統價值**：自動化單一判斷、範圍內可控，判斷成功率高。正確輸入即可完成。
- **確效狀況**：流程規則內 100% 判定正確，不需要再次篩選過濾錯誤；但無法學習，只能依照流程規則進行，超出流程定義範圍，則失去效能；故需要每年進行維護，進行升級。

第二階段為 1990-2000，『標籤式 AI』，以人為標示各系統中的每個標籤，作為系統查詢歸納依據和標示定義各種詞語，如 TESLA 自駕系統，藉由 AI 學習標籤給予更好的駕駛體驗、維基百科則是這世紀最偉大的百科全書，集結了所有願意分享的智慧結晶，這讓知識更好的傳達，截至目前為止主要是文字和資訊的傳達。

- **系統價值**：標準化標識、人為知識匯集，作為 AI 的底層知識集結。
- **確效狀況**：仰賴人為維護、若有超脫於範圍外或衍生或新增的知識，若要人為再度維護。

第三階段為 2000-2010，『機器學習 ML』，承上一階段開始用資料訓練模型，擁有各種決策樹、回歸模型以及固定邏輯推演；不局限於文字和資訊，可圖像識別。

- **系統價值**：圖像識別、知識集結於系統，只要底層資訊正確、持續產生不在需要人為維護，系統可以識別並給予正確判斷。
- **確效狀況**：AI 開始有具體的判斷，可以從 CAMERA 辨別，人臉識別技術大躍進。

第四階段為 2010-2020，『深度學習 DL』，承上一階段除了資料、圖片、影音也加入模型訓練，擁有卷積網路；AI 可以揉捏各種輸入的資料來源進行綜合判斷。

- **系統價值**：圖像、影像、語音 / 語意識別，可以揉和各種資料來源，回饋如同人類感知的回答。
- **確效狀況**：AI 的判斷開始細膩化，USER 可以使用正確的自然語句提問得到回應。

第五階段為 2020-至今，『生成式 AI/大型語言模型』，硬體算力的大型突破，訓練速度大幅提升，已經能夠理解自然語言，可以自行產生新內容並進行推理，突破各種知識領域的限制，讓 USER 能從茫然無知跨越到通材的程度，知識的啟發讓 USER 能夠快速進入所需的狀況，同時間 AI 已經突破單純的系統程式，上升到助理的程度。

- **系統價值**：知識型的工作大幅提升效率，縮短成長時間。
- **確效狀況**：降低重複性工作、決策與產出提升 3~10 倍、降低人為遺漏和疏失、快速發展衍生新知識以及個人 AI Agent 普及化。

參考資料

[1] 韋惟珊，2026/1/11，蒐集統整，數位時代，reel 必備的四大領域 AI 工具清單

[2] INSIDE，2025 台灣 AI Agent 生態系大調查，P.6