

以數據驅動優化

中小企業融資環境計畫 | 徵件說明會

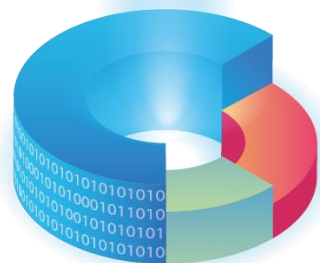
| 主辦單位 |



| 執行單位 |



解題構面與資料應用分析



講者：蔣居裕 計畫總顧問



目錄

壹 | 資料內容與應用方式

1. 管制資料來源
2. 各題目可能對應資料與欄位
3. 實作驗證環境說明

貳 | 提案重點與建議

1. 提案書撰寫建議
2. 常見缺漏與加分方向

1. 管制資料來源

以**政府投資方案資料庫**相關資料為基礎，經**去識別化與隱私強化處理**後提供使用，

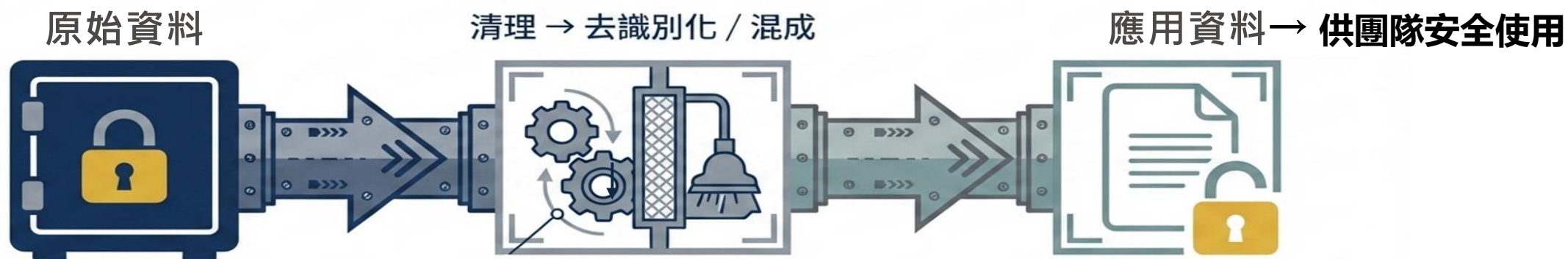
在兼顧**資料安全**與**合規**前提下，支援解題團隊進行模型訓練與驗證。

資料內容涵蓋**企業財務指標、投資損益、處分情形及營運狀態**等，

可協助掌握**投資案件整體輪廓、歷年經營變化與成長軌跡**，

作為**風險預警、投資評估及出場分析**等應用之基礎。

兼顧分析價值與機敏資料保護



由顧問事先處理，確保團隊最終

僅接觸轉換後之資料形式 → **保護機敏資訊**

2. 各題目可能對應資料與欄位(1/2)

題	出題構面	題目名稱	管制資料需求	主要外部資料
1	市場動態	產業資訊整合系統	✗ 不需	USPTO / TIPO / Crunchbase / arXiv
2	投資智慧	投資評估資料標準化與整合系統	✗ 不需	MOPS、商業司、LinkedIn、官網爬取
3	投資預警	★ 自動化動態營運監測與風險預警系統	✓	MOPS、月營收公告、產業景氣指標
4	出場預測	★ 產業出場策略案例庫建置與動態出場預測模型	✓	TWSE / TPEX 歷史、M&A 公告、國際資本市場估值
5	開放提案	開放式資料應用提案 (解題團隊自由發揮)	△ 任何整理後資料皆可作基底	自選

註：建議管制資料不做為解題唯一的私有資料來源

2. 各題目可能對應資料與欄位(2/2)

題目3 (自動化動態營運監測與風險預警)

營業收入淨額

營業淨利

稅後淨利

毛利率、稅前淨利率

流動比率、負債比率

每股盈餘 EPS

每股淨值

本季營收是否成長

是否提報投資損失

現金流量是否足以支撐半年

營運結果在股東權益的表現

grade 軌跡

題 4 (出場策略 + IPO 前哨指標)

處分情形

已處分損益

處分完畢日期

投資損益

預估總損益

上市櫃情形 (IPO 前哨指標基底)

產業別

企業現況

目前處分情況

出場前 3 年連續季報

歷次投資 / 處分進度 / 已完成處分 / 股票出售 / 投資損失 / 股利發放

題 5 (開放提案)

本案釋出的欄位資料皆可作為解題基底

比賽團隊可運用於：產業比較、出場路徑模擬、AI 投資助手、跨案聚類分析等

註：實際釋出欄位仍須待 6月正式資料完成反查測試與資料安全檢核後確認，並將於官網同步更新最終釋出欄位說明。

3.實作驗證環境說明

因本計畫應用金融領域資料涉及機敏性，為保障本次實證協作之數據於合規環境下應用，須請入選團隊於部分使用情境下，配合於安排之專屬環境下進行資料驗證與模型測試。

場域硬體設備

地端 AI 工作站

供資料查詢、運算與分析使用，具備常見開發工具，地端模型，如：Gemma 4 26B-A4B或相似模型

1. 顯示晶片(GPU)：GeForce RTX 5090 Laptop GPU 24GB GDDR6
2. 處理器(CPU)：AMD Ryzen9 9955HX (2.5GHz/Turbo 5.4GHz)
3. 記憶體(RAM)：64GB
4. 儲存空間(SSD)：2TB

地端 AI 實作工作站

供資料查詢、運算與分析使用，具備常見開發工具

1. 顯示晶片：獨立顯示晶片，具備至少 6GB GDDR6 專用顯存。
2. 處理器：搭載 Intel Core Ultra 7 155H (16核/22緒) 或同等級以上處理器。
3. 記憶體：內建 64GB DDR5-5600MHz (含) 以上，且須為雙通道配置。
4. 儲存空間 (SSD)：內建 2TB NVMe Gen4 固態硬碟。
5. 顯示螢幕：14 吋 (含) 以上。
6. 作業系統：預裝 Windows 11 Pro (專業版) 64-bit 繁體中文版。

是否須進入實體場域

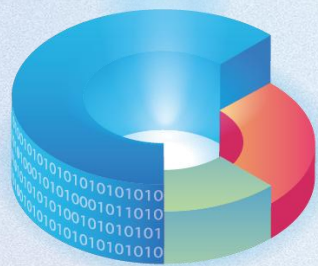
依據各團隊實際資料需求而定



若為下列需求者，須配合進場作業：須使用中企署內部合成資料或其他須在封閉環境下運行之資料。

若團隊僅使用自行準備資料，或以已釋出之非機敏欄位進行模型訓練與驗證，則可於自身場域或遠端環境完成，不須進入指定專屬場域。

註：實際協作設備規格將依最終提供內容為準，並同步公告於官網。



貳 | 提案重點與建議

1. 提案書撰寫建議(1/4)

三、欲解題構面：請勾選欲回應之解題構面，並填寫擬解決的問題及使用情境描述

✓題目一、產業資訊整合系統

請簡述擬解決的問題及使用情境：

我們欲解決投資機構於投前評估階段需耗費大量時間手動蒐集與整理分散產業資訊之問題，導致資訊不完整、更新不即時且分析效率低落。在使用情境上，投資人或分析師僅需輸入特定產業（如半導體、綠能）或目標企業名稱，系統即可自動整合並輸出該標的之完整產業分析報告，包含產業概況、供應鏈上下游位置、主要競爭對手、技術發展趨勢、國際對標企業與市場定位分析，作為投前初步篩選與評估依據。

在資料應用上，我們將整合公開產業報告、新聞資料、企業公開資訊、專利資料與市場研究數據，透過資料擷取（API/爬蟲）、資料清理與去重、自然語言處理（NLP）與語意分析技術，建立產業知識圖譜與結構化資料庫。在系統輸出上，將提供標準化產業分析報告與視覺化模組，包括產業鏈關係圖、競爭者定位矩陣、技術演進趨勢圖與市場分布分析，使使用者能快速掌握產業全貌並進行比較分析。

執行上採模組化系統架構，分為資料層、分析層與應用層逐步開發，並以可擴充資料管線設計支援不同產業。協作方面，將結合投資專業人員提供分析指標定義、產業顧問提供領域知識校正，以及技術團隊負責模型與系統開發，以持續優化分析準確度。

1. 提案書撰寫建議(2/4)

四、資料應用規劃：如何規劃、運用及整合資料資源來支撐解題策略	
(一) 欲使用之資料欄位	本案希望由主辦單位提供之資料欄位包含：企業基本資料（公司名稱、產業別、成立時間、地區）、財務摘要資料（營收、獲利、成長率、資本額）、投資與交易相關資訊（投資輪次、估值區間、投資時間點）、以及既有產業分類標籤。用途為作為產業分析模型之基準資料，用於標準化分類、產業比較分析與模型驗證，並作為知識圖譜之核心節點資料。
(二) 自備資料情形	本團隊預計使用之既有資料包含： <ol style="list-style-type: none">1. 產業研究報告（PDF/HTML，約數千份，涵蓋2018–2025年）2. 新聞資料（JSON/CSV，來源為公開新聞API與爬蟲資料，約數十萬筆）3. 企業公開資訊（HTML/PDF，如年報與公開說明書，涵蓋近5–10年）4. 專利資料（CSV/XML，來源為公開專利資料庫）5. 市場研究與產業分析資料（PDF/CSV，跨產業長期資料） 資料欄位類型主要包含：企業名稱、產業分類、時間標記、財務指標、技術關鍵字、事件描述與來源連結等。上述資料將用於建立產業知識圖譜、訓練語意分析模型與產業趨勢分析。
(三) 完成提案所需的其他資料來源及應用方式	本案亦將整合政府開放資料（如公司登記資料、產業統計資料）、國際資料平台（如世界銀行、OECD產業數據）、專利與技術資料庫，以及第三方市場研究平台資料。用途為補強產業外部環境資訊與國際比較分析，提升模型完整性與準確性。資料將透過API串接或定期爬取方式更新，並納入統一資料管線進行清理與標準化。

1. 提案書撰寫建議(3/4)

五、執行方式與協作安排	
(一) 欲導入之商品或服務是否已有運用實蹟	<p>✓否：本案為新建置之資料整合與分析系統，但團隊具備資料分析、NLP模型與系統開發相關專案經驗。</p>
(二) 協作期程規劃	<p>本案協作採短期驗證與快速迭代方式，於三個月內完成核心功能驗證與最小可行系統 (MVP) 建置：</p> <ul style="list-style-type: none"> ●第一階段 (第1個月)：需求定義與資料建置 完成使用情境與分析指標確認，並進行多源資料盤點、資料擷取流程設計 (API/爬蟲) 與基礎資料庫建置。 ●第二階段 (第2個月)：資料整合與分析模型建構 完成資料清理、去重與標準化處理，並導入NLP語意分析與基礎知識圖譜建構，建立產業關聯分析能力。 ●第三階段 (第3個月)：系統整合與成果驗證 完成報告生成模組與視覺化介面整合，進行測試驗證與使用者回饋調整，產出可展示之系統雛型與分析成果。
(三) 預計使用工具或平台	<p>資料處理工具：Python、Pandas 資料擷取：API 串接、Web Crawling工具 AI建模：NLP模型 (Transformers)、向量資料庫 (Vector DB)、知識圖譜工具 執行環境：雲端運算平台 (AWS/GCP/Azure擇一)、Docker容器化部署 是否使用指定場域：可配合主辦單位提供之地端或雲端實作環境進行部署與測試</p>
(四) 期待主辦單位提供之協助	<p>✓法規諮詢：協助釐清資料使用相關法規，確保應用合規。 ✓資訊安全：提供資安指引 (如去識別化、權限控管)，確保資料處理安全。 ✓資料應用：協助說明資料欄位與結構，利於資料串接與分析應用。 ✓技術開發：提供技術諮詢與輔導，協助優化模型與系統架構。</p>

1. 提案書撰寫建議(4/4)

六、對應評選指標之說明	
(一) 服務或技術創新性	本案整合多源異質資料 (產業報告、新聞、專利、企業資訊等) ，結合自然語言處理 (NLP) 、知識圖譜與語意分析技術，建立可自動生成產業分析報告之系統，突破傳統仰賴人工蒐集與判讀之模式。透過結構化資料與關聯分析機制，提供跨資料來源之整合洞察，並以視覺化方式呈現產業鏈與競爭態勢，具備高度創新性與應用價值。
(二) 提案內容完整性	本提案已明確說明欲解決之問題、使用情境、資料來源與應用方式，並具體規劃系統架構 (資料層、分析層、應用層) 及三階段執行期程，涵蓋資料建置、模型開發至系統整合與驗證流程。同時亦說明所需工具、技術方法與協作分工，整體內容具邏輯性與可操作性，能完整呈現解決方案之執行藍圖。
(三) 應用可行性	本案技術上採用成熟之NLP模型、資料處理工具與雲端架構，具備實作可行性。透過分階段開發與MVP驗證機制，可於短期內完成核心功能建置，並依實際使用回饋持續優化，具備良好落地與擴充潛力。
(四) 市場需求與潛力	隨著投資決策日益仰賴數據分析，投資機構對於快速掌握產業資訊與企業定位之需求持續提升。本案可有效降低資訊蒐集成本，提升分析效率與決策品質，具備明確市場需求。未來亦可延伸應用至創投、金融機構、顧問公司及政府單位，具備跨產業應用潛力與商業化發展空間。
(五) 社會影響力	本案透過提升投融資決策效率與透明度，有助於促進資金更有效配置至具潛力之中小企業，進一步支持產業創新發展。長期而言，將有助於優化整體投資生態系、降低資訊不對稱問題，並促進產業升級與經濟成長，具備正向社會影響力。

2. 常見缺漏與加分方向

常見缺漏

1. 問題、User Persona、價值主張定義不夠精準。
2. 提案方向過於通用，未凸顯解方特色與實際應用價值。
3. 提案範圍過大，時程估計過於樂觀。
4. 著重技術或方案介紹，缺乏具體應用情境、使用流程與導入方式說明。
5. 只想推既有產品或解決方案，缺乏對領域痛點的理解，硬套解題。
6. 未說明如何因應資料缺漏、資料量有限或資料品質不一致的情境。
7. 缺乏 PoC / MVP 實證導入規劃，未明確說明所需資源、執行方式與時程。
8. 未考量投資實務需求，例如人工審查流程、可解釋性或風險控管需求。
9. 團隊分工與合作模式說明不足，無法看出落地執行能力。

加分方向

1. 能提出在資料無法取得情境下之替代作法或落地機制。
2. 能清楚說明與市面既有解決方案或工具的差異化、優勢、創新點、應用價值。
3. 能兼顧模型效能、可解釋性與實務導入需求。
4. 具跨域合作規劃（如創投、金融、資服或產業場域合作）。
5. 能提出後續商業化、規模化、或擴充應用之可行模式。
6. 能考量資料安全、權限管理與地端部署等實務限制。

以數據驅動優化
中小企業融資環境計畫

徵件開跑！

即日起至 **06.15 (一) 中午12時**

至活動官網完成**線上提案**



更多資訊詳見官網！