金融科技產業聯盟 第三次會員大會

數位金融實務規範建議工作圈 玉山銀行

2025年12月16日

玉山銀行 E.SUN BANK

AGNDA

- 金融無塵室進度與年度回顧
- 可程式化AI 治理進度

金融無塵室進度與年度回顧

數位金融實務規範 - 工作圈說明

使命　建立**安心**且**普惠**的金融環境

目標　兩年有感。重點專注於科技防詐、多元資料應用、以及 金融關鍵基礎科技 的發展。我們將優先建立制度並進行專案規模的實證試驗。

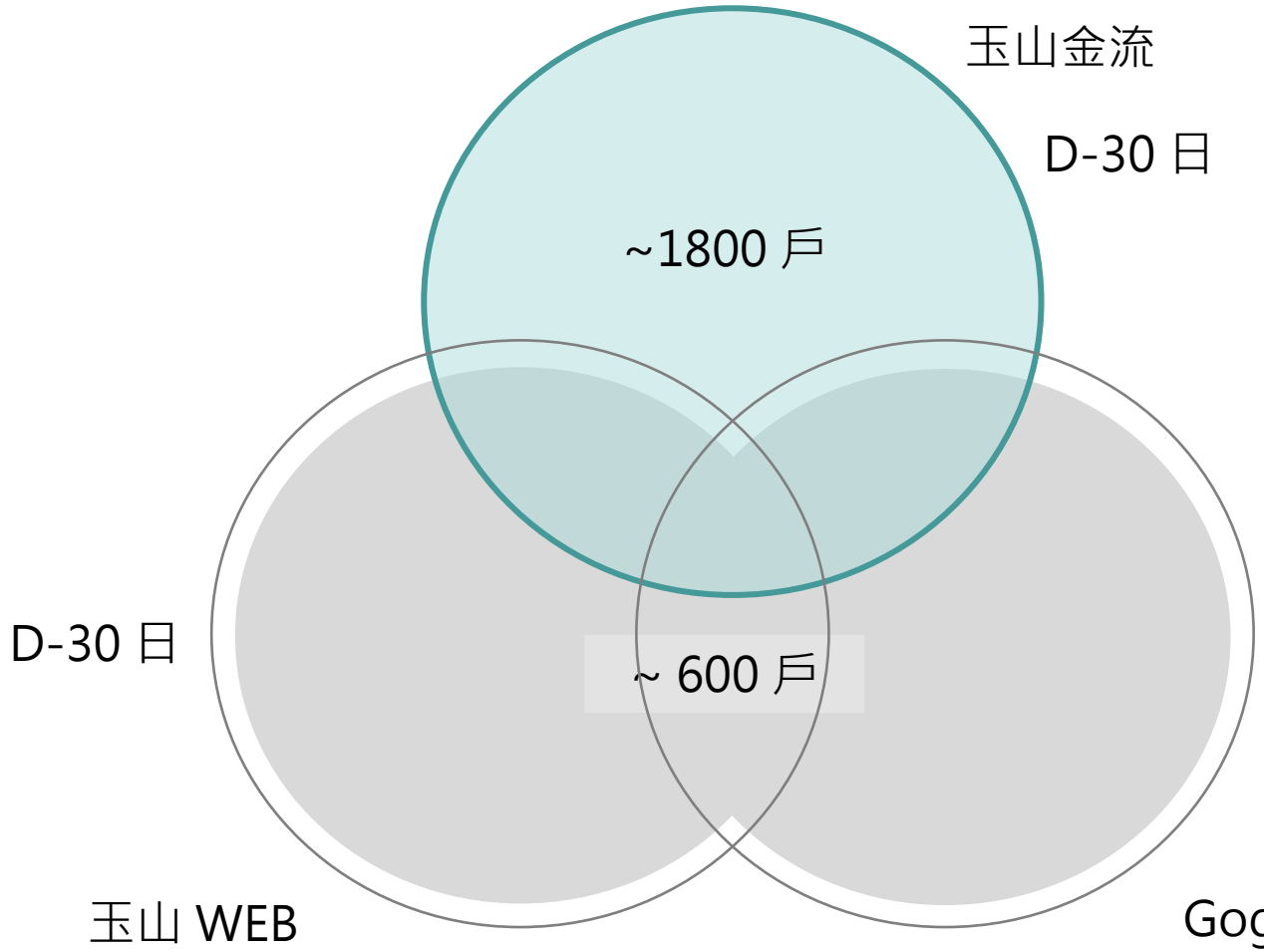
本次進度

金融無塵室

- **科技防詐實證案**
  - 智能資訊網 第一階段 - PoC
  - 建置常設防詐實驗室 第二階段 – 擴大參與方
  - 第三階段 – 跨業防詐
- **完善制度與環境**
  - 跨金融機構資料探勘可用
  - 跨機構模型服務

一邊保護隱私，一邊把詐騙線索找出來

1 整合電信數據多識別三成潛在警示戶



如果僅依賴金融機構自身數據，將錯失近三分之一的早期預警機會。跨業數據整合是防堵詐騙破口的關鍵

2 金融無塵室運作解析：兼顧數據協作與隱私保護

核心概念：資料可用，原始個資不可見

Stage1: 資料來源 (input)

銀行 A

銀行 B

電信 C

各家機構在內部將原始資料(如：顧客ID、交易資訊）進行加密與去識別化處理

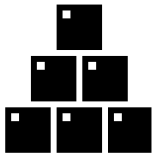
Stage2: 安全環境(Processing)

金融無塵室

加密後的資料上傳到無塵室。在此環境中，資料只能被聯合分析與模型訓練，任何人都無法直接查看或帶走資料。

Stage3: 協作產出(Output)

分析洞察報告



模型參數

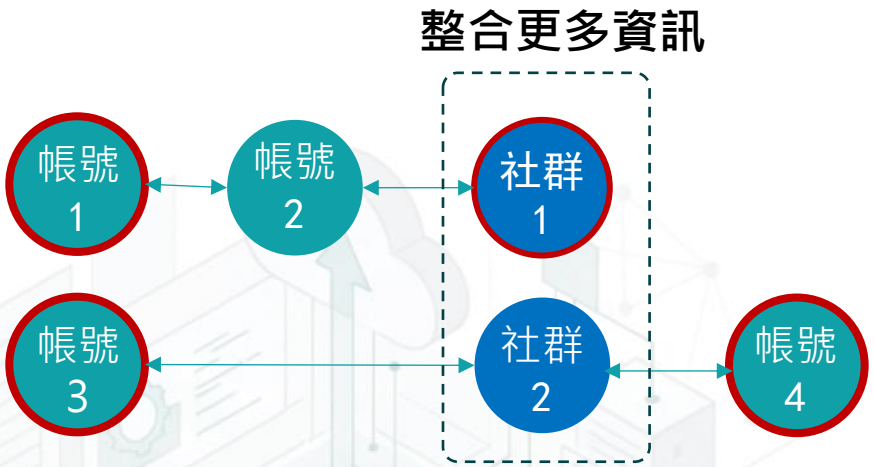
只有經過匯總、不含個資的分析結果或模型才能被參與方取回應用。其他資料與環境在任務結束後銷毀。

此機制確保了機構間的協作能符合個資法規與商業機密保護要求，是未來跨業數據合作的標準模式。

3 防詐實驗室方向

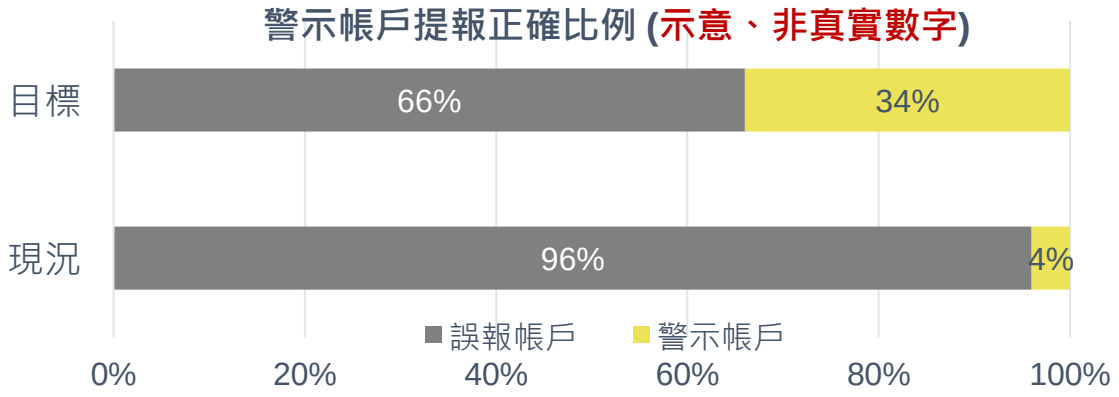
1. 找出更多潛在被害人

結合**多元跨業**資訊，捕捉更多**潛在**被害人



2. 提升警示精準度

更好的演算法與硬體架構，迅速且精準抓取警示戶



1. 整合通聯、社群、跨業資料，找更多潛在被害人
2. 加快演算法速度，更快捉出有問題的人
3. 持續調校，讓誤報更少、抓得更準

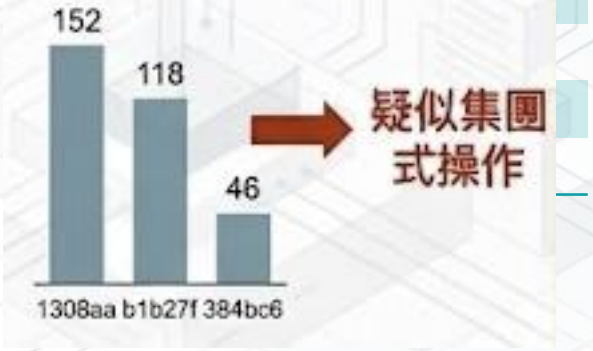
3 跨機構聯合防詐

宏觀：網路地理指紋

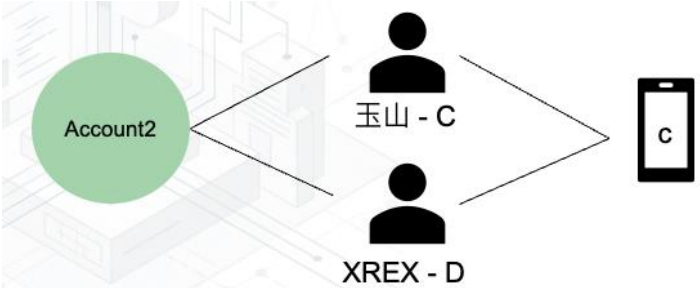
海外長時間多次連線紀錄

警示帳戶
懷疑是車手

國家	城市	ratio
ZA	BEN	0.50
JP	ABI	0.23
CN	DAQ	0.20
DE	ETT	0.20
VN	THỊ	0.16
ZA	KEM	0.16



微觀：跨機構帳號裝置關聯圖譜



人員 C & D
共用同一支手機 c

人員 C & 人員 D & Account
尚未有任何警示註記在 玉山 & XREX

價值：先機防禦，主動阻斷

4 實驗室目前進度

模型會重複以下步驟來更新每個帳戶的「嫌疑指數」

- 第 1 步：收集資訊

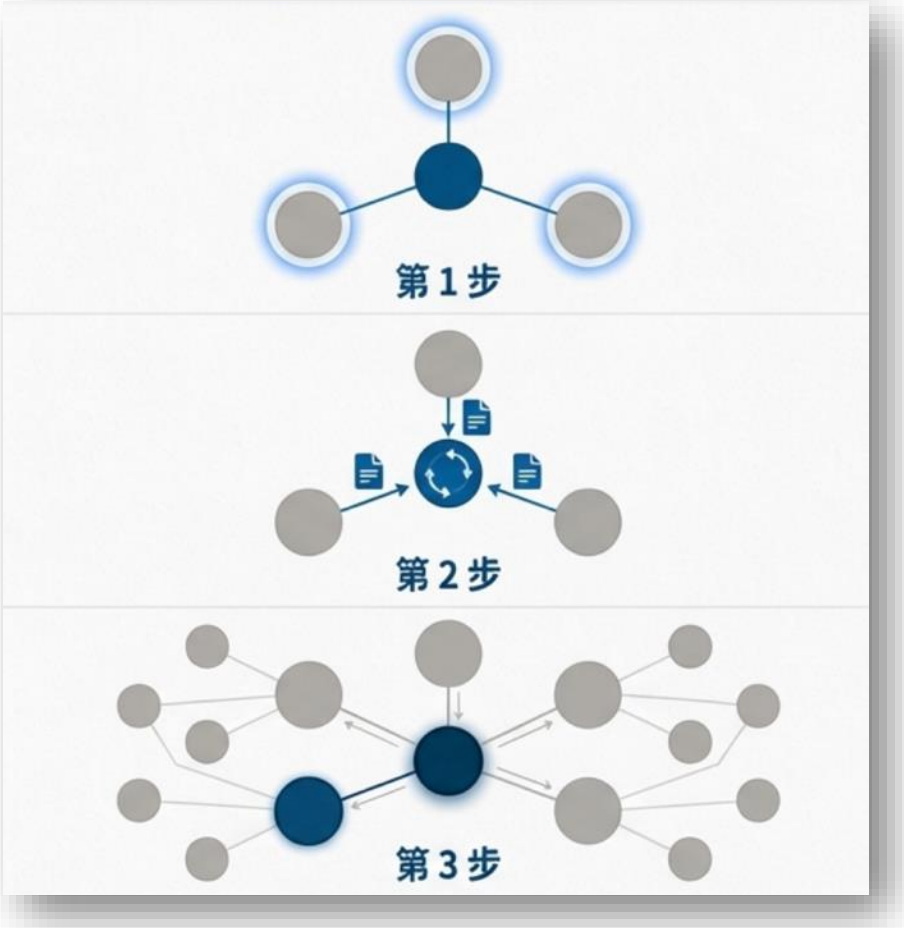
每個帳戶先看看自己「好朋友」（直接交易對象）的檔案

- 第 2 步：更新自己

綜合好朋友們的資訊，更新自己的特徵，形成新的認識。

- 第 3 步：擴大範圍

重複這個過程，資訊會傳得更遠，連「朋友的朋友」都會納入考量。



模型會自動學習如何整合這些來自鄰居的資訊，找出最重要的線索。

5

金融無塵室三階段發展藍圖

~2025.04

第一階段(已完成)
智能資訊網實驗

任務
完成初步概念驗證，確認跨行資料合作的可行性與價值。

國家資通安全研究院 National Institute of Cyber Security
財團法人資訊工業策進會 INSTITUTE FOR INFORMATION INDUSTRY
中華郵政
玉山銀行
FinTechSpace 金融科技創新園區
中國信託 CTBC

2025.05-2026.03

第一階段(進行中)
防詐實驗室
(金融無塵室2.0)

任務
完成初步概念驗證，確認跨行資料合作的可行性與價值。

金管會
moda 數位發展部

2026.03 – 2026.12

第三階段
AI研發基礎建設

任務
成為常態化的研發平台機制，整合政府資料匯流平台，不僅只用在防詐，更要成為國家AI發展的基礎架構之一。

可程式化AI治理進度

數位金融實務規範 工作圈與可程式化AI 治理目標

秉持工作圈使命建立安心、普惠、共好、有共識的金融環境

產業可以主動定義並實作出可量化的技術指標，形成產業層級的共識

促進金融業共好

保護利害關係人

內部健全發展

合規

共同的痛點

合規舉證困難 ：多為原則性的定義，是否有可操作技術框架

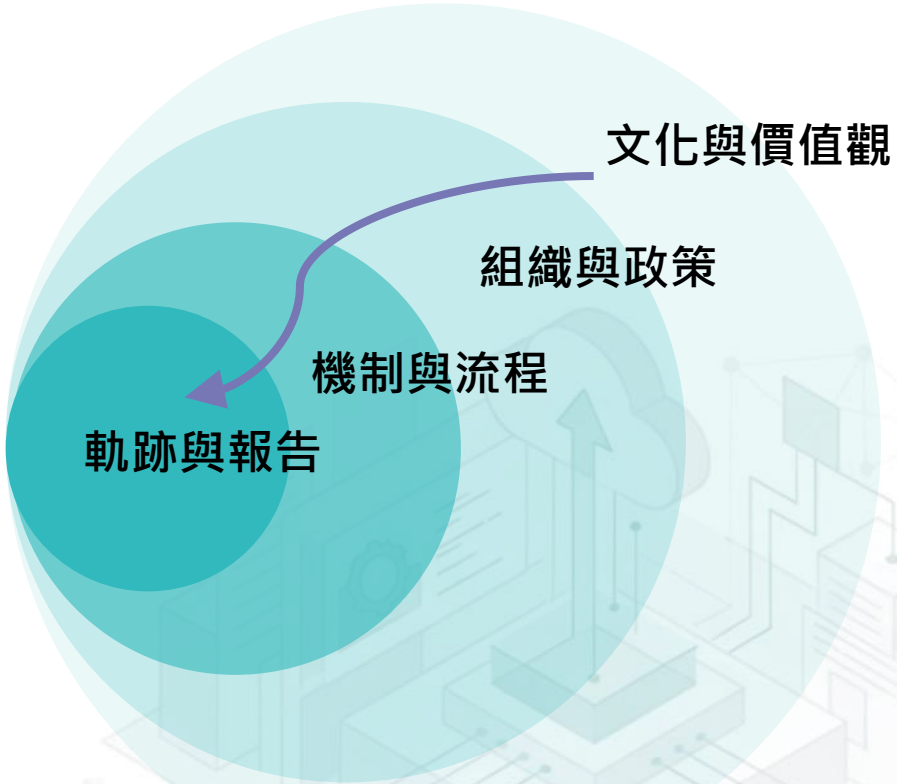
內部管理定義不一：是否透過交流形塑出共通語言

- 監管單位釋出相關內容可見高度重視風險控管、消費者保護與資訊安全
- 根據金管會公布金融業應用人工智慧調查結果已有超過百家金融機構在不同業務中導入 AI
- 金管會已於 2024 年 6 月正式發布「金融業運用人工智慧(AI)指引」，並預計將逐步完善 AI 基本法。

Taiwan FinTech Alliance
金融科技產業聯盟

11

第一階段從實做的pilot 案收斂各機構治理的『方向』



治理複雜且與多面向環環相扣的

文化與價值	**Why ：** 企業為什麼需要AI 治理？回應規範？對顧客負責？ 道德類型的原則要回應到什麼程度？
組織與政策	**How：** 企業要花怎麼樣的資源與成本做AI治理 什麼層級的團隊管理 是否要導入工具
機制與流程	
軌跡與報告	**What：** 有多少AI應用系統、如何盤點 留下什麼軌跡、如何監控

第一階段可程式化AI 專案以技術層的實作出發

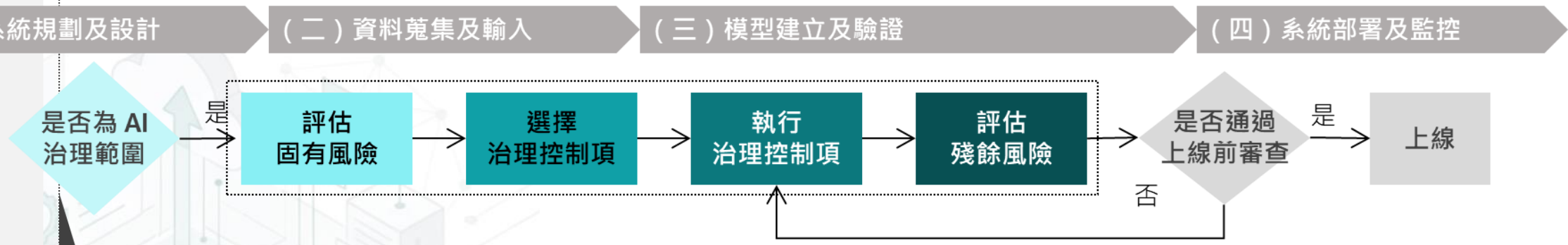
AI 治理框架

策略層

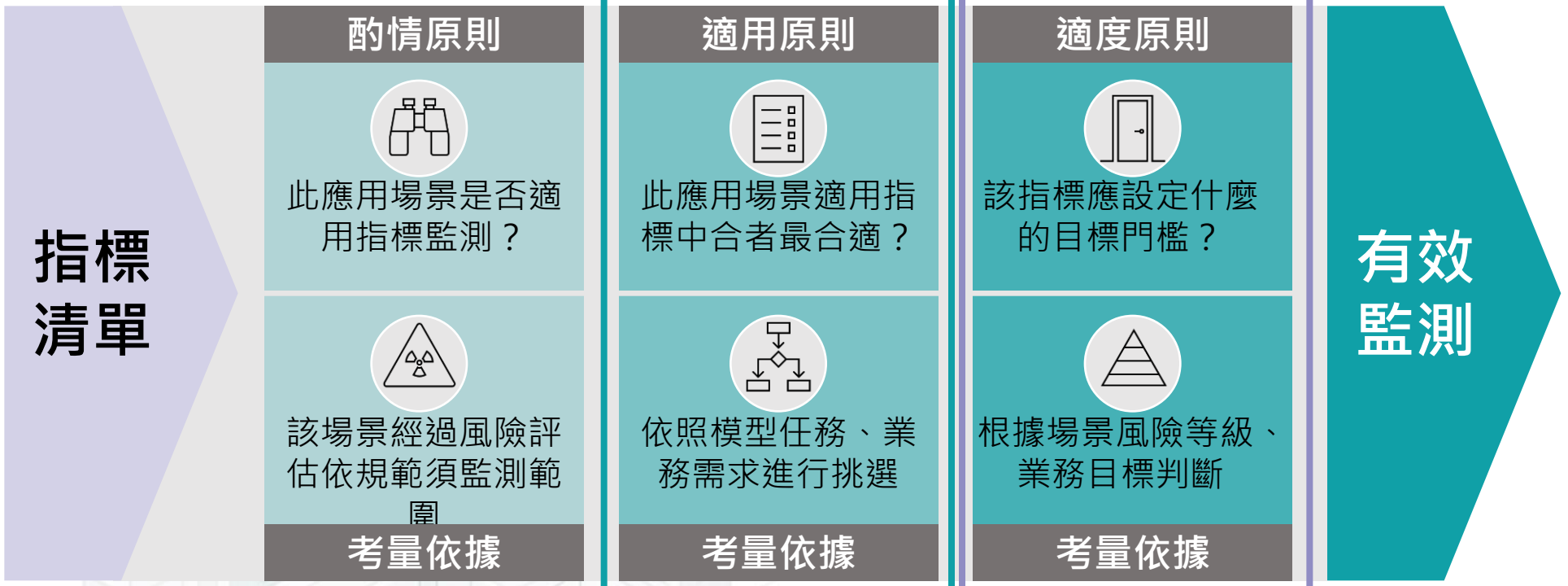
營運層

技術層

1. 第一階段由機構內部挑選的pilot 專案實作
2. 將定期交流會議分享監測指標



AI治理延伸的評估不是疊床架屋應回歸到既有內規作為『原則』的判斷



AI 治理所列示之原則	機構內可參考施作方向或門檻選用原則
隱私性	因應個資法所訂定之內規
系統穩健性	機構申請ISO27001 所定之規範
…	…

依照風險等級、模型應用類型取得技術治理手段

依風險篩選

風險層級

- 高風險
- 中風險
- 低風險

依模型應用挑選技術方法

不同的 AI 模型任務與技術適用不同的技術方法

ML[1] 數值預測	ML 模型將輸入資料經判斷後產生位於一連續範圍內的數值
ML[1] 分類	利用機器學習進行資料分類
檢索增強生成	利用資料庫擴充 AI 問答時的知識範圍（RAG）[1]
FM[1] 數值預測	基礎模型將輸入資料經判斷後產生位於一連續範圍內的數值
FM[1] 分類	利用基礎模型進行資料分類
……	

討論並確認

選擇對此場景最具代表性者

控制項	技術方法
計算穩健性指標（表現準確）	準確率（Accuracy）
	F1 分數（F1 Score）
	ROC 曲線下面積（AUC）
	召回率（Recall）
	精確率（Precision）

透過多樣化的AI 系統應用pilot 產出可供參考的報告書

示例 ：AI 輔助產品行銷

使用情境

透過機器學習分析客戶過去的交易與往來紀錄，並預測下一階段可能使用的產品，協助業務單位在適當時機進行跨售與推薦

系統設計

- 模型演算法：XGBoost
- 模型輸入：結構化資料
- 模型輸出：類別
- 自行訓練之模型
- 即時性：批次 - 每週推論
- 個人屬性資料：有使用

#	開發階段	任務 / 控制項名稱	控制構面	管控的風險類型
1	系統開發	計算穩健性指標（表現準確）-開發階段	穩健性	系統表現不具備準確性
2	系統開發	計算可解釋性指標-開發階段	可解釋性	系統缺乏運作可解釋性
		…		
10	營運監控	計算可解釋性指標-上線監控	可解釋性	系統缺乏當責與可問責性/系統缺乏運作可解釋性

治理面向	控制項	指標名稱	定義
穩健性	計算穩健性指標（服務可靠）	Error Rate	計算失敗的請求或是任務佔比
	計算穩健性指標（服務可靠）	Task Execution Duration	從任務啟動到結束所需的時間
	計算穩健性指標（表現準確）	AUC（Area Under the ROC Curve）	衡量 ROC 曲線（以真正率 TPR 為縱軸，假正率 FPR 為橫軸）面積

治理面向	控制項	指標名稱	結果	結果說明（建議）
穩健性	計算穩健性指標（服務可靠）	Error Rate		
穩健性	計算穩健性指標（服務可靠）	Task Execution Duration		

透過多樣化的**AI** 系統應用**pilot** 產出可供參考的報告書

1. 無須交換任何資料僅需再自己的環境執行技術層級指標
2. 無須提供商業細節，僅針對應用情境之生命週期及對應原則、指標討論
3. 透過交流可釐清不同面向的指標通過的 “門檻” 原則

維度			in/output	備註	
non-LLM	有個資	對外	輸入：圖片 輸出：文字	OCR model	•資料格式轉換
		對內	輸入：資料表 輸出：機率值	XGBoost	
	無個資	對內	輸入：資料表 輸出：機率值	XGBoost	
LLM	有個資	對外	輸入：文字 輸出：文字	自建架構(開源模型)+RAG	•開源模型建構 LLM
		對內 (規劃中)	輸入：文字 輸出：文字	GPT API + RAG	
	無個資	對外	輸入：文字 輸出：文字	Gemini API + RAG	LLM 經典應用專案
		對內	輸入：文字 輸出：文字/圖片	•GPT+RAG •Dall-E	•API Tool calling •多子系統架構 •多模態輸出

中長期規劃：
將分散的『合規成本』轉化為共有的『技術資產』

01. Pilot 案實作 → **02. 收斂工具包** → **03. 擴大協作與驗證機構** → **04. 主管機關協作討論**

2025/11 **招募協作夥伴** | **NOW** | 2026/Q3 | 2026/Q4

	01. Pilot 案實作	02. 收斂工具包	03. 擴大協作與驗證機構	04. 主管機關協作討論
階段目標	• 以pilot案實踐“可程式化”AI 治理 • 建立金融機構治理交流	• 涵蓋足夠技術選型的治理框架 • 定義初版工具包之使用說明	• 擴大驗證期 • 確認技術選型與涵蓋情境	
具體產出	• pilot案AI 治理之報告書 • 以pilot 案收斂基礎版治理框架	• 技術指標工具包 V1 • 工具包V1 技術操作說明書 • 工具包V1 使用規範	• 工具包 V2	

數據是防詐的關鍵，
AI治理是新的信任資產，
協作是勝利的唯一途徑。