AI 發展風起雲湧

國際倡議 AI 治理 AIMS 奠定基石

美國聯邦政府對於常見供應商威脅已經發展出供應鏈資安風險管理（C-SCRM）。在供應鏈韌性這必要前提下，企業宜預先完成合規準備並完善資安治理制度，納入 AI 風險管理（AIRM）與 AI 管理系統（AIMS），避免因 AI 資安的落差所造成的供應鏈排擠，而喪失商業機會。

文／梁日誠

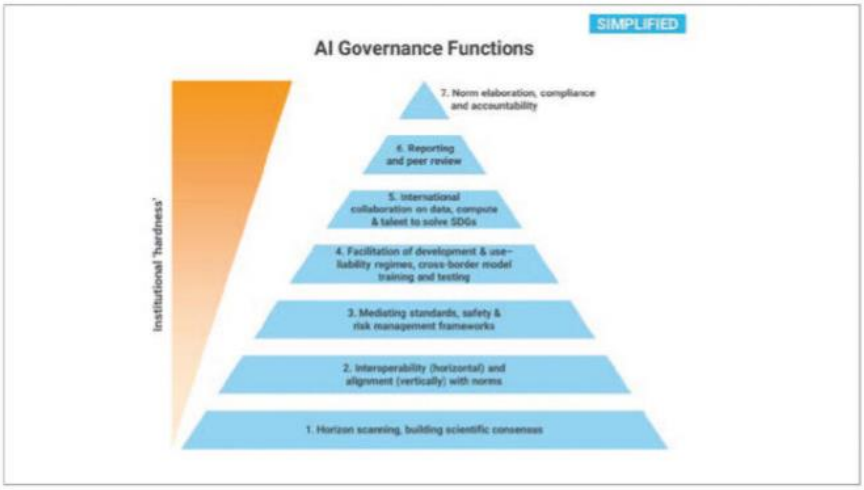
當 AI 業界對於 Responsible／Trustworthy AI 匯聚共識之際，聯合國（UN）大會更於 2024 年 3 月發佈「**抓住安全、可靠和值得信賴的人工智慧系統帶來的機遇，促進可持續發展**」的大會會議決議，決議中陳述 AI 治理（AI Governance）的重要，並回顧 2023 年 12 月由聯合國高級別 AI 諮詢機構（AI Advisory Body，AIAB）發佈了「**為人類治理人工智慧（Governing AI for Humanity）**」的暫行報告；至於 AIAB 其最終報告則預計將於 **2024 年 9 月的未來峰會（Summit of the Future）**召開前發佈，各會員國考慮通過全球數位契約（Global Digital Compact）來持續推動 AI 治理。

由 UN 帶領的國際 AI 治理倡議，正如火如荼地蔓延，國際標準組織（ISO）也適時的推出對應的 AI 治理與人工智慧管理系統（AIMS）相關的國際標準以支持國際社會所需，AI 的重要性與特性，使得 AI 治理的發展異於其他議題（如:資安）的「**管理先行，治理隨後**」的節奏，身處 AI 浪潮中的各利害相關者，宜適時的展現 AI 治理的績效，以確保國際 AI 競爭力的持續領先地位。

UN 的 AI 治理倡議

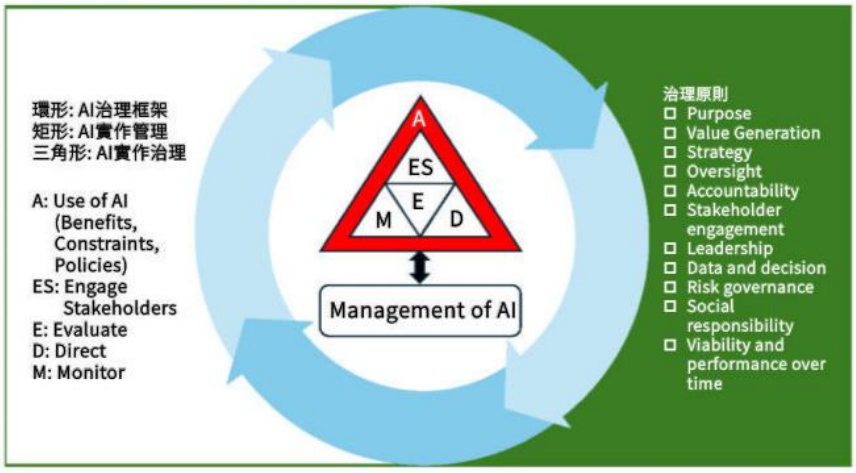
於 UN AIAB 的暫行報告中，識別了 AI 治理的指導原則，涵蓋包容性（Inclusivity）、公共利益、資料治理的中心地位、普遍性、網路化和多利害相關者（multistakeholder）、國際法（International Law）等與機構或組織功能（Institutional Functions），以指導建立新的全球 AI 治理機構。

AI 治理功能的多層式架構如 **<附圖一>**，AIAB



<附圖一> AI治理功能，資料來源: AIAB Interim Report

識別了 AI 治理須涵蓋的幾項功能，其中包括定期評鑑 AI 的狀況與發展軌跡、發展與協調標準、安全和風險管理框架、促進國際多方合作以增強全球南方（the Global South）的能力、監督風險和協調應急回應、有助於 SDGs，以及制定具有約束力的當責規範等。



<附圖二> AI治理模型案例

UN 之下的「聯合國教育、科學及文化組織（聯合國教科文組織，UNESCO）」自 2021 年便公布由 193 個會員國所採用的「**AI 道德建議（Recommendation on the Ethics of Artificial Intelligence）**」，並持續彙聚國際資源，以「**全球 AI 道德與治理觀察站**」推動並監督各項 AI 治理發展，例如：Readiness Assessment Methodology（RAM）、Ethical Impact Assessment（EIA）等。組織如何選擇適當的 AI 與 Data 相關標準與方法以達到合規，同時展現與 UN AI 治理的指導原則相校準，是各組織的重要課題。

國際標準組織的 AI 治理模型

國際標準於**組織治理**（Governance of organizations，**ISO 37000:2021**）、**IT 治理**（Governance of IT for the organization，**ISO 38500:2024**）與 **AI 治理**（Governance implications of the use of artificial intelligence by organizations，ISO **38507:2022**）等議題，均提供國際標準供採用。以最新 2024 年版 **ISO 38500 的 IT 治理模型**為基礎，考量 ISO 38507 中 AI 治理的特性，可形成如**<附圖二>的 AI 治理模型案例**。

其中，「AI 實作治理（Governance of AI Practice）」的運作主體為治理機構（Governing Body，GB），「AI 實作管理（Management of AI Practice）」以 **AIMS ISO 42001 實踐**，「AI 治理框架（Framework for the Governance of AI）」的遞迴式架構則包含了**方向（Direction）**、**容量（Capacity）**、**政策（Policy）**、**委派（Delegation）**、**績效（Performance）**、**當責（Accountability）**等 6 個元素。例舉的 AI 治理模型並包含了 11 項治理原則，治理原則中的社會責任（參考 ISO 26000 社會責任）可對應於 ESG 議題的各項指標，如: SDGs。

基於 ISO 國際標準的 AI 治理模型案例中，**AI 實作治理與 AIMS 須共存且充分互動**，包括: AI 實作治理對 AIMS 的策略與政策進行指導（Direct）、AI 實作治理對 AIMS 提出的計畫與建議進行評估（Evaluate）、AI 實作治理對 AIMS 提出的績效與符合性進行監督（Monitor）等。AI 國際標準的治理原則對應於 UN 的指導原則，ISO AI 治理模型則呼應於 UN AI 治理功能，為各組織展現 AI 治理的最佳機制。

梁日誠（GPM-bI CISSPI CCISMI CCISAI PII CCPI CCAI AIMPI FIAAISI FHCA-EU AI Actl CAIEI CDEI CDAI DSFMI CBAEI FHCA-GDPRI）現為加拿大 SCC／MC ISO／IEC JTC1／SC42、SC27、ISO／TC22／SC32、IEC／TC65 技術組成員，ISO 42001／ISO 27001／ISO 27701／ISO 22301／ISO 20000-1／IEC 62443-2-1 稽核師及講師，TCIC 環奧國際驗證公司全球營運總經理。