

經濟合作與發展組織發布《促進AI可歸責性：在生命週期中治理與管理風險以實現可信賴的AI》



經濟合作與發展組織（Organisation for Economic Co-operation and Development, OECD）於2023年2月23日發布《促進AI可歸責性：在生命週期中治理與管理風險以實現可信賴的AI》（Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI）。本報告整合ISO 31000：2018風險管理框架（risk-management framework）、美國國家標準暨技術研究院（National Institute of Standards and Technology, NIST）人工智慧風險管理框架（Artificial Intelligence Risk Management Framework, AI RMF）與OECD負責任商業行為之盡職調查指南（OECD Due Diligence Guidance for Responsible Business Conduct）等文件，將AI風險管理分為「界定、評估、處理、治理」四個階段：

1. 界定：範圍、背景、參與者和風險準則（Define: Scope, context, actors and criteria）。AI風險會因不同使用情境及環境而有差異，第一步應先界定AI系統生命週期中每個階段涉及之範圍、參與者與利害關係人，並就各角色適用適當的風險評估準則。
2. 評估：識別並量測AI風險（Assess: Identify and measure AI risks）。透過識別與分析個人、整體及社會層面的問題，評估潛在風險與發生程度，並根據各項基本價值原則及評估標準進行風險量測。
3. 處理：預防、減輕或停止AI風險（Treat: Prevent, mitigate, or cease AI risks）。風險處理考慮每個潛在風險的影響，並大致分為與流程相關（Process-related）及技術（Technical）之兩大處理策略。前者要求AI參與者建立系統設計開發之相關管理程序，後者則與系統技術規格相關，處理此類風險可能需重新訓練或重新評估AI模型。
4. 治理：監控、紀錄、溝通、諮詢與融入（Govern: Monitor, document, communicate, consult and embed）。透過在組織中導入培養風險管理的文化，並持續監控、審查管理流程、溝通與諮詢，以及保存相關紀錄，以進行治理。治理之重要性在於能為AI風險管理流程進行外在監督，並能夠更廣泛地在不同類型的組織中建立相應機制。

相關連結

- [Organisation for Economic Co-operation and Development \[OECD\], Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI \(Feb. 23, 2023\)](#)
- [Advancing accountability in AI, Organisation for Economic Co-operation and Development](#)

你可能會想參加

- 【2023科技法制變革論壇】AI生成時代所帶動的ChatGPT法制與產業新趨勢
- 「跨域數位協作與管理」講座活動
- 新創採購-政府新創應用分享會
- 【線上場】113年「新創採購機制及鼓勵照護機構參與推動」說明會
- 【北部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會
- 【南部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會
- 113年新創採購-照護機構獎勵說明會
- 【南部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會

- 【北部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會
- 【中部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會
- 【臺北場】113年度新創採購-招標作業廠商說明會
- 【臺中場】113年度新創採購-招標作業廠商說明會
- 【高雄場】113年度新創採購-招標作業廠商說明會



許嘉芳

法律研究員 編譯整理

上稿時間：2023年06月

資料來源：

Organisation for Economic Co-operation and Development [OECD], *Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI* (Feb. 23, 2023), <https://www.oecd-ilibrary.org/docserver/2448f04b-en.pdf?expires=1683078463&id=id&accname=guest&checksum=030F1E26850085E4A7AC6B1EB60266A1> (last visited May 3, 2023).

Advancing accountability in AI, Organisation for Economic Co-operation and Development, <https://oecd.ai/en/accountability> (last visited May 3, 2023).

延伸閱讀：

陳箴，〈美國國家標準與技術研究院公布人工智慧風險管理框架（AI RMF 1.0）〉，2023年2月，<https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8974>（最後瀏覽日：2023/05/03）。

張滌心，〈歐盟執委會通過關於《人工智慧責任指令》之立法提案〉，2023年1月，<https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8958>（最後瀏覽日：2023/05/03）。

文章標籤

人工智慧

 推薦文章