建構安全、可信任的數位治理生態

人工智慧、大數據分析、雲端服務、區塊鏈等科技發展逐漸成熟，帶動全球數位經濟大幅成長，隨著科技應用的進展、產業結構的轉型、生活方式的改變，使得既有數位生態更容易發展，也更具破壞性。資策會以「數位信任治理生態」為視角，建構數位基礎環境所須涵蓋「安全」、「信任」、「友善」等大原則。

近來大熱的生成式AI，急速拉近各界與AI等數位科技運用的距離，大家赫然驚覺「生活處處可AI」。放眼未來，生成式AI有各式各樣的可能，開發者可依需求和技術能量選擇適合的資源進行開發，以超大基礎模型帶動不同任務的多模態發展，有助於各產業的數位轉型與商模再造，擴大數位內容生成、行銷創意、健康醫療、金融服務、法律協助、生產製造、教育學習、及政府服務等各面向的應用。生成式AI將成為下一世代經濟發展的通用技術（General Technology）之一，將是推動新科技必須掌握的關鍵。

數位生活固然處處是機會，卻也處處是風險。像「深偽技術」若被有心人不當使用，製造不實資訊比過去更容易，加上網路傳播資訊便利，使得可能侵權的不實資訊氾濫成災。由人機共同協作的產製模式，透過資料、算力、算法產出成果，其中的數據資料取自網路所截取的圖表和文件，若使用者提出惡意指令、錯誤資訊或不實話術，AI模型的學習能力也可能讓生成內容產生隱私、侵權與歧視的風險。若政府未予以管控，將可能對隱私等重大基本權益保障產生威脅。

「如何治理」已成數位轉型避無可避的急迫議題，許多國家運用「數位治理生態」概念以為因應，例如制定AI必須遵守的倫理原則，在AI治理上納入生命週期概念，強調應該關注AI應用的風險與分級，並應重視AI生態系，而為落實相關措施，各國多有AI之政策諮詢或建議機構，在促進AI技術及產業發展的同時竭力避免風險。

面對數位科技帶來的機遇與風險，資策會觀察國際與台灣的相關治理規範走向。依據MIT Technology Review、IDC，以及Gartner等重要研究報告，二○二二至二○二三年最熱門且具應用可能性的科學技術，即是「值得信任的人工智慧」（AI TRiSM）。近期歐盟提出的人工智慧法草案（AI ACT），美國提出的「人工智慧風險管理框架」（AI RMF 1.0），中國大陸發布《互聯網信息服務深度合成管理規定》，都揭示相關規範內容。我國今（二○二三）年核定「台灣AI行動計畫2.0」，是建構兼顧科技創新及風險治理的可信任AI發展環境，包括人才優化及留才攬才、重視AI倫理法制、推動資料治理及流通。今年五月廣島舉辦G7峰會，與會國承認各國對規範AI的立場不同，然而考量生成式AI的急迫性，G7峰會聯合聲明將責成相關部長建立「廣島AI進程」（Hiroshima AI process）以討論不實資訊等AI議題，為產業發展建立可互通的AI治理架構。

為打造數位治理生態、促進可信任AI的實現，資策會有以下方向性建議：一、以政策推動為主軸，進而尋找關鍵或急迫優先的議題，適時修訂相關法制。同時也可視不同特定議題之需求及風險程度之高低而適時回歸各主管機關進行立、修法。二、宜從輔助產業創新發展的部分著手，在人工智慧管理等新興領域可參照先進國家作法，先以多方對話模式確認技術或應用發展的風險，同時亦應保留足夠的發展空間與彈性，形成符合我國產業發展特色的治理規範體系。而生成式AI是個絕佳契機，讓我國各界具體且實際地直面數位轉型下的治理挑戰及機會。

自由時報 Liberty Times Net

上稿時間： 2023年07月28日

新聞來源： https://talk.ltn.com.tw/article/paper/1596262

文章標籤