我國人工智慧治理策略：先指引、後法律

臺灣在2022年提出的科學技術白皮書「2035科技願景」中，將人工智慧、半導體等列為重點發展科技，以「前瞻創新、民主包容、韌性永續」為核心，透過科技力量來驅動國家整體轉型，並積極打造人工智慧跨部會協作平台（前瞻科技平臺）：臺灣AI卓越中心（AICoE），由國科會扮演協調角色。2023年行政院再核定「臺灣人工智慧行動方案2.0」，聚焦人才與技術發展、完善運作環境、提升國際影響力以及回應人文社會議題，期望能在發展產業與技術同時，妥善管理人工智慧，並以人工智慧解決國家與社會問題。

資策會科技法律研究所蔡宜臻研究員表示，在人工智慧政策與法制規範方面，可能涉及不同風險與倫理議題，例如深偽技術與假訊息問題、生成式AI產製內容之權利問題、人工智慧資料治理與資安問題、人力資源市場重組問題等。鑑於人工智慧技術創新之速度和可能面臨的挑戰，全球主要國家－以及臺灣－皆致力在不妨礙技術發展下，尋求建立可信任人工智慧之治理方針與原則。

臺灣政府一方面積極關注世界各國對於人工智慧的監管模式，例如歐盟提出的《人工智慧法》草案（Artificial Intelligence Act）、美國拜登總統簽署的AI安全行政命令，以及英國、日本等國提出的各領域監管指引、技術標準與產業促進白皮書，另一方面亦積極與產業界、法律界共同協商，尋求最適宜的人工智慧法制架構，以取得在促進產業發展與人權保障間的平衡。

蔡宜臻研究員舉例，臺灣在十月發布規範政府部門的「行政院及所屬機關(構)使用生成式AI參考指引」，重點包括禁止用生成式AI書寫機密文書、公務員不得向生成式AI提供公務機密、各機關使用生成式AI執行業務時應該適當揭露等。未來公營事業、公立學校等，都可以參考此指引，民間企業亦可參考此指引，制訂適合之自律規範。

除此之外，臺灣亦認識到在人工智慧發展過程中，資料治理是除了針對人工智慧技術本身的管理以外，另一需要重視的議題。因此，臺灣的目標之一即是建立資料應用法遵與合規機制，以完善可被信任、公平與具可解釋性的人工智慧應用環境，並拓展人工智慧所需之資料生態系。

總結而言，臺灣以建立可信任之人工智慧技術發展為目標，以指引先行，採取促進創新與不妨礙技術發展的政策手段，建構可信任人工智慧環境，並積極與各界溝通，尋求在最適當之時機推出可信任人工智慧法制，以兼顧產業發展與人權保障。

中央社

上稿時間： 2023年12月26日

新聞來源： https://www.cna.com.tw/postwrite/chi/360574

文章標籤