

美歐可信任AI監管合作之發展路徑

◎王煜翔／中華經濟研究院WTO及RTA中心 分析師

2022年12月美歐TTC部長會議通過了「可信任AI監管合作之發展路徑圖」，確立雙方未來推動AI監管合作的發展路徑。本文依據AI監管合作路徑圖所揭示的三大合作領域，依序說明當前先進國家發展AI監管合作的主要方向，包括：調和AI技術用語與分類方法、在AI國際標準領域共同建立領導地位，以及監控應用AI技術所衍生的新興風險。同時，本文進一步歸結當前先進國家之合作路徑，分析AI監管發展趨勢對我國業者可能產生之潛在影響，以及政府推動法規合作之政策意涵。

關鍵詞：可信任人工智慧、風險為基礎AI規範、美歐貿易和技術委員會

Keywords: Trustworthy AI, Risk-based AI Regulation, EU-US Trade and Technology Council

隨著人工智慧（Artificial Intelligence, AI）技術應用的不斷進步，AI應用可能在生產力、供應鏈、競爭力、市場力量、創新和就業等經濟層面產生重大影響。同時，AI更可以協助決策者制定出更有效率的政策措施。然而，AI在社會經濟層面的廣泛應用也衍生出新興的風險，包括：AI對個人資訊的控制可能侵犯隱私、創造反競爭環境、種族或其他形式的偏見，侵犯人權或是對人體健康生命安全造成危害等。這些AI所衍生的風險主要歸因於用於自主決策的演算法設計問題，必須透過監督管理機制來防範演算法開發

者故意或無意導致的AI設計問題。

為了因應AI發展所帶來的新興風險問題，美國、歐盟等先進國家均已開始研議實施AI監管的必要性，2020年2月歐盟率先提出「AI政策白皮書」（White Paper on Artificial Intelligence）揭示歐盟對AI管理政策之主要發展方向，規劃將採取風險分級管制方式建立可信任的AI法規環境。歐盟隨後於2021年4月提出「人工智慧法律調和規則草案」（Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence），以下稱「歐盟AI規章草案」，未