

世界經濟論壇發布《人工智慧公平性和包容性藍圖》白皮書



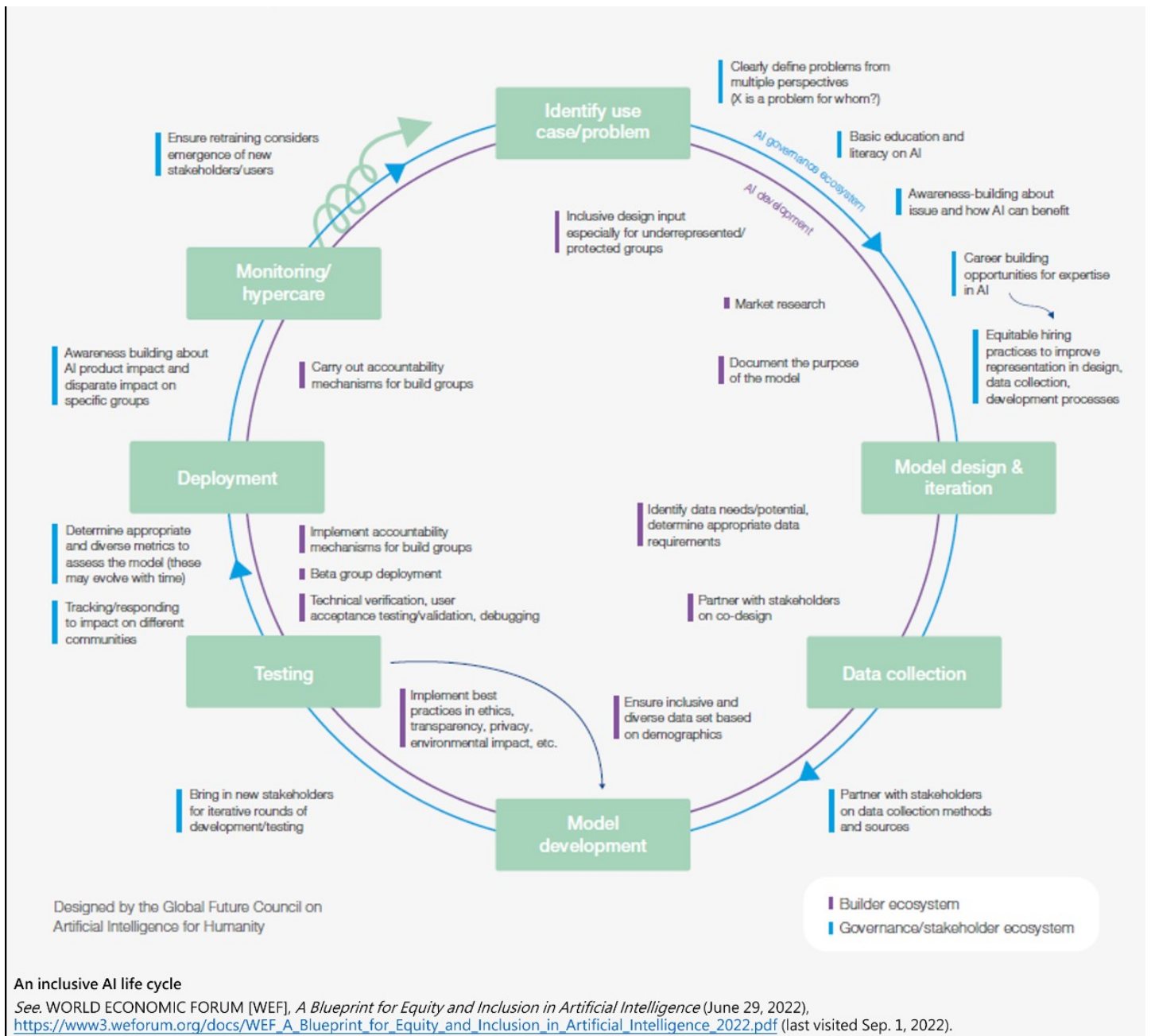
世界經濟論壇（World Economic Forum, WEF）於2022年6月29日發布《人工智慧公平性和包容性藍圖》白皮書（A Blueprint for Equity and Inclusion in Artificial Intelligence），說明在AI開發生命週期和治理生態系統中，應該如何改善公平性和強化包容性。根據全球未來人類AI理事會（Global Future Council on Artificial Intelligence for Humanity）指出，目前AI生命週期應分為兩個部分，一是管理AI使用，二是設計、開發、部署AI以滿足利益相關者需求。

包容性AI不僅是考量技術發展中之公平性與包容性，而是需整體考量並建立包容的AI生態系統，包括（1）包容性AI基礎設施（例如運算能力、資料儲存、網路），鼓勵更多技術或非技術的人員有能力參與到AI相關工作中；（2）建立AI素養、教育及意識，例如從小開始開啟AI相關課程，讓孩子從小即可以從父母的工作、家庭、學校，甚至玩具中學習AI系統對資料和隱私的影響並進行思考，盡可能讓使其互動的人都了解AI之基礎知識，並能夠認識其可能帶來的風險與機會；（3）公平的工作環境，未來各行各業需要越來越多多元化人才，企業需拓寬與AI相關之職位，例如讓非傳統背景人員接受交叉培訓、公私協力建立夥伴關係、提高員工職場歸屬感。

在設計包容性方面，必須考慮不同利益相關者之需求，並從設計者、開發者、監督機關等不同角度觀察。本報告將包容性AI開發及治理整個生命週期分為6個不同階段，期望在生命週期中的每個階段皆考量公平性與包容性：

- 1.了解問題並確定AI解決方案：釐清為何需要部署AI，並設定希望改善的目標變量（target variable），並透過制定包容性社會參與框架或行為準則，盡可能實現包容性社會參與（特別是代表性不足或受保護的族群）。
- 2.包容性模型設計：設計時需考慮社會和受影響的利益相關者，並多方考量各種設計決策及運用在不同情況時之公平性、健全性、全面性、可解釋性、準確性及透明度等。
- 3.包容性資料蒐集：透過設計健全的治理及隱私，確定更具包容性的資料蒐集路徑，以確保所建立之模型能適用到整體社會。
- 4.公平和包容的模型開發及測試：除多元化開發團隊及資料代表性，組織也應引進不同利益相關者進行迭代開發與測試，並招募測試組進行測試與部署，以確保測試人群能夠代表整體人類。且模型可能隨著時間發展而有變化，需以多元化指標評估與調整。
- 5.公平地部署受信任的AI系統，並監控社會影響：部署AI系統後仍應持續監控，並持續評估可能出現新的利益相關者或使用者，以降低因環境變化而可能產生的危害。
- 6.不斷循環發展的生命週期：不應以傳統重複循環過程看待AI生命週期，而是以流動、展開及演變的態度，隨時評估及調整，以因應新的挑戰及需求，透過定期紀錄及審查，隨時重塑包容性AI生態系統。

綜上，本報告以包容性AI生態系統及生命週期概念，期望透過基礎設施、教育與培訓、公平的工作環境等，以因應未來無所不在的AI社會與生活，建立公司、政府、教育機構可以遵循的方向。



相關連結

- World Economic Forum [WEF], *A Blueprint for Equity and Inclusion in Artificial Intelligence* (June 29, 2022)
- A Blueprint for Equity and Inclusion in Artificial Intelligence*, World Economic Forum (June 29, 2022)
- Global Future Council on Artificial Intelligence for Humanity, World Economic Forum (June 29, 2022)

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

許嘉芳

副法律研究員 編譯整理

上稿時間：2022年10月

資料來源：

World Economic Forum [WEF], *A Blueprint for Equity and Inclusion in Artificial Intelligence* (June 29, 2022),

https://www3.weforum.org/docs/WEF_A_Blueprint_for_Equity_and_Inclusion_in_Artificial_Intelligence_2022.pdf (last visited Sep. 1, 2022).

A Blueprint for Equity and Inclusion in Artificial Intelligence, World Economic Forum (June 29, 2022), <https://www.weforum.org/whitepapers/a-blueprint-for-equity-and-inclusion-in-artificial-intelligence> (last visited Sep. 1, 2022).

Global Future Council on Artificial Intelligence for Humanity, World Economic Forum (June 29, 2022), <https://www.weforum.org/communities/gfc-on-artificial-intelligence-for-humanity> (last visited Sep. 1, 2022).

延伸閱讀：

王柏蘆，〈美國國家標準暨技術研究院規劃建立「人工智慧風險管理框架」，並徵詢公眾對於該框架之意見〉，2021年10月，<https://stii.iii.org.tw/article-detail.aspx?tp=1&d=8727&no=64>（最後瀏覽日：2022/09/02）。

文章標籤

科法觀點

平衡AI發展優劣勢 歐盟發布人工智慧法律調和規則草案助企業

從臉部辨識到情緒辨識 資策會科法所：AI發展應尊重人權與民主價值

推薦文章