新加坡網路安全局發布人工智慧系統安全指南，以降低AI系統潛在風險

新加坡網路安全局（Cyber Security Agency of Singapore, CSA）於2024年10月15日發布人工智慧系統安全指南（Guidelines on Securing AI Systems），旨在強化AI系統安全，協助組織以安全之方式運用AI，降低潛在風險。 該指南將AI系統生命週期分成五個關鍵階段，分別針對各階段的安全風險，提出相關防範措施：

（1）規劃與設計：提高AI安全風險認知能力，進行安全風險評估。

（2）開發：提升訓練資料、模型、應用程式介面與軟體庫之供應安全，確保供應商遵守安全政策與國際標準或進行風險管理；並辨識、追蹤及保護AI相關資產（例如模型、資料、輸入指令），以確保AI開發環境安全。

（3）部署：適用標準安全措施（例如存取控制、日誌記錄），並建立事件管理程序。

（4）運作與維護：持續監控AI系統的輸入和輸出，偵測異常與潛在攻擊，並建立漏洞揭露流程。

（5）壽命終期：應根據相關行業標準或法規，對資料與模型進行適當之處理、銷毀，防止未經授權之存取。

CSA期待該指南發布後，將有助於預防供應鏈攻擊（supply chain attacks）、對抗式機器學習攻擊（Adversarial Machine Learning attacks）等安全風險，確保AI系統的整體安全與穩定運行。

相關連結

- Guidelines and Companion Guide on Securing AI Systems, Cyber Security Agency of Singapore
- Singapore's Cyber Security Agency issues security guidelines for AI systems, CMS Law-Now

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

翁嘉璟
副法律研究員 編譯整理

上稿時間：2024年11月

資料來源：

Guidelines and Companion Guide on Securing AI Systems, Cyber Security Agency of Singapore, https://www.csa.gov.sg/Tips-Resource/publications/2024/guidelines-on-securing-ai (last visited Oct. 27, 2024).

Singapore's Cyber Security Agency issues security guidelines for AI systems, CMS Law-Now, https://cms-lawnow.com/en/ealerts/2024/10/singapore-s-cyber-security-agency-issues-security-guidelines-for-ai-systems?utm_source=lawnow-realtime&utm_medium=email&utm_campaign=Singapore%e2%80%99s%20Cyber%20Security%20Agency%20issues%20security%20guidelines%20for%20AI%20systems&utm_id=4097&utm_term=read_more&utm_content=697002 (last visited Oct. 27, 2024).

延伸閱讀：

許嘉芳，〈美國國家標準暨技術研究院發布「人工智慧風險管理框架：生成式AI概況」〉，資策會科技法律研究所，2024年10月，https://stli.iii.org.tw/article-detail.aspx?tp=1&d=9263&no=64（最後瀏覽日：2024/10/28）。

文章標籤

人工智慧　人工智慧風險管理

推薦文章