返回列表 上一篇 下一篇

美國國家標準與技術研究院公布人工智慧風險管理框架（AI RMF 1.0）

美國國家標準與技術研究院（National Institute of Standards and Technology, NIST）於2023年1月26日公布「人工智慧風險管理框架1.0」（Artificial Intelligence Risk Management Framework, AI RMF 1.0），該自願性框架提供相關資源，以協助組織與個人管理人工智慧風險，並促進可信賴的人工智慧（Trustworthy AI）之設計、開發與使用。NIST曾於2021年7月29日提出「人工智慧風險管理框架」草案進行公眾徵詢，獲得業界之建議包含框架應有明確之衡量方法以及數值指標、人工智慧系統設計時應先思考整體系統之假設於真實世界中運作時，是否會產生公平性或誤差的問題等。本框架將隨著各界使用後的意見回饋持續更新，期待各產業發展出適合自己的使用方式。

本框架首先說明人工智慧技術的風險與其他科技的差異，定義人工智慧與可信賴的人工智慧，並指出設計該自願性框架的目的。再來，其分析人工智慧風險管理的困難，並用人工智慧的生命週期定義出風險管理相關人員（AI actors）。本框架提供七種評估人工智慧系統之信賴度的特徵，包含有效且可靠（valid and reliable）：有客觀證據證明人工智慧系統的有效性與系統穩定度；安全性（safe）：包含生命、健康、財產、環境安全，且應依照安全風險種類決定管理上的優先次序；資安與韌性（secure and resilient）；可歸責與資訊透明度（accountable and transparent）；可解釋性與可詮譯性（explainable and interpretable）；隱私保護（privacy-enhanced）；公平性—有害偏見管理（fair – with harmful bias managed）。

本框架亦提出人工智慧風險管理框架核心（AI RMF Core）概念，包含四項主要功能：治理、映射（mapping）、量測與管理。其中，治理功能為一切的基礎，負責孕育風險管理文化。各項功能皆有具體項目與子項目，並對應特定行動和結果產出。NIST同時公布「人工智慧風險管理框架教戰手冊」（AI RMF Playbook），提供實際做法之建議，並鼓勵業界分享其具體成果供他人參考。

本文為「經濟部產業技術司科技專案成果」

相關連結

National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) (2023)

NIST AI Risk Management Framework Playbook, National Institute of Standards and Technology

AI RISK MANAGEMENT FRAMEWORK, National Institute of Standards and Technology

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

陳箴
副法律研究員 編譯整理

上稿時間： 2023年04月

資料來源：

National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (2023), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (last visited Feb. 16, 2023).
NIST AI Risk Management Framework Playbook, National Institute of Standards and Technology, https://pages.nist.gov/AIRMF/ (last visited Feb. 16, 2023).
AI RISK MANAGEMENT FRAMEWORK, National Institute of Standards and Technology, https://www.nist.gov/itl/ai-risk-management-framework (last visited Feb. 16, 2023).

延伸閱讀：

對於該框架之意見，資訊工業策進會科技法律研究所，https://stli.iii.org.tw/article-detail.aspx?no=66&tp=1&d=8727（最後瀏覽日：2023/02/16）。
美國參議院於2022年4月提出《演算法問責法案》對演算法治理再次進行立法嘗試，資訊工業策進會科技法律研究所，https://stli.iii.org.tw/article-detail.aspx?tp=1&d=8912&no=64（最後瀏覽日：2023/02/16）。
美國情報體系發布「情報體系運用人工智慧倫理架構」，資訊工業策進會科技法律研究所，https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8520（最
瀏覽日：2023/02/16）。
加拿大政府提交予國會《人工智慧資料法案》，資訊工業策進會科技法律研究所，https://stli.iii.org.tw/article-detail.aspx?tp=1&d=8888&no=64（最後瀏覽日
2023/02/16）。
英國政府提交予國會「人工智慧監管規範政策報告」，資訊工業策進會科技法律研究所，https://stli.iii.org.tw/article-detail.aspx?tp=1&d=8891&no=64（最後
覽日：2023/02/16）。

文章標籤

人工智慧

推薦文章

你 可 能 還 會 想 看

美國新能源法案預定於2010年前興建新核電廠

美國總統布希於本（ 8 ）月 8 日簽署能源法案，法案目的除減少對國外能源依賴外，另亦授權興建一座新核能發電廠。布希政府希望於 2010 年前開始建造核能廠。　　儘管核能爭議大，但現今國際油價已飆高達每桶 63 美元，在美國參眾兩院日前通過、布希總統今簽署的能源法案中，同意興建的新核電廠，是美國自 1979 年三哩島事件以來，第 1 座預定興建的核能廠。　　能源法案的通過，被視為是布希政府一大勝利，也是相關利益團體石油公司的勝利。布希自 2001 年上台即大力鼓吹此法案，經 4 年多爭議，眾參院才分別在 7 月 28 、 30 日通過。　　除新建核電廠外，能源法案內容還包括：准許在海岸探勘石..

美國聯邦貿易委員會（FTC）提議加強兒童隱私規則，以進一步限制企業將兒童的資訊用來營利

美國聯邦貿易委員會（Federal Trade Commission, FTC）於2023年12月對《兒童線上隱私保護規則》（Children's Online Privacy Protection Rule, COPPA Rule）提出修法草案，並於2024年1月11日公告60日供公眾意見徵詢。 FTC依據兒童線上隱私保護法（Children's Online Privacy Protection Act, COPPA）第6502節授權，訂定COPPA Rule，並於2000年通過生效，要求網站或提供線上服務的業者在蒐集、使用或揭露13歲以下兒童的個人資訊之前必須通知其父母，並獲得其同意。本次提議除了限制兒童個人資訊的蒐集，亦限制業者保留此些資訊的期間，並要求他們妥善保存資料，相關規定如下：（1）置入固定式.

英國資訊委員辦公室（ICO）進行監理沙盒初步公眾意見徵詢

英國資訊委員辦公室（Information Commissioner's Office, ICO）2018年9月就監理沙盒為初步公眾意見徵詢，以瞭解其可行性。ICO監理沙盒之建立係依據英國2018-2021年科技策略（Technology Strategy for 2018-2021），並參考英國金融行為監理總署（Financial Conduct Authority, FCA）已成功發展之沙盒機制。ICO將提供組織於安全可控且不排除資料保護法規適用的環境下，以創新方式應用個資於開發創新產品與服務，並提供關於降低風險與資料保護設計（data protection by design）的專業知識和建議，同時確保組織採取適當安全維護措施。徵詢重點分為六部分： 障礙和挑戰（Barriers and…

電信產業號碼資料庫之應用與法制議題－以個人隱私保護為中心

最　多　人　閱　讀

- 二次創作影片是否侵害著作權-以谷阿莫二次創作影片為例
- 美國聯邦法院有關Defend Trade Secrets Act的晚近見解與趨勢
- 何謂「監理沙盒」?
- 何謂專利權的「權利耗盡」原則?

財團法人資訊工業策進會　統一編號：05076416