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

推薦文章

你 可 能 還 會 想 看

我國電子遊戲場業管理條例修法研析-參考美國佛州Family Amusement Games Act

我國電子遊戲場業管理條例修法研析 -參考美國佛州Family Amusement Games Act 資策會科技法律研究所 法律研究員　王凱嵐 105年04月06日　　去年七月，美國佛州政府簽屬一份Family Amusement Games Act 法案，條文部份針對電子遊戲機做完整的定義解釋，及遊戲機所擺放的場域做限制。特別的是此法案排除博弈遊戲，僅針對非博弈遊戲做獨立規範。　　新修法案之目的在於解決相關遊戲業者面對舊法時的疑惑。修法後將博弈性遊戲與技術性遊戲做出區分，另關於遊戲擺放的場域以及遊戲獎賞都有明確的立法條文。法案做出明確規範後，大幅度的降低遊戲機業者對於政府法規的不確定感，在遊戲產業上能...

Syngenta位於巴西Parana的基改研究機構遭到當地政府沒收

瑞士跨國種子及作物科技公司Syngenta AG (SYT)正與巴西政府為基改活動展開訴訟。去（2006）年11月9日，Syngenta在巴西境內基改作物研究機構被迫關閉，研究機構所在地的Parana州州政府並以Syngenta違反巴西聯邦環保法規為由，沒收其所有投資的資產。Parana州境內有一座自然保護區－伊瓜蘇國家公園，伊瓜蘇國家公園是舉世著名的伊瓜蘇瀑布（Iguacu Falls）的所在地。根據巴西聯邦環保法規規定，基改作物不得栽種於自然保育區的十公里以內。
Syngenta位於Parana州的基改研究機構佔地達123公頃，然而距離伊瓜蘇國家公園卻僅約有六公里。1986年以來，Syngenta即已擁有該研究區域的產...

日本公布《空中移動革命藍圖》

日本經濟產業省與國土交通省共同組成的「空中移動革命之官民協議會」（空の移動革命に向けた官民協議会），於2018年12月20日第4次會議中公布《空中移動革命藍圖》（空の移動革命に向けたロードマップ，以下簡稱「本藍圖」），期待飛天車（electric vertical take-off and landing, eVTOL）的實現可在都市交通阻塞時或欲前往離島、山間地區等情形下，提供新移動方式，也可運用於災害時的急救搬運及迅速運送物資等。　　本藍圖之「飛天車」係電動垂直起降型的自動駕駛航空機，外型近似直升機，並規劃三條發展路線：實際應用目標、制度及標準之整備、機體及技術之研發。從實際應用目標出發，本...

歐盟執委會規劃制訂「2050能源發展藍圖」

歐盟執委會（European Commission）於去（2011）年12月公布「2050能源發展藍圖（Energy Roadmap 2050: a secure, competitive and low-carbon energy sector is possible）」，主要係執委會承諾將推動歐盟於2050年前達成溫室氣體80-95%減量目標（相較於1990年排放基準），建立具競爭力之低碳經濟社會，所以規劃擬訂「2050能源發展藍圖」，期望能導引歐盟走向「無碳化目標（Decarbonisation Objective）」，同時並確保能源供應安全及保持國際競爭優勢。　　並且，奠基於之前「歐洲2020發展策略（Europe 2020）」所設立推動「20-20-20」溫室氣體減量及能源效率目標，歐盟執委會認...

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