

[← 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）；可解釋性與可詮釋性（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**

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>（最後瀏覽日：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>（最後瀏覽日：2023/02/16）。

文章標籤



推薦文章

你可能還會想看

日本特許廳利用人工智慧審查專利與商標申請

日本特許廳（Japan Patent Office, JPO）從去（2016）年12月開始，與NTT Data公司合作，使用人工智慧（Artificial Intelligence, 簡稱AI）來系統化的回答有關專利問題，且依成果顯示，與原先運用人力回復的成果相當；JPO因此決定於今（2017）年夏天開始，將AI技術分階段應用於專利及商標的審查案，並期望能於下一會計年度（2018年4月至2019年3月），在審查業務中全面運用AI技術。JPO指出，透過AI技術能有助於將專利及商標審查程序中繁冗的檢索程序簡化，以專利審查為例，可搜尋大量文件與檔案，進行專利先前技術檢索，以確保相關技術尚未獲得專利保護，同時也可以協助專利分類；此...

美國閒置頻譜發展近況

為了讓業者間服務不受干擾，政府在規劃頻段時，皆會設置護衛帶(Guard Band)，以維持服務品質。不過，隨著科技的進步，業者彼此干擾程度降低，頻譜的使用也較過去有效率，導致頻段常有閒置的情況。是故，FCC在2008年正式公告開放閒置頻譜(white space)，透過業者無須取得執照，以增加頻譜的活用與增加民眾網路接取。美國在2009年完成無線電視數位化後，亦從700MHz較低頻段留下成對5*5MHz，期望透過該頻段覆蓋率高特性，增加業者投資偏遠地區，使當地民眾享有網路帶來的便利性。閒置頻譜的開放利用，雖可增加公益性與頻譜使用性，但亦存有干擾無線麥克風、行動電話與廣播服務等服...

美國聯邦航空總署准許美國有線電視新聞網在人群中使用小型無人機進行拍攝

美國有線電視新聞網（Cable News Network, CNN）通過美國聯邦航空總署（Federal Aviation Administration, FAA）之審查，允許使用小型無人機（small Unmanned Aerial Vehicle, sUAS）直接穿越人群中（flying directly over a person or people）進行拍攝採訪，為美國目前第一件允許在商業目的中使用小型無人機自由穿梭人群之豁免核准案。美國於2016年8月通過聯邦法規第107篇（14 CFR Part 107）又稱小型無人機規則（small UAS rule），規定關於小型無人機之操作規範。其中該規則列舉7種操作禁止事項，須事前經由美國聯邦航空總署豁免方得進行操作（又稱Part 107 Waiver），分別為：1...

歐盟執委會提出《用電資料相互操作性要求及程序實施規則草案》促進電力服務相互操作性

歐盟執委會（European Commission）於2022年7月29日提出《近用電錶及用電資料之相互操作性要求及非歧視性與透明性程序實施規則草案》（Commission Implementing Regulation (EU) on interoperability requirements and non-discriminatory and transparent procedures for access to metering and consumption data），於2022年9月5日草案第二階段之公眾意見徵集結束。本草案以進一步落實《內部電力市場指令》（Directive (EU) 2019/944）中賦予用戶近用有關用電及包括行政手續費用、使用輸配電過路費等資料，促進智慧電錶系統（smart metering system）於資料模型階段及應用...

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？

➤ 隱私權聲明

➤ 聯絡我們

➤ 相關連結

➤ 徵才訊息

➤ 資策會

➤ 網站導覽

財團法人資訊工業策進會 統一編號：05076416

Copyright © 2016 STLI,III. All Rights Reserved.