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

IBM提出「人工智慧日常倫理」手冊作為研發人員指引

隨著人工智慧快速發，各界開始意識到人工智慧系統應用、發展過程所涉及的倫理議題，應該建構出相應的規範。IBM於2018年9月02日提出了「人工智慧日常倫理」（Everyday Ethics for Artificial Intelligence）手冊，其以明確、具體的指引做為系統設計師以及開發人員間之共同範本。作為可明確操作的規範，該手冊提供了問責制度、價值協同、可理解性等關注點，以促進社會對人工智慧的信任。 一、問責制度（Accountability） 由於人工智慧的決策將作為人們判斷的重要依據，在看似客觀的演算系統中，編寫演算法、定義失敗或成功的程式設計人員，將影響到人工智慧的演算結果。因此，系統的設計和開發.

荷蘭智庫提出發展永續性生質燃料的政策建議

面對解決氣候變遷及尋找替代石化能源的全球性問題，生質材料（biomass）的開發與利用深受期待，然而，生質材料的生產與利用是否適當，乃是複雜的決策工具，一國政府在推動與能源、溫室氣體減量有關之政策工具時，必須意識到這些政策工具背後所蘊藏之其他風險。 在面對生質材料的風險，荷蘭政府可謂最先有此問題意識，並嘗試在提出政策工具時盡可能作妥適規劃的先進國家之一。荷蘭是歐洲最大的棕櫚油進口國，以棕櫚油製成的產品在荷蘭超市到處可見，部分棕櫚油也用於能源供應，荷蘭甚至有業者打算興建專門使用棕櫚油運轉的電力供應站（power stations）。 為確保利用棕櫚油及其他生.

音樂串流服務網站鼻祖Grooveshark正式關閉

美國音樂串流服務網站Grooveshark於2015年4月30日在紐約聯邦法院與三家唱片公司(Warner Music Group, Universal Music Group, Sony Music Entertainment)達成和解協議，以避免由陪審團判決(jury verdict)所帶來高達7億3千6佰萬美金的侵權賠償金。Escape Media Group以5千萬美金、公開道歉及關閉經營將近10年的Grooveshark網站為代價結束了這起爭訟多年的著作權訴訟案。 Grooveshark網站的成立理念爲提供使用者上傳音樂的平臺，樂迷可透過平臺互相分享與檢索音樂，因此網站原本適用於數位千禧年著作權法(Digital Millennium Copyright Act)中的避風港原則。惟Grooveshark網站實質上透過.

日本中小企業廳與特許廳聯合施政，強化中小企業或新創企業智財資源之運用

日本於2021年12月27日公布「促進中小企業或新創企業智財運用之行動計畫」，該行動計畫是考量到面臨COVID-19疫情、數位化轉型、氣候變遷等背景下，中小企業或新創企業必須善用企業嶄新的技術或發想，以應對商業環境的變化，而智財權作為企業競爭力的動力來源，顯示出強化智財的管理及運用是不可欠缺的課題。 為了提升中小企業或新創企業的智財運用，日本中小企業廳與特許廳以提供一站式服務整合智財運用支援作為目標，制定行動計畫。施政主要重點如下： 強化與「智財綜合支援窗口」之整合：中小企業廳強化與特許廳聯合INPIT（National Center for Industrial Property Information and…

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