

## 歐盟議會發布《可信賴人工智慧倫理準則》

2019年4月9日，歐盟議會發布《可信賴人工智慧倫理準則》（Ethics Guidelines for Trustworthy AI）。此次內容大致延續歐盟人工智慧高階專家小組（High-level Expert Group on Artificial Intelligence）於2018年12月18日發布的《可信賴人工智慧倫理準則草案》（Draft Ethics Guidelines for Trustworthy Artificial Intelligence）之內容，要求人工智慧須遵守行善（do good）、不作惡（do no harm）、保護人類（preserve human Agency）、公平（be fair）與公開透明（operate transparency）等倫理原則；並在4月9日發布的正式內容中更加具體描述可信賴的人工智慧的具體要件，共計七面向概述如下：

1. 人類自主性和監控(Human agency and oversight)：AI係為強化人類能力而存在，使人類使用者能夠做出更明智的決策並培養自身的基礎能力。同時，AI應有相關監控機制以確保AI系統不會侵害人類自主性或是引發其他負面效果。本準則建議，監控機制應可透過人機混合（一種整合人工智慧與人類協作的系統，例如human-in-the-loop, human-on-the-loop, and human-in-command）的操作方法來實現。
2. 技術穩健性和安全性(Technical Robustness and safety)：為防止損害擴張與確保損害最小化，AI系統除需具備準確性、可靠性和可重複性等技術特質，同時也需在出現問題前訂定完善的備援計劃。
3. 隱私和資料治理(Privacy and data governance)：除了確保充分尊重隱私和資料保護之外，還必須確保適當的資料治理機制，同時考慮到資料的品質和完整性，並確保合法近用資料為可行。
4. 透明度(Transparency)：資料、系統和AI的商業模型應該是透明的。可追溯性機制（Traceability mechanisms）有助於實現這一目標。此外，應以利害關係人能夠理解的方式解釋AI系統的邏輯及運作模式。人類參與者和使用者需要意識到他們正在與AI系統進行互動，並且必須了解AI系統的功能和限制。
5. 保持多樣性、不歧視和公平(Diversity, non-discrimination and fairness)：AI不公平的偏見可能會加劇對弱勢群體的偏見和歧視，導致邊緣化現象更為嚴重。為避免此種情況，AI系統應該設計為所有人皆可以近用，達成使用者多樣性的目標。
6. 社會和環境福祉(Societal and environmental well-being)：AI應該使包含我們的後代在內的所有人類受益。因此AI必須兼顧永續發展、環境友善，並能提供正向的社會影響。
7. 問責制(Accountability)：應建立機制以妥當處理AI所導致的結果的責任歸屬，演算法的可審計性（Auditability）為關鍵。此外，應確保補救措施為無障礙設計。

### 相關附件

[Ethics Guidelines for Trustworthy AI \[ cfm?doc\\_id=58477 \]](#)

### 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](#)

**蔡宜臻**

專案經理 編譯整理

上稿時間：2019年06月

#### 資料來源：

1. EUROPEAN COMMISSION, Ethics Guidelines for Trustworthy AI, [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=58477](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477) (last visited May 1, 2019)

文章標籤



## 推薦文章

### 你 可 能 還 會 想 看

#### 中國大陸有關網路預約出租汽車之規範

中國大陸國務院法制辦公室前於2015年10月10日在網站上公告，有關其交通運輸部就《網路預約出租汽車經營服務管理暫行辦法（徵求意見稿）》對外徵求意見至同年11月9日止。該暫行辦法係因應利用網路建構服務平台，並提供非傳統之職業駕駛或營業車輛的運輸服務類型，如Uber等。由前述公開資料觀之，中國大陸預計對Uber或相關業者，只要符合從事網路預約出租汽車經營服務，即納入交通運輸主管部門之管制範圍。且依提供服務類型不同，區分為網路預約出租汽車經營服務（指平台）、及網路預約出租汽車經營者（實際提供服務之業者）二大類，並分別進行管理，如不得提供類似計程車之巡遊載。

#### 經濟部技術處研究機構智慧財產管理制度評鑑與台灣智慧財產管理規範(TIPS)驗證內容比較

#### 蕃茄醬瓶身相似設計引發商標侵權之爭

今年(2014)3月6日美國號稱蕃茄醬巨人的H.J. Heinz Co. (以下簡稱Heinz)於美國德州聯邦法院向一家德州公司Figueroa Brothers Inc. (以下簡稱Figueroa)提起商標侵權訴訟，主張Figueroa製造販售的蕃茄醬採用與其設計幾近相同的瓶身(ketchup bottle)，侵害其極具識別性、代表性的商業表徵(trade dress)。

Heinz目前針對該玻璃瓶設計已註冊取得3個聯邦商標，其除了主張聯邦商標法保護外，亦基於普通法(common law)提起商標侵權主張。然而，Heinz表示，在提起訴訟前，已數次嘗試與Figueroa私下解決此爭議，但未果，所以最後才會訴諸法律途徑，提起訴訟。Heinz於訴狀中表示從1890年代開始，...

#### 日本電業節能義務之介紹

### 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

> 徵才訊息

> 網站導覽

> 聯絡我們

> 資策會

> 相關連結

財團法人資訊工業策進會 統一編號：05076416

---

Copyright © 2016 STLI,III. All Rights Reserved.