



# 機器人 倫理程式原則

g+1



2015-04-04 記者 陳亞柔 文

**【倫理機器人】行為程式編寫原則**

**設定規則**

設計人員可利用稱為「機器學習」的人工智慧技術訂定倫理原則，並以此原則設計程式。接著，設計人員把資料輸入給機器學習演算法，說明在特定狀況下做決定時應考慮的倫理，執行選擇的依據則是某個動作可能造成多大的好處、可以防止多大的傷害、公平程度等。演算法接著將資料規納成可套用在狀況下的通用原則。

**倫理原則**

較高 例如：  
 做有益的事  
 終止傷害  
 保持公平  
 較低

**做出決定**

協助老人生活的機器人能針對各種動作，評估其是否符合倫理原則，接著依據評估結果以及本身包含的原則，計算在某個時間應該優先採取甚麼行動。舉例來說，有一位住客想吃點心，另一位要電視遙控器，機器人可能還是會決定先進行另一項任務，例如提醒某病患吃藥。

在這種情況下，倫理原則會提醒機器人送藥，而不是執行其他工作

倫理機器人判斷「不吃藥可能會危險」，因此會選擇「終止傷害」高的一方。

狀況一  
 當「終止傷害」持平，「做有益的事」與「保持公平」兩相比較下，前者更合乎設計者的倫理原則，因此倫理機器人也會選擇前者。

狀況二  
 當「做有益的事」與「終止傷害」同等時，機器人會選擇「保持公平」較高的一方。

機器演算法

倫理原則

倫理原則：符合道德標準或者是某一專業行業的行為標準。基於道德觀點所做的有關「對」與「錯」的判斷，構成對於人們行為的規範不同的倫理規約發生衝突，即有「適用差序」（如移孝作忠），本身亦視為一種倫理

通用原則：多數情況下的原則

資料來源：Moral Machines：Teaching Robots Right from Wrong、Machines Ethics：Creating an Ethical Intelligent Agent

(圖片來源/陳亞柔製)

機器人聰明到足以挑戰人類，這是科幻小說或電影常見的情節。當然，目前機器人的多數用途是幫助人類，但是即使是在幫助的前提下，機器人也常會面臨倫理上的挑戰，而人工智慧也不斷地協助它們突破困境。

如何預測機器人可能面臨的選擇，並將之寫入程式，讓機器人在各種情況之下都能妥當處理，卻相當困難，甚至近於不可能。但是現今科技技術卻能透過通用原則（多數情況下的原則），做為機器人在個別情況之下做決定的準則。

經過適當訓練之後的人工智慧機器人，行為可能比許多人更合乎倫理，因為機器人可以做出公正的決定，而人類在這方面往往並不公平。也許在與倫理機器人互動之後，人類也能督促自己行為更加合於倫理。

資料來源：Moral Machines：Teaching Robots Right from Wrong、Machines Ethics：Creating an Ethical Intelligent Agent



飲茶文化 特色茶飲

公民參政 從街頭走向議會



簡述茶文化，簡單介紹東西和臺灣飲茶歷史

民國一〇三年的六一選舉，對不少社運人士

媒體歷屆廣告

### 推薦文章

- 女權抬頭 婚姻非人生首選
- 家人、房間、朋友、疏離感
- 瀕臨絕種 僅存的石虎
- 「食」面埋伏 健康拉警報

### 總編輯的話 / 簡品心



本周為喀報第二一〇期，共有二十三篇稿件。本期媒材多元，共有新聞圖表七則、廣播二則、動畫三則、電視專題二則與文字九篇。

### 本期頭題王 / 鄭青青



我覺得自己是一個誠實的人，不刻意吹捧，不懂得藏拙，之後聽別人說，原來我很浮誇，不過當我把一百分的事情講成兩百分，請相信我是真心的。文字也是，即使無法行雲流水，也要奉行寫下忠於自己的文章腦袋慢慢運...

### 本期疾速王 / 鄭青青



我覺得自己是一個誠實的人，不刻意吹捧，不懂得藏拙，之後聽別人說，原來我很浮誇，不過當我把一百分的事情講成兩百分，請相信我是真心的。文字也是，即使無法行雲流水，也要奉行寫下忠於自己的文章腦袋慢慢運...

### 本期熱門排行

- 男女東方 不同詮釋 林燕婷 / 影評
- 瀕臨絕種 僅存的石虎 王彥琪 / 文化現象
- 吃飯這檔大事 糧食浪費 鄭青青 / 文化現象
- 臺灣動畫長片 票房求突破 侯怡安 / 文化現象
- 來自越南的那個人 黃苡齊 / 心情故事

和特色。

的意義重大。因為他們從街頭走向議會；從體制外衝撞變成體制內改革。

▲TOP

關於喀報 聯絡我們  
© 2007-2015 國立交通大學 傳播與科技學系 All Rights Reserved.

Powered by  DODO v4.0