



AI 驅動的 認知作戰與反制策略

◆ 中正大學傳播學系教授、臺灣事實查核中心董事長 — 羅世宏

近期，親俄羅斯的駭客團體「NoName057」與「RipperSec」對臺灣金融系統發動大規模的分散式阻斷服務（DDoS）攻擊，導致多家金融機構網站短暫癱瘓，包括臺灣證券交易所內的多個機構都受到影響。

資訊戰時代

隨著人工智慧（AI）的快速發展，我們正進入一個以技術為核心的資訊戰時代。根據 2023 年牛津大學網際網路研究所的報告，全球有超過 70 個國家曾經使用 AI 技術進行資訊操控。AI 驅動的「認知作戰」（cognitive warfare）已成為影響國際安全的關鍵因素。

AI 驅動的認知作戰：概念與現況

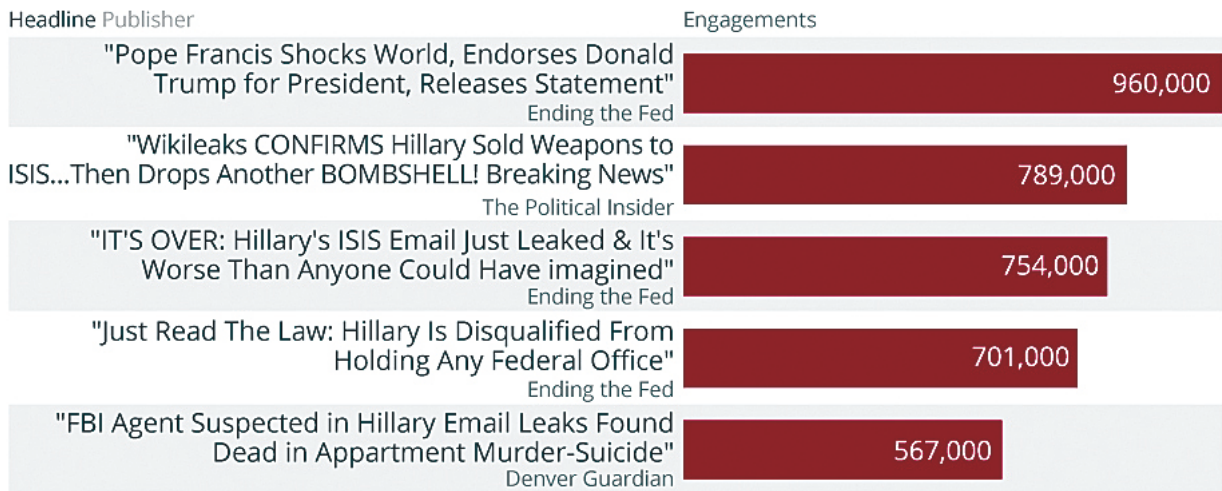
透過資訊操控、心理影響及情感煽動，認知作戰企圖達到影響決策、削弱士氣或瓦解敵人內部團結的目的。AI 的崛起使得認知作戰變得更加精密和高效。

一、深偽技術與虛假訊息的威脅

深偽技術（deepfake）利用深度學習生成或修改成看似真實的數位內容，常用

Fake News Is A Real Problem

Facebook engagement of the top five fake election stories*



Total Facebook engagement for top 20 election stories (August-election day)



* Engagement is measured as total number of shares, reactions and comments

@StatistaCharts Source: Buzzsumo via BuzzFeed

statista

2016 年美國大選期間，透過 Facebook 散布的假消息，單條觸及量最高可達數十萬，可見其影響之廣。
(Photo Credit: Statista, <https://www.statista.com/chart/6795/fake-news-is-a-real-problem/>)

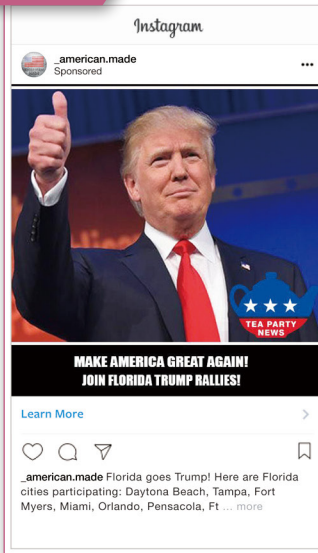
於製造虛假訊息、偽造影音等，能夠極大地誤導公眾，對民主社會中的公共討論和選舉過程造成嚴重影響。舉例來說，2018 年比利時一個政黨發布一段深偽影片，顯示美國前總統川普呼籲比利時退出《巴黎氣候協定》。根據 2024 年的一項研究，全球網絡上流傳的深偽影音數量年增率高達 900%，根據世界經濟論壇的資料，由於生成式 AI 技術製造的假圖像逼真且具有潛在的經濟收益，深偽技術在 2019 至 2020 年間增長了 900%。

二、運算宣傳與目標投放的精準攻勢

AI 演算法能分析大量數據，掌握個人心理偏好與行為模式，進行量身定制的運算宣傳（computational propaganda）。比



比利時某政黨發布美國前總統川普呼籲比利時退出《巴黎氣候協定》的深偽影片，藉此宣傳政黨政策，雖於影片結尾加上影片為假的說明，仍誤導多數民眾誤信。(Photo Credit: Vooruit, https://x.com/vooruit_nu/status/998089909369016325)



俄羅斯在 2016 年美國總統大選期間利用社交媒體投放政治廣告，且特意針對選民的恐懼和偏好設計，致使社會輿論嚴重分化。（Source: HPSCI, <https://democrats-intelligence.house.gov/news/documentsingle.aspx?DocumentID=330>）



各大社群媒體如 X、Facebook，皆發布公告表示在香港反送中期間，有許多假帳號肆虐、散播不實訊息，並經由調查發現其與中國政府有關。（Photo Credit: X, https://blog.x.com/en_us/topics/company/2019/information_operations_directed_at_Hong_Kong; meta, <https://about.fb.com/news/2019/08/removing-cib-china/>）

方說，俄羅斯在 2016 年美國總統大選期間，利用社交媒體投放針對不同選民群體的政治廣告，這些廣告由 AI 生成，針對選民的恐懼和偏好進行誘導，導致輿論嚴重分化。

三、機器人帳號與資訊洪流

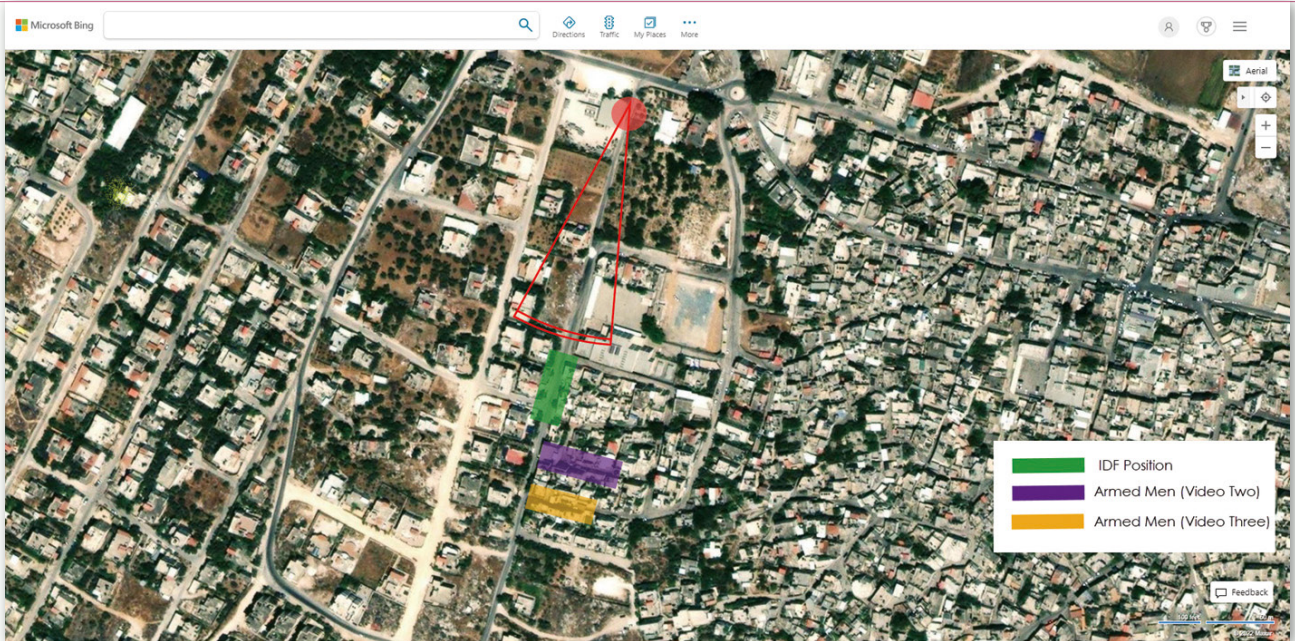
AI 驅動的自動化機器人帳號能快速在社交媒體上擴散資訊，形成資訊洪流，使真實資訊被虛假訊息淹沒。這種攻擊策略容易混淆視聽，削弱公眾對媒體和政府的信任。例如在香港「反送中」運動期間，

大量機器人帳號（bots）被用於製造支持香港政府的假聲量，並散播不利於示威者的虛假資訊。

俄烏戰爭與以巴戰爭中的事實查核案例

一、布查大屠殺揭露：「鈴鐺貓」的調查

2022 年俄烏戰爭期間，烏克蘭布查鎮發生大規模平民遭屠殺事件。俄羅斯政府否認責任，聲稱畫面是烏克蘭和西方媒體所偽造。獨立、第三方事實查核組織



有關阿爾賈茲拉記者希琳（紅點處）被槍殺一案，鈴鐺貓針對影片蒐證、分析、推算彈道軌跡，最終證明子彈是由以色列軍方（綠塊處）射擊。（Photo Credit: [bellingcat](https://www.bellingcat.com/news/mena/2022/05/14/unravelling-the-killing-of-shireen-abu-akleh/), <https://www.bellingcat.com/news/mena/2022/05/14/unravelling-the-killing-of-shireen-abu-akleh/>）

「鈴鐺貓」（Bellingcat）利用開源情報（OSINT）和衛星影像進行調查，最終證明這些平民是在俄軍控制期間被殺害。

二、俄羅斯散布北約士兵大傷亡的假訊息

隨著俄羅斯侵略烏克蘭的戰爭持續，社交媒體和網路空間裡出現大量有關北約軍人在烏克蘭戰場上傷亡的報導，如「90名波蘭雇傭兵在烏克蘭戰鬥中喪生」、「數百名法國軍人遺體從烏克蘭運回」。這些虛假訊息主要來自 SouthFront 與 Hal Turner Radio Show 等假新聞網站，經查是俄羅斯刻意操作所致，為的是打擊烏克蘭和北約的士氣。

三、以巴戰爭中阿爾賈茲拉記者遇襲事件

2022 年 5 月，以色列與巴勒斯坦衝突中，半島電視臺記者希琳·阿布·阿克萊赫（Shireen Abu Akleh）在報導西岸時被槍殺。鈴鐺貓與 CNN 合作使用影片分析

和彈道軌跡推算，證明子彈是來自以色列軍方方向的高精度射擊。此一查核結果引發國際對以色列軍方的譴責，以色列政府也更正其官方說法。

AI 認知作戰的風險與挑戰

大量虛假資訊和深偽內容增加公眾的認知負荷，導致決策疲勞。麻省理工學院研究顯示，虛假新聞比真實新聞更快被分享，尤其是政治虛假訊息，傳播速度高出 70%。這不僅影響大眾判斷，也導致資訊超載和心理疲勞。

AI 驅動的認知作戰正威脅民主制度基礎，如公平選舉和言論自由。例如 2022 年巴西大選期間，虛假訊息在社交媒體猖獗，AI 生成的虛假影像和文字資訊使得選戰充滿不實指控，製造社會混亂，影響社會團結與選舉結果。巴西高等選舉法院的統計資料顯示，選舉期間至少有超過 25% 的熱門政治資訊被證實為不實。



大量假訊息和深偽資訊不僅增加閱聽者的認知負荷疲勞，更致使決策產生偏誤。

針對深偽技術，需開發更先進的 AI 驗證工具，目前已有深偽資訊識別軟體，可檢測出影音的造假痕跡。（Photo Credit: Devfolio, <https://devfolio.co/projects/deep-detect-a-deepfake-shielding-tool-dbc3>）



反制策略與應對措施

面對 AI 驅動的認知作戰，反制策略必須涵蓋技術、政策、教育、事實查核及可信賴資訊生態等 5 個層面，才能有效因應此類新型態威脅。

一、技術反制：開發 AI 對抗技術

針對深偽技術，我們需要開發更先進的 AI 驗證工具，如 Google 於 2023 年推出的深偽識別軟體 DeepDetect，能夠在數秒內檢測出影音的造假痕跡，檢測準確率達 94% 以上。

二、政策管控：加強立法與跨國合作

政府應推動立法，嚴懲刻意製造和散布虛假訊息的行為。歐盟於 2024 年通過的《數位服務法》要求平臺對於系統性製造與散播的虛假內容採取有效反制措施。而

處於無國界的網路資訊時代，國際合作亦是關鍵，共同打擊假新聞的製造與傳播，才能遏止其持續滋長。

三、媒體素養教育：提升公眾防範意識

加強媒體素養教育，幫助公眾辨識虛假訊息，此為防範認知作戰的根本之策。聯合國教科文組織數據顯示，僅 30% 的年輕人具備基本的媒體素養教育，學校和媒體機構應加強此方面的訓練。

四、事實查核：揭露虛假訊息的利器

事實查核是反制認知作戰的重要手段，如「鈴鐺貓」、FactCheck.org 等組織透過數據分析和開源調查方法揭露虛假訊息，防止公共危害。又如，過去兩年多來，關於俄烏戰爭，臺灣事實查核中心發布超過 120 篇事實查核報告，其中多件涉及視



教育部設有媒體素養教育資源網，內有大量相關教材和活動訊息，適合不同學齡的民眾使用學習，藉此培養公眾辨識虛假訊息的能力。（資料來源：教育部媒體素養教育資源網，<https://mlearn.moe.gov.tw/>）

覺媒體的誤導使用，包括深偽技術、錯誤標示的影片與圖片等。

五、建立可信賴的資訊生態系統

我們應該鼓勵和支持可信賴媒體的發展，創建一個由獨立、透明和專業的新聞機構主導的資訊生態系統。這些機構應建立強大的事實查核機制，並持續強化自身的數位安全能力，以應對各類 AI 認知作戰的挑戰。

結論

AI 驅動的認知作戰對國際社會構成重大威脅，我們需要透過技術創新、政策法規、教育推廣與事實查核來因應此一新型態的戰爭威脅，因為只有提升整個社會對抗資訊操控的能力，才能更強韌地守護民主社會的未來。



臺灣事實查核中心曾發布多篇有關俄烏戰爭的查核報告，其中多件涉及視覺媒體的誤導使用，包括深偽技術、錯誤標示的影片與圖片等。（圖片來源：臺灣事實查核中心，<https://tfc-taiwan.org.tw/articles/10580>）