



「平權之路：數位時代的社會風險與治理」課群

用AI對抗AI：數位性暴力快速反應平台

國立中央大學 企管四 黃靖洋/企管三 范純彬/企管三 吳虹儀/法文三 溫子寧/經濟二 曾翎琍

專案簡介

本專案旨在設計一套「數位性暴力申報平台」，整合 AI 偵測、快速反應團隊（RRT）與 NGO 代理機制，目標是將案件處置時間從「週」大幅縮短至「小時」，實現從有法可管到即時執法。

背景動機

惡意內容散播速度極快，攻擊者可在不到 1 分鐘內將內容全網散播。

現行法規雖有基礎，但面臨判斷門檻高和流程缺口的問題，受害者需面對漫長的防守流程：報案、公文往返、平台審核等，形成「漫長障礙賽」，耗時過長，往往造成二度傷害。

研究目標

提供單一入口，整合被害人、NGO 與政府端的案件管理。

本平台結合自動化 AI 偵測與專家人工審核，大幅降低受害者申報門檻，減少二度傷害，並從以往的被動反應轉換為政府主動偵測與快速應對的能力。



申報網站示意圖



申報網站示意圖

課程名稱：性別科技與社會

授課教師：姜貞吟老師、楊燕枝老師

特別感謝：姜貞吟老師、楊燕枝老師、劉百耕助教、林樺霖助教、王儀涵助教

系統偵測結果

已發出通報。案件編號：TB-K1LH4S。RRT 將於短時間內人工覆核。若您為 NGO 代理，系統已將您列為聯絡窗口。

風險等級：HIGH（分數：100）

- 系統判定為高風險：建議請求立即通報與暫時下架。
- 若您為 NGO 代理且已上傳授權文件，請確認檔案已附上。

複製追蹤編號

關閉

[2025/12/10 下午2:14:04] Case created: TB-K1LH4S
[2025/12/10 下午2:14:04] Auto-action: immediate_takedown triggered
[2025/12/10 下午2:14:04] Notification sent to RRT and

AI偵測結果

系統偵測結果

系統證據片段：換臉, deepfake, deepfake, 裸露

風險等級：HIGH（分數：100）

- 系統判定為高風險：建議請求立即通報與暫時下架。
- 若您為 NGO 代理且已上傳授權文件，請確認檔案已附上。

請求立即下架 / 通報（自動）

聯絡 24/7 支援

AI偵測結果

數位人文內涵

本專案利用人工智慧與數據驅動的技術來應對數位性暴力風險，其核心價值在於透過科技工具落實人權與權利，強化社會治理能力。此提案彌補了法律在數位傳播速度下的執行缺口，使法律賦予的權利得以實質賦權於受害者和支持組織，體現了數位工具在改善社會公平與正義方面的跨域應用。

貢獻/心得

製作過程中，我們最大的體悟是：與數位性暴力的擴散速度賽跑，時間就是生命線。看到 AI 成功識別高風險內容，並立即啟動 RRT 緩解時，那種使命感無可取代。這不僅是技術堆疊，更是對信任的承諾，我們必須讓申報流程同時具備匿名、加密、和 NGO 介入的溫暖。這系統的真正價值，在於將冰冷的程式碼，轉化為受害者最需要的即時支援與法律公義。

作品連結

