University of Taipei

Computer Science

透過爬蟲與人工智慧進行數位性暴力影片之下架

Student ID: U10916024

Student: Cheng-Hao, Zhang

張呈顥

November 2023

目錄

第一章 緒論

一、 提案動機

性犯罪一直是台灣社會的重要議題。科技進步使犯罪方式變得更多樣，尤其是網路。韓國的「N 號房事件」中，犯罪者利用加密通訊傳播性犯罪影片，嚴重侵害未成年受害者的性自主權；同樣，台灣的「小玉 Deepfake 換臉事件」中，嫌犯透過人工智慧技術將被害者臉龐加入色情影片，並販賣，引起了社會廣泛譴責。

這些性自主犯罪不僅侵犯受害者的權利，還上傳至網路造成二次創傷。社會需要在法律面前保持穩定，以應對現代挑戰。性犯罪的打擊需要更強有力的規範，以確保受害者得到公平對待，並嚴懲加害者，以此守護社會的安寧。

台灣的法律在處理數位性暴力影像下架的情況時，並沒有明確的、嚴格的規定。如果數位性暴力影像是經雙方同意拍攝的，那麼通常不會適用窺視、竊聽、竊錄罪，而只會適用於刑度較輕的散布猥褻物品罪。

誠如前揭所述，數位性暴力影像之下架十分不易，牽扯到包含但不限於法源、人權、網路治理乃至技術面，這需要政府與網路供應商共同協商並實作，此乃本文之提案動機，下架數位性暴力影片以避免受害者遭受二次、甚至三次傷害。

二、 提案目的

本文主要目標是在探討相關方案，以使用爬蟲和人工智慧技術有效下架數位性暴力影片，並探討法律上之可行性與相關漏洞。再者，此計畫還需要考慮政府與網路產業業者之立場，試圖在不侵犯網路自由之前提下，有效防止犯罪影片之流傳以及下架。

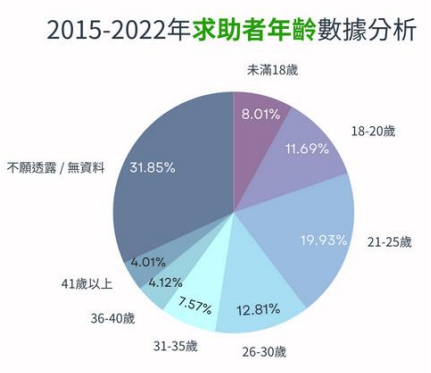
第二章 觀察與發現

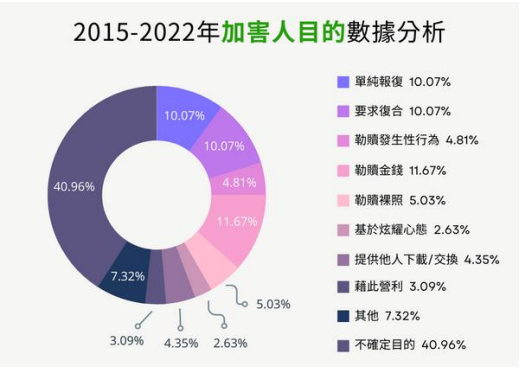
一、 現行情況

有鑑於網路流通快速，數位性暴力影片一旦廣為散播後，想要一抹而淨實乃天方夜譚；惟技術上，可根據搜尋引擎規則或是透過 IPS[1]進行限制，但此舉需要有相當之法源進行法律保留後，得為之。

就目前台灣法律而言，處理未經同意散布數位性暴力影片的問題，只有在涉及兒童和少年色情內容方面有相關規定。根據《兒童及少年性剝削防制條例》第 4 章，明文禁止散布、播送、販售兒童和少年的猥褻照片；因此，主管機關 iWIN[2]網路內容防護機構可以根據該條例第 8 條的規定，要求 ISP 在涉及兒童和少年色情內容方面先行移除內容 (沈忻儒, 2021)。

然而，對於成人內容，iWIN 並未被授予裁處的公權力，僅能進行與業者的溝通，協助業者下架內容。根據婦女救援基金會之網站可見，性私密影像外流不是妳/你的錯至今服務人數高達 888 人，此外還有多少隱藏數據則不得而知。因此，目前台灣在處理未經同意散布成人數位性暴力影片的情況時，法律規定相對不夠明確，需要進一步的法規和機制來保護受害者的權益。





資料來源：https://www.twrf.org.tw/info/title/846 (2023/09/19)

[1] Internet Service Provider
[2] Institute of Watch Internet Network

二、 社會關懷

正如前文所述，加害者將犯罪影片散布到網路上，對受害者造成終身陰影；又因網路傳輸速度快，受害者容易遭受二度甚至三度的傷害。國家法律除了要懲治加害者外，更應該將受害者的保護、社會關懷和醫療協助納入重要政策要點。本文建議以下架和清除相關犯罪影片為基礎，利用爬蟲和人工智慧技術。由於這一提案牽涉到相當程度的人權和隱私侵犯，以及網路治理議題，因此需要有相應的法源依據來執行。本文還將探討法律研究、政府部門與民間團體如何進行協商和合作，此乃本提案實行可行性的關鍵。

第三章 解決方案

一、 法源監管與人權隱私

要透過人工智慧進行數位暴力影像、影音之下架需要有相當充足的法源依歸，否則與法不合並且會違反法律保留原則。根據憲法第 12 條：「人民有秘密通訊之自由。」意指，政府不得窺探人民之秘密通訊，網際網路之通訊亦然；然查，憲法第 23 條：「以上各條列舉之自由權利，除為防止妨礙他人自由、避免緊急危難、**維持社會秩序**，或**增進公共利益**所必要者外，不得以法律限制之。」此條對於憲法人權之保障劃出例外範圍，意即此條四項之事項有限制之必要者，政府得以法律限制之。數位性犯罪影片有涉及社會秩序以及公共利益，因此於憲法層級有制定法律限制人民權利之必要且可行，符合憲法第 23 條之意旨。

- ⊙ 刑法相關條文：刑法中包括了一些條文，規定了性犯罪、散布淫穢物品等犯罪行為，這些可能與數位性暴力有關。另外，此些項目涉及**刑事訴訟**流程以及**檢察**體系偵查流程，有必要探討刑事訴訟法相關流程之修正。

- ⊙ 網際網路提供商相關法律：媒體、網路服務商相關法律規定了網站經營者的責任和法源，特別是關於數位內容的監管。特別是針對 iWIN 網路內容防護機構之權限給予，因為已有未成年影音之撤除權限，但成年數位犯罪影音之規範尚未完善，有檢討之必要。

- ⊙ 兒童及少年性剝削防制條例：這個條例明確禁止散布兒童和少年的色情內容，但對成人內容的規定相對較少。

- ⊙ 個人隱私保護：我們必須確保在任何數位性暴力相關活動中，不侵犯任何人的個人隱私。這包括不隨意收集、儲存或公開個人敏感訊息。

- ⊙ 資料保護：任何收集的個人資料都必須受到適當的資料保護，以確保資料的安全性和隱私性。這包括加密、安全存儲和限制訪問。

- ⊙ 知情同意：如果我們要處理敏感訊息或數位性暴力相關內容，應該獲得相關當事人的知情同意，除非有法源規定允許例外情況。

- ⊙ 匿名化：在可能的情況下，我們應該採取措施匿名化或去識別化數據，以保護受害者的隱私。

二、 爬蟲技術的應用

1. 確定目標網站：

首先，需確定哪些網站或社群平台可能包含數位性暴力影片。這些地方可能包括社交媒體、影片分享網站、討論區等。

2. 建立爬蟲程式：

使用程式語言（例如 Python）建立一個爬蟲程式，定期檢查目標網站上的新內容。這個爬蟲程式需要能夠搜尋並識別數位性暴力相關的關鍵字、標記或特徵。

3. 設定辨識準則：

制定一組標準或規則，協助爬蟲程式辨識數位性暴力影片。這些標準可以包括特定的關鍵詞、影像辨識技術、數位指紋等。需考慮不同平台和語言可能使用的不同用語和符號。

4. 定期巡視網站：

爬蟲程式應定期巡視目標網站，以搜尋新內容。這可以根據需要設定，例如每小時、每日或每週。

5. 辨識和追蹤：

當爬蟲程式發現符合辨識標準的內容時，應該記錄下相關資訊，並追蹤其來源。這可能包括網址、上傳者的資訊、上傳日期等。爬蟲程式也可以自動發出警示，通知相關機構或組織。

6. 報告和處理：

當數位性暴力影片被辨識出來時，相關機構可以採取適當的行動，例如報告給網站管理者、執法機關或民間組織，以便採取進一步的法律或技術措施，包括影片的移除。

7. 持續更新：

數位性暴力的形式和傳播方式可能會不斷變化，因此需要定期更新和改進爬蟲技術，以確保能夠因應新的挑戰。

三、 人工智慧模型

1. 資料收集與建立模型：

首先，我們需要蒐集大量的數位性暴力相關資料，包括文字、圖片、影片等。這些資料將被用來訓練人工智慧模型。模型可以使用深度學習技術，如卷

積神經網絡（CNN）或遞歸神經網絡（RNN），來學習辨識數位性暴力的模式。此法牽涉到隱私與受害者感受，需要有相當授權、法源依據方可執行。

2. 模型訓練：

接下來，我們使用蒐集到的資料來訓練人工智慧模型。這個過程需要大量的運算能力，但它可以讓模型學會辨識數位性暴力內容的特徵，如露骨的性內容、侮辱性的言語等。

3. 實時檢測：

訓練完成的模型可以被嵌入到網站、社交媒體平台或應用程式中，用來實時檢測用戶上傳的內容。當有人上傳可能包含數位性暴力的內容時，模型可以立即偵測到。

4. 警報和處理：

當模型檢測到數位性暴力內容時，它可以自動發出警報，通知相關機構或管理者。這讓他們能夠快速處理該內容，可能包括封鎖用戶、刪除內容或報警給執法機關。

5. 持續改進：

人工智慧模型的效能需要持續改進，以應對變化多端的數位性暴力形式。這可以透過定期的模型更新和重新訓練來實現，以確保模型能夠不斷適應新的內容和威脅。

6. 隱私保護：

在實施這些模型時，必須嚴格遵守隱私和資訊保護相關法規，確保用戶的個人資訊不受侵犯。應用人工智慧模型來自動檢測數位性暴力內容。

四、 政府與網路業者

提出與社交媒體平台、網路服務提供商和政府合作的方式，以實現有效的內容下架。

第四章 執行策略

一、 階段性實施計劃

⊙ 開發階段

在開發階段，我們將著手開發所需的爬蟲技術和人工智慧模型。這個階段的主要任務包括：

- 建立爬蟲程式：我們將使用程式語言（例如 Python）建立一個強大的爬蟲程式，能夠定期檢查目標網站上的新內容。
- 收集相關數據：我們需要蒐集大量的數位性暴力相關資料，包括文字、圖片、影片等，以用於後續模型訓練。
- 訓練 AI 模型：同時，我們也將開始訓練人工智慧模型，這可能涉及到深度學習技術，例如卷積神經網絡（CNN）或遞歸神經網絡（RNN），讓模型學習辨識數位性暴力內容的模式。

⊙ 測試階段

一旦開發完成，我們將進行測試以確保解決方案的正常運作。這階段的工作包括：

- 模型性能測試：我們將對 AI 模型進行性能測試，確保其能夠正確辨識數位性暴力內容。
- 爬蟲程式的有效性測試：同時，我們也將測試爬蟲程式的有效性，確保其能夠正確檢查目標網站上的內容。

⊙ 擴展階段

在成功測試後，我們將解決方案擴展到更多的網站和平台，以擴大其覆蓋範圍。這將包括：

- 整合解決方案：將解決方案整合到不同的網站、社交媒體平台或應用程式中，以實現實時檢測和處理。
- 持續改進：因為數位性暴力的形式可能不斷變化，我們將確保解決方案能夠持續改進，應對新的內容和威脅。

二、 資源需求

⊙ 技術資源

- 軟體：我們將使用程式語言和開源工具來開發爬蟲程式和 AI 模型。
- 硬體：為了支持大規模的數據處理和訓練，可能需要適當的硬體設備。

⊙ 人力資源

- 軟體開發人員：需要具有爬蟲和 AI 開發經驗的專業人員。
- 數據分析師：負責收集、處理和準備相關數據。
- AI 專家：協助設計和訓練人工智慧模型。

⊙ 財務資源

- 開發成本：包括軟體開發、硬體購置和訓練成本。
- 運營成本：包括持續維護、監控和更新的成本。
- 人力成本：工程師和專家的薪資和福利。

三、 監督與評估

⊙ 監督方法

- 監控工具：使用監控工具來追蹤解決方案的運行狀況，以便及時檢測問題。
- 報告機制：設定報告機制，允許相關機構和團體隨時接收解決方案運作的報告。

⊙ 評估方法

- 性能評估：定期進行性能評估，測試 AI 模型的辨識能力和爬蟲程式的有效性。
- 測試結果分析：分析測試結果，以確保解決方案符合預期效果。

⊙ 定期報告

- 定期向相關利益相關者提交報告，以便他們了解解決方案的進展和效能。這有助於及時調整策略和改進解決方案。

第五章 預期效益

透過爬蟲和人工智慧（AI）來下架數位性暴力影片的預期效益是多方面的，它們有助於保護受害者的權益、減少數位性暴力的傳播，並提高社會安全性。以下是一些預期效益：

⊙ 受害者保護：最重要的效益之一是保護受害者的權益。數位性暴力影片的下架意味著受害者的私人內容不再流傳和被濫用，這有助於減輕他們的痛苦和恐懼。

⊙ 預防二次創傷：當數位性暴力影片被下架後，受害者不再需要擔心其私人內容不斷被觀看或分享，這有助於減輕他們的心理壓力和二次創傷。

⊙ 社會安全：下架數位性暴力影片有助於維護社會安全。這些影片的存在可能激勵模仿者或鼓勵更多的犯罪行為，因此有效下架可以減少此類行為的發生。

⊙ 法律合規：使用爬蟲和 AI 技術下架非法內容有助於確保法律的遵守。這有助於維護社會秩序並防止違法行為的擴散。

⊙ 社交媒體和平台的改進：社交媒體和網路平台可能會更積極地合作，以提高其內容審查和監管機制。這有助於改善整體網路環境，使其更加安全和

受保護。

- 教育和意識：解決方案的實施可能促使社會更關注數位性暴力的問題，並提高人們對其嚴重性的認識。這有助於提高社會的教育水平和意識。
- 減少數位性暴力的傳播：下架數位性暴力影片將減少這些內容的可訪問性，降低其在網路上的傳播，進而減少更多人受害的風險。

透過爬蟲和人工智慧技術下架數位性暴力影片，有助於保護受害者的權益、改善社會安全、合規法律、改進網際網路平台、提高教育和意識水平，並減少這類內容的傳播，這些效益對於建立更安全、更公平的網路環境至關重要。

第六章 結論

本文著重探討了使用爬蟲和人工智慧技術來有效下架數位性暴力影片的解決方案，以及相關的法律、社會影響和執行策略。這個問題在當今數位時代尤其重要，因為性犯罪和數位性暴力已經成為全球社會面臨的嚴重挑戰之一。

數位性暴力對受害者的侵害，以及網路上的散播如何導致受害者遭受二次創傷。這強調了為什麼需要更強有力的規範和技術來應對這個問題。此外，台灣法律在處理這類問題時的不足之處，需要更明確的法律規定和機制來保護受害者的權益。

本文主要目標即探討使用爬蟲和人工智慧技術來下架數位性暴力影片的可行性，並探討相關的法律漏洞。同時，我們強調了政府和網路產業業者的合作重要性，以在不侵犯網路自由的前提下，有效地防止犯罪影片的流傳和下架。此外，相關法律對於成人性犯罪影片內容的規定相對於未成年不足，以及社會需要更多關懷和保護受害者的措施都是需要持續關注追蹤之事項。

在解決方案部分，我們提到了使用爬蟲技術和人工智慧模型的方法，以及需要考慮的法律和隱私問題。同時，我們強調了政府、網路業者和社會團體之間的合作，以實現這一解決方案的可行性。

最後，在執行策略部分，我們提出了階段性實施計劃，資源需求，監督和評估方法，以及預期的效益。我們認為這個解決方案將有助於保護受害者權益、減少數位性暴力的傳播，提高社會安全性，符合法律合規，改進網路平台，提高教育和意識水平，並減少這類內容的傳播。

本文的目的是提出一個全面的解決方案，以應對數位性暴力的問題，並保護受害者的權益。這是一個複雜的問題，需要**政府、業者和社會**共同努力，以建立更安全和更公平的網路環境。希望這個提案可以引起更多人的關注，促使相關利害關係人採取行動，實現這一目標。

第七章 參考文獻

沈忻儒. (2021). 數位性暴力台日韓現行法規分析. 台北市.

張呈顥. (2021). 淺談中華民國刑法第 222 條修正案（N號房條款）. 台北市.

Beautiful Soup: https://www.crummy.com/software/BeautifulSoup/bs4/doc/

Requests: https://docs.python-requests.org/en/latest/

Scrapy: https://docs.scrapy.org/en/latest/

Selenium: https://selenium-python.readthedocs.io/

W3Schools: https://www.w3schools.com/python/

Convolutional Neural Networks (CNNs) - Stanford CS231n:

http://cs231n.github.io/convolutional-networks/

RNN and LSTM - Stanford CS231n:

http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture10.pdf

TensorFlow RNN：https://www.tensorflow.org/guide/keras/rnn

PyTorch 中的 LSTM 模型：

https://pytorch.org/tutorials/intermediate/char_rnn_generation_tutorial.html