

全民資訊素養推廣計畫

【Deepfake 真可怕？】延伸學習文章

- 教學對象：國中七至九年級
- 主題：網路識讀

隨著數位資訊科技及人工智慧的發達，要製作出幾可亂真的偽造照片、聲音或影片，已經不再是遙不可及的事，在面對五花八門的網路資訊時，有時會陷入何者為真、何者為假的盲點中，因此我們每一個人都需要培養思辨能力。

● 什麼是 Deepfake？

深偽技術 (Deepfake) 是深度學習 (deep learning) 和偽造 (fake) 的混合詞，泛指各種藉由人工智慧達成的合成照片、聲音或影片之技術。近年來也有不法分子運用電腦合成或深偽技術 (Deepfake) 進行犯罪，例如製作造謠影片、詐騙取財、數位／網路性暴力，除了侵害他人的隱私權和名譽之外，也對社會安定與安全造成危害。

● Deepfake 技術會有破綻嗎？

當然有，資訊業界和學術界在持續發展 Deepfake 技術的同時，也在研發偵測 Deepfake 造假的技術，但發展至今 Deepfake 技術的破綻已經越來越少、越來越難偵測。Deepfake 影片產生破綻的主意原因是素材的不足，這是由於現階段的 Deepfake 通常需要大量的訓練資料 (影片或照片) 才能達到理想的偽造成果，例如不只需要正面的人臉，還需要各個角度、喜怒哀樂的表情、各種的姿勢以及高解析度的畫面才有可能產生品質良好的成品。因此，我們可以以放大畫面、觀察臉部的或頭髮邊緣有沒有不自然的參差感，以及觀察眼睛的方向是不是眼珠無法轉動、永遠望著固定方向的方式來偵測照片或影片是否為造假。

● 常見利用 Deepfake 技術的犯罪類型

1. 造謠影片：透過換臉技術製作或假冒公司高層或知名人物的影片，藉此散布虛假的言論或話題。因此，若影片內容或人物出現與日常言行不一致的影片，就有必要進行查證，判斷是否為名人在社群平臺認證唯一帳號，確認影片是否為所屬公司或公部門官方網址所發佈的。
2. 詐騙取財：歹徒可能利用合成影片與語音，誘騙被害人聽從指示。因此，民眾在採取任何付款或其他行動前應格外警惕。
3. 數位／網路性暴力：不法分子可能利用 Deepfake 技術合成色情影片，即使影片中的人物非受害者，但這樣的行為已對受害者造成無法彌補的傷害。

● **防止變成 Deepfake 的主角，我可以怎麼做？**

1. 替個人照片與影片加浮水印，雖無法完全避免被盜用，但可增加不法分子移除浮水印需要的後製時間成本，因此不法分子可能會放棄盜用你的資料，改選擇其他可以輕易取得影片的對象。
2. 不輕易點選陌生人傳送的奇怪網址，容易中木馬病毒被盜取個人資料與帳號，生活照與影片還可能被當成詐騙集團的素材，間接造成親友受害受騙。
3. 網路購物就算交易失敗、驗證有問題等，也有其平臺的官方客服帳號來聯絡處理，若有陌生買家、賣家假裝熱心引導處理，一定要提高警覺，勿點選任何網址或聽從指示操作銀行帳戶。
4. 如果遇到公務通知，請手機手動直接撥打官方網站公開的洽公業務電話，以及手動輸入官方網域名，切勿貪圖方便不查證就點擊簡訊、對話框中的電話號碼與網址。

● **萬一變成 Deepfake 的主角，該怎麼辦呢？**

1. 熟記五個步驟：冷靜、截圖存證、申訴、報警，以及諮商及輔導。
 - (1) 冷靜：急著爭吵可能會激起加害者情緒、讓影片散步得更廣，許多網路霸凌者最期待看到受害者氣急敗壞的樣子。
 - (2) 截圖存證：將相關內容全部截圖保存，對之後的申訴或訴訟程序非常有幫助。
 - (3) 申訴：若不幸變成 Deepfake 影片主角，應蒐證後向網路平臺申訴或向 iWIN 網路內容防護機構(<https://i.win.org.tw/>)檢舉。iWIN 機構會協助要求網路平臺將影片、照片移除。
 - (4) 報警：電信警察可根據電信公司提供的 IP 使用時間與地點資料，逐步縮小搜查範圍，找出犯罪嫌疑人。警方也會協助要求網路平臺將影像移除。
 - (5) 諮商及輔導：可透過教育部性平教育資訊網及校園輔導通報機制尋求協助；亦可撥打衛生福利部安心專線 1925、張老師生命專線、1980、生命線專線 1995，對外尋求支援避免沉溺於受害者的情境中。
2. 對外尋求支援：若不慎成為 Deepfake 的影片主角，並一直處於沮喪及低落的情緒時，建議尋求教師或家長的協助，或搜尋多元的協助管道，找到一個能相信與傾訴的對象。

身處在新興科技與人工智慧發達的時代，我們更要對網路上諸多事抱持「停看聽」的態度，對於網路資訊，更需要保持警覺心。若看到任何不對勁的圖文或影片，盡量做到求證當事人，並查證訊息的來源；若接到疑似親朋好友聲音的電話，建議先抱持著保留態度，不要輕易地跟對方聊天；遇到無法即時求證的事，建議先秉持謹慎小心的態度，不任意散布或肆意評論，以免不小心成了散布謠言的幫兇。