

AI深伪色情急增，如何才能制止新型性暴力？ | Whatsnew

影像是假的，伤害是真实的；一种新型性暴力正在冒起，而监管仍然滞后。



2023年10月19日，德国汉堡，一名男子在电脑上浏览色情网站。摄：Marcus Brandt/AP/达志影像

特约撰稿人 艾迪

刊登于 2024-03-28

[#AI监管](#) [#人工智能](#) [#性暴力](#) [#数据隐私](#) [#色情产业](#) [#人工智能](#)



在3月19日-20日，全球首个“深伪侵犯”（deepfake abuse）[网上研讨会](#)召开，期间[公布](#)一个震撼的数据：至今全球已知的互联网深伪色情影片超27万条，比起2019年的1.4万条激增18倍，而这些影片的总观看次数多达40亿次。

深伪色情是一种新形态的影像性暴力，指未经本人同意将脸“移植”到性影像。在2017年，深伪移植技术开始进入公众视野，主要指透过人工智能（AI）中的深度学习（Deep Learning）技术合成的伪造图像、影片、声音。只需特定对象的大量影音素材，Deepfake 技术就能制作出拟真的影片。

然而，其诞生很快就伴随针对女性的影像性暴力。根据位于荷兰的身份验证公司 Sensity AI 的[调查](#)，96%的深伪视频是女性性影像，她们并未同意创建这些内容；从2018年到2020年，网络上的深伪色情影片数量大致每六个月就翻一倍。

据研讨会的多名专家指出，公开可用的人工智能图像生成工具，以及可将女性“数码脱衣”的“Nudify”廉价且易于使用，令深伪影片无须像以往那样要收集大量高解像度的图像并掌握机器学习的知识。如今，甚至只需一张真实的图像便可生成极真实的性暴力影像。

2024年1月，音乐天后 Taylor Swift 成为深伪色情的受害者，她的深伪裸照在社交媒体 X 疯传，引起广泛讨论。

“人生的每个面向都受到牵连”

深伪色情产生的影像虽然是假，但对人的伤害是真实而巨大。而深伪色情正日渐普及，无论是普通女性还是名人，都有机会受到深伪色情的伤害。

华盛顿大学科技与政策实验室的 Noelle Martin 在研讨会指出，深伪色情“剥夺我们在世界中自主决定的权利和能力。其影响贯穿一生，影响就业能力、未来收入、人际关系和亲密关系。人生的每个面向都受到牵连。”

对一些公众女性来说，深伪技术也打压她们的言论自由，令她们暴露在这种新型性暴力的风险。据《WIRED》的[报导](#)，实况主播成为深伪色情的受害者，她们感到被侵犯、骚扰，有些内容更传到家人。在台湾，“理科太太”“白痴公主”“地球”“奎丁”、黄捷和高嘉瑜等网络名人也成为被换脸至色情片的被害人。

另外，这种新型性暴力也损害女性的政治参与。如美国白宫性别政策委员会 Cailin Crockett 所言，“从起底（doxxing）到虚假信息宣传，再到深伪视频和非自愿私密影像，这些都威胁女性的政治参与和领导地位，进而削弱民主。”

搜索引擎逃脱法律责任？

就这种新型性暴力，现有的法律和制度未能有效应对，尤其搜索引擎常常免于法律责任。

根据 NBC News 的[报导](#)，搜索引擎是接触深伪色情的主要途径。现时搜寻“deepfake”等相关字眼，Google 和 Bing 的搜索结果并未禁止相关搜寻，甚至在顶部放置深伪色情的网站。虽然 Google 有表格供受害者填写，进而移除相关网址，但 Google 并不主动出击和下架深伪影片。Google 曾在一份声明中表示：“我们明白这些内容对受害者是痛苦的，正积极努力为搜索引擎提供更多保护。像其他搜索引擎，谷歌索引存在于网络上的内容，但我们积极设计排名系统，避免出现令人震惊的有害内容。”

对此，多名专家都呼吁 Google 和其他主要技术公司阻止搜寻结果中出现用于深伪侵犯和其他非自愿色情的网站和应用程式。

在美国，1996年通过的“通讯端正法”（Communications Decency Act）保护科技公司（如Meta、X）不因用户贴文而遭到诉讼和起诉，这条法案如今受到注视和批评。柏克萊加州大学 Hany Farid 主张取消平台公司的法律豁免权，他认为现行法律不鼓励科技公司采取有意义的行动来对抗深伪色情。

“在任何其他行业，如果你受到一家公司的损害，你都可以在法庭上维护自己的权益，希望解决问题。科技行业是唯一一个我们无法在法庭上维护自己权益的行业。”Goldberg 律师事务所的 Norma Buster 如此说道。



2020年3月24日，韩国当局将“N号房”网络性犯罪案主谋赵主彬由警局移送至检察官办公室。

迎头而上的法律监管

除了针对搜索引擎，各国正检视法律，希望监管这种新型性暴力。

在美国，只有大约十几个州的法律禁止制作或散播性暴力深伪影片，很多州正逐步修改他们的非自愿色情法律。

在英国，自今年1月31日起，《线上安全法（Online Safety Bill）》将未经同意分享 AI 生成的私密图像视为非法行为。

在韩国，在2019 年底“N 号房事件”爆发后，性暴力犯罪条文有所修订。2021年1月21日，韩国修正《性暴力犯罪法》相关规定，明文禁止利用深度造假制作虚伪影像等数位性暴力行为。然而，这条法律只处罚“制作、散布虚伪影像者”一方，并未针对“购买、消费虚伪影像”的另一方。

在台湾，2021年警方拘捕百万 YouTuber 小玉（朱玉宸），指其在 Telegram 群组“台湾网红挖面”中利用 AI 换脸技术，将许多女性名人换脸成色情片主角，并将相关的影片售卖谋利，犯案一年多至少获利1000万元台币。

现时针对深伪色情犯罪，台湾受害人可诉诸《刑法》妨害名誉罪及妨害风化罪，另受害人也可依《民法》求偿，因名誉、肖像遭受侵害，提起精神上的损害赔偿。在2023年1月，立法院亦通过防制性暴力的《刑法》部分条文，当中意图散布、播送、交付、公然陈列，或以他法供人观览，以电脑合成或其他科技方法制作关于他人不实的性影像（即例如 Deepfake），足以生损害于他人者，处5年以下有期徒刑、拘役或科或并处50万元以下罚金；若是意图营利，处7年以下有期徒刑，得并处70万元以下罚金。

另外，在台湾，Deepfake 也可构成“非公务机关非法利用个人资料罪”。可处 5 年以下有期徒刑得并科100万元以下罚金；或可构成性骚扰。此外，2023 年1月新修的《性侵防治法》还规定网路业者对于性侵害犯罪嫌疑，应先行限制浏览和移除有关网页资料，并通知警察机关及保留相关资料，若无正当理由而不愿配合，将可处新台币6万至60万罚金。

在香港，有案性利用 deepfake 作诈骗，传转载有事主样貌的移花接木色情短片勒索，另有案性利用 deepfake 技术骗取贷款。至于法例方面，在2021年9月，立法会通过《2021年刑事罪行（修订）条例草案》，当中引入四项新罪行，其中一项“[未经同意下发布私密影像，或威胁如此行事](#)”针对未经同意下发布或威胁发布私密影像的行为，不论该影像是否获当事人同意下拍摄。而发布或威胁发布的罪行同时涵盖移花接木以显示某当事人私密部位的影像（如：DeepFake Porn、深伪色情影像），或显示其进行私密行为的影像。

尽管全球正迎头赶上法律监督深伪色情，但消灭这种新型性暴力并不容易。[有律师指](#)制作人可能使用浮动 IP，难以追查，加上影片是在国外网站流通，社群平台未必配合执法部门调查，很难制裁犯罪者。就算抓到人，影片可能还在网上流通，移除影像需要平台配合。

[#AI监管](#) [#人工智能](#) [#性暴力](#) [#数据隐私](#) [#色情产业](#) [#人工智能](#)

本刊载内容版权为端传媒或相关单位所有，未经[端传媒编辑部](#)授权，请勿转载或复制，否则即为侵权。

華文世界不可或缺的深度報導和多元聲音

端傳媒尊享會員 8 折

含華爾街日報和當今大馬會籍

點擊訂閱

端傳媒的下一程，需要你的守護。今天就成為訂閱會員，支持我們走下去，支持華文世界不可或缺的深度報導和多元聲音。點擊了解更多[會員計畫](#)