

台湾台阶上



温伟中
woonwj@sph.com.sg

“深伪”（Deepfake）技术为视频换脸易容，不仅自娱娱人，也可自欺欺人。随着人工智能（AI）造假技术升级，网络世界进入“有图没真相”的荒谬时代，藏在键盘后的黑手，是否已能借此“抹黑、抹黄或抹红”受害者，进行人格谋杀、影响选举，甚至制造国族对立、挑动战争？我们的社会准备好了吗？且看台湾的近况，思索如何打假求真。

深伪以假乱真 民众如何求真？

你看过几乎能以假乱真的“挖脸”造假影片吗？比如美国前总统特朗普突然开口大唱“我爱你中国”、英女王站在桌子上跳舞？

更恶劣的造假影片“挖脸”或“换脸”，不单开开玩笑，还会让受害人“丢脸”；在社会层面，可以增添敲诈勒索；在政治层面，可能对内制造对立，对外激化仇恨。在这场防打假的战役中，没有人应置身度外。

台湾警方刑事局10月18日逮捕了网红“小玉”朱玉宸（26岁）与两名共犯，以涉嫌散布猥亵物品与妨害名誉罪移送法办。小玉被指用深伪技术造假，偷换上百名女艺人、网红、政治人物和素人的脸，合成并嫁接到成人影片（AV）女优的身体，制成色情片在数千人的群组中贩售。

初步调查显示，小玉的团伙非法获利1100万元（新台币，下同，53万新元），招募会员超过8000人。小玉过后以30万元保释在外。

此案震惊社会，遭殃的政界名人包括民进党立委高嘉瑜和高雄无党籍市议员黄捷，两人都已到警局报案。

小玉被捕当天，蔡英文总统也在面簿贴文呼吁公众关注深伪科技犯罪行为，严正表态将修法反制，也希望网络平台业者出力，一同阻止假假影像伤害无辜者。

虽然蔡英文没有被小玉恶搞，但她也在这一波换脸潮中成了受害者。社交媒体已流传多个视频，将她换脸并合成到清凉装女郎身上，梳着两条小辫子，戴上兔耳朵，完全颠覆总统庄重、拘谨的形象。有评论形容，恶搞黑手简直是“太岁头上动土”。

深伪技术（Deepfake）也称深度伪造，俗称“挖脸”或“换脸”，专指通过人工智能“深度学习”和“伪造”的人体图像合成技术。2017年网上已流传将名人移花接木的性爱影片，过后陆续有假新闻、恶作剧的视频广传，目前也有多个手机应用让公众自娱，瞬间换脸成明星。

小玉原是点阅量达百万的YouTuber，据专家估计，百万YouTuber每月单靠点阅量和互联网平台分润，每个视频可月收30万元。小玉去年4月宣布退出YouTuber，过后悄悄从网红转战“深伪制作人”，靠网友会员费或另外付费指定制作深伪色情片，收入直接飙升三倍到年收千万元。

文化大学国际企业系副教授郭正亮在政论节目中指出，年收千万元只是“刚开始”，小玉估计还有其他收入来源，若未来继续把深伪影片卖到全球华人圈，预计又是一轮暴利。

侦办小玉造假案的警官林建隆（刑事局科技犯罪防制中心主任）接受《联合早报》采访时说，他5月收到举报小玉违法的报案，侦办与搜证五个月后，10月将她逮捕。

他表示，警方当然希望尽快捉人破案，但毕竟侦办速度取决于案件复杂度，更重要的是别让造假影片和假信息扩散疯传。

传播学专家：深伪造假留下永恒印记 事后澄清也很难减轻伤害

“警方可以在辨识和鉴识影片的真伪后，向社会大众说明。但我认为迅速澄清和下架，才是阻止事态变严重的最重要方式。澄清需要当事人出面，下架需要科技平台、政府主管机关去处理。尤其在攸关社会、族群和政治对立



深伪技术俗称“挖脸”，专指通过人工智能“深度学习”和“伪造”的人体图像合成技术。虽然可以减低电影制作成本，但若广泛被用于造假，就会对个人和社会造成困扰和负面冲击。（iStock）

的公共议题上，必须由各方齐头并进，合力防打假。”

台湾传播与政治学者认为，深伪造假的伤害有如“数位刺青”（数位本地称数码），会留下永恒印记，即使事后澄清也很难减轻伤害。要对付深伪“病毒”，社会须加强立法和媒体素养，形成全社会防堵机制，个中挑战是既要保障言论自由，也要精准打击犯罪者。

资深媒体人出身的铭传大学广播电视学系主任杜圣聪受访时说，从2009年科幻电影《阿凡达》（Avatar）的外星巨人，到2015年《玩命关头7》（Furious 7）用技术合成法让已故男主角保罗沃克“复活”续演，到自娱娱人的变脸手机应用，都能看到深伪技术的优点，既可减低电影制作成本，也让人留下美好身影。

针对深伪技术的负面效应，他指出，只要明显是猥亵、伤害名誉、敲诈勒索，都有相关法令可依法办理。

但杜圣聪坦言，对于纯属深伪、恶搞的行为，以受害者受到的伤害、加害者获得的巨大利益来看，台湾现有法律完全不符合比例。这包括相关刑罚都可以“易科罚金”，以缴钱取代坐牢，或被捕后付远低于犯罪所得的保释金，一转身就重获自由。

他说，深伪带来的伤害“基本上是不可逆的”，有如“数位刺青”，就算要求下架，也往往已被广传，在网上留下“永恒的印记”。

他更形容深伪技术在法律漏洞中生存，有如德尔塔变种病毒，“因为是轻症，更容易感染和广泛传播，令人防不胜防”。“面对病

毒，就要打疫苗，推动数位法治、媒体素养教育的‘疫苗追加剂’。打一剂不会终生免疫，社会必须持续对话找共识，不能延宕立法，也要持续提升社会的媒体素养和分辨能力。”

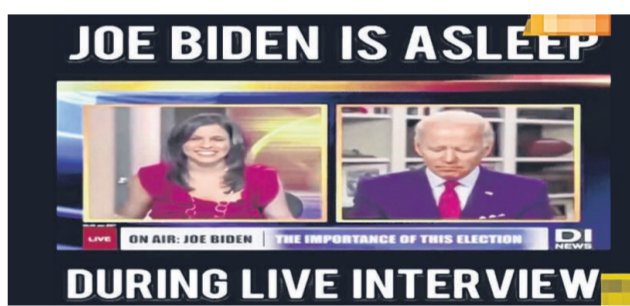
学者：须加强防治机制 但不能变质为钳制言论自由

中国文化大学广告学系主任钮则勋教授说，传播科技一般都能促进人类互动，增添乐趣和便利，但目前深伪科技的负面影响比较明显，包括剥削女性尊严、歹徒借此勒索欺诈，尤其在选举时会冲击选情，也可能造成社会冲突对立、激化跨国跨族的仇恨。

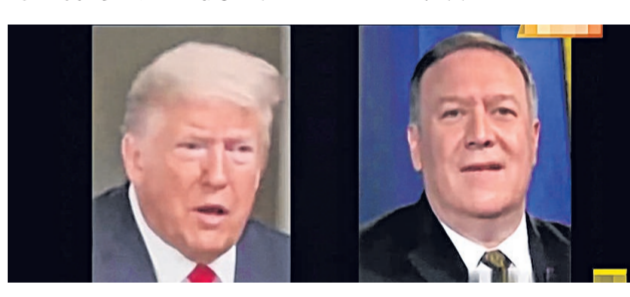
他认为，政府必须深刻思考和厘清这项科技的利弊，制定适切的法令。因为，若在选举期间有人利用深伪科技造假影片，向对手“抹黑、抹黄或抹红”或为自己漂白，“绝对可能影响选举，要防控与控管也很难”。

“不明就里的网民会觉得是真的，影响他的观感。因为影片很劲爆、冲击力强，分享广传的效应会很强。这将导致选举竞争、政党与候选人的攻防冲突都加大，政党或候选人的侧翼网军和水军，也会热衷于这样的操作。如此一来，政党与候选人会疲于奔命做澄清，但刻板印象已经形成，澄清的力度已无法抵消深伪影片第一次上架的冲击力。”

钮则勋指出，必须加强防治机制，精准打击犯罪，但也不能变质为钳制言论自由。如果政府借由打击和管制深伪，把手伸到管制网络言论自由，反而会制造寒蝉与噤声效应。



去年8月疯传的视频显示，时任美国民主党总统候选人、现任总统拜登接受电视台女主播专访时打瞌睡，过后真相大白，视频是嫁接多年前另一人受访时睡着的尴尬一幕。（互联网视频截图）



打着反华大旗的美国前总统特朗普（左）和时任国务卿蓬佩奥，在深伪抖音（TikTok）短片中合唱《我爱你，中国》。（互联网视频截图）

深伪视频政治恶搞 美大选前丑化拜登形象

深伪视频用于政治恶搞，对政治人物形象和选情的冲击也引起关注。

去年11月初的美国总统大选，在8月就已疯传一个视频，显示时任民主党总统候选人、现任总统拜登接受电视台女主播专访时打瞌睡，女主播还接连喊了几声“Wake up（醒醒）”。如今真相大白，原来视频是嫁接多年前另一人受访时打呼的尴尬一幕。这个造假视频反映出拜登敌对阵营试图塑造和强化“瞌睡乔”（sleepy Joe）的昏庸形象。

到了9月，轮到时任美国总统特朗普的深伪抖音（TikTok）短片作品出炉。只见他和时任国务卿蓬佩奥合唱《我爱你，中国》，澎湃激昂地诉说“我爱你春天蓬勃的秧苗”“你家乡的甘蔗，好像乳汁滋润着我的心窝”。笑果十足之余，难免让一些选民感到混乱，打着反华大旗的两名美国政要，怎么会这么爱中国？

许多分析指出，无论拜登打瞌睡，或特朗普高唱爱中国，这些深伪影片的动机并不单纯。它们既可能来自敌对阵营，也可能是候选人自导自演，精心打造被霸凌的“受害者形象”，从而收获政治利益。

我国通过POFMA管制社媒假消息

作为对国际媒体全面开放的国家，新加坡也面对社媒散播大量假消息，以及被社媒平台操控的威胁，无论是面簿或它旗下的WhatsApp平台，都充斥着无数假消息和谎言。对此，国会在2019年通过了《防止网络假信息和网络操纵法令》（POFMA）加以应对，管制互联网平台在本地地的行为。

根据POFMA，针对影响大且广泛传播的假信息，政府可发出广泛更正指示，要求主流媒体和社交媒体平台等刊登相关更正，向广大民众澄清事实真相。

创办“事实查核中心” 台10人团队助民众辨真伪

在台湾，民间在三年前创立的“台湾事实查核中心”，如今已成为市民查证信息真伪的“好帮手”。学界也努力开发共享辨伪工具，帮助公众“以AI打AI”辨识深伪真面目。

记者走访位于台北中山区的台湾事实查核中心，发现这个10人团队充满“打爆假新闻”的战斗力的。为了贯彻“排除政府资助”、只靠民间捐款的宗旨，中心选择贴近市民的中路线，但也因此遭到不同政治阵营的攻击。

在政治色彩浓厚的台湾社会，这样一个组织会不会被质疑“站哪边”？尤其是在澄清不利政府的消息时，会不会被怀疑“跟政府走得最近”？

事实查核中心总编审：有时蓝的说你绿 有时绿的说你蓝

台湾事实查核中心总编审陈慧敏回答《联合早报》时说：“好题目都有争议。有时蓝的说你绿，有时绿的说你蓝。无论蓝绿我们都有中枪，做哪一题都会被骂。有时还被抹红成共同路人，我们被贴的标签还蛮多的。所以基本上是成功的。这也代表你要在枪林弹雨中端出查核报告。”

针对深伪事件，她透露今年年初美国前总统特朗普承认败选时，有传言说是深伪造假事件，她马上请科技专家帮忙，最后查证影片如

假包换，并非伪造。

陈慧敏说，目前中心收到的假消息以移花接木的照片和事件为主，但海内外的事实查核界、科技平台都密切关注深伪技术发展，认为该技术的成本一旦降低，影像疯传的杀伤力就更大，打破多数人赖以辨识真假的眼见为实原则，令公众真假难辨。

她透露，深伪技术也会打击新闻自由和民主发展，曾有印度独立调查记者揭弊，该女记者的头像过后被制作成色情影片，对她进行人格毁灭。“选举紧逼时，若出现政治人物的深伪影片，也会来不及澄清而影响选情。”

中心目前与社交媒体面簿、Line与科技平台谷歌都有合作，也为市民开发出获得新闻奖肯定的“查核工具箱”“谣言捕手”等工具，会在受询时查证谣言，或在重大公共议题如选举做票、化学兵进城喷毒气等传言发生时，独立撰写查核报告，主动推送给用户与媒体。

开发辨伪工具在网页上与公众分享

曾协助陈慧敏辨识深伪影片的成功大学数据科技研究所助理教授许志仲，是研究人脸辨识及合成技术多年的学者。他正开发辨伪工具，并已在网页上与公众分享。

许志仲建立了“以AI打AI”的新模式，分析现有深伪技术，找出共同伪造线索，让防伪AI



台湾事实查核中心总编审陈慧敏认为，一旦成本降低，深伪影像的杀伤力就更大，打破眼见为实的原则，令人真假难辨。（温伟中摄）



成功大学数据科技研究所助理教授许志仲建立了“以AI打AI”的新模式，开发与市民共享的辨伪工具。（受访者提供）

系统主动侦测与辨识真伪。

“要让科技帮到社会，必须学界、业界和政府三方合作，也建制至少半公开的系统，让公众可以使用，迅速传进影片，让系统侦测其真伪。我认为必须平心静气看待新科技，制止负面效应，也让它发挥最大的正面效应。”