

國立臺灣大學法律學院法律學系

碩士論文

Department of Law
College of Law

National Taiwan University

Master Thesis



論深偽（deepfake）不實影像之刑事管制

—以保護法益為核心

Criminal Regulation of Deepfake

-Focus on Legally Protected Interests

秦偉翔

Wei-Hsiang Chin

指導教授：謝煜偉

Advisor: Yu-Wei Hsieh, Ph.D.

中華民國 112 年 6 月

June 2023

謝辭



終於啊，也把這條路給走完了。

三年的研究生生涯，說顛簸有些言重，但也絕對算不上風平浪靜。

放榜的當下身處悲劇、入學的時刻帶著極致的抑鬱開局、為了司法官多兜轉備考了一年、讀書讀膩所以離經叛道地跑去實習、一度決定休學卻又在一週內反悔，宛若迷茫了整路，卻終於在最後半年返璞歸真，開始真心感受到研究的快樂。但也慶幸，最終交出的成果能無愧於心吧。

感謝寶寶師的指導跟溫柔，讓我即便帶著一些徬徨，仍然能安定地前行。在撰寫論文的旅途中，其實自己有過無數次不安，自覺不是那種能夠真正醉心於學術、熱愛研究的人，因此偶爾擔心自己會德不配位，也惶恐自己無法符合老師的期待。但真的感謝老師即便公務傍身、案牘勞形，卻總是充滿耐心的導引，讓我最終能順利完成研究所的學涯。

感謝在實習時候照顧過我的諸多前輩們，尤其是 Nigel 老師、信瑩律師與訢慈，感謝你們的存在讓我在未知的人生迷霧中，逐漸能夠勾勒出一幅藍圖，從一些偉岸的身影中，找到了自己應該邁步向前的方向。也感謝你們在實習旅途中對我的溢美與期待，這些對我而言都是重要的動力，足以督促自己不能停下腳步，必須鞭策自己到心之所向的彼方。

感謝我的家人，爸、媽、弟弟、爺爺、阿嬤、姨、舅舅、舅媽跟所有愛我的人，感謝你們讓我在不缺愛的環境中成長，讓我始終無須懷疑自己值得被愛，因此有信心能夠克服人生的所有難關，而我也愛著你們。

感謝吾友海莉，13 年的相識已經超過我們一半的人生，雖然在各自人生各個階段的交友圈中，我們或許很少是彼此最親暱的存在，但值得慶幸的是，你我都能夠是彼此最後的、不會崩塌的港灣，那個無論風雨飄渺都能安身立命之所，承接著彼此的悲喜與苦痛，脆弱與不堪。我珍惜著過往每個與你相濡以沫的時刻，也期待未來每個攜手前行的風景，願你我都能活成心目中理想的模樣。

感謝武陵演辯的學弟妹們，閔、達、小玥、思昀，感謝有你們陪我歷經人生的榮光，也慶幸是你們伴我走過人生的晦暗。作為一個驕傲又自卑的玻璃心，我自知自己的那些或鋒芒畢露，或抑鬱 emo，都不是容易消化與承接的情緒，還好你們都是如此聰明的存在，有些是正道的光，有些是在道德的長程邊緣反覆試探，才能讓我成功活過魔法大戰的地獄，總之，感謝你們。

感謝瑪瑪的大家，U、Ben、Phie、艮，感謝有你們代表著我的青春，讓我每每在想起高中回憶時，總慶幸曾經經歷過一些青春的甜酸。真的既喜歡也珍惜不定期的飯局與每年的小旅行，能夠陪伴彼此一起走過人生的各個階段，看著彼此成長，真的是太好了！希望我們可以一起去到更多更酷的遠方。

感謝忙內這段旅途的陪伴。其實對於該不該留下這段掙扎了一陣子，但最終細想，在我開啟論文副本模式之後，一路上有著最多交流的人類毫無疑問的

絕對是你，就覺得還是滿值得留下一個段落紀念的吧。與你的交流跟互動多半時間都是有趣的，夾雜著一些我對於未竟之路的遺憾與寄託、一些對於靈動機敏的嚮往、一些可以放心凡爾賽的自在、一些證明尚能飯否的展示、一些小作文，還有很多很多的偷臭。總之感謝有你讓這段旅途變得繽紛，也祝福你的肝也能始終保持繽紛（有夠真摯的祝福無誤（燦笑。

感謝成日被我摯友限動 emo 轟炸的各種人類們 XDD 我自知有著一躁起來就有夠愛碎念的爛個性，也有著一旦有了成就尾巴就想翹起來的擊敗模樣，所以特別謝謝大家包容我的各種敏感與不安（合掌）。也感謝好多人不時的找我出門吃飯，很喜歡聽著你們分享人生故事，讓我對於這個世界有著更多的觀點與視野，希望將來的人生依然能與你們時常聚首。

最後一段感謝自己吧，感謝你總是如此的驕傲又任性、好強而不服輸，總是覺得自己被低估而渴望變得強大。雖然你這種得隴望蜀又毛一堆的爛個性大概貢獻了你八成的負面情緒，但也是因為這些執念與慾望，好高騖遠又不安於室，才促使你得到現在所擁有的一切。希望在未來你能好好跟自己的脾氣和平共處，找到適當的平衡與舒服的出口，期待你能走到一個比起現在的自己所能想像得，更美好遠方。

該開啟下一場戰鬥了呢！

2023 年 6 月，於長興 BOT

摘要



近年來隨著科技發展，deepfake 科技已經使一般人都能夠輕易地產出高度擬真的不實圖影，在帶給世人便捷與效益之餘，也同時因為不肖人士的濫用而為社會帶來了許多負面衝擊，其中包含許多與性相關之應用，然而與性無關之濫用所帶來的麻煩也不容忽視。

為了因應 deepfake 帶來的衝擊，我國於 2023 年通過刑法修正案，新增了若干針對不實影像之刑事規範。然而自修法歷程中的討論內容觀之，立法者對於不實影像可能侵害的法益為何缺乏定見，而學說上對此的討論亦莫衷一是，甚至對於修法前之既存法規是否足以因應 deepfake 帶來的衝擊都存在爭議。前述對於保護法益的爭論，除了與是否應另立專門之法規處理此類議題之決策有關外，尚會影響條文之構成要件應如何設計。

因此，本文旨在以探討不實影像可能涉及之保護法益為核心，進而分析出應該如何以刑法對不實影像加以規制較為適當，期待能藉此提供我國適當的立法建議。

本文首先將討論標的區分成「有混淆之虞的不實性影像使用」(類型一)、「無混淆之虞不實性影像使用」(類型二)、「有混淆之虞的與性無關之不實影像使用」(類型三)與「無混淆之虞的與性無關之不實影像使用」(類型四)等 4 種類型。

隨後透過文獻回顧作為方法，本文逐一分析各類型不實影像可能涉及的保護法益，並在最終發現前述之類型二與類型四基於客體欠缺足夠的欺詐性，並未侵害法益，因此不應受刑法規制；而針對類型一，對於目前主流意見所圍繞的名譽法益與隱私法益兩種見解本文皆採取反對立場，而認為應該以「形象法益」作為其保護法益；最後針對類型三，本文認為其確實為社會帶來前所未有，且既存之法規難以充分處理的新衝擊，其影響力與現行法下選擇處罰之無形偽造類型相當，因此應該修法加以因應以維護公共信用法益。

在釐清了保護法益後，本文以美國州法為主要對象，透過比較分析的方法，發現在規範客體、客觀構成要件行為、主觀意圖等若干要件設計上，各法規之間存有取捨上的歧異。本文在逐一分析所有歧異之後，以此為基礎批判了 2023 年度我國關於不實影像之修法存有不恰當之處，並在最終提出具體的條文設計作為修法建議。

關鍵詞：深度偽造、deepfake、隱私權、名譽權、形象法益、公共信用

Abstract



In recent years, with the development of technology, deepfake technology has made it easy for ordinary people to create highly realistic false images. While providing convenience and benefits to society, it has also brought about many negative impacts due to the misuse by unscrupulous individuals. This includes numerous applications related to pornography, but the troubles caused by non-sexual misuse should not be ignored.

In response to the impact of deepfake technology, our country passed an amendment to the Criminal Code in 2023, which introduced several criminal regulations targeting false images. However, based on the discussions during the legislative process, it is evident that lawmakers lack a clear understanding of the legal interests that false images may violate. There is also a lack of consensus in academic discussions regarding this matter, and even the controversy surrounding whether the existing regulations prior to the amendment are sufficient to address the impact of deepfake technology. The aforementioned debates on legal interests not only relate to whether to establish separate regulations to address such issues but also to how the constituent elements of the relevant provisions should be designed.

Therefore, this article aims to focus on legal interests involved in false images and analyze how the criminal law can appropriately regulate them. The intention is to provide suitable legislative recommendations for our country.

The article first categorizes the subject of discussion into four types: "the use of sexual false images with potential confusion" (Type 1), "the use of sexual false images without potential confusion" (Type 2), "the use of non-sexual false images with potential confusion" (Type 3), and "the use of non-sexual false images without potential confusion" (Type 4).

Subsequently, through a literature review, this article analyzes the legal interests that may be involved in each type of false image. It concludes that Type 2 and Type 4, due to their lack of sufficient deception, do not violate legal interests and should not be subject to criminal regulation. Regarding Type 1, this article takes a position opposing the mainstream views surrounding reputation and privacy interests and argues that "image interests" should be the protected legal interest. Finally, concerning Type 3, this article recognizes its unprecedented impact on society and the inadequacy

of existing regulations to address it adequately. Therefore, the law should be amended to protect the legal interests of public credit.

After clarifying the protected legal interests, this article primarily focuses on state laws in the United States. Through a comparative analysis, it identifies disparities in the design of several elements, such as the regulated object, objective constituent elements, and subjective intent, among different jurisdictions. Building on the analysis of these disparities, the article criticizes the 2023 amendment in our country regarding false images for its inadequacies and ultimately presents specific provisions as legislative recommendations.

Keywords: Deepfake, Privacy, Reputation, Image, Public Credit

目次



摘要

ABSTRACT

目次

第一章 緒論

第一節 研究背景與問題意識	I
第二節 文獻回顧	IV
第一項 文獻簡介	VI
第二項 文獻檢討	1
第三節 研究架構與方法	8

第二章 DEEPFAKE 應用之類型化與既存之實定法規制

第一節 DEEPFAKE 應用之類型化	8
第一項 正面之應用	8
第一款 娛樂機能	8
第二款 教育機能	8
第三款 自我實現機能	9
第四款 治療機能	9
第二項 負面之應用	9
第一款 是否與性相關	10
第一目 DEEPFAKE 技術實際應用情況	10
A. 與性相關之應用	10
B. 與性無關之應用	11
第二目 類型化之緣由	11
A. 學說見解	12
B. 立法體例	12
第二款 是否有致使理性一般人混淆之虞	13
第一目 DEEPFAKE 技術實際應用情況	13
第二目 類型化之緣由	15
A. 學說見解	15

B. 被害人感受	15
C. 立法體例	16
第二節 既存實定法下之規制狀況	18
第一項 有混淆之虞的不實性影像使用（類型一）	19
第一款 學說見解	19
第二款 刑法第 315 條之 2 第 3 項散布竊錄影音罪要件檢討	19
第三款 刑法第 310 條誹謗罪要件檢討	19
第四款 刑法第 235 條散布猥褻物品罪要件檢討	21
第一目 客觀構成要件不該當	21
第二目 無法充分評價法益侵害	24
第五款 個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪要件檢討	25
第一目 學者見解	25
第二目 本文見解	26
A. 「意圖損害他人之利益」之要件應僅限於損害財產上利益之情形	27
B. 未營利之散布者難以構成「意圖為自己或第三人不法之利益或損害他人之利益」之要件	28
C. 本罪應僅限於「具組織性、系統性的資料蒐集、處理或利用行為」	29
D. 後續散布者之行為難以被認為係散布「個人資料」	30
第六款 結論	32
第二項 無混淆之虞的不實性影像使用（類型二）	33
第一款 刑法第 310 條誹謗罪要件檢討	34
第二款 個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪要件檢討	34
第三款 結論	35
第三項 有混淆之虞的與性無關之不實影像之使用（類型三）	36
第四項 無混淆之虞的與性無關之不實影像之使用（類型四）	37
第五項 總結	37

第三章 各類型不實影像使用所涉之保護法益與應罰性 39

第一節 入罪化之標準	39
第二節 有混淆之虞的不實性影像使用	42
第一項 被害人受傷情況	42
第二項 所涉及之法益	45
第一款 隱私權	45
第一目 實務見解	45
第二目 學說見解	46
第三目 本文見解	47



A.隱私權的類型化	47
B.隱私權應僅限於「對於秘密資訊的刺探與洩露」	48
第二款 名譽權	52
第一目 實務見解	52
第二目 學說見解	53
第三目 本文見解	54
A.實然層次	55
B.應然層次	57
C.結論	60
第三款 性自主權	62
第一目 實務見解	62
第二目 學說見解	62
A.積極面向的性自主	62
B.消極面向的性自主	63
第三目 本文見解	63
A.積極面向的性自主	63
B.消極面向的性自主	65
第四款 性相關的冒犯性/仇恨性言論	68
第一目 仇恨性言論未必使弱勢群體處境更為不利	69
第二目 定義仇恨性言論之標準困難，且可能因此導致弊害	71
第三目 不實性影像之使用難以被認為是仇恨性言論	73
第五款 形象法益	74
第一目 實務見解	74
第二目 學說見解	75
第三目 本文見解	75
第三節 無混淆之虞的不實性影像使用	81
第一項 實務見解	81
第二項 學說見解	82
第三項 本文見解	84
第一款 學說基礎	84
第二款 比較法對人格法益之立法模式	85
第三款 被害人自述之傷害狀況	86
第四項 結論	86
第四節 有混淆之虞的與性無關不實影像之使用	89
第一項 實務見解	89
第二項 學說見解	92
第三項 本文見解	93
第一款 新興技術確實對社會帶來有別以往之新衝擊	93

第二款 現行法規不足以充分因應新衝擊	97
第一目 偏重於處理「對於被描繪者之負面描述」，難以處理特定案件類型	97
第二目 僅是正面表列劃定保護範圍，因而有所缺漏	99
第三目 偏重於單次事件所生的影響，忽略系統性之傷害	100
第三款 所保護之法益	101
第五節 無混淆之虞的與性無關之不實影像之使用	104
第一項 實務見解	104
第二項 學說見解	105
第三項 本文見解	105
第六節 總結	105

第四章 不實影像規制之美國州法介紹與我國立法評析 106

第一節 比較法介紹	106
第一項 簡介	106
第二項 TEXAS STATUTES ELECTION CODE SEC. 255.004. TRUE SOURCE OF COMMUNICATION. (德州選舉法)	106
第三項 CALIFORNIA CIVIL CODE SECTION 1708.86 - CAUSE OF ACTION FOR DEPICTED INDIVIDUAL. (加州民法)	107
第四項 CODE OF VIRGINIA 18.2-386.2. UNLAWFUL DISSEMINATION OR SALE OF IMAGES OF ANOTHER; PENALTY. (維吉尼亞州刑法)	111
第五項 NEW YORK 52-C. PRIVATE RIGHT OF ACTION FOR UNLAWFUL DISSEMINATION OR PUBLICATION OF A SEXUALLY EXPLICIT DEPICTION OF AN INDIVIDUAL. (紐約州民法)	112
第六項 CALIFORNIA 2021 CALIFORNIA CODE ELECTIONS CODE SECTION 20010 - CAMPAIGN MATERIAL WITH SUPERIMPOSED IMAGE OF PERSON PROHIBITED. (加州選舉法)	112
第七項 MASSACHUSETTS PETITION (ACCOMPANIED BY BILL, HOUSE, NO. 1755). (麻州草案)	114
第八項 THE DEEP FAKES ACCOUNTABILITY ACT (DEEP FAKES 法草案)	115
第九項 比較法整理	120
第一款 主觀要件	120
第一目 用以限縮所規範之不實影像的種類範圍	120
第二目 對於不實性影像之使用，用以限縮成立不法之範圍	120
第三目 對於不實性影像之使用，用以調整法律效果	120
第四目 調整不同種類行為人間，成立主觀不法之要件	121
第二款 所規範之客體之欺詐性	121
第一目 自製作時之主觀意圖加以限縮	121
第二目 自對於客體之形容詞描繪加以限縮	121
第三款 所規範之客體之種類限制	122
第一目 載體形式之區別	122



第二目 載體主題之區別	122
第四款 客觀上構成要件行為	123
第一目 僅製作階段即可能成立不法	123
第二目 需達揭露或散布階段始成立不法	123
第三目 除揭露或散布外，尚需符合其他構成要件始成立不法	123
第五款 警語、免責聲明或浮水印之定性與效果	124
第一目 明文規定加註後得以阻卻不法	124
第二目 明文規定即便加註亦不能阻卻不法	125
第六款 明文之豁免條款	125
第二節 2023 年修法評析與建議	125
第一項 刑法第 319 條之 4 與其相關規定	126
第一款 主觀要件	126
第一目 以「脅迫、騷擾或恐嚇」作為主觀意圖要件	126
第二目 以「脅迫、騷擾或恐嚇」作為加重處罰之要件	127
第三目 針對不同種類行為人，設計相異之主觀不法要件	128
第二款 所規範之客體之欺詐性	129
第一目 自製作時之主觀意圖加以限縮，不足以排除無混淆之虞的散布不實性影像行為成立犯罪	129
第二目 對於客體之描述	129
第三款 所規範之客體之種類限制	131
第一目 應包含圖像	132
第二目 不應包含音檔	132
第四款 客觀上構成要件行為	133
第一目 不應於製作階段即予以處罰	133
第二目 除散布行為外之其他客觀構成要件	135
第五款 明文之豁免條款	138
第一目 因欠缺混淆潛能而豁免	138
第二目 因有正當目的而豁免	139
第六款 不應以「意圖營利」作為加重刑罰事由	140
第七款 結論	141
第二項 刑法第 339 條之 4	142
第三節 其他修法建議	143
第一項 客觀構成要件行為，須達「行使」階段始有應罰性	144
第二項 仍以「有使人混淆之虞」作為要件	145
第三項 「足以生損害於公眾或他人者」之認定	145
第四項 結論	146
第四節 總結	146
第一項 針對刑法第 319 條之 4 與其相關規定	146

第二項 針對刑法第 339 條之 4 第 4 款規定

第三項 新增刑法第 220 條之 1 規定

第五章 結論

第一節 研究成果

第二節 研究限制、反思與展望

第一項 關於有混淆之虞且與性無關之不實影像（類型三）

第二項 關於無混淆之虞且與性相關之不實影像（類型二）

參考文獻

一、中文文獻

（一）專書

（二）期刊論文

（三）政府文書

（四）網路文獻

二、英文文獻

（一）期刊論文

（二）網路文獻



147

147

148

148

152

152

152

154

154

154

154

156

157

157

157

159

第一章 緒論



第一節 研究背景與問題意識

早自 2018 年起，在美國即出現了利用 deepfake 技術製作不實性影像的社會現象，其中大多是針對女性為描繪對象之非合意色情（nonconsensual porn）作品¹。而其所影響的對象早已不限於名人，在著名社群平台如 Reddit 上即曾出現子論壇（subreddit），或 Discord 也曾經出現聊天室，供 deepfake 性影像製作者間相互交流，使彼此在製作各自重要他人之不實性影像時，技術得以切磋進步²。研究顯示，有高達 70% 的 deepfake 色情影音受害人都只是一般的個人，而非著名公眾人物³。

而 deepfake 所生之不實影像的影響範圍絕非僅止於與性相關，例如藉此登入他人以臉部辨識機制驗證的系統，擅自進行金融交易⁴，亦或是偽造出肉票正受自己控制的不實影像，藉以成功向關係人勒贖⁵，都是此類新技術所可能帶來的社會風險。更遑論在政治場域上，無論是在馬來西亞曾發生偽造內閣大臣與其敵對政治人物之性愛影像之事件⁶，亦或是在烏克蘭正逢烏俄戰爭之際，曾流出不實影像，偽造總統出面向烏克蘭士兵喊話，要求他們放下武器停止戰鬥的影音⁷。如此種種皆顯見 deepfake 作為具革命性之新技術，其確實對社會帶來過去不曾想像之衝擊。

之所以使 deepfake 之影響力有別於過往既存之影音編輯工具，核心緣由無非在於其高度擬真的特性。其運作原理係透過「大規模生成對抗網絡高真度自然圖像合成訓練」，利用人工智慧學習所製成，以至於其真實性甚至與可以與實際之照片難以區別⁸。

¹ Haoarchive (2021), Deepfake porn is ruining women's lives. Now the law may finally ban it, in <https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/>. (最後瀏覽日：07/13/2023)

² Harris (2019), Deepfakes: False Pornography Is Here and the Law Cannot Protect You, *Duke Law & Technology Review* 17, 101.

³ Ray, Bot Generated Fake Nudes Of Over 100,000 Women Without Their Knowledge, *Says Report*, in: <https://www.forbes.com/sites/siladityaray/2020/10/20/bot-generated-fake-nudes-of-over-100000-women-without-their-knowledge-says-report/?sh=7507150b7f6b>. (最後瀏覽日：07/13/2023)

⁴ 許恒達 (2022)，〈深度偽造影音及其刑法規制〉，《法學叢刊》，第 265 期，頁 5。

⁵ Chesney, Citron (2019), Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review*, Vol. 107, Issue 6, 1772.

⁶ Thombre (2021), Deconstructing Deepfake: Tracking Legal Implications and Challenges, *International Journal of Law Management & Humanities*, Vol. 4 Issue 4, 2270-2271.

⁷ 雷喻翔 (2022)，〈臉，為何會被挖走？深度偽造(DeepFake)介紹〉，《清流雙月刊》，第 42 卷，頁 29。

⁸ 余和謙 (2019)，〈人工智慧之治理—以深度偽造為例〉，《科技法律透析》，第 31 卷第 8 期，頁 53-56。

為了因應如此衝擊，世界多國都陸續修法因應。以美國為例，即有德州、加州、紐約州、維吉尼亞州陸續修正 5 部法規，新增規定以處理 deepfake 技術所帶來之社會影響。

而在我國，與 deepfake 相關之議題被大眾關注之濫觴，應是有知名媒體於 2021 年 5 月報導，指稱國內知名 Youtuber 涉嫌於 2020 年至 2021 年間，將我國多名女性藝人、政治人物、Youtuber 等之臉孔，「換臉」至成人影像上藉以營利⁹。在該事件經檢調偵查後引起全國性的高度關注，於 2021 年 10 月時時任行政院長蘇貞昌、法務部長蔡清祥即於國會備詢時表明欲對此修法因應¹⁰，而於該年 11 月間法務部即通過修法草案，增訂「製作或散布他人不實性影音罪」陳報行政院審查¹¹。而此波修法意向獲得了朝野立委跨黨派之共識，最後共計有四大政黨之委員共提出 19 份修法草案版本，與法務部所提之版本一併交付至國會進行審查，最終於二讀黨團協商階段始決定僅討論法務部提交之版本¹²，而於 2023 年 1 月三讀通過刑法修正案，新增了刑法第 319 條之 4 規定，處罰製作與散布不實性影像之行為。

除了針對與性相關之不實影像新增處罰規定外，其實針對與性無關之不實影像，在 2023 年度之刑法修正討論過程中，亦有共計 7 個版本之修法草案試圖增設規定對其做出規範，惟僅因為擔心立法效率等技術性理由而未於本次修法一併通過¹³。惟從此亦可見得，與性無關之不實影像使用，亦是立法部門有意關照之社會現象。針對與性無關之不實影像使用進行規制的一般性規定，截至本文完稿之 2023 年 6 月時，皆尚未於國會通過。惟於 2023 年 5 月時，立法院三讀通過之刑法修正案，新增了刑法第 339 條第 4 項第 4 款加重詐欺罪之規定，將利用不實影像完成詐欺行為作為新的詐欺罪加重事由，亦是此波因應 deepfake 技術衝擊現象下的產物。

然而在前述看似順利、平穩的修法歷程中亦非全無異音。例如在修法前、後皆有國內學者明確指出應以既存之刑法規範加以因應 deepfake 之使用即可，無須特別為此另訂新法¹⁴。又如於修法歷程中，提案機關法務部認為刑法第 319 條之 4 所保護之法益為隱私權¹⁵，惟此見解經司法院於其審查加註意見中明確

⁹ 鏡傳媒 (05/06/2021)，〈臉被偷走之後—無法可管的數位性暴力？台灣 Deepfake 事件獨家調查〉，<https://www.mirrormedia.mg/projects/deepfaketaiwan/>（最後瀏覽日：07/13/2023）

¹⁰ 中國時報 (10/19/2021)，〈法務部祭「小玉條款」1 個月內提修法重罰科技犯罪〉，<https://www.chinatimes.com/realtimenews/20211019002132-260407?chdtv>（最後瀏覽日：07/13/2023）

¹¹ 三立新聞網 (11/17/2021)，〈防再有小玉「深偽」換臉 法務部公布修法 5 大重點〉，<https://www.setn.com/News.aspx?NewsID=1028157>（最後瀏覽日：07/13/2023）

¹² 立法院公報第 111 卷第 84 期委員會紀錄頁 123。

¹³ 立法院公報第 111 卷第 84 期委員會紀錄頁 117。

¹⁴ 相關論述：許恒達 (2022)，前揭註 4，頁 16；林琬珊 (2023)，〈「性影像」與隱私之刑法保護——新修刑法「性影像」規定之評析〉，《月旦法學雜誌》，第 333 期，頁 51；古承宗 (2023)，〈刑法作為抵制「數位暴力」之手段——以惡意影音為例〉，《月旦法學雜誌》，第 333 期，頁 19；陳俊偉 (2023)，〈重思深度偽造影音的入罪化——以人工智慧的科技風險規制為思考方向〉，《月旦法學雜誌》，第 333 期，頁 37。

¹⁵ 立法院公報第 112 卷第 14 期委員會紀錄頁 164。

反對，而認為應以名譽權作為該條文之保護法益為宜¹⁶。於黨團協商階段，與會委員間亦各持己見，莫衷一是，最終法務部代表甚至指出¹⁷：「這個保護法益到底是多少，有沒有包括名譽權，這個部分可能得慢慢靠學說與實務上的發展。」，而在未有明確之結論狀況下即倉促通過修法。

而學說上，對於不實性影像使用可能涉及何種法益侵害之議題上，亦存在一定程度的意見分歧。有明確認為係屬名譽權侵害者¹⁸，亦有指出應屬「性隱私或性人格」者¹⁹，此外尚有認為係屬「形象法益」者²⁰。整體而言，在學界對於不實性影像使用涉及之保護法益為何，仍處於意見百家爭鳴之狀態。此外，關於若欲以專法對不實性影像之使用加以規制，構成要件之設計上應如何因應為宜，亦有深思之必要。在比較法上，美國於已頒布之各州立法中，關於所規範之客體應如何以法條文字加以描述、所規範之構成要件行為為何，皆存在若干差異，其中某些的價值選擇更是南轅北轍。然而在我國目前學說討論上，由於比較多的學者皆採取不應另立專法，而主張以既存之實定法加以規制即可之態度，因此關於若欲另立專法時其構成要件應如何設計，亦相對欠缺討論。

最後，目前國內學說討論之量能，高度集中在針對與性有關的不實影像規制上，而對於與性無關之不實影像可能帶來之衝擊、刑法應否或應如何應對，討論上皆相對匱乏。惟如前所述，即使是與性無關之不實影像，在比較法經驗上業已發現會對社會帶來新的衝擊，而我國亦有立法者正蠢蠢欲動可能正準備推動此部分之修法，因此本文認為存在對其詳加討論之必要。

準此，本文之問題意識即集中在回答以下三個問題：「對於各種類之不實影像，若不採特別訂立專法之模式，單以既存之實定法是否得以充分應對 deepfake 技術所帶來之衝擊？」、「對於各種類之不實影像之使用，可能涉及之保護法益為何？」、「對於各種類之不實影像，若需訂立專法加以因應，其條文應如何設計較為適當？」。此外，必須特別說明的是，雖然由於當前立法趨勢明顯係針對 deepfake 此一技術所作之因應，惟在將來亦可能出現其他新興的製作不實影像之技術，且我國當前修法所使用之文字亦未特別將技術手段限縮在 deepfake 此一種類上，故本文研究所得之結論，於將來其他具備高度擬真特質之製作不實影像技術上亦可能有適用之餘地。

¹⁶ 立法院公報第 112 卷第 14 期委員會紀錄頁 166-168。

¹⁷ 立法院公報第 111 卷第 84 期委員會紀錄頁 123。

¹⁸ 林琬珊（2023），前揭註 14，頁 51；古承宗（2023），前揭註 14，頁 19；前揭註 14，頁 37。

¹⁹ 廖宜寧（2022），〈從「妨害風化的猥褻物品」到「妨害性隱私的性影像」——散布性關聯影音圖像之不法性質〉，《月旦法學雜誌》，第 327 期，頁 168-169。

²⁰ 許恒達（2022），前揭註 4，頁 22-25。



第二節 文獻回顧

第一項 文獻簡介

許恒達同時討論了與性相關和與性無關之不實影像規制。其撰文於刑法第 319 條之 4 之增訂尚未三讀通過前，指出 deepfake 製成之不實影像未對現有刑法管制架構有全面性的翻轉，依照當時既存之實定法應即足以因應，因此上策之選擇不應為此增訂專法加以規範。惟若立法者礙於社會壓力而必須增訂專法因應，其能接受之中策為僅有針對與性相關之不實影像進行規制，並借鏡德國法上立法體例，認為此類型影像所可能侵害之法益應是「形象法益」，且在構成要件文字上應強調客體「模擬」或「仿真」之特質，而不宜以「不實內容」一詞來描述所規範之客體。至於針對與性無關之不實影像部分，其雖同意既存之實定法下「處罰必須從既有法益與相應構成要件進行判斷，這導致其制裁範圍相對受限」，惟其仍認為現存之法規已足以對應此類影音，而無另行增訂刑法規範必要²¹。

吳芳毅僅針對與性相關不實影像之規制進行討論。其撰文時刑法第 319 條之 4 之增訂亦尚未三讀通過，指出刑法第 310 條之誹謗罪、刑法第 235 條之散布猥褻物品罪於散布不實性影像之行為中雖皆有所適用，惟認為這些規制的手法都需仰賴其他相關的法益來串連，形成「保護法益紊亂之嫌」，因此主張應另立新法加以完善保護。至於此類行為所涉之法益，吳芳毅先以一般性之「色情報復」（即非限於以不實性影像為之，利用真實性影像者亦包含在內）做為討論基礎，指出「色情報復」所侵害之基本權包含性隱私權、名譽權、性自主權。惟在從一般性之「色情報復」，將論理嫁接回「不實性影像」時，其認為「其屬於造假影片，應認不具有私密性問題」，因此承認保護法益有所差異。惟其仍主張實質結果上與傳統色情報復相類似，皆會使被害人遭受巨大痛苦，而認為不宜將其二行為之評價區別對待²²。

許昭元主要著眼於與性相關之不實影像規制進行討論。其撰文之時刑法第 319 條之 4 尚未修正通過，故主要是針對各機關與委員提出之草案版本進行評析。其指出性幻想無論多麼低俗或骯髒，皆應受憲法思想自由之保障，而單純製作不實性影像之行為，雖已有外顯行為展現於外，仍屬於「思想輔助」或「思想輔助」而同屬思想自由保障之範疇。而僅在製作階段之行為，既尚未散布，對於臉孔被盜用者難以被證實有法益侵害之危險，因此不應以刑法規範加以介入管制²³。

²¹ 許恒達（2022），前揭註 4，頁 22-25。。

²² 吳芳毅（2020），〈深度偽造為色情報復之侵害與規制〉，《檢察新論》，第 28 期，頁 190-197。

²³ 許昭元（2023），〈論以深度偽造（deepfake）技術製造性私密影像是否受憲法思想自由之保障——以我國近期刑法修正草案為對象〉，《憲政時代》，46 卷 4 期，頁 455-470。

陳俊偉所討論之對象，則同時涵蓋了與性相關和與性無關不實影像之規制。其撰文時現行刑法第 319 條之 4 規定業已修正通過，而其對該條文則抱持批判觀點。在與性相關之不實影像使用方面，其認為所涉法益應是名譽法益，因此既存刑法中之妨害名譽罪章即足以因應，並於此部分明確反對前述許恒達所提之可能保護「形象法益」的觀點，理由在於其指出許恒達認為形象法益與名譽法益的區分標準，即前者涉及個人私領域而後者則係因個人公領域生活歷程而生，事實上難以操作而無法區別，也因此指出許恒達所援引之德國法實務中，對於形象的解釋業已導向公然侮辱罪靠攏。另外在對於現行新法之評析上，其亦批評立法者並未交代何以僅有與性相關之不實影像才會侵害人格法益，而其他種類之不實法益卻無保護必要性。同時現行法下在「意圖散布而製造」之行為階段即須承擔刑事責任，亦受其批評有處罰過度前置之問題。而在與性無關之不實影像方面，其則指出既存之實定法下，透過妨害名譽罪與散布各類不實訊息之犯罪處理即可，至多考量此手段「逼真與難以辨識，再加上網際網路的傳播作用」而可以考慮作為刑法第 310 條第 2 項加重誹謗罪之新事由，惟並無為此另立專法之必要²⁴。

古承宗則主要針對與性相關之不實影像規制進行討論。其撰文時刑法第 319 條之 4 業已修正通過，而其對該條文亦持批判立場。其認為不實性影像可能涉及之法益係屬名譽法益，因此得以透過既存實定法中的妨害名譽罪章加以處理。其明確反對將此類行為與德國法上侵害形象法益之犯罪相提並論，理由在於德國法上該罪係以影像本身為真實為前提，而不實性影像本身性質係屬虛偽，而因此應劃入妨害名譽的誹謗罪範疇。其最後指出，一般社會活動、政治、國家安全、金融市場等皆可能受到 deepfake 技術下之不實影像攻擊，新法將可罰範圍限制在與性相關之影像欠缺合理的依據。且其亦批評新法將處罰時點前置至製作行為並不合理，蓋其本質上屬於誹謗的預備行為，而一般的誹謗犯罪並不會處罰行為人的預備行為，此舉將造成刑法內在體系的衝突²⁵。

廖宜寧所討論之對象包含了此次修法關於「性影像」之所有條文，而不實性影像之規制僅是其論述的一部分。其認為此類行為涉及之法益不宜以「性隱私」或名譽權作為規制，而屬於「性人格」之範疇²⁶。其認為非屬性隱私之理由在於，不實性影像所影響者非私密領域資訊。而認為非屬名譽權之理由在於，避免性道德觀念中潛在的歧視與污名擴大，並且鑑於近來刑法上存在限制妨害名譽罪章適用之呼聲，為避免此類行為之規制影響對於妨害名譽罪章之檢討，結論上其認為獨立條文規制相關行為是可以接受的選項²⁷。

林琬珊所討論之對象，亦主要係整部「妨害性隱私及不實性影像罪章」，其論述主力集中在檢討此次修法於刑法第 10 條中增設關於「性影像」之概括定義

²⁴ 陳俊偉 (2023)，前揭註 14，頁 36-38。

²⁵ 古承宗 (2023)，前揭註 14，頁 12-19。

²⁶ 該文提及「性人格」一詞時，所援引之註腳係前述許恒達之文，故似可推論此處所稱之「性人格」可能即是許恒達所稱之「形象法益」。

²⁷ 廖宜寧 (2022)，前揭註 19，頁 164-170。

性規範，認為新法就「性影像」之規定存在不少問題，行為客體的規範方式無法充分突顯法益的需保護性，建議透過修法進一步將性影像的規範內容與隱私法益相連結。其中關於不實性影像之部分的論述，則主要認為由於裸體影像主人並不是臉譜擁有者，因此實難謂對於臉或頭部之影像主人而言，有性的隱私受侵害的問題²⁸。

第二項 文獻檢討

回顧完前述文獻後，回頭檢視本文所欲回答之三個問題意識，本文認為仍有些許不足而尚待填補之處。

首先是對於與性無關之不實影像使用議題上，多數文獻皆未將其所著墨之重心著眼於此，而少數有提及之文獻，亦皆僅以相對少之篇幅和討論量能，概略地指出其認為以現行既存之法制即足以充分因應。惟在此類文獻中，亦有承認「由於必須回歸各特別法中條文個別檢驗，因此適用上有所侷限者」。如此一來，則是否真如其所說的，全然無須以專法加以規制？又，如要以專法加以規制，其保護法益為何？條文應如何設計？則皆有進一步探究之必要。事實上，於 2023 年刑法修正歷程中，就已有 7 部修法草案訂定有針對此類不實影像之刑事規制，而法務部亦表明其並不排斥相關立法。因此，在目前立法者已經展現出修法意向後，此問題更是有積極討論之必要。

再者，在與性有關之不實影像使用議題上，關於既存之實定法制是否足以因應衝擊，學者於肯否見解間呈現出了紛雜之意見。此問題之根源，本質上源自於各學者間對於此類影像之使用所可能涉及之法益侵害，有著截然不同的看法。簡言之，認為保護法益係屬名譽法益者，較傾向認為現行法下妨害名譽罪章即足以處理之，反之則會認為須額外立法因應。

然而在這些紛雜的學說意見之間，較多的是學者正面論述自己對於所採取之保護法益的論理基礎與想法，而欠缺對於彼此看法之檢驗與交鋒。若往下深究，可被認為有所交流者僅有陳俊偉與古承宗明確附具理由，反對許恒達提出之形象法益的觀點，以及廖宜寧明確反對採取名譽法益的見解。惟此兩股勢力目前皆僅止於第一層的交鋒階段，因此皆有值得進一步深究之必要。

此外，本文亦認為過往即已有許多能夠使影像不實之編輯技術，而 deepfake 所帶來之衝擊之所以令人震撼，即是在於其高度仿真、宛若真實影像的特質。因此，是否具備「足以使人混淆」之潛能，應是此一議題下重要之影響因子，亦可能對於「是否成立既存實定下之犯罪」，以及「個別法益之侵害與否」皆有所影響。然而在前述所回顧之文獻中，學者皆未特別將此因子納入討論，亦是本文認為較為可惜而欲透過此研究與以填補之處。

最後，由於比較多數的學者皆認為此類影像無須以專法形式加以規範，因此對於若欲以專法規範時，其條文結構應如何設計為宜亦較欠缺討論。僅有零

²⁸ 林琬珊（2023），前揭註 14，頁 50-57。

散的主張，如許恒達認為不宜以「不實影像」作為構成要件文字，又例如許昭元、陳俊偉、古承宗皆有提及把「製作階段」之行為納入規制係刑罰過度前置化的展現等。惟面對現行法已經選擇以專法模式加以規制之現實下，究竟其條文應如何調整始為適當，亦有探明之必要。



第三節 研究架構與方法

本文第二章為「不實影像之類型化與既存之實定法規制」，主要旨在介紹本文以「是否與性相關」與「是否有致使理性一般人混淆之虞」兩個判斷基準將不實影像進行類型化區分，並解釋說明選擇前述2種判斷基準之正當性基礎與目的，結論上將不實影像區分為4種類型。此外，在本章之後半段，本文將依序檢討4種類型之不實影像若不以專法處理，單在既存之實定法規制下能否成立特定種類犯罪？在法規適用下是否足以充分評價法益侵害，對當事人給予適足之保護？

本文第三章為「各類型不實影像使用所涉之保護法益與應罰性」，主要旨在依序以4種類型之不實影像為核心，分別介紹依實務見解、學說見解會如何認定其涉及之保護法益，本文並會對於前述實務與學說所提出之保護法益見解與以檢討，而在最終給出結論，指明本文認為適宜之保護法益。

本文第四章為「不實影像規制之比較法介紹與我國立法評析」，主要旨在介紹美國州法現行已經通過針對不實影像進行規範之5部法規，並額外介紹2部本文認為值得參照之修法草案。除單純介紹各部法規個別之特色外，本文尚會整理出其在條文文字設計上有所異同、價值取捨之要件結構，以供我國修法建議時得以參照。而在本章後半段，本文將會以第三章討論之法益判斷為基礎，針對我國已針對不實影像規制完成立法之刑法第319條之4與刑法第339條之4規定予以修法之建議與評析，此外針對與性無關之不實影像規制部分，亦會給予獨立的修法建議，最終於第五章進行本文結論。

在研究方法上，本文在第二章與第三章皆會以文獻回顧做為主要之研究方法，依序檢討4種類型之不實影像使用分別應如何適用法規與認定保護法益。而在第四章中，則主要會透過比較分析之方法，借鏡比較法之經驗提供我國修法上之參照，而在若干條文設計之方案選擇上，亦會透過些許文獻回顧之方法，藉以論證本文作出特定立法建議之選擇的正當性。

至於在比較分析之對象選擇上，則囿於筆者語言能力之限制，以及資料之可觸及性，選擇以美國州法作為參照之對象，併予敘明。

第二章 deepfake 應用之類型化與既存之實定法規制



第一節 deepfake 應用之類型化

第一項 正面之應用

本文欲探討關於 deepfake 技術使用之刑事管制議題，所針對者當然是其所帶來的負面影響。然而，技術本身作為一種價值中立的工具，在可以被用以為非作歹之餘，若是能夠被加以善用，自然也會創造出正面的價值。因此本文欲先自介紹 deepfake 技術正面之應用出發，以求不會對其有先入為主的負面偏見。

第一款 娛樂機能

目前 deepfake 技術已經被廣泛使用在娛樂產業當中，特別是針對於拍攝過程中遭逢演員過世，亦或是希望重製已逝演員曾扮演之經典腳色時，deepfake 技術即可派上用場。例如知名電影「神鬼戰士」(Gladiator)以及「星際大戰」(Star Wars)都曾使用過此手法使已逝演員在作品中重現²⁹，而電影「玩命關頭 7」男主角之一的 Paul Walker 於拍攝途中意外車禍過世，劇組即以類似技術使其弟弟出演完成電影之拍攝³⁰。而韓國亦有電視台透過此技術，以該電視臺主持人為原型製作出虛擬的 AI 主播。電視台表示，此種應用可在突發災難狀況時，迅速向觀眾播報新聞內容，同時節省人力、時間、費用等資源³¹。

此外，deepfake 技術亦可應用在諷刺或諧擬的喜劇 (satire and parody) 當中，如同我國過去曾出現過讓藝人模仿政治人物或娛樂名人一般，以此針砭時政、監督政府。更甚者，deepfake 技術的應用亦可讓外語片的配音變的更加真實，得以提升觀影體驗³²。

第二款 教育機能

deepfake 技術亦可以被使用在教育目的上，例如可以透過技術使已逝之歷史人物出現在影音中，直接與學生進行對話，以達到更生動與活潑的教學功能。

²⁹ 吳芳毅 (2020)，前揭註 22，頁 188。

³⁰ 雷喻翔 (2022)，前揭註 7，頁 30。

³¹ 雷喻翔 (2022)，前揭註 7，頁 32。

³² Brown (2020), Deepfakes and the Weaponization of Disinformation, 23 *VA. J.L. & TECH*, 33.

而教育功能亦可能延伸到教室之外，例如 2018 年時，美國媒體 BuzzFeed 即刻意使用 deepfake 技術製作出前總統 Barack Obama 談話的虛偽影音，希望藉以提醒民眾加強對於 deepfake 技術的防範意識³³。另外亦已經有使用者透過 deepfake 技術之使用，藉以加強對大眾的公衛教育，防患疫情傳播³⁴。

第三款 自我實現機能

目前的 deepfake 技術使用，亦有應用在特定種類的癱瘓患者上，使其能夠以自己的聲音開口講話，進而完成更好的自我實現。例如科技公司 CereProc，即透過技術創造出數位語音，讓因為健康狀況而無法發聲的患者得以表現自我³⁵。此外，對於此類患者而言，亦可以透過將自己以及知情且同意的伴侶之臉孔，置換至成人影音當中，使其得以體驗現實世界所無法達成的性愛實踐，達成對於患者的自我實現³⁶。就性愛的方面，目前亦已經有至少一家成人影音公司特別針對此市場，提供使用者參與特定性愛場景，特別是對於那些物理與生理上受到限制的人們來說，可以讓自己置身於現實世界無法實現的性愛場景當中³⁷。

第四款 治療機能

最後，亦有論者提及，deepfake 之技術可以被用以針對特定的戀童癖患者，透過製造出虛擬的兒少性影音，藉以作為患者復健療程之用，以防止更多真實的兒童受害³⁸。

第二項 負面之應用

談完了對於 deepfake 之正面應用後，仍然必須面對目前有許多的濫用，進而可能造成傷害的事實。本節將簡單介紹兩條作為類型化標準的軸線，並以此為經緯，將不實影像之使用類型化成 2*2 的 4 種類型。以下將簡述選擇軸線標準的緣由，並說明做出類型化區分的正當性。

³³ Chesney, Citron (2019), *supra* note 5, 1770.

³⁴ Langa (2021), Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes, *Boston University Law Review*, Vol. 101, Issue 2, 767-768.

³⁵ Langa (2021), *supra* note 34, 767-768.

³⁶ Chesney, Citron (2019), *supra* note 5, 1771.

³⁷ Brown (2020), *supra* note 32, 33.

³⁸ Olson (2022), The Double-Side of Deepfakes: Obstacles and Assets in the Fight against Child Pornography Notes, *Georgia Law Review*, 882.

本文認為做出以下的類型化區分，將有助於本文隨後之討論，包含各類型不實影像之使用於既存實定法下受規制之狀況、可能涉及侵害之法益為何、有無刑事處罰之正當性、最終構成要件應如何設計等，皆與類型化之區分有密切之關聯。



第一款 是否與性相關

第一目 deepfake 技術實際應用情況

a. 與性相關之應用

如前所述，deepfake 議題在我國引起注目主要是源自於與性相關的使用，事實上這也並非孤例，根據資安公司 Deeptrace 於 2020 年所做的研究顯示，在其所識別的 14,000 個 deepfake 影片中，有高達 96% 都與色情有關，其中以女性受害者佔大宗，不乏知名影星與音樂家³⁹。無獨有偶，美國科技研究單位 Sensity AI 於 2021 年釋出報告顯示，其自 2018 年 12 月開始追蹤網路上 deepfake 之使用狀況後發現，有 90~95% 的 deepfake 影片是非合意的色情（nonconsensual porn）影片，其中又有 90% 左右是針對女性所為⁴⁰。此外，根據網路監管媒體調查發現，在其研究的 10 個色情網站當中，就有 8 個提供觀眾管道得以觀看 deepfake 技術製作而成的影片，且有一個名叫「Deepnude」的網站，提供用戶自行製造 deepfake 的色情影音，隨後因遭受嚴厲批判而下架。然而在下架前 24 小時內網站的瀏覽次數高達 5,45,000 次，且即便在網站下架後，該網站所產生的 deepfake 色情影音素材，仍大量透過網路公開資料庫或暗網等途徑散播⁴¹。

更甚者，deepfake 色情影音所影響的對象早已不限於名人。例如，在美國著名社群論壇 Reddit 上即曾出現子論壇（subreddit），主題旨在提供人們交換情報的溫床，使彼此在以 deepfake 技術製作各自暗戀對象，或者前任的重要他人之不實性影像的技術得以切磋、進步。而另一個知名的社群軟體 Discord 也曾經有類似的聊天室，供 deepfake 性影音之製作者間相互交流⁴²。此外，根據一項 2020 年來自美國的研究顯示，有高達 70% 的 deepfake 色情影音受害人都並非名人或有影響力的公開人物，只是一般的個人⁴³。

從以上種種資料都可以看出，目前 deepfake 的使用上，很大程度確實集中在與性相關的題材。

³⁹ Mihalcik (2020), California laws seek to crack down on deepfakes in politics and porn, in <https://www.cnet.com/news/politics/california-laws-seek-to-crack-down-on-deepfakes-in-politics-and-porn/>. (最後瀏覽日：07/13/2023)

⁴⁰ Haoarchive (2021), *supra* note 1. (最後瀏覽日：07/13/2023)

⁴¹ Thombre (2021), *supra* note 6, 2270-2271.

⁴² Harris (2019), *supra* note 2, 101.

⁴³ Ray, *supra* note 3. (最後瀏覽日：07/13/2023)

b. 與性無關之應用

然而，不可忽略的是仍有許多 deepfake 的使用雖然與性無關，卻仍然可能導致個人、社會，乃至於國家法益的侵害，因此有被關注或甚至規制的必要。

例如，有學者指出 deepfake 技術可以被應用在攻擊他人的金融系統或是電腦系統，透過製作他人的臉孔表情，藉以登入他人以臉部辨識機制驗證的系統，從中竊取檔案資訊，亦或是擅自進行金融交易等⁴⁴。亦有學者警告，deepfake 可能作為一種很有效的工具，被利用在綁架詐欺案件中，讓犯人可以偽造出肉票正受自己控制的不實影像，藉以成功向關係人勒贖⁴⁵。在具體的個案上，例如美國曾發生一名 50 歲的中年人利用 deepfake 製作不實影像，假扮年輕女性以替自己的頻道獲取流量⁴⁶。另外亦曾發生嫌犯透過 deepfake 製作不實影像，創造假身分冒充彭博社的記者，成功騙取許多投資人的個資。同一年亦有人利用 deepfake 技術，在 LinkedIn 上創造假身分，聲稱自己是來自美國智庫的研究者，向多名政府官員接觸，詐取機密資訊⁴⁷。

在國家層級的影響部分，在馬來西亞，曾發生有心人士利用 deepfake 技術，偽造內閣大臣與其敵對政治人物之不實性影像的事件，該影像在國內快速病毒式的傳播，雖然很快就被澄清是偽造，然而卻業已造成信賴的破壞⁴⁸。在比利時，曾有政黨利用 deepfake 技術，偽造時任美國總統 Donald Trump 發言的影像，該內容被認為足以明顯地影響（apparently interfering）國家的內部事務⁴⁹。在烏克蘭，2022 年正受到俄羅斯入侵的危急存亡之際，亦出現駭客利用 deepfake 技術，偽造烏克蘭總統 Vladimir Zelenskyy 出面向烏克蘭士兵喊話，要求他們放下武器停止戰鬥的不實影像⁵⁰。

凡此種種都得以看出，在與性無關的領域上，從個體權利，乃至於到社會利益與國家安全，都可能受到 deepfake 技術濫用所帶來的衝擊，而在此類與性無關的 deepfake 使用當中，據研究顯示主要仍是以名人作為描繪的對象，諸如娛樂產業、政治人物或 CEO 等⁵¹。

第二目 類型化之緣由

⁴⁴ 許恒達（2022），前揭註 4，頁 5。

⁴⁵ Chesney, Citron (2019), *supra* note 5, 1772.

⁴⁶ Kugler, Pace (2021), Deepfake Privacy: Attitudes and Regulation, *Northwestern University Law Review* 116, 623.

⁴⁷ Thombre (2021), *supra* note 6, 2271.

⁴⁸ Thombre (2021), *supra* note 6, 2273.

⁴⁹ Galston, Is seeing still believing? The deepfake challenge to the truth in politics, *BROOKINGS*, in: <https://www.brookings.edu/research/is-seeing-still-believing-the-deepfake-challenge-to-truth-in-politics/>.（最後瀏覽日：07/13/2023）

⁵⁰ 雷喻翔（2022），前揭註 7，頁 29。

⁵¹ Kugler, Pace (2021), *supra* note 46, 622.

如前所述，目前 deepfake 技術之應用以與性相關的使用佔最大宗，且影響範圍已經從早期僅限於名人，擴散到一般個人都可能淪為受害者。然而，與性無關的應用也絕非罕見，對社會的影響亦不容忽視。而對於此兩種類型 deepfake 不當使用之法律規範，學說與各法體系之立法例上，多有認為應該區分對待的聲音。

a. 學說見解

在學說方面，即有學者介紹表示，德國法主流意見認為 deepfake 使用從內容上可以區分成色情與非色情兩類，且兩者涉及不完全相同的法律議題⁵²。且該學者亦表示，其認為我國或可考慮對「散布具色情意味的深偽影音」之行為設立獨立的規範要件，但並不支持對「無關色情的深偽影音」另立新設的規範要件⁵³。

另外在 2021 年，有學者針對美國人民進行一項研究調查，結果顯示多數美國人看待與性相關和與性無關的兩種類型 deepfake 應用時，有著截然不同的態度。整體而言，美國民眾認為與性相關的 deepfake 相較於與性無關的 deepfake 會導致更大的傷害 (harmfulness) 與可責性 (blameworthiness)；且前者相較後者，更不會隨著散布者明確註記 (labeled) 該影片是偽造的作品，即因此得以降低傷害或可責的程度⁵⁴。

此外，亦有學者明確指出，近年來學界討論 deepfake 影音帶來的傷害，主要從兩個面向出發，其一是對於個體而言的人性尊嚴或情感福祉，另一是對社會而言的國家安全或民主體制之影響⁵⁵。

自前述學說意見可知，無論是比較法上抑或是我國學者，甚或是美國針對人民對於 deepfake 使用之態度觀之，與性相關和與性無關之 deepfake 使用，確實都可能存在討論脈絡上之差異。

b. 立法體例

此外在立法體例方面，例如我國於 2023 年 1 月針對刑法進行修正，其新增之條文僅針對於不實性影像之使用，至於與性無關之使用雖然於立法審議過程中出現於多個修法草案版本中，於黨團協商程序中亦曾被提及，然而確實未見於本次之修法成果當中。而同年 5 月刑法之修法結果，則是僅針對將利用不實影像進

⁵² 許恒達 (2022)，前揭註 4，頁 9。

⁵³ 許恒達 (2022)，前揭註 4，頁 25。

⁵⁴ Kugler, Pace (2021), *supra* note 46, 639-654.

⁵⁵ Kugler, Pace (2021), *supra* note 46, 660.

行詐欺之行為，納入刑法第 339 條之 4 加重詐欺罪之規範類型，而並非針對與性無關之不實影像使用進行一般性規範。

而此選擇性立法的狀況亦並非孤例，即有學者介紹美國並未訂有一般性的保護性規範，提供一般市民對抗與性無關的不實影像使用。在美國現行法下，相對於傳統上誹謗相關罪名更傾向保護普通老百姓，針對 deepfake 的不當使用已然形成只有公眾人物反而更受保護的現象⁵⁶。

與之相對的，加州⁵⁷、維吉尼亞州⁵⁸、紐約州⁵⁹都分別針對與性相關的 deepfake 不當使用訂有民事或刑事上的保護規範，而德州⁶⁰與加州⁶¹則分別針對選舉期間的 deepfake 不當使用訂有民事或刑事上的保護規範。另外亦有學者指出，在美國絕大多數州都已經設有禁止復仇式色情（revenge pornography）規範的如今，其中絕大多數的州卻仍未將禁止範圍觸及覆蓋至以 deepfake 製作之非經合意性相關影音⁶²，且直到 2022 年 6 月為止，美國在聯邦法層級皆無管制 deepfake 相關立法規範出現⁶³。

關於上述我國草案與比較法上之詳細介紹與分析，本文將於第四章進行說明。本文於本段落進行如上說明的目的主要在於解釋，不僅是國內外學說有主張對於與性相關和與性無關之 deepfake 的濫用不能等同視之，就實際立法例上觀察，事實上對於兩種類型之 deepfake 的濫用亦存在明顯之差別待遇。本文無意在此階段即斷言前述學說主張、美國人民素樸的法感情，抑或是法律規範體例對於與性相關和與性無關之不實影像使用的差別對待必然合理正當。然而作為討論之伊始，既然兩種類型如今社會在客觀上發生的頻率、樣態、性質確有落差，且無論是學界或民間普遍存在將其差別以觀的現象，則或許兩種類型間確實存在著保護法益之不同、既存現行法保護密度之落差或構成要件設計上需求的歧異。

準此，本文選擇以「是否與性相關」作為第一條劃分類型化的區分基準，以求後續討論的完整與順暢。

第二款 是否有致使理性一般人混淆之虞

第一目 deepfake 技術實際應用情況

⁵⁶ Kugler, Pace (2021), *supra* note 46, 632.

⁵⁷ CAL. CIV. CODE § 1708.86.

⁵⁸ VA. CODE ANN. § 18.2-386.2.

⁵⁹ NY. Private right of action for unlawful dissemination or publication of a sexually explicit depiction of an individual § 52-C.

⁶⁰ TEX. ELEC. CODE ANN. § 255.004.

⁶¹ CAL. ELEC. CODE § 20010.

⁶² Brown (2020), *supra* note 32, 54.

⁶³ Nagumotu (2022), Deepfakes Are Taking over Social Media: Can the Law Keep up?, *IDEA: The Law Review of the Franklin Pierce Center for Intel* Issue 2, 141.

deepfake 的本質其實是一種透過影音編輯來製造偽作的技術，然而此類技術並非橫空出世，濫用亦非毫無先例可循。

例如在佛羅里達州曾發生校園槍擊案，其中的一名學生因此大力遊說槍枝管制的法案，並為此拍攝了一張撕毀射擊場靶紙的形象照以供宣傳理念之用。然而此舉卻引來擁槍陣營之不滿，遂使用拙劣的影音編輯技術將照片中的靶紙換成了美國憲法副本，意圖透過其撕毀美國憲法副本的虛偽照片來指控其對國家不忠誠，藉以傷害其形象。所幸，此舉很快就因為拙劣的技術而被識破、澄清，然而在此之前該照片卻也已經被病毒式的大量傳播，且論者也認為事件中一些本就傾向相信該假訊息的民眾，會因此忽略了照片的瑕疵⁶⁴。

又好比在美國，曾出現有心人士透過被稱作「Shallow fake」或「Cheap fake」（與 deepfake 相對的低度偽造）的影音編輯技術，刻意將時任眾議院議長 Nancy Pelosi 發言的影片降低速度並且模糊影像中的聲音，以營造其醉後神智不清的形象，最終雖很快的被識破、釐清，但在此之前也業已被點閱超過百萬次⁶⁵。

其實即便是使用 deepfake 技術本身，也都可能因為技巧之拙劣，或整體使用之脈絡，而未必存在致使一般人混淆的風險與可能性。例如在我國著名案件中，即有被害人指出當中一些影片：「因為手法非常拙劣，完全看得出就是合成的，圖像的邊緣都還有黑黑的線，就像只是把我的臉貼上去一樣。」⁶⁶。此外，2022 年亦發生過一起以 deepfake 技術偽造 Elon Musk 之推薦投資影片藉以詐財之案件，然而即有論者指出該影片其實技術相當拙劣，聲音與畫面時常不同步，嘴型與其口述之內容也有明顯落差⁶⁷。又好比在前述 Youtuber 與實況主遭換臉的案件中，亦有資料顯示其中所使用的背景成人影音素材，有一部分是來自日本的成人影片，或是在流通者之間本就理解、註明是利用 deepfake 所做的偽作⁶⁸。本文認為基此，至少就這類影片而言，很可能因為前述的使用脈絡（明顯與被換臉之知名被害人使用不同語言且屬於不同國籍、明顯註記是使用 deepfake 製成之偽作等），而欠缺使理性一般人混淆之可能。

事實上，亦有學者於討論 deepfake 之刑事管制議題時提及：「對『被害人』的素樸意識而言，……，我們或可試想：重要的政治人物頭像，被貼附到人獸

⁶⁴ Chesney, Citron (2019), *supra* note 5, 1755-1757.

⁶⁵ Brown (2020), *supra* note 32, 15-16.

⁶⁶ 鏡傳媒 (05/04/2022)，〈【換臉挨打怒提告 2】遭換臉不雅片曝光 陳沂坦言「身心留下陰影」〉，<https://www.mirrormedia.mg/story/20220503ent012/>（最後瀏覽日：07/13/2023）

⁶⁷ wepro180 科技新聞 (05/26/2022)，〈【Video Hunt】Elon Musk 投資平台保證 3 成回報率 三個破綻識破 Deepfake 騙局〉，https://www.wepro180.com/deepfakescam_220526/（最後瀏覽日：07/13/2023）；Abrams, Elon Musk deep fakes promote new BitVex cryptocurrency scam, *bleepingcomputer*, in: <https://www.bleepingcomputer.com/news/security/elon-musk-deep-fakes-promote-new-bitvex-cryptocurrency-scam/>（最後瀏覽日：07/13/2023）

⁶⁸ 鏡傳媒 (05/06/2021)，前揭註 8（最後瀏覽日：07/13/2023）

性交的影片上，該影音除了滿足特定購買者的性想像之外，沒有任何人會相信該政治人物會作出類似行為，還被拍下影片，要說妨害了被害人的客觀社會名望，似有論理扞格。」⁶⁹

準此，我們應該可以承認，在 deepfake 之使用，甚或更廣的影音編輯製造偽作技術中，事實上亦存在該偽作不足以致使理性一般人混淆的可能。

第二目 類型化之緣由

做出此類型化區分之重要性在於，無論是觀察學說意見主張濫用 deepfake 技術應該受刑事處罰之立論基礎，亦或是從相關事件受害人自身揭露的受害心路歷程觀之，「可能因高度擬真性而遭誤會、曲解」，似乎都是一項重要的影響因素。

a. 學說見解

例如我國學者即指出，其原則上認為應該維持現行法架構，無須針對 deepfake 議題增設獨立的處罰要件，惟若立法者礙於社會壓力必須修法因應，其能接受之底線係增設散布與性相關之 deepfake 影像的處罰規範。究其保護法益和處罰的正當性基礎，學者認為係「侵害被害人的社會評價」、「對『個人形象』（Image）造成干擾」、「產生的「（他人不知性活動為何）的外在形象」等⁷⁰。然而，倘若所散布者係本文於前述的「無致使理性一般人混淆之虞」的 deepfake 偽作，對觀眾而言可能根本不會產生「該被描繪對象確實從事影像中之性行為」之印象，此際，「是否足以侵害被害人社會評價？」、「有無對個人形象造成干擾？」、「能否產生錯誤的外在形象？」，皆不無疑問。

無獨有偶，在美國亦有學者明確主張，不現實且令人難以相信（unrealistic and implausible）的 deepfake 作品，不具備與超級寫實（hyper realistic）的作品同等級的，足以傳遞假資訊、濫用社會隔閡與分裂，以及製造政治動盪⁷¹的能力。也因此，學者認為立法者在處理對於 deepfake 的惡意使用時，最重要的事情即是在立法時著重於針對 deepfake 作品的說服力與可信度（persuasiveness and believability）⁷²進行規範。

b. 被害人感受

⁶⁹ 許恒達（2022），前揭註4，頁16。

⁷⁰ 許恒達（2022），前揭註4，頁22-25。

⁷¹ disseminate false information, exploit social division, or create political unrest

⁷² Langa (2021), *supra* note 34, 788.

除學者看法外，自受害者角度觀之，亦可發現「是否使一般人混淆其真的拍攝該性相關影像」，是其高度關注的重點之一，甚至也很大程度影響其所受傷害之程度。例如於我國發生著名案件後，即有媒體專訪其中之受害人。其中許多人皆表示「可能被誤會自己真的從事影音中之性活動」是其很在乎的顧慮⁷³。

本文並非要主張撇除前開顧慮後，對於這些被害人而言就不存在傷害可言。然而若此顧慮事實上確實是影響被害人傷害程度的重要因子的話，於評估立法規制時，就不得不將其視作重要的參考指標。

c. 立法體例

另外在各立法例上，亦可看出在嘗試管制 deepfake 之濫用時，對於是否要求「有致使理性一般人混淆之虞」此一要件，呈現了截然不同的態度。例如⁷⁴加州在 2019 年通過一項針對選舉時 deepfake 惡意使用之立法，其使用「具有實質欺騙性的聽覺或視覺素材」（materially deceptive audio or visual media）一詞來描述所欲規制的客體，並且要求前述客體必須同時符合以下兩種狀況：（1）該圖像、音檔或影像會使理性一般人誤以為真⁷⁵（2）該圖像、音檔或影像會使理性一般人，相較於看到或聽到未經變造的原始版本，對其所欲表達的內容產生根本性的不同理解或印象。⁷⁶若立法例採取此立場，則本文前述所提之，因技術拙劣有明顯破綻、因使用脈絡而使觀眾知悉影像中所描繪之人並非真正，或透過明確註明是利用 deepfake 技術做成之偽作的案例，皆會落在所欲規制的範圍之外。

又好比 2019 年聯邦眾議員 Yvette Clarke 提出的 The DEEP FAKES Accountability Act 草案（至今尚未經美國國會通過）則另闢蹊徑，其並未針對任何單純散布 deepfake 不實作品之行為進行處罰，與之相對的，其處罰的行為是當散布特定種類（意圖羞辱或騷擾的性相關影音、意圖製造暴力、激起武裝或外交衝突、干預選舉等）deepfake 作品時，並未確實揭露其係使用 deepfake 製成之偽作⁷⁸或未依照規定於作品加註浮水印（watermarks）之行為。且該法案亦明確規定，若「依照其情境，理性之人不會將偽造的物質活動誤認為是其中出演者所為

⁷³ 鏡傳媒（05/06/2021），前揭註 8（最後瀏覽日：07/13/2023）

⁷⁴ 詳細關於比較法上各立法例之分析，本文將於第四章詳細介紹，本章僅擷取對論述有必要之部分說明。

⁷⁵ The image or audio or video recording would falsely appear to a reasonable person to be authentic.

⁷⁶ The image or audio or video recording would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.

⁷⁷ CAL. ELEC. CODE § 20010 (c).

⁷⁸ fails to disclose that audiovisual content is a deepfake

真正的物質活動」⁷⁹，則可以構成無須被處罰的例外。⁸⁰若立法例採取此立場，則只要散布者明確註明是使用 deepfake 製成之偽作且加註浮水印，即便是作品內容涉及性相關活動，或是干預選舉等，都不在其管制範圍之內。

此外，法國刑法第 226-8 條規定：「以任何方法，發行對描述他人圖像或言詞進行編輯改造（montage）的出版物，且未經該他人同意者，應處以 1 年監禁和 15,000 歐元的罰款。若該出版物明顯（obvious）係經編輯改造，亦或其係經編輯改造之事實被明確指明者，不在此限。」⁸¹⁸²本條文係於 2000 年時通過、2002 年時生效，斯時顯然立法者不可能已經想像 deepfake 的存在，因此本條文構成要件亦未包含影像（video），也無法通盤處理、對應如今的 deepfake 議題，然而仍然得以一定程度於討論 deepfake 議題時提供借鑑。本條文規範在法國刑法典第 6 章的「侵害人格罪章」（OFFENCES AGAINST PERSONALITY）其中的第 2 節「侵害個人形象」（OFFENCES AGAINST THE IMAGE OF PERSONS），此與前述我國學者認為 deepfake 的性相關使用可能「對『個人形象』（Image）造成干擾」的說法不謀而合、相互輝映，而此立法例下明顯認為倘若該出版物明顯可以被發現是經編輯的產物，亦或是明確標明是經編輯之產物時，則即欠缺以刑法規制之必要性。

反之，亦有採取截然不同的立法方向者，例如德州於 2019 年通過針對選舉期間 deepfake 濫用規範之法案中，對於所規範之對象僅描述道：「本節中所稱的『deepfake 影像』係指帶有欺騙性意圖所製作，而宛若描繪了一個真實人物從事未曾實際發生行為之影像」⁸³，而並未對於其「擬真程度」做出要件之限制。

自此觀之，此立法例對於所欲規制之 deepfake 作品範圍有著相當寬泛的認定，並未要求有「有致使理性一般人混淆之虞」的相關要件，導致即便是極度拙劣之仿作，亦會落入該法打擊範圍。

此外，甚至亦有立法例中規定「行為人做出本節中所規範之行為因而涉訟時，即便已在該露骨的性素材中註明免責聲明，表明將該被描繪之個體納入該素材是未經授權的，或表明該被描繪之個體並未參與該素材的創作與開發，仍不得執此作為訴訟上之抗辯。」⁸⁴⁸⁵顯見，相較於前述立法例認為只要具體敘明係使用

⁷⁹ a reasonable person would not mistake the falsified material activity for actual material activity of the exhibited living person

⁸⁰ H.R. 3230 (2019), 116th Cong..

⁸¹ “A sentence of one year's imprisonment and a fine of €15,000 apply to the publication by any means of any montage made that uses the words or the image of a person without the latter's consent, unless it is obvious that it is a montage, or this fact is expressly indicated.”

⁸² FRENCH PENAL CODE ARTICLE 226-8.

⁸³ In this section, "deep fake video" means a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality.

⁸⁴ It shall not be a defense to an action under this section that there is a disclaimer includedthat the depicted individual did not participate in the creation or development of the material.

⁸⁵ CAL. CIV. CODE § 1708.86.

deepfake 之偽作且加註浮水印後即可完全免責，本立法例採取截然相反的態度，明確表明了即便附上的相關免責聲明仍不得執此作為訴訟上之抗辯。此外，紐約州在 2021 年的新修法亦有完全近似之規定⁸⁶。自前述各立法體例可知，關於「是否要求有致使理性一般人混淆之虞」的相關要件上，不同的立法者有著截然不同的價值選擇。

綜上所述無論自學者意見、被害人受訪陳述，亦或是各立法者的構成要件設計觀之，「致使理性一般人混淆之虞」與否在對於處理 deepfake 濫用之刑事規制時，是一項值得討論的重要變因，也因此，本文選擇此一條件作為類型化區分的軸線，以求隨後討論的完整與便

至此，本文已經簡單介紹 deepfake 於現實世界中被使用的現況，在簡述了其可能帶來的正面效益之餘，亦說明了在負面應用方面，以「是否與性相關」與「是否致使理性一般人混淆之虞」兩條軸線去切分，各自可能在客觀現象上、學者看法上、人民態度上產生歧異，因此有類型化區分之必要。

以此兩軸線去做劃分，將可以把所有的 deepfake 使用區分成「有混淆之虞的不實性影像使用」（類型一）⁸⁷、「無混淆之虞不實性影像使用」（類型二）、「有混淆之虞的與性無關之不實影像使用」（類型三）與「無混淆之虞的與性無關之不實影像使用」（類型四）等 4 種類型。

在隨後的章節當中，本文將依序討論 4 種類型的 deepfake 使用，各自在既存實定法下受規制的現況、是否有保護不足之問題，並且討論其各自可能涉及的保護法益為何、有無以刑法規制之正當性？

第二節 既存實定法下之規制狀況

本文此處所稱之「既存實定法」，指的係本文撰文當下，扣除 2023 年修法新增定之刑法規定以外的其他刑事法規。之所以欲討論既存實定法下之規制狀況，一方面係因為針對與性無關之不實影像應用上，至今仍無為此專定之法規加以規範，因此有必要加以確認現行法對此規制之狀況，進而探討有無額外另定專法之必要。

⁸⁶ NY. Private right of action for unlawful dissemination or publication of a sexually explicit depiction of an individual § 52-C.

⁸⁷ 本文之所以選擇使用「不實影像」一詞，主要係因 2023 年刑法修正後，新增之刑法第 28-1 章、刑法第 319 條之 4，以及刑法第 339 條之 4 規定，皆以此文字描述所欲規範之客體。為求討論上得以與我國法制接軌，不產生額外誤會，本文隨後皆會以此詞彙進行討論。

另一方面，即便現行法下已經對不實性影像訂定專法加以因應，惟探討除該專法外，其餘既存實定法對於此行為之規制狀況，一來能夠釐清 2023 年度修法時之環境背景、探明立法部門選擇另訂新法之緣由，亦可從中再省思為此訂定專法規範之正當性。

第一項 有混淆之虞的不實性影像使用（類型一）

第一款 學說見解

於不實性影像之使用上，我國學者有認為行為人利用 deepfake 製成虛假影音，以虛偽的裸體或性行為的內容，妨害了被害人於社會上受到評價的名聲，因此得以成立刑法第 310 條之誹謗罪。且該散布之行為，也會構成刑法第 235 條之散布猥褻物品罪，因而主張基於刑法本身已有為數甚多的構成要件可以適當管制散布不實性影像的行為，因此無須另訂專法做為刑事規範⁸⁸。惟亦有認為，雖同樣承認前述的刑法第 310 條之誹謗罪、刑法第 235 條之散布猥褻物品罪於此類行為中有所適用，然而這些規制的手法都是仰賴其他相關的法益來串連，形成「保護法益紊亂之嫌」，因此主張應另立新法完善保護者⁸⁹。

據本文觀察，前述學說於討論時皆並未考量「有無致使理性一般人混淆之虞」之類型化區分，本文認為此將導致討論散布不實性影像是否成立現行法下之犯罪時產生一些偏誤。準此，本文於隨後討論前述問題時，對於本文認為有所歧異之處特別著墨說明。

第二款 刑法第 315 條之 2 第 3 項散布竊錄影音罪要件檢討

刑法第 315 條之 2 第 3 項規定：「製造、散布、播送或販賣前二項或前條第二款竊錄之內容者，依第一項之規定處斷。」，其構成要件將散布之客體限縮在違反未得同意竊錄罪所得的犯罪產物，因此以 deepfake 技術所製成之不實性影像顯然不符合構成要件所述之內容，因而散布該影像之行為自然無法成立刑法第 315 條之 2 第 3 項散布竊錄影音罪。

第三款 刑法第 310 條誹謗罪要件檢討

對此，學說有認為因為此行為形同傳述被害人「拍過色情影音」的不實事實，因此會造成客觀名聲貶低的效果，從而得以成立刑法第 310 條誹謗罪⁹⁰。另也有

⁸⁸ 許恒達（2022），前揭註 4，頁 16。

⁸⁹ 吳芳毅（2020），前揭註 22，頁 190-192。

⁹⁰ 許恒達（2022），前揭註 4，頁 20。

認為此行為會導致公眾得到「被害人曾讓人拍攝過性相關影像」與「被害人曾與人發生性行為」的錯誤印象，從而「易使被害人於社會上遭受訛笑、譏諷」，因此足以認定成立名譽之傷害而可以成立刑法第 310 條誹謗罪⁹¹。

然而細究刑法第 310 條構成要件限定在「指摘或傳述足以毀損他人名譽之事者」，因此，單純導致公眾得到「被害人曾讓人拍攝過性相關影像」與「被害人曾與人發生性行為」的印象即便不實，惟其是否「足以毀損名譽」仍應被探究。對此，有實務見解指出被告散布被害人真實性影像之行為，形同「指謫告訴人之行為開放、隨便，為不在意社會道德觀念之人，客觀上確足使一般人對告訴人產生行為不檢、道德倫常觀念低落之負面觀感。」，因而認為此舉得以成立刑法第 310 條之誹謗罪⁹²

然而面對前述見解，即有學者指出普通公眾接觸到這些性相關影音時，至多僅會得到「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」等資訊，而是否會因此即認為影音中之參與者放蕩、隨便、任意讓人拍攝性相關影音道德上有瑕疵，完全取決於主觀想像或社會道德價值。準此，學者批判法官逕以自身之道德價值，推導出「讓人拍攝私密照片者，在性方面之價值觀有問題」負面價值判斷結果並不合理⁹³。

對此，本文完全同意前述批判實務見解的學者看法。一來是在客觀的實然面上，「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」等資訊，是否就如法院所述「足使一般人對告訴人產生行為不檢、道德倫常觀念低落之負面觀感」的現象，即非無疑。

退步言之，縱使實然上社會觀點真有歧視、汙名化此類行為之傾向，在應然上法院逕自將此行為定性、認證為「足以毀損名譽之事」是否適宜亦非無疑。就好比假設我國社會目前實然上確實存在極端歧視原住民的狀況，認為與原住民結婚是件有損名譽之事。若有人公開對外傳述、錯誤指摘他人曾與原住民結婚，此際，此事在實然上確實導致被害人主觀感受上極度不快，同時導致知悉此事之群眾對該被害人的評價有所影響。然而即便如此，本文仍認為並不宜逕認此舉係屬妨害名譽。蓋一旦法院作出了如此的判決，形同站在官方的立場，「認證」了「與原住民結婚是件丟臉、羞恥、有損名譽之事」，可能加劇歧視。

誠然，只要妨害名譽罪章還存在的一天，法院作成的每一個有罪判決，都一定會對特定詞彙，或特定行為形成「認證」，標記其足以毀損名譽，使其有被「標

⁹¹ 吳芳毅（2020），前揭註 22，頁 191。

⁹² 臺灣臺中地方法院 111 年度訴字第 330 號刑事判決、臺灣高等法院 111 年度上易字第 920 判決、臺灣高等法院臺中分院 110 年度上易字第 800 號判決、臺灣高等法院 108 年度上易字第 1379 號判決皆同旨。

⁹³ 張天一（2019），〈散布私密影像之刑事責任〉，《月旦法學教室》第 198 期，頁 29。

籤化」的可能。然而，此際法院即更應該深思，自己所做的價值判斷不僅是被動的反應社會的價值觀，而係應該積極的導引社會的觀念型塑。準此，本文認為法院公開的「認證」「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」等資訊是足以毀損名譽之事是絕對不合宜的。

在討論散布真實性影像刑事責任的脈絡中，即有學者擔憂以刑法第 235 條散布猥褻物品罪相繩，形同將被害人所拍攝之性相關影像貼上「猥褻物」的標籤，可能會因此對被害人形成二度傷害⁹⁴。如此一來，則法院作出本文前述的「認證」，其傷害可能有過之而無不及，當應慎之。刑法的重要機能之一無非在於法益的保護，倘若做成前述的此類判決結果，在該個案中雖看似替被害人「討回公道」，然而同樣的，該個案中的被害人也可能再次因為被法院「認證」的事實，因而強化內心的羞恥或自我價值低落感，結論上是否確實保護到被害人即非無疑。

誠然，本文完全相信前述實務判決或學者見解認可此類行為可以以誹謗罪處理，並非有意傷害被害人抑或是強化汙名，毋寧是在既存之實定法下欠缺其他保護被害人、懲罰行為人的管道，因此才只能訴諸此取徑希望至少能發揮些許懲戒、嚇阻或彌補之功效，然而事實可能卻適得其反。然而，這正在在顯示了既存的實定法下其實沒有適切的法規可以貼合此類行為所造成的法益侵害，硬是死搬硬套現行法框架欲處理此一社會現象無疑是削足適履。

綜上所述，本文以為在實然面上，目前可能尚欠缺足夠的證據足以支持「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」等資訊是足以毀損名譽之事；在應然面上，法院積極認證前述行為是足以毀損名譽之事，反而可能因為「法院認證」而導致的標籤化，無法達到保護法益的刑法目的。因此本文認為現行法下誹謗罪並無法適用於此類型之案件，釜底抽薪之道應該是重新設計出得以貼切評價此類行為侵害法益之條文才是上策。

第四款 刑法第 235 條散布猥褻物品罪要件檢討

本段落將分成「散布不實性影像是否可能成立刑法第 235 條散布猥褻物品罪」，以即「若可以成立本罪，其是否已充分評價法益侵害」兩個層次討論。

第一目 客觀構成要件不該當

⁹⁴ 蕭郁唐 (2016)，〈性隱私內容外流風波——從美國立法例論我國違反本人意願散布性隱私內容之入罪化〉，《科技法律透析》，28 卷 10 期，頁 47。

在第一個層次上，我國已針對不實性影像之刑事管制議題發表見解的學者皆肯定此等行為會有刑法第 235 條散布猥褻物品罪之適用⁹⁵，惟本文認為此說法並不完全精確。主要是對於那些「非屬公然散布」或「有經適當隔絕之不實性影像散布」個案中，本條文應無適用餘地。

對於「公然」此一要件，實務法院穩定指出：「刑法第 235 條第 1 項之供人觀覽猥褻物品罪，屬侵害社會法益之罪，其中散布、販賣、公然陳列乃例示規定，均屬供人觀覽方法之一，所謂散布，乃散發、傳布於公眾而言，考其立法目的，以此等行為，使猥褻物品流傳於社會公眾，足以助長淫風，破壞社會善良風俗，其可罰性甚為顯著，此與猥褻物品僅供己或僅供極少數特定人觀覽，未達危害社會秩序而屬個人自由權限範疇之情形有別，故設刑罰規定，以資禁制。從而本罪所稱以他法供人觀覽之補充概括規定，雖未明定為公然，實與上開例示規定相同而含有公然之意，必係置於不特定人或特定多數人可得觀賞、瀏覽之狀態下，始足當之（最高法院 84 年度台上字第 6294 號判決意旨參照）。⁹⁶」

自此可知，倘若行為人於製作性相關之 deepfake 影音後，僅供「少數特定人」觀覽者，很可能自始即構成要件不該當。如臺灣臺中地方法院 104 年度訴字第 1028 號刑事判決中，法院即以被告散播裸照對象僅特定一人，因而認定其不成立刑法第 235 條散布猥褻物品罪⁹⁷。

又，對於「非屬硬蕊且經適當隔絕」之相關要件，釋字第 407 號解釋謂：「惟猥褻出版品，乃指一切在客觀上，足以刺激或滿足性慾，並引起普通一般人羞恥或厭惡感而侵害性的道德感情，有礙於社會風化之出版品而言。猥褻出版品與藝術性、醫學性、教育性等出版品之區別，應就出版品整體之特性及其目的而為觀察，並依當時之社會一般觀念定之。」，又釋字第 617 號解釋謂：「刑法第二百三十五條第一項規定所謂散布、播送、販賣、公然陳列猥褻之資訊或物品，或以他法供人觀覽、聽聞之行為，係指對含有暴力、性虐待或人獸性交等而無藝術性、醫學性或教育性價值之猥褻資訊或物品為傳布，或對其他客觀上足以刺激或滿足性慾，而令一般人感覺不堪呈現於眾或不能忍受而排拒之猥褻資訊或物品，未採取適當之安全隔絕措施而傳布，使一般人得以見聞之行為。」

自此，實務法院便依循釋字意旨，將所謂猥褻物之認定區分成「含有暴力、性虐待或人獸性交等而無藝術性、醫學性或教育性價值」的硬蕊猥褻物，與其他的軟蕊猥褻物。其中刑法第 235 條散布猥褻物品罪所處罰者僅限於公開散布硬蕊

⁹⁵ 吳芳毅（2020），前揭註 22，頁 190-191；許恒達（2022），前揭註 4，頁 8。

⁹⁶ 此已幾乎成為實務法院例稿，如臺灣高等法院 110 年度侵上訴字第 109 號判決、臺灣高等法院臺中分院 111 年度軍侵上訴字第 4 號判決、臺灣高等法院高雄分院 109 年度上易字第 27 號判決等均採之。

⁹⁷ 臺灣高等法院高雄分院 105 年度上訴字第 609 號判決亦同旨。

猥褻物與未經適當隔絕的軟蕊猥褻物之行為，而不及於散布軟蕊猥褻物且經適當隔絕之情形。

而所謂適當隔絕，實務法院有明確揭櫫：「然而，『Twitter』所認定之『敏感內容』不只包括猥褻之內容，也包括暴力（戰爭）、政治、藝術等內容，也就是說『Twitter』上之『敏感內容』標示不足以讓使用者辨別該內容是否為猥褻內容而加以避開。使用者可能是為了觀看猥褻內容以外之原因，而點選「查看」或更改其隱私權設定為顯示敏感內容，因而在沒有辦法事前預料的情況下，觀看到『Twitter』網路平台上所顯示之猥褻內容。因此，『Twitter』對於敏感內容之顯示方式，難認為適當之安全隔絕措施。」⁹⁸「使用者在『訂閱』被告之帳號之前，可以充分考慮是否想看這些猥褻內容再行『訂閱』被告之帳號，不會在沒有預期的狀況下看到屬於猥褻影像之前述宣傳片。因此，……，應認已採取適當之安全隔絕措施，而無刑法第 235 條第 1 項規定之適用。」⁹⁹自前述同一判決內，法院對於行為人將猥褻影像上傳至兩不同平台，分別認為「難認為適當之安全隔絕措施」與「應認已採取適當之安全隔絕措施」可知，其認定是否採取適當隔絕之標準主要係以「會否使大眾、任意第三人在沒有預期的狀況下看到猥褻影像」作為標準。

此外，另有判決指出：「上開網路設有警告標語，點選已滿 18 歲後，才會進入網路頁面，一般人應可輕易判斷，網頁內容並不適合未滿 18 歲之人觀覽，有意願者，才會點入，即……，無此嗜好者，不會被迫接受。更何況本案亦採取「實名制」，經過認證後，被告 2 人才會寄送「無修正寫真圖包」的連結，可以有效掌控訂閱者，避免外流。可見被告 2 人於本案已經採取「適當之安全隔絕措施」，自與刑法第 235 條第 1 項之構成要件不合。⁹⁹」此判決亦是以「觀眾是否會被迫接受」作為有無採取「適當之安全隔絕措施」的判斷標準。

準此，倘若行為人於製作不實性影像後，若該影音不涉及「含有暴力、性虐待或人獸性交等而無藝術性、醫學性或教育性價值」的硬蕊物，且將該影音以前述法院揭櫫的「適當之安全隔絕措施」供人觀覽，亦不會成立刑法第 235 條之散布猥褻物品罪。也因此，本文以為前述學者直接認為散布不實性影像之行為，必然會有刑法第 235 條散布猥褻物品罪之適用，有些不夠精確。

何況，以發生在我國的知名案件為例，其運作樣態即包含了「在需經驗證的私人群組」、「客戶點單後一對一傳送」的特性¹⁰⁰，又，即便渠等影片被外流到成人影音平台，這些平台亦很可能有前述判決所揭櫫的「適當之安全隔絕措施」。

⁹⁸ 臺灣高等法院 111 年度上易字第 297 號判決。

⁹⁹ 臺灣高等法院臺中分院 109 年度上訴字第 2799 號判決。

¹⁰⁰ 鏡傳媒 (05/06/2021)，前揭註 8（最後瀏覽日：07/13/2023）

因此事實上，甚至可以積極一點的說，此類案件類型有很高的可能無法有刑法第 235 條散布猥褻物品罪之適用。



第二目 無法充分評價法益侵害

退步言之，縱使承認在部分個案中仍可能有刑法第 235 條散布猥褻物品罪之適用，單以此條文規範散布不實性影像之行為是否足以完善地保護法益亦非無疑。

抽象的來說，刑法第 235 條散布猥褻物品罪屬於保護社會法益之犯罪，且自本文前述所引判決可知，本條文所著重無疑是偏重於「潛在之公眾是否會遭逢道被迫看見猥褻物品的風險」，而沒有絲毫保護性影像中被描繪者之意圖。

具體一點的說，前述保護法益的錯位可能影響諸多被害人在實體上或刑事程序上的權益。例如，由於本罪係屬保護社會法益之犯罪類型，因此若僅有此條文得以規制行為人，被害人很可能因此無法對此行正當防衛。舉例而言，當有某職員甲使用同事乙之電腦時，發現其未登出的成人影音網站頁面上正以其自己帳號名義發表之不實性影像，職員甲若直接擅自自該網站刪除該影片，可能涉及刑法第 359 條無故刪除電磁紀錄罪，卻不得主張正當防衛。

又好比，臺灣臺北地方法院 110 年度聲判字第 311 號刑事裁定指出：「刑法第 235 條第 2 項之意圖散布而製造、持有猥褻之影像等罪嫌，屬侵害社會法益之犯罪，自難認聲請人為上開犯罪之直接被害人。¹⁰¹」倘若不被認為是直接被害人，此將直接影響到的是刑事訴訟法第 232 條所規定之告訴權、刑事訴訟法第 319 條所規定之自訴權。

此外，刑法第 235 條散布猥褻物品罪被認為屬於保護社會法益之性質，參酌前述實務判決明顯偏重在意「潛在之公眾是否會遭逢道被迫看見猥褻物品的風險」，性影像中被描繪者很可能被認為並非本罪下之被害人，此時可能受影響者包含刑事訴訟法第 248 條之 1 至刑事訴訟法第 248 條之 3 關於受害人受偵訊時特別保護之相關規定、刑事訴訟法第 271 條受害人於審判期日陳述意見之權利、第 271 條之 2 至第 272 條之 4 關於受害人審判期日特別保護之相關規定、刑事訴訟法第 289 條對於量刑表示意見的權利、刑事訴訟法第 344 條之請求上訴權、刑事訴訟法第 451 條之 1、第 455 條之 2 關於被告行簡易程序的表示意見權，與

¹⁰¹ 按：本件之告訴人即聲請人並非性相關影音中被描繪之當事人，故精確地來說，法院並未直接評價若是性相關影音中被描繪之當事人是否屬於直接被害人。本文以「刑法第 235 條&直接&被害人」作為關鍵字，於法學搜尋引擎檢索後查無其他法院明確表態之紀錄。惟此不妨礙刑法第 235 條之罪確實可能被實務法院認為係保護社會法益之罪名，因此否定特定個人作為直接被害人之可能。

刑事訴訟法第七編之三關於被害人訴訟參與之相關規定，影響範圍可謂是相當巨大。

準此，本文以為於散布以 deepfake 製作之性相關影像案件中，首先基於刑法第 235 條散布猥褻物品罪之構成要件，與我國實務在釋字 407 號解釋、釋字 617 號解釋後對本罪名之解釋，很可能使大量的此類案件無法成立本罪。又，縱使承認部分案件可能可以成立本罪，由於本罪係保護社會法益之犯罪，保護對象也顯非性影像中被描繪之人，因此對於這類受害人而言，將可能因此失去許多實體上與訴訟上之權限，準此，本文以為單以刑法第 235 條散布猥褻物品罪規制此類行為，對於法益之保障難謂充分。

第五款 個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪要件檢討

第一目 學者見解

按個人資料保護法第 41 條規定：「意圖為自己或第三人不法之利益或損害他人之利益，而違反第六條第一項、第十五條、第十六條、第十九條、第二十條第一項規定，或中央目的事業主管機關依第二十一條限制國際傳輸之命令或處分，足生損害於他人者，處五年以下有期徒刑，得併科新臺幣一百萬元以下罰金。」，另按個人資料保護法第 20 條第 1 項規定：「非公務機關對個人資料之利用，除第六條第一項所規定資料外，應於蒐集之特定目的必要範圍內為之。但有下列情形之一者，得為特定目的外之利用：一、法律明文規定。二、為增進公共利益所必要。三、為免除當事人之生命、身體、自由或財產上之危險。四、為防止他人權益之重大危害。五、公務機關或學術研究機構基於公共利益為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人。六、經當事人同意。七、有利於當事人權益。」

關於散布不實性影像是否可能成立前開罪名，在我國目前曾公開對 deepfake 刑事管制議題發表看法的學者當中，有完全並未提及關於個人資料保護法之適用可能性者¹⁰²，亦有引用其他學者之見解，進而主張「惟其『識別性』及『為自己或第三人不法之利益或損害他人之利益』之構成要件似乎於適用上尚不明確」，因此認為應先僅就刑法部分進行討論者¹⁰³。然而，關於前述所稱「構成要件適用尚不明確」之學者見解¹⁰⁴發表於 2016 年，細譯其論理，該文之所以認為個人資料保護法在構成要件適用上欠缺明確性，主要是基於該學者使用「猥褻、個人資料保護法」或「猥褻、裸照、名譽」或「裸照、個人資料保護法、名譽」等關鍵

¹⁰² 許恒達（2022），前揭註 4。

¹⁰³ 吳芳毅（2020），前揭註 22，頁 191-192。

¹⁰⁴ 蕭郁澹（2016），前揭註 94，頁 49-50。

字經判決檢索，發現在當時僅有兩則判決於相關案件審理中提及個人資料保護法，且對於要件適用上仍有疑義，據此使該學者做成了個人資料保護法在構成要件適用上欠缺明確性的結論。



第二目 本文見解

然而，筆者於撰寫本文之時，同樣以「猥褻、個人資料保護法」進行檢索，搜尋結果已有 206 則搜尋結果，而其中不乏認定此類行為成立犯罪之判決¹⁰⁵。又，後經最高法院刑事大法庭做成 109 年度台上大字第 1869 號刑事裁定，明確指出「個人資料保護法第 41 條所稱『意圖為自己或第三人不法之利益』，應限於財產上之利益；至所稱『損害他人之利益』，則不限於財產上之利益。」，故，如今法秩序環境已經與學者撰文之 2016 年明顯不同，因此實務對於個人資料保護法要件適用之態度，著實有重新討論之必要。

觀察實務上對此類案件以個人資料保護法審理時著重之處，確實如前述學者所說，法院基本都接受裸照本身構成個人資料保護法第 2 條所稱之個人資料，而將爭執焦點集中在「識別性」與「為自己或第三人不法之利益或損害他人之利益」兩個要件上。

關於「為自己或第三人不法之利益或損害他人之利益」之要件，如本文前述業經最高法院 109 年度台上大字第 1869 號刑事大法庭裁定明確化。於此之後，實務見解基本穩定認為「意圖恫嚇、羞辱、霸凌而分享他人性相關影音（無須公然、只針對特定少數人亦然）」本身即成立意圖損害他人利益。如臺灣高等法院臺中分院 109 年度上訴字第 2532 號判決即指出：「個人資料保護法第 41 條所稱『意圖為自己或第三人不法之利益』，應限於財產上之利益；至所稱『損害他人之利益』，則不限於財產上之利益，有最高法院 109 年度台上大字第 1869 號刑事大法庭裁定可參。本件被告未經告訴人之同意，擅自將上開關於告訴人身體隱私部位內容之照片傳送給證人陳正雄，雖未取得任何財產上之利益，然其所為，造成告訴人精神上受到莫大之痛苦，顯已損害告訴人之人格利益，至為明確。¹⁰⁶」

而關於「識別性」之要件部分，實務於此類型案件之態度基本是認為直接或間接方式識別該個人之資料即可，並不需要達到「一般第三人可以直接識別」的程度。例如臺灣高等法院臺中分院 109 年度上訴字第 2532 號判決即指出：「換言之，只需該等資訊得以直接或間接方式識別該個人之資料即為已足，並不以須達到『一般第三人可以直接識別』之程度為要。而『身體特徵』具有生物上之識

¹⁰⁵ 如臺灣高等法院花蓮分院 110 年度上訴字第 48 號判決、臺灣高等法院臺中分院 109 年度上訴字第 2532 號判決、臺灣高等法院臺中分院 109 年度上訴字第 2532 號判決等皆然。

¹⁰⁶ 臺灣高等法院 111 年度上訴字第 1182 號判決、臺灣高等法院 109 年度上訴字第 3876 號判決同旨。

別性，『臉部』固為身體常見可識別之特徵，但除『臉部』以外，因每個人之身體外觀仍具相當差異性，若結合人之身體其他多個部位外觀特徵，已足以個別化而具有識別性，當同屬該法所欲保護之標的，不得非法擅自利用。¹⁰⁷」反觀，若該裸照不足以具備特定到特定個人之識別性，即可能被認為不成立本罪，如臺灣高等法院 104 年度上訴字第 1393 號判決指出：「經查，本案被害人乙○○之猥褻影像固經被告將之散布於網際網路，惟觀諸上揭影像，一般第三人並無法因此直接識別究為何人？本案被告所散布者，亦非姓名、國民身分證統一編號、護照號碼、指紋等資料具有直接特定個人識別性之個人資料，一般人亦無法單以觀覽照片之方式，即可直接識別係本案之被害人乙○○；再參以本案之個人資料，僅有被害人乙○○之照片及被告加註照片之標題，客觀上結合此複數之個人資料，亦無法識別究為何人、究竟是否為特定之人，即非屬於個人資料保護法所規範之對象。」

準此，當今司法實務在「識別性」與「為自己或第三人不法之利益或損害他人之利益」兩個要件認定上，都與前述學者於 2016 年撰文時不可同日而語，業已相對明晰。然而前述判決所處理者皆是影音中受描繪者係屬真實的案件類型，個人資料保護法在不實性影像傳播案件中應如何適用，仍有值得討論的空間。

目前實務上唯一針對此類型案件做成判決的案件，即係針對我國知名藝人、實況主與 Youtuber 遭換臉至成人影片之事件所做。臺灣新北地方法院 111 年度審訴字第 356 號判決指出：「合成被害人等臉部特徵之猥褻影片至 Google Drive 雲端供會員或購買者付費瀏覽，被害人等之人臉特徵、姓名、藝名及網路暱稱等均足以直接或間接識別被害人等，屬個人資料保護法所規範之個人資料無訛，又渠等所為之利用行為，顯已逸脫蒐集上開個人資料特定目的之必要範圍，使瀏覽影片之人得與其他資料對照、連結而識別特定個人，且該等影像目的係在於凸顯淫穢、不雅之猥褻行為。」、「在其粉絲專頁內散布、播送之方式，吸引更多粉絲或追蹤者之加入或關注，藉此增加其網路流量，並從中牟取不法利益，而共同意圖為自己不法之利益及損害他人（即如附表所示之被害人）之利益。」等語，因此認定被告成立個人資料保護法第 41 條之罪。

本文對於個人資料保護法在不實性影像傳播案件中能否被適用，有以下幾點顧慮：

a. 「意圖損害他人之利益」之要件應僅限於損害財產上利益之情形

首先是在主觀意圖部分，本文認為前述刑事大法庭裁定指出「個人資料保護法第 41 條中，『意圖為自己或第三人不法之利益』，應限於財產上之利益；至

¹⁰⁷ 臺灣高等法院花蓮分院 110 年度上訴字第 48 號判決、臺灣高雄地方法院 108 年度簡字第 1059 號判決同旨。

所稱『損害他人之利益』，則不限於財產上之利益。」其在同一條文中，同樣就「利益」一詞所指涉的文義，作出截然相異的解讀並不適當，兩者仍應皆限定在損害「財產上利益」始足當之。

如即有學者檢討該裁定指出，本裁定之所以做出該見解，論理上係將個人資料保護法提案委員等同於「立法者」，並將其特定發言內容作為具拘束力的解釋準則，此論理在法學方法論上可能並不合理¹⁰⁸。且此見解也會使這個特殊之主觀構成要件，「失去用以表徵行為與法益侵害之間具有密切關聯的客觀事態之可能，而僅流於伴隨犯行乃至整體犯行的修辭。」

同時學者亦爬梳近年來「資訊自決權」與「資訊隱私權」之發展，認為個人資料保護法保護法益之正當性應並著眼於「現代高度資訊化社會中，基於商業目的趨動有組織、有系統且大規模的進行蒐集、處理或利用資料的現況」，並且基於行為類型明確化的要求，因此認為個人資料保護法第 41 條中，特殊主觀構成要件要件所稱的「利益」，應從循著前述「資訊自決權」與「資訊隱私權」發展脈絡下的「財產上之利益」加以理解¹⁰⁹。

如此一來，則過於個案中被告透過販賣不實性影像販賣營利，因此被認為成立「意圖為自己利益」之要件當屬無疑。然而若於不涉及販賣營利的案件類型中，則僅能從主張被告「意圖損害他人之利益」下手，此際被害人若無明顯財產受侵害之際，則個人資料保護法第 41 條之罪即難以適用。

b. 未營利之散布者難以構成「意圖為自己或第三人不法之利益或損害他人之利益」之要件

其次，退步言之，縱使承認前述大法庭裁定見解並無違誤，而認為「意圖損害他人之利益」包含損害他人人格權亦屬當之，於特定個案中亦可能難以認定被告存有此意圖。

本文前述所引之數則針對「真實性影像」散布的案件中，法院雖然都認為單純散布該影像的行為本身即屬於「意圖損害他人之利益」，然而這些判決的原因事實中被告與被害人間往往是分手後的情侶或職場上有嫌隙的同事，且有其他事證得以輔以論證被告散布被害人之私密照係為恫嚇、威脅、報復、嘲諷而做，因此得以被認定存在主觀意圖。然而，此見解能否想當然耳的援引到不實性影像散布案件中，並非毫無疑問。直言之，即便客觀上該行為真的會對被害人造成侵害，

¹⁰⁸ 簡至鴻（2022），〈個人資料保護法第 41 條「違法侵害個資罪」之基本課題——以最高法院 109 年度台上大字第 1869 號刑事裁定的檢討為契機〉，《刑事政策與犯罪防治研究》，第 32 期，頁 180。

¹⁰⁹ 簡至鴻（2022），前揭註 108，頁 174。

也未必代表被告主觀上對此有所認知，尤其，許多散布名人不實性影像的群眾甚至可能是該名人之粉絲，製造、散布或取得該影像的動機都只是自娛，而無法被連結、論證到具備「損害他人之利益」的主觀意圖。

就如在我國爆發知名名人受害事件後，有媒體專訪購買影片的民眾後指出¹¹⁰：「但他們多數未意識到自己正從事嚴重犯罪。有人無辜地說：『說真的，我們這群都還算是紳士，頂多自己尻尻。』也有人說，『我們是學術交流。』有時他們也討論倫理問題，群組裡不乏有人自承『我們有些人真沒法控制自己想看啥，壓制不住自己的欲望』，但隨即又給自己找台階下，『但我們絕對沒有想過要去『故意』傷害（被挖面的）本人。』」無論客觀上該名人所受損害有多巨大，立法者既然額外於本罪中設定特殊主觀構成要件，司法部門即應完全尊重，不能僅以被害人客觀上所受之損害，即論證被告主觀上具備損害的意圖，否則將完全架空立法者設立特殊主觀構成要件之意旨。

準此，縱使被害人客觀上遭受侵害，只要被告主觀上欠缺積極製造損害之意圖（例如主張「自己覺得不是『真實的性影像』就不會傷害到本人」、「大家都知道是假的影音，有什麼好受傷的」等語。此在新興領域的行為類型，整個社會對於此行為帶來的影響都欠缺共識時，未嘗不是合理的抗辯），即可能因為主觀上根本難以被認定具有損害被描繪者的意圖，而不構成「意圖損害他人之利益」之要件，導致個人資料保護法在使用不實性影像散布案件中無法被適用。

c. 本罪應僅限於「具組織性、系統性的資料蒐集、處理或利用行為」

再者，談到在客觀構成要件部分，亦有論者指出¹¹¹本罪之成立應限定在「具組織性、系統性的資料蒐集、處理或利用行為，而排除一次性、偶發性等單一資料的利用行為」。其立論基礎主要在於，關於個人資料保護法所欲保護之法益通說皆多肯認係以「資訊自決權」與「資訊隱私權」作為保護對象。然而，此保護法益內涵具有高度抽象性質，因此為了確保刑事立法手段的正當性，並且合理的劃分應罰性界限，必須實質的審視並檢驗其不法內涵的建構。

學者爬梳了隱私權在學術上發展的兩次轉折、三個階段後，認為如今對隱私權的保護意旨應該著眼在有別於過往以特定對象的監控、曝光其私領域資訊為主的侵害理解，被稱之為「資料庫監控」（dataveillance）的新型態隱私侵害態樣。亦即係在大量數位化資訊當中，分析出特定個體行為模式的狀況。並且，這種新型態的侵害隱私的態樣，往往會因行為大規模的掌握資訊流通系統的運作，因此

¹¹⁰ 鏡傳媒（05/06/2021），前揭註8（最後瀏覽日：07/13/2023）

¹¹¹ 劉定基（2020），〈個人資料保護法〉，《月旦法學教室》，第218期，頁65-66；簡至鴻（2022），前揭註108，頁180。

使本人無法察覺。基於此種態樣下的隱私侵害，才足以支撐個人資料保護法中將此等侵害入罪的刑事立法手段正當性，並合理的劃分應罰性界線¹¹²。

準此，若以「具組織性、系統性的資料蒐集、處理或利用行為，而排除一次性、偶發性等單一資料的利用行為」作為標準，個人資料保護法第 41 條之罪於使用 deepfake 的性私密影像傳播案件中或許對於製造且大規模散布之源頭有適用的可能，然而對於後續被化整為零的每個個體所為的散布行為，則很可能即無適用的餘地，然而這些每個零星個體的散布行為，往往仍然是對被害人持續擴大傷害的風險源頭，應仍有控管之必要。

d. 後續散布者之行為難以被認為係散布「個人資料」

最後，即便退步言之認為一次性、偶發性等單一資料的利用行為亦可受個人資料保護法第 41 條規制，本文也懷疑個人資料保護法中所謂「個人資料」的涵蓋範疇，可能會在要起訴後續散布者時有所爭議。

申言之，目前實務認定散布「真實性影像」時，基本肯認臉譜、身體特徵係屬「個人資料」無疑，因此若影像中足以識別到特定個人，即可成立本罪。而在唯一一起實務認定散布「不實性影像」案件中，由於所審判的對象是製造影片的源頭，其所使用的是被害人「原始照片」中的臉譜做出編輯後再散布，且在犯罪過程中明確以藝名、網路常用暱稱去連結被害人本人，因此被認為是「個人資料」亦無問題。惟，倘若是對於後續再散布該「不實性影像」之人，是否仍屬於利用「個人資料」，本文認為容有爭執空間。本文認為癥結點在於原始的「個人資料」，在經過多大程度的編輯後，仍得被認為與原始的「個人資料」具備同一性，值得懷疑。

舉比較極端的例子，若有行為人甲利用藝人 A 之「原始照片」，將起臉譜做成一般大眾都可明確識別是該藝人 A 之卡通化公仔或虛擬人像（avatar）銷售，再經行為人乙二度散布。就行為人甲之部分，基於其使用之「原始照片」中有藝人 A 之臉譜，其將其編輯後再販售之行為或可被認為是利用「個人資料」無疑，惟，就行為人乙而言，其所利用者僅是「經編輯後的臉譜」，雖仍可識別是特定個人，然本文認為此際已經跟「個人資料」失去同一性，難認其再次散布之行為可以成立個人資料保護法之犯罪。甚至，行為人甲、乙二人，在利用名人之餘，亦可較不招搖的將其與本人連結，而係與消費者之間建立一種「看破但不說破」的關係，彼此心中皆知道所描繪之對象即是藝人 A，但明面上皆否認經編輯後的成品係其本人。

¹¹² 簡至鴻（2022），前揭註 108，頁 161-172 頁、頁 180。

事實上此類的商業操作亦非罕見，例如 NBA 商標中運球的剪影，被普遍認為是以湖人隊知名球星 Jerry West 之照片作為基底，因此 Jerry West 也被球迷戲稱是「Logo Man」。然而 NBA 官方卻因為不願面對可能產生的授權費用問題，因此不曾官方承認商標中係利用 Jerry West 之照片，隨後即便 Jerry West 本人也曾對此不滿希望更換，卻不曾得到官方正式回應¹¹³。

僅以此例套用回我國法制說明本文適用個人資料保護法的立場，若 NBA 官方使用的不是剪影，而是直接大喇喇地使用 Jerry West 本人之原始照片作為商標，則因為其臉譜、身體特徵可供辨識，屬於「個人資料」無疑，其商業使用的行為也符合「意圖為自己利益」之要件，因此可能成立個人資料保護法第 41 條之犯罪無疑；惟於 NBA 官方如現況下使用了剪影，但若確有證據顯示其係使用某張 Jerry West 本人之原始照片進行編輯後散布，由於該原始照片仍可被認為是「個人資料」，因此其行為仍有可能可以成立個人資料保護法第 41 條之犯罪；但若是對於再散布剪影的其他球迷來說，由於該剪影是「原始照片經編輯」之產物，與原先之「個人資料」應以失去同一性，而難以被認為成立本罪。

如此一來，則回歸到使用 deepfake 的性私密影像傳播案件中，亦可能出現主謀確實以某名人之原始照片作為基礎去編輯製作影片後散布，再經他人二度散布的狀況。於此案件類型中，主謀本人因為犯罪過程中確實有接觸到可以被認為是「個人資料」的原始照片，因此在論理上確有認為其可以成立個人資料保護法第 41 條之犯罪的空間。然而，對於後續轉傳、二度散布之行為人而言，若該影片並不特別附著名人之姓名或其他資訊，整個交易、散布過程如本文前述處於一個「看破但不說破」的關係，則，其所二度散布之影片中「高度擬真、看似某名人」的臉譜，是否真的可以被認為是該名人之「個人資料」，即非無疑。

如果接受了散布可被識別是本人的被卡通化的名人虛擬人像、散布被剪影化的 Jerry West 之 Logo 不成立個人資料保護法之犯罪的話，那麼散布不實性影像，或許在論理上亦有可能被認為並非對「個人資料」之使用。本文當然承認「被剪影化」、「被卡通化」因此所失去同一性的程度，可能與「被 deepfake 技術編輯」有所落差，同時也同意並非原始照片一經編輯即必然失去同一性，否則只要稍微修圖、調整光澤即可規避責任，也會架空個人資料保護法的規範意旨。然而在各種編輯行為之間，究竟該以什麼標準去認定同一性仍未喪失因此仍可認定是「個人資料」，在實務與學說意見上確實欠缺表態。

該標準應該必須與實務所述的可否連結回特定個人的「可識別性」無關，畢竟在上述所說的「被剪影化」、「被卡通化」的案例中，社會透過整體脈絡去理

¹¹³ Dator, The NBA logo's complicated history, explained, SB Nation, in: <https://www.sbnation.com/nba/22300942/nba-logo-history-jerry-west-kobe-bryant>. (最後瀏覽日：07/13/2023)

解都仍然可能得以識別是使用特定名人之形象，然而本文仍認為這種案例當中難以認定被害人之「個人資料」會因為「被剪影化」、「被卡通化」的影音散布而遭侵害。則，若不以「可識別性」作為基準，又該以什麼標準區分「個人資料」的界線？該標準下，「使用 deepfake 技術編輯後」之影音又是否仍屬於「個人資料」，應存有疑義。

誠然，本文以上從主觀面與客觀面開展，試圖分析關於在不實性影像散布案件中，個人資料保護法可能無法適用的各種理由。目前由於欠缺足夠個案累積，因此也無從探知將來實務對此之態度。然而本文單純從法學方法的推理出發，認為 1.基於個人資料保護法第 41 條之特殊主觀構成要件，應該限於「意圖損害他人財產上之利益」之情形 2.基於不實性影像散布案件中，不若典型復仇式色情案件中往往有事證得以證明損害他人的主觀意圖，倘散布者並未藉此營利，則很可能基於特殊主觀要件的不該當而無法成立個人資料保護法第 41 條之罪 3.基於個人資料保護法第 41 條罪所保護之法益推導而出的實質不法內涵觀之，本罪之適用應限於「具組織性、系統性的資料蒐集、處理或利用行為，而排除一次性、偶發性等單一資料的利用行為」4 由於後續散布者所散布的是「原始圖片使用 deepfake 技術編輯後」的產物，是否仍可認為與原始圖片具備同一性，因此仍得以認定是利用「個人資料」，仍有爭執空間。準此，欲單以個人資料保護法第 41 條之罪來規制、對應不實性影像散布案件，會有許多可能造成被害人傷害的行為無法成立該罪，仍存有規範漏洞的風險。

退步言之，縱使在那些有個人資料保護法第 41 條之罪適用的案件中，抽象一點的談仍可能有法益保護不夠完整之狀況。蓋個人資料保護法所保障之法益通說多認為係「資訊自決權」與「資訊隱私權」已如前述，惟實務見解在解釋本罪條文中「損害他人之利益」的構成要件時，亦有明確指出：「立基於限縮解釋之立場，個資法第 41 條所稱之『損害他人之利益』，應限定在個人之資訊隱私及自主權利益『以外』之生命、身體、自由、名譽或財產等利益損害。」顯見，實務亦明確區分「資訊自決權」與「資訊隱私權」的侵害，與藉由侵害「資訊自決權」與「資訊隱私權」所衍伸出的其他基本權（如名譽權、人格權）侵害係屬二事，本就不能畫上等號。

因此，即便在那些有個人資料保護法第 41 條之罪適用的案件中，不能僅以法體系得以評價其對「資訊自決權」與「資訊隱私權」之評價即已足，而忽略了其後續的其他法益侵害（就如行為人也可以透過違反個人資料保護法第 41 條之罪的手段，大規模的進行誹謗、詐欺等行為，也不能僅以「資訊自決權」與「資訊隱私權」之侵害已經被評價，即謂得以取代對於後續其他法益侵害的保護）。

第六款 結論

綜上所述，關於「有混淆之虞的不實性影像」案件中，基於第 315 條之 2 第 3 項散布竊錄影音罪構成要件對於所散布之客體有所限制，因此難以成立本罪。

基於法院實然面上難以認定向公眾傳遞「他人曾從事性行為」或「從事性行為並記錄成影音」之行為，構成對於名譽之減損；又，應然面上亦不應認定前開行為構成名譽之減損，因而難以成立刑法第 310 條誹謗罪。

基於依實務法院以及大法官解釋之意旨，於許多不實性影像案件中只要傳遞對象係一對一為之，或有做出適當之隔絕措施，則將難以論以本罪。且，縱使可以成立本罪之個案中，也將因為抽象上法益評價並非保護影片中被描繪之人，因此導致對其保護有所欠缺，甚至導致其實體上或程序上利益無法充分受保障。

基於個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪中，「意圖損害他人之利益」之特殊主觀構成要件，應僅限於損害財產上利益之情形，若散布不實性影像之行為人並未以此營利，則可能難以論以本罪；又，縱使承認「意圖損害他人之利益」之特殊主觀構成要件，尚包含侵害他人人格權之情形，於散布不實性影像之行為人並無營利事實之案件中，可能因為其僅是為了自娛而無傷害影像中被描繪之人的意圖，而無法論以本罪；又，據學者研究本法所保障法益，即隱私權的發展脈絡認為，本罪之適用應僅限於「具組織性、系統性的資料蒐集、處理或利用行為」，準此，則零星的散布行為可能難以成立本罪；又，後續散布者所傳遞的係「經編輯後的原始圖片」，此際是否仍能認為係利用「個人資料」並非無疑。基於以上種種理由，散布不實性影像案件中，許多個案可能都難以成立個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪，且縱或得以成立，其保護之法益評價亦並非充分。

第二項 無混淆之虞的不實性影像使用（類型二）

在此種行為類型當中，適用原先既存之實定法的狀況大抵與本文前述對於類型一之分析相同，應無刑法第 315 條之 2 第 3 項散布竊錄影音罪與刑法第 310 條誹謗罪之適用；關於刑法第 235 條散布猥褻物品罪與個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪論罪時，僅有部分案件適用，將產生規範漏洞。且縱使於有所適用之案件中，其保護規範對於被害人而言仍嫌不足，此部分皆如前文所述，在此茲不贅述。

惟本文仍欲對其「無混淆之虞」之特質，因此影響其在原先既存之實定法規適用的狀況提出進一步的說明：

第一款 刑法第 310 條誹謗罪要件檢討

如前文所述，本文認為「有混淆之虞」的不實性影像使用不成立刑法第 310 條誹謗罪之理由主要在於，法院難以確認單純使公眾得到或誤認「被害人曾讓人拍攝過性相關影像」與「被害人曾與人發生性行為」的資訊在實然上確實會導致被害人名譽減損，且在應然上亦不應由官方「認證」渠等行為係貶損名譽之事，以免強化汙名化與偏見。

而於「無混淆之虞」的不實性影像使用案件中，除了如上理由之外，退步言之，縱使如實務法院認為散布「真實性影像」有本罪之適用，於此種案件類型中行為人亦應仍是無法構成本罪。

蓋縱使如學者所說，認為導致公眾得到「被害人曾讓人拍攝過性相關影像」與「被害人曾與人發生性行為」的錯誤印象，從而「易使被害人於社會上遭受訕笑、譏諷」，因此足以認定成立名譽之傷害而可以成立刑法第 310 條誹謗罪¹¹⁴，其重點仍在「錯誤印象」一詞。白話來講，就算真的如同目前實務法院見解一般，承認被社會知道自己曾經拍過性愛影片是可恥的，但在「無混淆之虞」的不實性影像使用案件中，根本沒有理性一般人會認為影中之人真的有拍攝性愛影片，因此也欠缺形成「錯誤印象」之可能。且不僅是客觀面上構成要件可能難以成立，在主觀面上，倘行為人已經明確標記影片中之人不曾真正從事該行為的情況下，亦應難以認為行為人存在使社會對被害人形成「錯誤印象」之故意。準此，散布「無混淆之虞」的不實性影像應係縱使採取實務見解，仍應得出無法成立本罪之結論。

基此，本文亦認為前述提及我國目前唯一一起關於散布不實性影像之判決（即臺灣新北地方法院 111 年度審訴字第 356 號判決）漏未考量「是否致使人混淆」之變因，即依循過往處理「散布真實性影像」之判決前例，認定有刑法第 310 條誹謗罪之適用，其認事用法並不妥適，併予敘明。

第二款 個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪要件檢討

個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪之構成要件，除違反同法第 20 條規定的客觀要件外，尚要求行為人符合「意圖為自己或第三人不法之利益或損害他人之利益」之特殊主觀構成要件，以及「足生損害於他人者」此等學說上所稱「適性犯」之客觀構成要件。

¹¹⁴ 吳芳毅（2020），前揭註 22，頁 191。

前述提及之我國目前唯一一起關於散布不實性影像之判決(即臺灣新北地方法院 111 年度審訴字第 356 號判決)意旨謂：「渠等所為之利用行為，顯已逸脫蒐集上開個人資料特定目的之必要範圍，使瀏覽影片之人得與其他資料對照、連結而識別特定個人，且該等影像目的係在於凸顯淫穢、不雅之猥褻行為，用以刺激、滿足或挑起他人之性慾，自屬猥褻影像無疑，進而毀損或貶抑被害人等之人格名譽及社會評價。」此段落係判決理由欄中，唯一似與前述提及的特殊主觀構成要件和客觀構成要件有關的涵攝。

如此一來，姑且不論本文前已敘明本罪所要求之特殊主觀構成要件中，應僅限「財產上利益」始符合個人資料保護法所欲保障的隱私權發展意旨，縱使肯認最高法院刑事大法庭之見解可採，則亦可見於此類案件中，法院欲說明行為人在主觀上具備「損害他人利益」或客觀上「足生損害於他人者」時，必須訴諸於「侵害名譽與社會評價」的見解。

惟就如本文前述，就算順應實務見解之堅持，認定被社會知悉或誤會行為人曾經拍過性愛影片是可恥的，但在「無混淆之虞」的不實性影像使用案件中，根本沒有理性一般人會認為被描繪之人真的有拍攝性愛影片。亦即，就算「拍攝性影像」是可恥的行為，但於此類型不實影像應用上，公眾皆知悉被描繪之人並未真正從事可恥的行為。準此，則亦應難以論證行為人的作為如判決所說，足以「毀損或貶抑被害人等之人格名譽及社會評價」。本案判決漏未考量「是否致使人混淆」之變因，即依循過往處理「散布真實性影像」之判決先例，認定有個人資料保護法第 41 條罪名之適用，本文認為其認事用法並不妥適。準此，則比起散布「有混淆之虞」的不實性影像較可能在現行法實務見解架構下成立個人資料保護法第 41 條之罪，散布「無混淆之虞」的不實性影像應係縱使採取實務見解，仍無法得出應成立本罪之結論。

第三款 結論

綜上所述，本文前述認為散布「有混淆之虞」的不實性影像行為不成立或難以成立各罪名之理由，於散布「無混淆之虞」的不實性影像案件中皆仍會遇到同樣之障礙。

況且，在此類案件類型中，在刑法第 310 條誹謗罪與個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪的構成要件檢驗上，尚會因為行為人所散布之影像，並不會使公眾誤會影音中被描繪之人過去曾經從事其未曾從事之性行為，因此更難被認定係侵害該當事人之名譽，亦難被認為有個人資料保護法第 41 條所要求之特殊主觀構成要件，或符合其客觀要件中

關於適性犯足生損害於他人之要求。因此，此類型之行為於既存實定法下更難受規制。



第三項 有混淆之虞的與性無關之不實影像之使用（類型三）

就此種類型而言，我國學者有認為現行法下已經有了充分的規制，足以對應此種 deepfake 濫用之社會衝擊，因此並無另行增訂刑法規範的必要¹¹⁵。

具體來說，此處可能涉及所有以「詐術」作為構成要件之刑法條文，例如刑法第 339 條詐欺取財罪、刑法第 313 條詐術妨害信用罪，甚或是詐術妨害考試公正、詐術妨害集會、詐術妨害投票等罪；或以「謠言或不實資訊（或內含此意旨）」作為構成要件之刑法條文，例如刑法第 251 條散布民生必需品不實資訊罪、刑法第 310 條誹謗罪¹¹⁶、選舉罷免法第 104 條意圖使人當選或不當選而散布謠言罪，或陸海空軍刑法第 72 條散布軍事謠言罪等。亦有認為此類行為應該於現行法下分別朝向妨害名譽罪與散布各類不實訊息之犯罪處理，其認為針對利用 deepfake 技術製成特別擬真之不實影像的新興議題，於妨害名譽法益時得為此修法增訂為加重誹謗之事由，然而在關於謠言相關之犯罪時，則因為可能牽涉言論自由之權衡，而主張於此部分無須修法因應¹¹⁷。

對於前述法規條文各自能在特定領域發揮一定作用遏止非性相關不實影像之惡意使用，本文敬表贊同。然而，本文以為透過前述條文所建構之規制網路，對於此類型之不實影像使用仍有法益保護上之漏洞。

具體來說，本文以為現行法建構的規制體系有以下幾點問題，包含「偏重於處理『對於被描繪者之負面描述』，難以處理特定案件類型」、「僅是正面表列劃定保護範圍，因而有所缺漏」、「偏重於單次事件所生的影響，忽略系統性之傷害」。此部分攸關於類型三之不實影像使用究竟涉及何種法益之侵害，故本文將於第三章進行詳細之論述。

綜上所述，本文並不否認藉由使用類型三之影像，得以作為一種犯下既存實定法既存規範之「手段」，進而可以針對該犯罪之「結果」依照既存之實定法加以處罰。惟本文認為該「手段」本身即應存在應罰性，因此僅以既存之實定法加以規範仍有所缺漏。

¹¹⁵ 許恒達（2022），前揭註 4，頁 15。

¹¹⁶ 技術上來說，本罪之構成要件事實上並未要求所指摘或傳述之資訊必須是謠言或不實資訊，然而由於刑法第 310 條第 3 項規定：「對於所誹謗之事，能證明其為真實者，不罰。但涉於私德而與公共利益無關者，不在此限。」，因此在討論本罪是否成立的層次上，指摘或傳述之資訊是否屬於謠言或不實資訊仍係重要的決定性因素。

¹¹⁷ 陳俊偉（2023），前揭註 14，頁 33。

第四項 無混淆之虞的與性無關之不實影像之使用（類型四）

於此類型的不實影像使用中，無論是基於影片本身技術過度拙劣，或透過影片使用的脈絡，導致理性一般人不至於混淆，亦或是透過特別標註的方式註明係透過 deepfake 技術所製造而成的偽作，總的來講都根本性的抹平了 deepfake 技術作為新興議題的特殊性，而與過往我們習慣的普通影音並無二致。

因此，本文主張此種類型之 deepfake 使用僅須被視作普通言論處理即可。本文並非認為此類作品之散布完全不可能成立犯罪，例如若利用 deepfake 技術製作出明顯拙劣、有瑕疵的影片，聲稱自己賣的商品連 Elon Musk 都推薦，因此使消費者向其購物，可能成立刑法第 339 條之詐欺罪（欠缺使理性一般人混淆之虞之欺騙能量，並不保證所有人都不會因此受騙，畢竟再拙劣的謊言都還是會有人買單）；又好比利用 deepfake 技術製作不實影像，借 Donald Trump 之口辱罵我國總統，即便明確標明是經編輯而成的偽作，仍可能成立刑法第 309 條之公然侮辱罪。

第五項 總結

針對類型一方面，本文認為散布其之行為無法成立刑法第 315 條之 2 第 3 項散布竊錄影音罪、刑法第 310 條誹謗罪；難以成立刑法第 235 條散布猥褻物品罪與個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪，且縱使成立前述二罪，基於所保護法益之錯置，以至於對被害人而言可能存在保護不足之問題。準此，既存之實定法規範可能不足以對應此類型之不實影像使用。

針對類型二方面，本文認為除了前述類型一討論脈絡下，既存之各實定法規難以適用的障礙皆會發生之外，於刑法第 310 條誹謗罪與個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪之論罪上，更會因為此類影音並不會讓公眾誤會影像中被描繪之人真的有從事影像中的性行為，因此更難以論證此類影片足以侵害被描繪之人的名譽，或證明行為人主觀上具備損害被描繪之人的特殊意圖、客觀上有足生損害於他人之潛能。

針對類型三方面，本文雖一方面肯認既存之實定法下有許多零星的法規得以在某些案件類型中對應到此類行為。然而，基於此類法規皆旨在處理個案中之「結果」，而忽略了使用此類型不實影像之「手段」本身即可能具備應罰性，因此本文認為於此類案件中，既存之實定法規制仍有不足。

針對類型四方面，本文認為在失去「易致使人混淆」之特色後，deepfake 相較其他傳統載體即失去特殊性，並無須另立新法加以規制，以現況下既存之實定法處理可能之犯罪議題即屬適當。



第三章 各類型不實影像使用所涉之保護法益與應罰性



第一節 入罪化之標準

基於本文旨在探討各類型之不實影像使用所涉及的刑事法規制議題，因此無可避免的必須面臨「究竟什麼樣的行為才值得被刑事處罰？」的問題。對此，筆者嘗試借鑑學說上關於法益論的討論，以及比較法上學者對於「入罪化」（Criminalization）之正當化基礎所提出之標準，希望得以覓得可供操作的標準或方向。

誠如學者所言，刑事制裁會加諸特別嚴重的法律責任於人民身上，其嚴屬性（harshness）最先，也最明顯的可以體現在對於罪犯的污名化（stigmatizing）效果上¹¹⁸。也因此，作為立法者更應該細究將特定行為加以入罪化的正當化基礎究竟為何。我國學者亦有指出，基於國家刑罰權之嚴屬性，因此並非一切憲法上應受保護之利益皆須動用刑法加以保護，而是必須憑藉法益的概念標示重要的利益，始有以刑法介入之正當性¹¹⁹。刑法因其具備最後手段性，唯有在行為侵害或違反該當社會核心價值時，始能動用¹²⁰。法益概念被期待於立法論上發揮「可罰性限縮機能」，以控制國家刑罰權的範圍。立法者在將某種行為態樣入罪化時，必須能夠合理的解釋該行為侵害了何等利益，而不能僅依憑該行為「不道德」即加以刑事處罰，否則即屬違憲¹²¹。民法與刑法之規範目的並不相同，因此兩者對於「不法」所設定之要件，自然也有差異。不得僅因民事上違反義務、侵害權利，即因此逕認其同時具備刑事不法性¹²²。

自前述學說整理可知，基於刑罰的嚴屬性，必須將其視為最後手段，也因此必須透過法益概念加以限縮。不是所有憲法上保障之基本權都值得動用刑法保護，不是所有民法上保護的權利被侵害時都需要以刑法介入、不能僅因一個行為不道德就將其視為刑事犯罪，而是只有在「重要的利益」、「行為侵害社會核心價值」時，始能正當化刑法的發動。然而，所謂的「重要的」、「核心的」終究是抽象的形容詞，存在高度的不確定性，究竟是何種程度的利益才值得動用刑罰保護，仍需要更多的概念加以充實與具體化其標準。

對此，學者有認為因為保護法益的內容是作為不法構成要件的解釋依據，若

¹¹⁸ Bendor, Dancig-Rosenberg (2016), Unconstitutional Criminalization, *New Criminal Law Review* 19 New Crim. L. Rev., 172.

¹¹⁹ 許恒達（2020），〈行為規範、保護法益與通姦罪的違憲審查——評釋字第791號解釋〉，《月旦法學雜誌》，第305期，頁28-30。

¹²⁰ 黃國瑞（2014），〈法益論之解構〉，《輔仁法學》，第48期，頁25。

¹²¹ 黃宗旻（2019），〈法益論的侷限與困境：無法發展立法論機能的歷史因素解明〉，《臺大法學論叢》，第48卷第1期，頁161。

¹²² 葉啟洲（2020），〈通姦罪違憲論的初始觀點——大法官釋字第554號解釋的聲請典故〉，《台灣法學雜誌》，第392期，頁45。

是該法益的概念「過於抽象且一般化」，則其就會難以發揮引導解釋的功能，因此不具備法益適格¹²³。有認為超個人法益需得還原成個人法益，才能確保刑罰作為維護個人良善生活環境而形塑的規範機制所欲達成的目的，同時避免法益喪失批判體系功能¹²⁴。亦有學者自釋字第 791 號解釋大法官所揭示的意旨，推導認為國家刑罰權只處理「侵害公益」與「具備反社會性之行為」，亦即重大憲法保障之利益，需對社會產生一定的傷害，此部分與英美上上所稱之「傷害原則」相互呼應。其援引美國學者 Joel Feinberg 提出之「傷害原則」與「冒犯原則」作為判斷動用刑罰正當性之依據。其指出所謂的傷害係指「利益的受挫」(setback to interest)，而「利益」係指「有關個人的健康幸福，且一般人對此會長期性地就其投入本錢，進而獲得相關利潤」的事項。傷害(harm)與冒犯(offense)並不相同，後者僅是「以錯誤的方式嚴重地造成他人的精神上的不快狀態」，與前者相比對當事人的損害較小，因此需要具備到「一定的條件」方能夠入罪化，所謂「一定的條件」之判斷，需參酌冒犯行為產生的「負面影響程度」，像是：冒犯的嚴重性、散播範圍、可否避免性及社會價值等¹²⁵。

比較法文獻上亦有指出單僅是製造傷害本身亦仍不足夠支撐刑罰介入之正當性，尚需該行為是侵害了「特定受保護之利益」(interest protected)，亦即強調保護「他人的權利」(rights of others)之概念。單純的冒犯(mere offense)亦不足以作為使特定行為被入罪化的正當化理由，唯有在行為會造成創傷性反應、長期傷害或反應性傷害(traumatic effects, remote harm or reactive harm)等情況，才足以成為將特定行為入罪化的正當性基礎¹²⁶。

自前述文獻回顧能夠發現，學說上對於「刑法的發動需有節制」、「只有足夠重要的利益才值得被刑法保護」是有相當共識的，也試圖對於何謂「重要利益」的標準提出更具體的刻畫，然而就所有討論以觀，其實還是難免流於「以抽象概念嘗試具體化另一抽象概念」的窘況。即便此一問題主要聚焦在法益概念一般定性之討論上，惟本文認為此問題的不明確，於判斷個別利益是否具備法益適格時，亦可能因此受到影響而同樣陷於模糊不清的困境。

事實上，近年來即有學者反思檢討法益概念，認為其最為人詬病之處在於定義不清、內涵不明。其指出自其觀察國內一般教科書之經驗，刑法學者多是以「法律所保護之利益」此種純粹複誦實證法規定的形式性方式說明，或只是進一步強調「生活利益」，或加上「重要」、「有價值」、「基本倫理價值」等抽象的限制，或訴諸人們的感受及共識。且學者亦指出這種難以描摹法益內涵的困境並非我國所獨有，比較法上的學者對於何謂法益的說法也是眾說紛紜。其最後引用多位德國學者所共同感嘆道：「直到現在，對於法益概念之澄清，縱使只想更清楚

¹²³ 古承宗(2023)，前揭註14，頁13。

¹²⁴ 黃國瑞(2014)，前揭註120，頁26。

¹²⁵ 林政佑(2020)，〈刑法第一百四十條侮辱公務員及公署罪之保護法益之探討〉，《國立中正大學法學集刊》，第74期，頁99-101。

¹²⁶ Hornle(2016), Theories of Criminalization, *Criminal Law and Philosophy* 10 Crim. L. & Phil., 303-306.

一點都做不到」¹²⁷。亦有學者直截點明法益標準不易拿捏，於刑法學發展至今還沒有形成受到絕對且共通的法益觀點¹²⁸。

面對如此的困境，學者有認為僅需所保護之利益沒有違反憲法價值即無須根本地否定其法益適格，而係於後階段以比例原則審視手段與目的間之互動，判斷其刑事立法是否適當即可¹²⁹。對於前述觀點，則有學者提出反駁與批判，理由主要基於「不在上游控管，勢必要加重比例原則其他階層的負擔」以及「無異於一般性的管制手段，這會使得刑法逐漸失去較高實體及程序保障的價值意義」¹³⁰，其提出具體因操作的標準可以參照德國學者 Roxin 所提出之標準，指出：「刑法只能作為最後手段保障國民本於自由與和平的社會生活，從而實現個人的基本權，為了實現該憲法利益的事實情狀」，因而得以負面排除「未侵害他人的道德禁忌」，或是「不可捉摸其具體內涵的保護內容」，作為適格之法益¹³¹。

對於前述學者提出之反駁與批判，本文較難以認同。針對「加重比例原則其他階層的負擔」之論點，本文認為需再細究其所在乎的該「負擔」所帶來的惡害究竟為何？若是擔心「因為加重負擔，因此可能『錯放』掉一些理應違憲的刑事法規範，導致刑罰過度擴張」，則本文之質疑在於，若系爭保護利益微不足道到學者甚至認為根本應該不成為適格之法益，同時又如學者所言刑法具備重大之嚴屬性對人民基本權干預最深，則在此狀況下究竟何以會在比例原則的必要性或衡平性層次上「錯放」掉系爭規範？本文殊難想像。而若其擔心之部分係「因為加重負擔，因此可能增加規範的審查者在論理上之困難」，則本文之質疑在於，對於一個規範的審查者而言，「硬是要操作一個模糊之標準去論證法益不適格」的論理難度，是否真比「在比例原則其他階層論證該規範並不恰當」更為輕鬆？本文對此亦持保留態度。

針對「使刑法逐漸失去較高實體及程序保障的價值意義」之論點，本文認為要強調刑法的特殊性，未必非得透過特別的法益要求始能展現其特別。在後階層的比例原則操作上，其特殊性亦會當然得有所體現。直言之，在必要性的階層上，基於刑罰的嚴屬性，自然會成為為了達成相同目的的一切手段中對人民基本權干預最深者，也因此，立法者唯有證明其他更輕微的行政、民事規範不足以達成與刑法相同的效果，始足以正當化刑事立法的管制，此處即體現了刑法的「特殊價值意義」。亦或是在衡平性的階層上，考量到刑罰的嚴屬性，在利弊比較的天平上，天然的將使弊害那端的砝碼相較其他管制手段更為傾斜，也因此立法者在刑事管制中必須負擔更大的義務，去證明天平上利益那端具備足夠重量，使能證明刑事管制係屬正當，此處亦能體現了刑法的「特殊價值意義」。更甚者，刑法的「特殊價值意義」應該是結果而非目的，在既存法律體系操作下自然能形成的現象，而無須、不必為了追求特殊而刻意改變操作。

¹²⁷ 黃宗旻（2019），前揭註 121，頁 164。

¹²⁸ 許恒達（2020），前揭註 119，頁 24。

¹²⁹ 薛智仁，〈會台字第 12664 號聲請釋憲案法律鑑定書〉，頁 7。

¹³⁰ 許恒達（2020），前揭註 119，頁 25。

¹³¹ 許恒達（2020），前揭註 119，頁 29。

準此，基於認定法益適格性時的標準模糊，且通常越是模糊不清、不正當、難以說明利益合理邊界的邊際案例，於常規的違憲審查自然會難以通過檢驗，對於避免刑罰過度擴張的結論並無不同，因此本文並不傾向從根本上的認為某種利益自始不足以作為法益受刑法保護。

誠然，前述整理的學說意見並非全然無意義，本文亦完全同意「並非所有受憲法或民法保護之利益都必須動用刑法保護」、「單純使人不快、單純只是違反道德或禁忌而未對他人造成侵害、不可捉摸其具體內涵等的利益，並不適宜動用刑法保護」等結論，只是本文於隨後討論所涉保護法益為何，以及該法益是否適合動用刑法加以保護時，將更著重在該法益內涵為何、與所干預之基本權如何比較權衡，而不根本性否定某些利益具法益適格性，合先敘明。

第二節 有混淆之虞的不實性影像使用

第一項 被害人受傷情況

承前所述，就「有混淆之虞的不實性影像使用」案例中，無論訴諸刑法第 315 條之 2 第 3 項散布竊錄影音罪、刑法第 310 條誹謗罪、刑法第 235 條散布猥褻物品罪或是個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪，都會導致某些個案無法被規制，抑或是縱能被規制仍可能基於所欲保護法益之錯位，因此不足已充分評價被害人之傷害。

然而本文亦同意，並不能僅從「既存法規無法處罰特定行為」的前提，即得出「該行為應被另外立法處罰」的結論，有可能正確的結論應是該行為並不值得被刑事處罰。因此正面的討論「究竟此行為侵害了什麼利益」仍有必要。準此，本段落將討論此行為究竟對於影音中被描繪的個體是否造成傷害？此傷害又屬於何種刑法所保護的法益？衡量該保護之法益與所侵害之基本權後，是否適合對於該行為加以刑事處罰？

關於「有混淆之虞的不實性影像使用」可能造成的傷害方面，學者有研究斷言該傷害就宛若是非經合意公開性私密影像¹³²；另有研究者調查指出，有足夠代表性數量的美國人認為未經合意的散布不實性影像影音，是極端有害（extremely harmful）的，且壓倒性的希望對此行為進行刑事制裁¹³³。且該研究者亦指出，不實性影像會造成一種類似於「扭曲」（distortion）的傷害，操縱（manipulate）他人被公眾認知、評價的方式，並且使他人以不準確的形象暴露（be inaccurately imposed）在公眾前¹³⁴。尚有學者指出，無論製作者的動機是什麼，不實性影像中

¹³² Harris (2019), *supra* note 2, 124.

¹³³ Kugler, Pace (2021), *supra* note 46, 611, 657-660.

¹³⁴ Kugler, Pace (2021), *supra* note 46, 611, 670.

的被害人在發現自己被編輯成影像中腳色時，都會受到劇烈的心理傷害，甚至因此感到羞辱與恐懼。同時其亦指出，鑒於社會上對於女性涉及性相關影像的汙名化，不實性影像中的被害人很可能會遭逢到後續的附隨傷害，例如在就業市場或其他地方遭逢不利益¹³⁵。亦有學者明確指出，散布不實性影像將影響影像中被害人就業可能性（employability），產生下游效應（downstream effects）¹³⁶。

而在具體個案上，華盛頓郵報的印度籍記者與專欄作家 Rana Ayyub 曾報導過一起印度當地的教徒屠殺行為，並且披露了當地行政官僚的腐敗與欠缺作為，因此導致引發當地一些人士的不滿，進而對其恐嚇，甚至以 deepfake 製作不實性影像來對其羞辱。透過其自述的心路歷程¹³⁷，我們能夠清楚的感知到此類事件對於受害人的具體影響是什麼：

「（當我在咖啡廳，第一次看到我的不實性影像後）我立刻開始嘔吐。尤其在印度這樣的地方，這可是一件大事。我不知道該如何是好，只能開始大哭。」

「在我整理好思緒之前，手機突然開始作響，開始有數百則 Twitter 的通知都是在分享那則影片。我朋友希望我刪掉 Twitter 別再看了，但我辦不到，我不希望那些人相信影片中的人真的是我。」

「開始有人會留言騷擾我，說了一些『想不到你有著那麼迷人的肉體』、『這是我的手機號碼，我現在在家裡有空』之類的話，開始向我詢問從事性交易的開價，甚至稱呼我是 ISIS 的性奴隸。在我每則 Instagram 貼文底下，都湧入了一堆留言內容，充斥著那則影片的截圖。」

「我完全不敢露面。我可以自詡是一名記者、一個女性主義者，但在此刻，我卻無法面對這些羞辱。後來我因為心悸跟焦慮症而被送醫，即便醫生開了藥，我還是忍不住的嘔吐、血壓飆高，我的身體因為壓力而有了劇烈的反應。」

「連我哥哥特地從孟買飛來德里探望，我都不敢見他，我不敢面對所有家人。直到事發後已經超過一年的如今，這些效果都仍影響著我，自那影片散布後，我就已經不再是同樣的一個我。過去的我曾經是個勇於發表意見的人，但現在的我卻會進行不必要的自我審查。現在的我已經完全不敢在 Facebook 上發表任何言論，因為我怕同樣的事情再度發生在我身上。我曾經是那麼的直言不諱，所以現在會變成這樣，真的是一個劇烈的轉折。」

¹³⁵ Chesney, Citron (2019), *supra* note 5, 1773.

¹³⁶ Kugler, Pace (2021), *supra* note 46, 625.

¹³⁷ Ayyub, I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me, *HUFFPOST UK*, in: https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316.（最後瀏覽日：07/13/2023）

「自那之後，我正常的工作生活也受影響。當我發文報導一起受到高度關注的謀殺案件時，人們就會貼出編輯後的我從事性行為的照片；當我發表新書時，就有警官發文質疑我是透過跟線人陪睡來獲取資訊；甚至有人指稱我跟內閣成員發生性行為。這些時間以來我不敢跟人說出我的感受，因為我怕聽眾不僅不會願意同理我、同情我，反而只會想從我身上挖掘更多細節。」

自以上自述之細節可知，此類行為對受害人而言當下的屈辱感、對於可能被誤會的恐懼、社會公眾後續的騷擾，乃至於對於隨後正常工作的干擾、歷經長時間仍無法復原的自我質疑與譴責，都是真真切切的存在。其實在我國發生著名案件後，亦有媒體專訪被害人，得出相似的結論，例如¹³⁸：

「球球最終發現，對被害人來說，世上沒有『還好』這件事。親近的人，她也許可以解釋『這不是我』，但其他人呢？訪談不到 10 分鐘，她眼淚大滴大滴落下。」

「我要怎麼證明那不是我？我走在路上、搭捷運，可能有一些陌生人，他看我或交頭接耳的時候，我都會覺得，是不是他們看過那個影片？覺得我是那樣的女生？覺得我是很糟糕的人？」

「遭到挖面之後，接連好幾個網友問她：『球球妳為什麼拍這種東西？』她也僅能解釋，自己沒拍色情影片，此事已經交給經紀公司處理。她主動避開與性有關的所有新聞，『連現在想起來我都覺得很不舒服，形象要建立起來是一件很困難的事，可是一部影片就把我毀掉了。』」

「先遭威脅、又無端被換臉做成 A 片，兩件事加起來，把她的精神狀況逼到極限，近一年她陷入憂鬱，焦慮失眠，也傷害自己，曾在枕頭下放美工刀，醒來卻沒有印象。『吃藥很想吐，不吃藥也很想吐，過了一段滿低潮的時期，但我的工作就是拍影片，還是要對鏡頭講話、表現得很開心...』YouTube 頻道上的她依舊到處吃喝玩樂，但離開鏡頭，在精神科醫師面前崩潰大哭，『醫生說不是我的錯，可是我很混亂，明明不是我，為什麼是我遇到這種事情？是不是我不要當 YouTuber 就好了？還是我做錯什麼？要被這樣對待？』球球哽咽說，她至今沒看完那支片，也不願再看到。」

「『如果未來真的做了一部身形很像我的 A 片，我到底要怎麼反駁？』如果、如果、如果，每一種如果的背後都是恐懼，她試想：『如果連我家人都不信，如果網友不相信，我該怎麼辦？如果我站出來講，有人會相信我嗎？要用聲音來判別嗎？要用場景來判別嗎？我要跟人家說『這 A 片的場景不是我家』嗎？我要把原始 A 片的番號查出來，跟大家說這是哪部片、身體本來是哪個女優嗎？』她

¹³⁸ 鏡傳媒 (05/06/2021)，前揭註 8 (最後瀏覽日：07/13/2023)

又想，但如果今天被挖面，臉被合成到不知名的影片和素人身上呢？那受害者能怎麼辦？」

自前述學者意見與被害人自述心路歷程皆可知，散布可能致使理性一般人混淆的 deepfake 性相關影音，會導致被害人莫大的傷害，其程度甚至可能不亞於現行刑法既存認為值得保護的法益類型。

第二項 所涉及之法益

第一款 隱私權

第一目 實務見解

2023 年度刑法關於規範不實性影像相關刑事管制之修正，在於必須增訂專門之刑事規範一節上獲得朝野高度共識，因此除了行政院法務部有提出院版的修正草案外，連同民進黨、國民黨、時代力量、民眾黨在內，共計有另外 19 件由跨黨派立法委員分別提出之修法草案，惟最後於二讀黨團協商時，立法委員以各版本過於紛雜難以整合為理由，決議僅針對院版進行討論已足¹³⁹。

於院版立法總說明時，針對刑法第 319 條之 4（下稱「deepfake 條款」）修法理由¹⁴⁰即指出「因網路資訊科技及人工智慧技術之運用快速發展，以電腦合成或其他科技方法而製作關於他人不實之性影像，可能真假難辨，易於流傳，如有意圖散布、播送、交付、公然陳列，或以他法供人觀覽而製作，或散布、播送、交付、公然陳列，或以他法供人觀覽他人不實之性影像，對被害人造成難堪與恐懼等身心創傷，而有處罰必要，爰參考韓國性暴力犯罪處罰特別法第十四條、美國維吉尼亞州刑法道德與風化罪章第十八點二之三八六點二條規定，增訂本條規定，以保護個人之隱私權及人格權。」其中明白指出本條規定所欲保障之法益係「個人之隱私權及人格權」。

惟前述院版之立法理由其實受到司法院高強度之挑戰，包含其審查時加註意見指出¹⁴¹：「草案立法理由認為本條保護法益包括『隱私權』，容有誤會，建請刪除。」、「司法院釋字第六〇三號解釋闡釋隱私權包含……保障個人生活私密領域免於他人侵擾及個人資料之自主控制……可知隱私權所保護者，為真正個人私生活領域免於遭受侵擾及個人資料自主控制權。因此，若非個人真正之私密資訊、活動，即非隱私權保護之對象。」、「行為客體為『不實』性影像，非被害人真正之性隱私資訊，其真正私密領域未受到侵擾，故『隱私權』應非本罪保護之法益。」，且於二讀黨團協商時司法院代表副秘書長周占春亦發言表示¹⁴²：「本院認為本條跟隱私權並無太大關連，它應該是名譽權受到影響，連同本章的章名也有這樣的

¹³⁹ 立法院公報第 111 卷第 84 期委員會紀錄頁 123。

¹⁴⁰ 立法院公報第 112 卷第 14 期委員會紀錄頁 164。

¹⁴¹ 立法院公報第 112 卷第 14 期委員會紀錄頁 166-168。

¹⁴² 立法院公報第 111 卷第 84 期委員會紀錄頁 11。

問題。」亦可見得自司法院的立場而言，明確反對 deepfake 條款係為保護隱私權所制定之規範。

面對此質疑，草案的提出機關即法務部列席之代表次長林錦村於黨團協商發言指出¹⁴³：「**最主要是保護性隱私**及人格權的完整性。剛才司法院提到名譽權是否會間接受到影響，這部分是否可以在立法理由中附帶提及，我們會再研究。」，且最終通過之立法理由亦全文照抄原先之立法總說明，並未有所更動，顯見法務部的立場始終堅持 deepfake 條款所保護之法益包含隱私權，然而全然並未回應司法院所提出之質疑。

而在二讀黨團協商與二讀廣泛討論兩次討論階段，與會之立法委員亦未對司法院提出之質疑有任何表態。

至此，簡要整理迄今實務上在立法階段各方的態度，法案提出者即法務部明確肯認 deepfake 條款保護對象包含隱私權、司法院明確表態基於隱私權之內涵僅包含保障個人「真實」之私密資訊，因此 deepfake 條款保護對象不能包含隱私權，而立法委員部分對此並無表示意見。

第二目 學說見解

在此議題上，國內學說見解高度一致，皆不認同 deepfake 條款所涉及之保護法益係隱私權，且論理上大抵與前述司法院意見相同。

例如有認為「虛偽事物理論上不會侵害隱私權」、「就將保護法益定位為隱私權，把假的影像當作是侵害他人的真實私密，這樣的立法理由與技術，恐怕反而才是假可亂真、真假不辨」者¹⁴⁴，亦有認為「由於裸體影像主人並不是他，因此實難謂就臉或頭部之影像主人有性的隱私受侵害的問題」者¹⁴⁵等¹⁴⁶。

惟在比較法上，亦有學者指出，雖然有論者基於性影像中被描繪之人不曾真實從事該行為，因此不可能侵害該個體之隱私權，進而認為虛擬復仇式色情無法帶來與真實復仇式色情相同等級的性隱私侵害。然而，學者認為此論點並不適用於那些「不特定個人的臉被用以創作成不實性影像」的狀況，因為該性影像的閱聽人可能難以理解該影像是虛假的，進而會相信影像中所描繪之人是真實從事該行為¹⁴⁷。學者進一步指出在大陸法系國家的理解中，普遍認為虛擬復仇式色情所侵害之基本權即是隱私權，並以斯洛維尼亞等國之立法例作為其論理之依據¹⁴⁸。其顯然是有意識並考量到前述反對意見後，仍然認為此類濫用不實性影像之行為會侵害被描繪之人的隱私權。

¹⁴³ 立法院公報第 111 卷第 84 期委員會紀錄頁 125。

¹⁴⁴ 許昭元（2023），前揭註 23，頁 479 頁、頁 485。

¹⁴⁵ 林琬珊（2023），前揭註 14，頁 51。

¹⁴⁶ 相似見解：前揭註 14，頁 37；許恒達（2022），前揭註 4，頁 19；廖宜寧（2022），前揭註 19，頁 168。

¹⁴⁷ ŠEPEC, LANGO (2020), VIRTUAL REVENGE PORNOGRAPHY AS A NEW ONLINE THREAT TO SEXUAL INTEGRITY, *Balkan Social Science Review*, 121.

¹⁴⁸ ŠEPEC, LANGO (2020), *Supra* note 147, 123-125.

第三目 本文見解

a. 隱私權的類型化

本文認為前述爭議必須回歸檢視隱私權本質究竟為何，始可能探求出合理之解答。

有學者爬梳隱私權概念發展史後指出，由於隱私權是為因應近代工業革命、資訊時代革命後才開發出現、受承認之新興基本權，在本質上並不屬於傳統上天賦人權中之一種，其本身內涵與基礎亦仍處於發展階段，因此其性質具有一定程度的開放性及不確定性¹⁴⁹。「隱私權」概念正式作為法律概念被提出之濫觴，首見於 1890 年，由學者 Samuel D. Warren 及 Louis D. Brandeis 在 Harvard Law Review 中所發表之“The Right to Privacy”一文中。隨後，自從二十世紀初期以降，隨著隱私權概念的發展，美國各州的實定法，以及法院判例皆陸續放寬對於隱私權的承認，然而此現象也使得許多乍看之下並不具共同特徵之案件，皆被歸納進「隱私權」的範疇中，進而使得隱私權之概念更加混亂。

此狀態造成許多學者開始爭執、探討隱私權究竟指涉的是什麼，直到 1960 年，加州大學 Berkley 分校法學教授 William L. Prosser 將隱私權分成四種分類，並在隨後成為學說上之通說，且同時獲得學界與實務界相當程度支持，才讓前述的爭端稍作平息。

所謂的四分類包含了 1.對個人之獨居、獨自性或個人性事務之入侵（下稱「入侵型」；）¹⁵⁰。2.對使個人難堪的私人事務之公開揭露（下稱「個人資訊不當揭露型」）¹⁵¹。3.將被害人置於錯誤之公眾理解下（下稱「公眾誤導型」）¹⁵²。4.為被告利益而未經同意使用被害人姓名或其他特徵（下稱「盜用身份型」）¹⁵³。¹⁵⁴

所謂的「入侵型」，原先指涉的係物理性的侵入、干預專屬於個人的空間，諸如住宅、房間，或是不法搜索隨身行李等。隨後擴張至較為抽象的利用電子設備竊聽，亦或是對其他私人資訊如銀行帳戶等的刺探皆屬之。

所謂的「個人資訊不當揭露型」要件有三，即係 1.針對未曾公開之私人資訊，2.對公眾公開揭露，且 3.該資訊在一般通念上係具冒犯性（offensive）或可抗議性（objectionable）即屬之。

所謂的「公眾誤導型」，指的係因加害人之「不實的」行為，以致被害人因而受到「公眾之誤解」，具體事例諸如：1.將某種意見或陳述之公開錯誤地歸於被害人所為（例如將劣質的詩作，錯誤的冠上他人之名，使該他人被誤會為該詩作之作者）2.將無關之被害人照片用以說明其他事項（例如使用他人相片做商業廣告，使該他人被誤會為該商品之代言人）3.將被害人錯誤地列入刑事或其他犯罪

¹⁴⁹ 林建中（1997），〈隱私權概念初探—從美國法之觀點切入〉，《憲政時代》，第 23 卷 1 期，頁 53。

¹⁵⁰ Intrusion upon the plaintiff's seclusion or solitude, or into his private affairs.

¹⁵¹ Public disclosure of embarrassing facts about the plaintiff.

¹⁵² Publicity which places the plaintiff in a false light in the public eye.

¹⁵³ Appropriation, for the defendant's advantage, of the plaintiff's name or likeness.

¹⁵⁴ 林建中（1997），前揭註 149，頁 56-59。

前科犯或高危險群之記錄中等。學者並指出此種類型之隱私權侵害，與誹謗（defamation）高度相似，兩者同樣涉及針對該他人相關資訊做出不實之訊息傳遞，差別僅在於誹謗案件中尚要求該資訊傳遞使被害人名譽下降，而此類型隱私權侵害則以傳遞錯誤形象為核心，誠然可能導致被害人名譽減損，然而即便導致被害人名譽持平未受影響，甚或名譽進而提升，亦不妨礙此種隱私權類型受到侵害之事實。

所謂的「盜用身份型」係指，基於商業性營利之目的，對被害人身份（identity，主要係指其姓名及其他特徵）之盜用，例如攝影師擅自將模特兒之照片，於同意範圍之外，用以商業宣傳等¹⁵⁵。

值得注意的是，前述關於隱私權的類型化，討論範圍主要是針對民事侵權行為（tort privacy），而在憲法層次上，美國法院實務曾經承認的隱私權種類包含婚姻生活內容、使用避孕藥、決定生育、決定死亡、性偏好、職業或教育之選擇、個人外觀表現等之自由，亦承認對個人生活相關資訊之控制，以及資訊流通過程中及資訊本身內容之秘密與不受侵擾，甚至是個人身體完整及行動之自由，都被認為屬於憲法上隱私權之一環¹⁵⁶。

b. 隱私權應僅限於「對於秘密資訊的刺探與洩露」

之所以需要特別開啟此段落進行討論，係因為如前所述，無論是司法院或是我國眾學者反對 deepfake 條款保護對象包含隱私權的論理，都不約而同的隱含了潛台詞，即認為所謂隱私權指涉的範圍僅限於類似「個人資訊不當揭露型」的狀態。唯有如此，才能得出因為不實性影像中，臉的部分欠缺私密性、身體的部分雖具私密性但並非真實個人資訊，因而不屬於對隱私權之侵害的結論。

然而，此種將他人之資訊、臉譜，透過不實資訊加以誤導大眾，甚至進而商用的行為，事實上即屬於前述「公眾誤導型」與「盜用身份型」的典型案列。在學者介紹此兩種案件類型所舉的典型範例係「將被害人照片用以商業宣傳，使人誤會被害人係屬商品代言人」，即構成形象扭曲與身份盜用，而足以侵害隱私權。則在製作、散布不實性影音案例中，使公眾誤會被害人從事特定性行為，甚至進而營利的行為，自然也屬於隱私權之侵害。

直言之，若認為隱私權之保護範圍應及於 Prosser 提出的四種類型，則即不能僅以散布不實性影音不構成「個人資訊不當揭露型」之隱私權侵害，就想當然耳得出該行為不侵害隱私權之結論。因此，本文認為前述司法院或是我國眾學者反對 deepfake 條款保護法益包含隱私權的論理，於此部分範圍內有所缺漏。

針對此議題，雖說前述 Prosser 提出的四種隱私權類型確實成為被普遍接受的學說，然而隨後亦有學者對其提出挑戰，認為所謂之隱私權應僅限於「個人資訊不當揭露型」與部分之「入侵型」。究其論理，主要係基於四種類型中摻雜了大量不同行為類型、難以提取出共同的核心要素去解釋與定義隱私權的內涵，這

¹⁵⁵ 林建中（1997），前揭註 149，頁 56-59。

¹⁵⁶ 林建中（1997），前揭註 149，頁 59-60。

導致實務法院無法建構出一個具說服力的隱私權概念，反而持續做成錯誤而不可調和的定義，隱私權的定義與範圍界定猶如「颶風中的乾草堆」一般混亂¹⁵⁷。學者認為隱私權所涵蓋之範圍僅限於個人資訊的不當刺探與洩露，並指出在「入侵型」中，可能出現噪音或氣味物理性入侵私人領域的案件類型，此種不涉及個人資訊刺探之類型，不應被視為隱私權之侵害。此外，「公眾誤導型」應該以誹謗相關法規（the law of libel or slander）對其處理，而「盜用身份型」則僅涉及個人財產上利益之侵害，而與隱私權無涉¹⁵⁸。

事實上，在我國學說討論中亦曾出現相似之見解。於討論司法院釋字第 791 號解釋關於通姦違憲之議題時，學者即有討論，通姦罪禁止婚外性行為之行為，除了干預個人積極面向上之性自主權外，是否亦同時干預了個人之隱私權？直言之，司法院釋字第 585 號解釋提出：「隱私權雖非憲法明文列舉之權利，惟基於人性尊嚴與個人主體性之維護及人格發展之完整，並為保障個人生活秘密空間免於他人侵擾及個人資料之自主控制，隱私權乃為不可或缺之基本權利，而受憲法第二十二條所保障」¹⁵⁹，其所稱之「侵擾」涉及之範圍究竟為何？若依照本文前述介紹的美國實務見解，對於個人性活動的自主決定，早已被明文確定是屬於隱私權保障之一環。然而學者最終提出結論，認為此處所稱之「侵擾」應僅限於對於個人生活秘密空間的刺探，不包括介入他人對此生活秘密空間的安排。其理由主要係基於美國法上藉由隱私權概念納入保護的諸多權利，諸如人格發展自由、一般行為自由等，在我國早已由大法官明文發展出特定類型的基本權，而無須非得透過隱私權概念加以保護。將隱私權之概念擴張的過大，不僅無法加強對該項基本權保護之程度，反而徒增隱私權與該基本權間之區分困難，因此認為在我國的法體系下有將隱私權限縮解釋之必要¹⁶⁰。

本文基本上同意前述學者意見，認為在我國法體系下應該將隱私權的理解限縮於對於私密資訊之刺探與洩露。

事實上，美國法上之所以將隱私權理解成多元紛雜之面貌，有其法學發展史上，基於獨特文化土壤的不得不然。直言之，其是因為制憲於相當早的時期，因此無法充分預見社會變遷後伴隨而生的新興權利保障需求，同時因為判例法國家能夠以實定法滿足權利保障需求的能動性相對遲緩，才形成了美國法下透過隱私權概念大包大攬進行各種權利保障之實況。反之，大陸法系國家則傾向在面臨個別權利保障需求時，分別創設獨立之基本權加以保障人民權利。亦即，在美國法上必須透過隱私權概念加以處理的問題，在我國法體系下其實都得以以一般人格權之概念加以對應¹⁶¹，並無須效法美國將一切權利保障的希望，都寄託於對隱私權的擴大解釋。

¹⁵⁷ PARENT (1983), A NEW DEFINITION OF PRIVACY FOR THE LAW, Law and Philosophy 2, 305-306..

¹⁵⁸ PARENT (1983), *supra* note 157, 322-325..

¹⁵⁹ 司法院釋字第 585 號解釋理由書。

¹⁶⁰ 薛智仁（2021），〈通姦罪之憲法審查—評司法院釋字第七九一號解釋〉，《成大法學》，第 42 期，頁 20-22。

¹⁶¹ 林建中（1997），前揭註 149，頁 66。

依照前述說明，本文能夠得出隱私權概念於我國無須過度擴張之結論。然而為何限縮之範圍應該僅及於關於「私密資訊之刺探與洩露」而非其他種類，論者有提出論點認為係基於「與社會上慣習用語之配合」。直言之即是，「絕大多數人提到隱私權時，想到的都會是信件被偷看、住居被竊聽或其他資訊被公開，而不是墮胎」¹⁶²。此種「因為『大家』覺得隱私權是這範圍，所以隱私權就應該是這範圍」的觀點乍看之下有些過於「素樸」，然而深思後卻也並非全無道理。蓋如前所述，之所以相較於美國法上的寬泛認定，希望減縮、限制隱私權所涵蓋範圍主要之理由係基於「讓隱私權乘載太多不同概念會造成解釋上的紊亂」、「避免在嘗試追求單一明確定義時，反倒因為涵蓋種類過多而治絲益棼」、「許多美國法上必須透過隱私權概念加以保護之基本權，在我國已有其他基本權足以承擔保護的責任」，準此，則透過「大家」怎麼看待隱私權作為決定隱私權保護範圍之標準，以此避免解釋上的紊亂、成功追求單一明確的定義，在應然面上不失是一條可行的取徑。

然而即便「應然面」上可以以此作為決定保護範圍之標準，仍然必須回答的問題是，在「實然面上」是不是真的如學者斷言所說的：「事實上，國內法律實務或學說上，對於隱私權之理解或關注多半集中在『對於個人之相關資料之控制』，亦即使其保留在私密之裝態而不為他人所知之意義。」、「『絕大多數人』想到隱私權時，都會想到被竊視、竊聽、刺探隱私資訊」¹⁶³？對此，本文認為可以從幾個面向去佐證前述學者所述之主張。

首先是如前所述，面對 deepfake 條款修法，從司法院加註意見，到本文撰文當下爬梳過國內曾對此議題發表意見的 5 名學者意見，全都不約而同的直覺式的僅以「是否有秘密資訊之洩露」作為判斷是否涉及隱私權之標準，從中足見在前述學者發表關於「國內學術界對隱私權的探討只聚焦在『對於個人之相關資料之控制』」之個人斷言式觀察的 26 年後，其主張仍未隨著時代遞嬗，而有必須修正更改之必要。此外，根據中華民國刑法第二十八章妨害秘密罪章的體系架構觀之，亦可以發現我國至少在刑法體系下，對於隱私之保護確實聚焦在對於秘密資訊的刺探與洩露。直言之，第 315 條、第 315 條之 1、第 315 條之 2 前 2 款規定，係針對刺探秘密資訊之行為進行處罰。而自第 315 條之 2 第 3 款以降，連同第 316 條至第 318 條之 2 規定，則係針對洩露秘密資訊之行為進行處罰。由此可知，在我國刑法體系「預設值」中，對於隱私權之想像確實聚焦在對於秘密資訊的刺探與洩露。

綜上所述，基於若效法美國法上將隱私權範圍涵蓋進各種紛雜之類型，可能導致難以明確掌握隱私權之核心內涵，以至於導致解釋論上之困難，甚至導致司法實務上認定流於恣意。同時考量我國司法發展史與美國有不同的成長背景與脈絡，因此我國法體系中關於隱私權之認定確實應該限縮。而決定限縮至何種範圍之依據，以「『社會通念』如何看待、理解隱私權」作為標準，應能有效規避前

¹⁶² 林建中（1997），前揭註 149，頁 70。

¹⁶³ 林建中（1997），前揭註 149，頁 70-72。

述對於隱私權過於紛雜所生之擔憂，故可以做為合理之判斷依據。

而在客觀上，無論從我國學者近兩年討論不實性影像所展現之態度，亦或是我國刑法規範體系下之預設值，應都能認為在我國國內，認定隱私權時較無爭議的核心部分確實即是「對於秘密資訊的刺探與洩露」，準此，本文認為將對隱私權的理解限縮在此範圍內，應屬正當。

然而對於此類將隱私權侵害限縮於「與個人相關資訊之保密」的見解，亦有學者提出擔憂。細究其論點，其一是認為在現代社會運作過程中，難以期待對於個人的一切資訊都無須對社會公開，從而，要認定哪些與個人有關的資訊係該個人具有「獨佔性」，因此可以被認為是受隱私權保障之範圍，可能仍有判斷模糊標準之虞；其二是其擔憂若將隱私權的概念作限縮解釋，某些在過去透過隱私權概念加以保障之基本權可能會失所附麗，從而導致規範完整性的欠缺¹⁶⁴。對此，本文認為其所提出之擔憂，並不妨礙本文認為隱私權範圍應該做如上限縮的結論。

關於第一點的質疑，誠然「何種個人資訊始屬於具獨佔性，因此可被認為受隱私權保障」確實需要進一步提出解釋與判斷的標準，同時亦絕對可能產生模糊空間。然而在我國現行法體系下本就透過「合理隱私期待」等概念，去建構出可操作的判斷標準。雖說在這些標準下仍然存在邊際案例挑戰眾人的判斷，然而此部分之判斷是否屬於隱私權的模糊程度，與進行限縮前使隱私權必須涵蓋包山包海的案件類型相較，終究還是使判斷標準朝向更為清晰的方向前進。

至於有關第二點的質疑，在美國法下或許有此擔憂確屬合理，然而在我國法體系下本就如前述學者所言，實務上透過建構個別基本權的方法，去處理美國法下全部交給隱私權涵蓋的案件類型。準此，即便對隱私權做出較為限縮的解釋，也應不至於導致使基本權保護產生真空的狀況。

從而，亦可回應前述比較法上學者認為「即便考量到不實性影音中私密部分並非真實，鑒於其有可能造成公眾誤解，因此仍然屬於隱私權侵害」的主張。此主張在思路或許受到 Prosser 所稱之「公眾誤導型」隱私權侵害影響，才會得出「因為公眾會誤解，因此仍屬隱私權侵害」之結論。然而基於本文並不認為該種類型之行為屬於我國法體系下隱私權應保障之範圍，故該學者之主張自然亦不可採。

此外，亦有學者在討論「資訊隱私權」時，強調其意義在於「在沒有通知當事人並獲得其書面同意之前，資料持有者不可以將當事人為某特定目的所提供的資料用在另一個目的上」，並認為其核心理念係「個人不僅是個人資料產出的最初來源，也是其正確性與完整性的最後查核者，以及該個人資料的使用範圍的參與決定者」¹⁶⁵。若按此理解，於不實性影音案件類型中，光是被害人真實的臉，未獲同意被用在其他目的上時，就其真實臉部的資訊控制被侵擾，似乎就屬於對隱私權之侵害。

¹⁶⁴ 林建中（1997），前揭註 149，頁 68-69。

¹⁶⁵ 廖緯民（1996），〈論資訊時代的隱私權保護—以「資訊隱私權」為中心〉，《資訊法務透析》，第 8 卷第 11 期，頁 22。

然而該文討論「資訊隱私權」主要係針對憲法上基本權、干預基本權在行政上應負擔之責任進行討論，而並非與刑法直接掛鉤。又觀察刑法體系中對於使個人資訊免於被刺探與揭露的保護上，對於該資訊都特別強調其「秘密性」，無論是刑法第 315 條強調「封緘」、第 315 條之 1 與第 315 條之 2 於實務發展上形塑出了「合理隱私期待」標準作為判斷客觀構成要件的一環、第 316 條至第 318 條之 2 甚至明文強調秘密作為構成要件之一部分。準此，在現行的刑法架構下並非所有的資訊刺探與洩露皆被保障，而是很明確地強調須具備秘密性的資訊始受保障，如此一來，則於不實性影音案件類型中，被害人應不能僅因自己的臉譜資訊被未經同意的使用於其他目的上，即主張隱私法益受侵害而主張應受刑法保護。

至此，本文認為基於隱私權保障範圍應僅限於「對於秘密資訊的刺探與洩露」，則前述司法院加註意見與我國眾多學者所提之見解應為可採，於不實性影像案件中被被害人被洩露之臉譜資訊不具備秘密性，而可能具備秘密性的身體部位並非被害人之真實資訊，故不應認為有隱私權之侵害。

第二款 名譽權

第一目 實務見解

於 2023 年刑法修法歷程中，法務部提出之草案總說明¹⁶⁶，於關於新設第二十八章之一「妨害性隱私及不實性影像罪」章之介紹部分，提及「而**名譽權**旨在維護個人主體性及人格之完整，為實現人性尊嚴所必要，受憲法第二十二條所保障。」然而，在其後該章節個別條文之立法總說明中，皆未提及究竟是何者法條涉及名譽權之侵害，此舉亦遭司法院於審查時加以批評。

司法院加註意見¹⁶⁷中提及：「惟對照本章各罪所保護之法益，……，均未提及保護『名譽權』，則究竟『名譽權』係哪一罪名所保護之法益？尚有未明。」、「本院認為刑法第三百十九條之四之罪（未經同意散布他人不實性影像罪），保護之法益包括『名譽權』」、「散布以深偽科技方法製作之不實性影像（如將商業 A 片主角置換為被害人之臉），傳述個人不實之生活事實（誤以為被害人拍過色情影音，或該不實性影像為本人真實之性影像），造成個人社會評價貶低、社會形象減損之結果，除人格法益受到侵害外，亦侵害個人名譽權。從而，建議對刑法第三百十九條之四之罪，於立法理由補充說明保護法益包括『名譽權』。」，而司法院代表於二讀黨團協商時亦發言表示¹⁶⁸：「另外也應該有保護名譽權，因為在這種狀況之下會使被害人遭到別人的歧視、嘲笑、攻擊或騷擾等等，所以他在社會上的人格權及地位評價都有受到貶損。」

面對司法院的挑戰，法務部代表於二讀黨團協商階段亦有所回應，其指出¹⁶⁹：「第一點，如果規定在這個章名是比較不合適，因為已經有妨害名譽章節，這裡

¹⁶⁶ 立法院公報第 112 卷第 14 期委員會紀錄頁 130。

¹⁶⁷ 立法院公報第 112 卷第 14 期委員會紀錄頁 131-130。

¹⁶⁸ 立法院公報第 111 卷第 84 期委員會紀錄頁 125。

¹⁶⁹ 立法院公報第 111 卷第 84 期委員會紀錄頁 124。

又多一個名譽，不是重複了嗎？」、「第二點，如果你認為是屬於包括妨害名譽的保護法益……法院實務上的適用就認為你還要證明妨礙秩序的故意，所以很多檢察官辦案就不起訴處分，因為起訴都是判無罪。……搞不好檢察官認為還要證明這一點很困擾……將來這個罪要成立的話，你就要證明他也有妨害名譽的故意。」等語，簡言之是基於1.擔心有重複立法之虞2.擔心實務上難以證明妨害名譽之故意，因而明確反對整個章節保護之法益包括名譽權。隨後在各立法委員分別針對此議題發表亦，法務部代表最終甚至表示¹⁷⁰「是新興的犯罪型態，這個保護法益到底是多少，有沒有包括名譽權，這個部分可能得慢慢靠學說與實務上的發展。」

綜觀法務部立場，其代表於黨團協商時之發言，與其自己所提出之草案總說明明顯衝突，究竟哪個立場才是其真意其實難以判斷。又，其作為提出草案之機關，最終居然也無法確定保護法益之範圍，而將確認保護法益之責任推卸予實務與學說之發展，展現出了一種「我還不知道處罰這行為係為保護什麼法益，但總之我就是想要先處罰再說」的態度，此部分之表態明顯不負責任，應該受到相當譴責。

而在立法委員態度方面，在黨團協商階段有認為係屬名譽權侵害者，如江永昌委員發言指出¹⁷¹：「今天要修法不外乎就是想要規範關於偽造、甚至是冠名影像，尤其是在網路上散播，對於名譽造成重大損傷。」；有認為不僅是屬於名譽權侵害者，如林楚茵委員發言指出¹⁷²：「假設我的臉被變了之後，不只是個人的名譽損失，甚至有可能是金錢上、甚至於其他更大的損失，而你只用妨害名譽去偵辦是不夠的。」、亦有認為不屬於名譽權侵害者，如范雲委員發言指出¹⁷³：「有人認為名譽有受損，其實是因為污名化造成他的名譽受傷，本來不是名譽的問題，他就是一個受害者，我們被車撞了不會說是名譽受傷，對不對？為什麼我的性私密影像被揭露，最後會傷害名譽？」

總之在此次修法審理過程中，對於保護法益是否涉及名譽權之議題，從法務部到司法院，甚至各立法委員之間，立場呈現相當的紛雜。然而最終基於討論時間的緊迫，並未達成任何共識，法務部代表雖於黨團協商中主張¹⁷⁴：「名譽權是否會間接受到影響，這部分是否可以在立法理由中附帶提及，我們會再研究。」，惟最後終究仍以法務部原先提出的版本送出委員會審查通過，且並無附上任何理由說明研究後何以仍做出維持原狀之決定。

因此此議題事實上很諷刺的，如同法務部代表所言，呈現出了一個完全無法探求立法者意旨的狀態，被迫僅能透過實務與學說之討論與發展，進而探明保護法益究竟是否包含名譽權之保障。

第二目 學說見解

¹⁷⁰ 立法院公報第 111 卷第 84 期委員會紀錄頁 123。

¹⁷¹ 立法院公報第 111 卷第 84 期委員會紀錄頁 19-20。

¹⁷² 立法院公報第 111 卷第 84 期委員會紀錄頁 124。

¹⁷³ 立法院公報第 111 卷第 84 期委員會紀錄頁 124。

¹⁷⁴ 立法院公報第 111 卷第 84 期委員會紀錄頁 125。

對於此爭議，針對 deepfake 條款訂定專法之議題進行討論的學者當中，多數皆認為其保護之法益包含名譽權，甚至認為既存實定法下之妨害名譽罪章即足以處理相關問題。例如有認為「如因而有名譽受侵害的問題，透過妨害名譽罪章即可加以處理。就此種情況，似無須在以保護性的隱私為主的罪章中，另外以最高法定刑 5 年以下有期徒刑的罰責來加以應對。」¹⁷⁵、「純粹以電腦合成或其他科技方法製作不實性影像的行為，實質上得以定性為誹謗的預備行為。就算今天此罪被刻意安排在新增的妨害性隱私罪章，還是無法否定立法本身就是另外一種形式的刑罰前置。」¹⁷⁶、「其實可以妨害名譽罪取代的情況，難以突顯本罪的獨立規範價值。」¹⁷⁷等。

相對於此，亦有少數意見認為不宜認為其保障法益涉及名譽權，學者即有指出：「之所以要針對此類案件另訂條文，而不直接沿用目前實務作法，即援引妨害名譽罪處理，原因或許可以理解為，……，正值社會性道德觀感快速變動的時期，不宜將之完全等同於個人外在名譽的評價。詳言之，……，但其中隱含的性道德內涵，卻也使得對性私密狀態的維護，有承認被害人『污名』的危險。」¹⁷⁸、「此外，鑑於近來刑法上的妨害名譽罪章亦不乏限制適用的呼聲，將性要素的社會複雜考量排除於一般名譽犯罪之外，也可避免其間接影響將來針對妨害名譽罪的檢討。據此，若本草案通過，則諸般散布性影像之行為，宜論以單一的妨害性隱私或性人格犯罪，而不再援引妨害名譽罪相關規定。」¹⁷⁹直言之，其論點主要係基於若認為 deepfake 條款保護法益涉及名譽權，一來可能導致性的汙名化，二來可能影響到對於妨害名譽罪檢討之浪潮。

第三目 本文見解

綜觀前述司法院之立場，以及支持 deepfake 條款所保護之法益係名譽權之學者見解，其立論的底層邏輯都是斷然的認為不實的傳遞「他人曾從事性行為」或「他人曾在從事性行為時拍攝影音」的虛假事實，必然導致該他人之名譽減損。

事實上這也與學說與實務長期以來的主流見解相符。例如有學者指出散布不實性影像之行為形同傳述被害人「拍過色情影音」的不實事實，因此會造成客觀名聲貶低的效果，從而得以成立刑法第 310 條誹謗罪¹⁷⁹。另也有認為此行為會導致公眾得到「被害人曾讓人拍攝過性相關影像」與「被害人曾與人發生性行為」的錯誤印象，從而「易使被害人於社會上遭受訕笑、譏諷」，因此足以認定成立名譽之傷害而可以成立刑法第 310 條誹謗罪¹⁸⁰。

¹⁷⁵ 林琬珊（2023），前揭註 14，頁 51。

¹⁷⁶ 古承宗（2023），前揭註 14，頁 19。

¹⁷⁷ 前揭註 14，頁 37。相似見解：許恒達（2022），前揭註 4，頁 16。這邊必須注意的是，許教授雖認為此類行為得以構成誹謗罪，然而其根本上認為此類行為侵害的並非名譽權，而係個人形象權。其中差別在於前者保護者係「公領域生活歷程所連結的社會名聲」，而後者則是「個人私領域生活史衍生的個人形象」。本文將於隨後詳述之。

¹⁷⁸ 廖宜寧（2022），前揭註 19，頁 168-169。

¹⁷⁹ 許恒達（2022），前揭註 4，頁 20。

¹⁸⁰ 吳芳毅（2020），前揭註 22，頁 191。

而過去實務在討論復仇式色情案件中，散布他人真實性影像是否構成誹謗罪時，亦有指出「被告……將告訴人之裸照 7 張及 GIF 檔 1 個上傳至臉書網站，並於 110 年 3 月 9 日以臉書帳號『Yen Huang』分享告訴人先前在臉書網站使用之大頭照，暗指其上傳之裸照、GIF 檔內之人即為告訴人，復於 110 年 7 月 5 日將 B 影片上傳到國外色情網站上，同時加註告訴人之真實姓名，顯係以散布圖畫之方式，指謫告訴人之行為開放、隨便，為不在意社會道德觀念之人，客觀上確足使一般人對告訴人產生行為不檢、道德倫常觀念低落之負面觀感，而對告訴人之品德、身分、人格、地位造成相當貶抑。」¹⁸¹之結論。

然而本文認為前述見解從實然面或應然面上觀察之皆不恰當，以下分述之：

a. 實然層次

在實然面上，即有學者指出普通公眾接觸到這些性相關影音時，至多僅會得到「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」等資訊，而是否會因此即認為影音中之參與者放蕩、隨便、任意讓人拍攝性相關影音道德上有瑕疵，完全取決於主觀想像或社會道德價值。準此，學者批判法官逕以自身之道德價值，推導出「讓人拍攝私密照片者，在性方面之價值觀有問題」負面價值判斷結果並不合理¹⁸²。亦有指出「未經被害人同意散布被害人之性私密影像，……，則比較容易會有證明上之困難。」¹⁸³

本文認為前述學者批判之見解擲地有聲。例如，美國於 2022 年最新的研究顯示，大多數的人民都認為成年人合意分享裸照的行為，無須被太多的汙名化（stigma）¹⁸⁴。本文以「裸照/私密照」、「民調/研究/調查/民意態度」等關鍵字於 Google 的搜尋引擎、碩博士論文查詢系統與法源或月旦等法學搜尋系統等處檢索，並無發現我國有做成類似之研究，然而自前述美國的調查可知，渠等行為之群眾評價會隨著地域、個人感受與生命經驗，甚至可能因為時空環境變遷而有所巨大的差異，公眾對於此類案件的被害人是否真的傾向展露負面的、敵意的態度，需要更多的證據加以證明，不得僅以立法者之直覺即妄加論斷。

看回台灣的經驗，退步言之，縱使證明公眾對於此類案件的被害人傾向加以敵意的態度，而此態度也確實導致被害人情感上受傷，然而這樣的受傷是否係屬名譽上之侵害亦非無疑。所謂的名譽，是對特定人表現於外的尊敬，亦是對他人

¹⁸¹ 臺灣臺中地方法院 111 年度訴字第 330 號刑事判決。臺灣高等法院 111 年度上易字第 920 判決、臺灣高等法院臺中分院 110 年度上易字第 800 號判決、臺灣高等法院 108 年度上易字第 1379 號判決皆同旨。渠等判決所指涉者都是散布真實之性影像之案件，惟不妨礙本文說明「實務見解認為使公眾知悉拍攝性影像，足以構成對名譽之妨害」。

¹⁸² 張天一（2019），前揭註 93，頁 29。

¹⁸³ 法思齊（2021），〈論以刑事法規範復仇式色情（Revenge Porn）之可能——以日本情色報復受害防治法為中心〉，《月旦法學雜誌》，第 314 期，頁 213。

¹⁸⁴ SWNS, Shocking number of people have sent nude photos, poll finds, *New York Post*, in: <https://nypost.com/2022/09/13/shocking-number-of-people-have-sent-nude-photos-poll-finds/>.（最後瀏覽日：07/13/2023）

存在價值的承認，是社會對於個人品格、能力所為的價值判斷¹⁸⁵。而所謂名譽向來在學說上認為有「內部名譽」、「外部名譽」與「主觀名譽」之區別。其中「內部名譽」是人格內部的價值本身，被認為不存在自外部被侵害的可能；「主觀名譽」是又稱「感情名譽」，係指特定人對於社會或他人就其人格價值之評價的主觀感受或反應，對於自己所具有之名譽感情。此種名譽被認為標準隨著每個個體浮動，具有高度不確定性。基此，我國學說與實務通說皆認為刑法所保護之法益僅限於「外部名譽」。所謂「外部名譽」係指社會對於特定人之「人格價值」所賦予，事實上得以加以檢驗之評價¹⁸⁶。以前述標準檢討此類性影像洩露之案件，在此類案件中被受害人遭受到的敵意很多時候是來自「檢討被害人」的風氣，例如被警察質疑「誰叫你當初要拍？」等語，而在此過程中被受害人也會因此感到受傷¹⁸⁷。

本文當然承認這種受傷的感覺是存在的，然而並不是被受害人有所受傷即代表係名譽權受侵害。如前所述，名譽權保障的核心應係「對人格價值之否定」，然而檢討被害人的風氣縱使不討喜，其說到底針對的仍是被害人的行為是否適當、是否係造成傷害的肇因之一作出評價與非難，而非針對被害人的人格本身做出否定。誠如范雲委員在黨團協商時的發言，一般通念下我們並不會認為車禍被害人的名譽受損，即便車禍被害人亦可能被「檢討被害人」，在事發後被指責行人過馬路不應走太慢、機車不應騎在內線車道等，在這些情境下受害者亦可能感到憤慨、受傷，然而這種情緒應更多的來自「我沒做錯事卻被指責的委屈」，而與名譽的受損未必相關。

具體來說，「檢討被害人」的本質是於事實層次（明明沒有違規被錯誤指摘有違規）或歸責層次（確實有違規，但與肇因無關）錯誤地指摘被害人具有過失、違反特定義務。本文認為若被錯誤指摘的是違反對他義務，例如過失導致他人受有損失，固可能屬於名譽侵害，蓋此時形同指摘被害人對他人的權益有所輕忽、不尊重，具有道德非難性。惟若被錯誤指摘的僅是違反對己義務，例如拍攝性影像案件中的「過失未能保護好自己」，此際是否仍屬名譽侵害本文即有所質疑，蓋「過失未能保護好自己」並不具備「過失傷害他人」一般的道德非難性。

同理，如前所述於此類案件中被受害人所受之不適感來源，亦可能是被眾人凝視與意淫的不適感。本文亦無意否定被害人感受到的此種精神壓力與受傷，然而同樣以是否存在「對人格價值之否定」作為標準檢討之，此類案件亦可能不必然與名譽權受損直接相關。直言之，即便在此際，假設那些觀覽影音的閱聽人對於被害人全是稱讚，表態認為被害人面容姣好、性能力很好、拍攝性影像的行為很值得肯定，也很難想像會因此減緩被害人的不適感或傷害。誠然，本文也同意此種「被凝視」的感受，某種程度上也可能雜揉著憂慮凝視者可能對自己抱持負面觀感（例如擔心他人覺得自己很淫賤等）的不適感，此確實可能與名譽法益侵害

¹⁸⁵ 張麗卿（2016），〈刑法對誹謗言論的合理規制〉，《臺大法學論叢》，第45卷第4期，頁2003。

¹⁸⁶ 林琬珊（2018），〈妨害名譽罪與負面標籤〉，《月旦刑事法評論》，第9期，頁69-71。

¹⁸⁷ 婦女救援基金會，〈有錢就能散布別人裸照，誰讓加害者有恃無恐？〉，<https://twrf.org.tw/info/title/151>（最後瀏覽日：07/13/2023）

有關，因此本文並非主張不實性影像全然不可能侵害名譽法益。然而，本文認為此類案件中被被害人所感受到的傷害來源可能很多，本質上並不完全來自「對人格價值之否定」，因此不宜將被害人感受到的傷害都想當然地歸因於名譽權受侵害。

綜上所述，在實然層面上，就美國的經驗而言，公眾未必認為「被害人曾讓人拍攝過裸照」或是「被害人曾與人發生性行為」是必須承受汙名的行為。台灣雖無相似研究，然而至少說明了社會面對此類資訊之態度可能隨時間與觀念改變而有所流動，因此實務法院過往僅依一般通念所為之價值判斷，也應與時俱進，時刻檢討於我國當前時空的社會環境是否仍屬於同樣的情況。退步言之，縱使從台灣過往性影像洩露的案件觀之，其中確實可能導致被害人面對一些來自公眾的敵意或不友善，而這些公眾的態度也確實使被害人感到受傷，然而此種受傷是否即屬名譽權的侵害並非無疑。所謂名譽權的核心應聚焦在「對人格價值之否定」，準此則若公眾的不友善係「檢討被害人」或是「凝視被害人的意淫」，本文完全同意被害人與此時可能感到受傷與不舒服，然而此種傷害應與對人格價值的評價未必全然相關，而不應完全地歸因為名譽權之侵害。

b. 應然層次

本文前一個段落已經對「洩露性相關影像之行為，是否在實然層次上會使被害人名譽減損」做出檢討與批判。退步言之，縱使承認實然層次上社會態度真的仍會因為該影像之流出，而降低對於當事人的評價，此際應該思考的下一層次命題是，應然層次上是否應該即因此肯認此屬名譽權之侵害？

對於 deepfake 條款所保護者是否為名譽權之爭議，重新檢視前述來自范雲委員的反對觀點，其指出：「就是不實性影像罪，如果能夠界定這一章的重點在這裡的話，本席認為這樣是比較好一點。有人認為名譽有受損，其實是因為污名化造成他的名譽受傷，本來不是名譽的問題。」¹⁸⁸「為什麼我的性私密影像被揭露，最後會傷害名譽？」¹⁸⁹；而來自學術界的反對觀點則指出：「不宜將之完全等同於個人外在名譽的評價。詳言之，雖然一個人的性活動與其人格形成相關，在規範上具有高度個人專屬性，但其中隱含的性道德內涵，卻也使得對性私密狀態的維護，有承認被害人『污名』的危險。」¹⁸⁹「此種規範判斷傾向，將與當前學說、社會運動者致力於性的去污名化期待不符。」¹⁸⁹

綜觀前述兩種聲音，可以發現其所關注者已非聚焦在實然上「社會如何看待性影像被散布之被害人」，而係進一步在討論「作為立法者應該如何看待性影像被散布之被害人」更甚者是「作為立法者應該如何導引社會看待性影像被散布之被害人」，也因此才會特別強調「汙名化」作為關鍵字。

¹⁸⁸ 立法院公報第 111 卷第 84 期委員會紀錄頁 124。

¹⁸⁹ 廖宜寧（2022），前揭註 19，頁 168-169。必須注意的是，依照該文脈絡，學者於此所主要討論的對象應係涉及到真實性影像的復仇式色情，惟本文認為其論理之脈絡於不實性影像之部分亦無明顯不同，併予敘明。

對此，本文持高度肯定的觀點。我國學者有介紹到，同樣是「外部名譽」，尚得以「事實名譽」與「規範名譽」兩個概念加以區分。所謂「事實名譽」係指事實上來自社會對於特定人所加諸的社會評價，相對於此，「規範名譽」指的是有規範性、觀念性的名譽。直言之，「事實名譽」即係討論實然上社會公眾係如何看待該事實，而「規範名譽」則著重在應然上應如何看待該事實。

依據學者介紹，近年來日本法上近來開始有認為，應該將名譽理解為「受到正當社會評價的權利」，嘗試從規範面來理解名譽權。具體操作即是，外部名譽之事實應被限縮在「於該人之責任所能變更之事實」範圍之內，主張應該將有關身體障礙、精神障礙、疾病、血統、階級等之事實，亦即將「單靠自己的努力也無法變更的事實」排除於名譽的範圍之外。學者之所以認為應從規範面來理解名譽的概念，主要原因即是擔心若認為這些事實會導致名譽之減損，則反而會追認、僵化、固著化社會之偏見，對於人格的保障並無益處¹⁹⁰。

舉例來說，若在種姓制度的社會中，兩個婆羅門於爭執時，一人斥責對方「曾與賤民結婚」，純粹依照「事實名譽」來理解，由於公眾實然上確實對「賤民」一詞抱持著高度負面的理解，因而很可能得出站在被害人之人格確實因此被否定、貶抑，而得以成立名譽權之侵害。反之，若是依照前述規範面的理解，則會認為「賤民」係天生即確定的階級，非屬「於該人之責任所能變更之事實」範圍內，而並不成立名譽權之侵害。同理，罵人「神經病」、「智障」等語，在規範面的理解下，亦會得出非屬名譽權侵害之結論。

此觀點曾面對一些學術上的挑戰，主要質疑點在於 1.被以這些事實加以指責，客觀上就是有受到名譽減損 2.認為這些指責是名譽的減損，未必代表著固著化與追認。面對前述反對見解的挑戰，學者指出：其「只是僵硬的劃分立場與操作形式邏輯」、「如果不承認社會偏見，也無法認為是名譽毀損，倘若一併考慮判決可能發揮的社會機能，則實難謂法院判決不會追認或固著化社會偏見。」，且學者亦指出，若某意思表示於客觀上真會導致事實上名譽之減損，真正導致貶低意涵者應係社會之觀念，而非該意思表示本身。也因此，該學者肯認：「有必要考慮在以社會名譽為判斷基準的同時，亦應將規範名譽當作名譽的底限來思考」，政府即便無法積極地消除社會對於特定群體貼上的標籤，至少也不能反而去認證該刻板印象，為其提供正當性。

同時值得注意的是，其該著作所針對的案例事實係「指責他人有如特定名人一般，是否成立妨害名譽」之討論，故其動用規範面思考的範圍似亦不僅限於「身體障礙、精神障礙、疾病、血統、階級等之事實」等界線內¹⁹¹。準此，本文認為其適用範圍應擴及至一切「本不應受負面評價，僅係因社會脈絡始承擔負面評價」之一切情況，而「從事性行為」、「拍攝性影像」皆應屬之。

同樣涉及「指責他人有如特定名人一般，有無侵害名譽權而涉及民事侵權」之案件中，另有學者指出：「基於多元社會的價值相對論來說，司法權並不適合

¹⁹⁰ 林琬珊（2018），前揭註 196，頁 69-72。

¹⁹¹ 林琬珊（2018），前揭註 196，頁 72-73。

對於任何個人的社會評價或名譽價值進行判斷，因為法院僅被賦予權限闡釋憲法及法律所預設之價值，而無權以法律外之價值觀審判他人。若在判決中認定將他人比擬為特定第三人之行為係構成名譽權之侵害時，無異於以判決否定該特定第三人之人格或名譽，實非妥適。」¹⁹²雖然該學者著作旨在討論「司法權」於「民事侵權案件」中應如何認定名譽權侵害的問題，與本文欲討論「立法者」於「刑事案件」中應如何認定名譽權侵害似有區別，然而考量到刑事制裁的嚴厲性，其認定侵害權利的標準理應較民事法上不法更為嚴格，若連民事侵權都不構成，則無在刑事上構成不法之理。此外，其論理基礎在於保護「多元社會的價值相對論」，於此目標下無論是司法部門抑或是立法部門，都理應認知，並審慎動用自己對社會的價值引導機能，是故學者見解於本文之討論中仍有其可以被援引之價值。

根據以上討論，針對「名譽權侵害之認定是否應採納進規範面的判斷」此一爭點上，主要需討論的核心應是 1.以國家身分去認定某個事實會減損名譽，是否將導致固化偏見之效果？2.若前者答案是肯定的，然而客觀上被害人也確實會因為該評價而遭逢事實面上之名譽傷害，此際應如何權衡？

關於第一個問題的答案，本文採取肯定的立場。即便在理想模型中，國家可以「技術性」的聲稱自己只是知悉該事實的社會評價、並且依照社會評價進行判決或立法，本身並不對該事實做成自己的評價。然而在現實面下，人民對於國家高權介入此類事件的理解，顯然不僅是消極的服膺社會評價，而係更為積極的「認證」。若以「法院 認證 公然侮辱」作為關鍵字搜尋，會發現大量案件，即便各自判決中可能有合理的論證脈絡，例如除了以特定名人指摘被害人外尚帶有其他負面批評的字眼，然而大眾傳媒與人民所接收到的訊息即是去脈絡化的「法院認證了某個人是侮辱人的用詞」¹⁹³。準此，如前述學者所言，「如果不承認此般社會偏見，也無法認為是名譽毀損，因此可能導致偏見的固化」應是可資認同的結論。

準此，則須進一步討論第二個問題，即若實然面上確實導致系爭個案中個體名譽被貶抑，但應然面上基於擔心加固刻板印象而認為理應不該認為認為有名譽權之侵害，此相衝突二價值應如何權衡？本文認為此應回歸刑事處罰名譽侵害的正當性基礎觀察之。直言之，學者有認為公然侮辱罪的正當性基礎在於避免標籤理論的效應，使個案中的被害人不至於因為遭受負面評價而逐漸修正自我印象，致使產生更嚴重的違反社會規範行為¹⁹⁴。如此一來，既然透過刑罰保護名譽權之

¹⁹² 葉啟洲（2012），〈罵人「共匪」或「胡錦濤」與名譽權侵害——臺灣板橋地方法院一百年度訴字第一三五—號民事判決〉，《月旦裁判時報》，第 13 期，頁 91-92。

¹⁹³ 例如：自由時報（05/05/2019），〈注意言辭！法院認證罵人「張淑晶」屬公然侮辱〉，<https://news.ltn.com.tw/news/society/breakingnews/2780500>（最後瀏覽日：07/13/2023）；華視新聞（03/12/2021），〈納豆身高遭法院認證侮辱？本人傻眼...大咖藝人相挺〉，<https://news.cts.com.tw/cts/entertain/202103/202103122034442.html>（最後瀏覽日：07/13/2023）；CTWANT（09/01/2021），〈嗆「長得像黃安」法院認證公然侮辱！黃安怒飆髒話回應了〉，<https://www.ctwant.com/article/137187>（最後瀏覽日：07/13/2023）

¹⁹⁴ 曾思堯（2017），〈從公然侮辱罪的存廢談保護法益的實態〉，《國際刑事法學之新脈動：余振華教授六秩晉五壽誕祝賀論文集》，第 1 期，頁 702-703。

目標即是避免標籤化，然而卻在個案中會為了保護名譽反而導致更多的標籤化現象，自然屬於手段無助於達成目的而違反了比例原則中的適當性原則，因而不應做成如此之立法。

舉例來說，如果處罰散布性影像行為的目的是為了保護被害人的名譽，然而法院認證「被公眾得知從事性行為還拍攝影像是可恥的」一事，卻反而可能是傷害該被害人的根源之一，日後的每個人都可以持著法院認證的結論去指摘當事人是可恥的，如此下來不僅在個案中對當事人而言並非有利，還會讓傷害擴及到與該案無關的，其他同樣屬於該群體的人民，自然是不適當的。

誠然，前述學說將認為應以規範面加以限縮的範圍，侷限在「於該人之責任所能變更之事實」，或是「以特定人作為侮辱用詞」的狀況，然而本文認為「被害人曾讓人拍攝過性影像」或是「被害人曾與人發生性行為」亦應有所適用。蓋之所以要用規範名譽的概念加以限縮，其目的係屬保護「多元社會的價值相對論」、避免固化刻板印象反而傷害人民等，則於散布性影像相關的案件中亦存在這些考量，而無允許國家「認證」拍攝性影像是可恥之理。更何況，在學說所舉的典型案例中，例如對著一名白人辱罵其猶如黑人一般，若被認為「猶如黑人」是侵害名譽，可能導致非裔群體被汙名化、對其的刻板印象與偏見被固化，然而至少個案中的被害人即該白人不致於因此被波及。

反觀，在散布性影像類型的案件中，國家出手「認證」拍攝性影像是可恥的，除了會對於社會中潛在有意從事此行為的人加以汙名外，連個案中的被害人都是被「認證」的主要影響對象，可能存在因為法院的行為更強化被害人的自我認同感低落的風險。於此範圍內相較於前述的以不實的「身體障礙、精神障礙、疾病、血統、階級等之事實」辱罵他人相較，規範手段無助於達成目的之程度無疑是更嚴重的。

c. 結論

綜上所述，本文認為在於 deepfake 條款所欲規範的案件類型中，一來欠缺紮實的證據，足以證明令人錯誤的相信「被害人曾讓人拍攝過性影像」或是「被害人曾與人發生性行為」在實然面上確實會導致被害人名譽受損。縱認在實然面上認會導致被害人事實名譽受損，然考量到以刑事規範保護名譽權目的之一在於避免標籤化、國家對特定事實做成價值判斷可能形成追認與固化偏見之效果以及，「多元社會的價值相對論」應積極被確保等理由，本文認為在認定名譽權侵害時須納入規範面的考量。

準此，國家不應「認證」「被害人曾讓人拍攝過性影像」或是「被害人曾與人發生性行為」是導致名譽減損的行為，以免個案中被害人反而因此受到更多之傷害，從而與刑法保護法益之機能有違。也因此，本文認為 deepfake 條款所保護之法益並非名譽權。

然而本文亦認為，同樣是「未侵害名譽法益」，此種「事實名譽確實有所減損，而僅因考量規範名譽概念，為避免標籤化而不認定係屬名譽法益之侵害」，

與傳統型態之「事實名譽未有減損」，仍有本質上之差別。亦即，本文承認被害人確實可能因事實名譽之減損而受有特定傷害，且認為此部分之減損與傷害亦應有受法律保護之必要。為求同時兼顧「適切地評價與保護事實名譽受減損之情況」與「避免標籤化帶來之社會偏見固著化」兩個目標，本文認為可能可行的方案係找尋到其他有別於名譽權之保護法益類型，且在劃定該保護法益之範圍時，亦將事實名譽受減損之評價一併涵蓋在內。

誠然，將這樣的操作細緻拆解，本質上其實是將「事實名譽雖有減損，然而依照規範名譽不應減損」的案件類型，從名譽法益保護的範圍內剪除，而另外依託其他法益，藉以評價事實名譽之減損。此主張仍可能面臨前述「只是僵硬的劃分立場與操作形式邏輯」的挑戰。亦即，此操作宛若把公眾看做朝三暮四的猴子，一方面認為若政府傳遞出「『散布他人拍攝性影像之謠言』是侵害名譽法益的行為」此一訊息，公眾會難以理解「政府只是承認了事實層次上確實名譽有減損，惟應然層次上不代表政府認可此類行為應該減損名譽」之潛台詞，因而仍然會產生固化刻板印象的後果。然而另一方面，卻又認為若政府傳遞出「『散布他人拍攝性影像之謠言』不是侵害名譽法益，而是侵害其他法益之行為，且在該其他法益內涵中，尚包含了事實名譽減損之評價」的訊息，公眾即不至於固著對於該行為之刻板印象或標籤化。

惟本文在考量如上顧慮之後，仍然認為做出這樣的操作有其意義與必要性。蓋之所以要引入規範名譽之概念，係為了避免某本不應遭受負面評價之族群或行為，被政府公權力貼上了負面形象的標籤，但此一考量本身並非有意替「明知公眾會給予事實上的負面評價，卻仍以不當用詞傷害他人」之行為人開脫，亦必須正視「被害人之事實名譽確實受損」之現象。

以前述婆羅門吵架的例子來說明，即是此過程中無論是自行行為人的主觀不法性，亦或被害人客觀上所受之侵害觀之，相較於典型的妨害名譽案件可能都並未有太大不同，僅係因為為了避免使「賤民」一詞被標籤化，才認為不宜以妨害名譽罪章處理之，惟此過程中的主客觀不法仍應有與以處理之必要，否則最終若政府傳遞給公眾的資訊變成「政府劃設出了一系列式實上可以傷害到人，但卻不會有法律上責任的詞彙庫」，亦應並非倡議採規範名譽觀點之學者所樂見。且那樣的社會中，若最終形成了所有人一言不和就把「賤民」一詞當作合法的髒話進行辱罵，而使這個詞彙大量與負面的脈絡聯繫在一起，終局上對於此一族群之地位改善也並非正面的幫助。例如「支那」一詞本無貶意，清末思想家梁啟超亦曾自稱「支那少年」，惟經過特定使用情境後，在日本形成了對中國之蔑稱，二戰後經駐日盟軍司令部調查認為屬於歧視性用語，要求日本停止使用。本文想藉此說明的是，讓一個詞彙由官方貼上會貶損名譽的標籤，固然可能加劇刻板印象。然而過猶不及的是，若使一個詞不受節制的被泛用，而持續與負面情境相連結，終局上也可能使本來無貶意的詞彙反而與負面情境緊密貼附而無法脫鉤。

站在「此類主客觀不法仍應有加以處理之必要」之前提下，本文認為尋求一個與傳統公眾觀念中「名譽」有一定距離之不同種類的法益，藉此涵蓋進對於事

實名譽侵害之評價，使該法益作為處理「事實名譽有減損，但因為規範名譽概念而認為不宜被認為是名譽侵害」案件類型之避風港，應是值得試之方案。蓋本文認為，若欲期待人民得以細緻的理解同樣使用「名譽」一詞的概念下，其實包含了事實名譽與規範名譽兩個概念，且實然層面的事實認定應該與應然層面的價值判斷做脫鉤等等，依照筆者的生命經驗來說這無疑是困難而不現實的。蓋就連要求人民理解「『A 指控 B 貪汙，B 因而對 A 提告誹謗，最終法院判決認為 A 無罪』此一狀況，未必代表法院『認證』B 有貪汙，可能僅是認為 A 未具備真實惡意」一事，依照目前的媒體和社會輿論實務看來都有本質上的困難，在許多的情形下都會直接被貼上「法院認證 B 貪汙」之結論。

然而若徹底讓對於事實名譽之保護，徹底與傳統上公眾認知的名譽概念脫鉤，在個案操作上至少可以提供公眾「誣指他人『曾與賤民結婚』，並非妨害名譽，因為我們不應認為『曾與賤民結婚』會貶損名譽」此一簡單、直白的結論。但與此同時，法院還是能夠過傳遞「誣指他人『曾與賤民結婚』可能侵害其他法益」之結論，來適當評價、處理此類行為中被害人所受之傷害，以及行為人主觀上之不法性。同樣是「操作形式邏輯」，本文認為此種操作之取徑，應該相較於前述學者所批判的對於規範名譽理論之反對見解，在成功率上更有期待之可能。

第三款 性自主權

第一目 實務見解

於立法歷程中，綜觀所有文件素材，均未發現任何說明 deepfake 條款係為了保障性自主權的依據。唯一可能的取徑在於法務部之立法總說明對於刑法第 319 條之 4 部分提到「增訂本條規定，以保護個人之隱私權及人格權。」此處所說除了隱私權之外的人格權是什麼容有解釋空間，惟在欠缺其他端倪的狀況下，實難得出立法者有意保護性自主權之結論。

第二目 學說見解

a. 積極面向的性自主

論者有認為此次 2023 年之刑法修正，由於不再強調性影像的未公開性或私密性，因此隱私權保護實際上卻已經不再是規範的重點，反而更趨近於所謂的「性的表現自由」。¹⁹⁵

亦有認為規制未經同意散布性私密影像（此處討論脈絡應係涉及真實性影像的復仇式色情議題）得以保障性自主權，其內涵包含使人能夠自由地展現並實踐其個人情慾的積極面向，並且兼有能保有免於受他人干擾和侵犯的消極面向。

所謂的積極面向，直言之，指的是性參與者樂於從事性活動、在性活動中拍攝與紀錄，都是性實踐的一部分。然而，若讓性參與者面臨到所記錄的性活動存在可能被散布之風險，則可能使性參與者無法徹底自在的從事前述的性實踐。進

¹⁹⁵ 古承宗（2023），前揭註 14，頁 18。

而，透過懲罰未經同意散布性私密影像的行為，得以保障此一積極面向的性自主¹⁹⁶。



b. 消極面向的性自主

進一步談到散布性私密影像是否侵害性自主的消極面向，論者有採取肯定見解，如有認為：「當參與者決定提供其性隱私內容供特定人欣賞，則其特定範圍內的散布自應受到保護，出於自願拍攝並傳送給特定人不表示對不特定之任意第三人之散布有默示的同意，法律應當保護性自主免於受他人干擾和侵犯。」（按：此處討論脈絡應係涉及真實性影像的復仇式色情議題）¹⁹⁷；亦有指出：「儘管 deepfake 沒有真正碰觸或揭露被換臉人的身體，但對於被換臉人來說，已侵害其『想不想要被公眾作為性聯想的對象』的性意思意願，當然侵害性自主權。」¹⁹⁸或有認為人應享有「免於成為他人性客體」的消極自由¹⁹⁹。比較法上亦有指出，在英美法系國家中多將此類不實性影像之散布，視為一種性冒犯（sexual offense）來對待²⁰⁰。相反的，有持否定見解者認為，性自主應該著重在人際之間面對面的互動，必須對被害人有所壓抑、妨害或利用既有外在環境，從而達成身體接觸的性交，或者雖未有身體接觸，但至少須有個人現時性互動的猥褻行為，而排除欠缺真實互動的散布手法²⁰¹。對此學者甚至指出，倘若肯定被害人對於嗣後有關性行為的資訊仍有控管的性自主權限，將導致行為人與他者發生性行為後，未經該他者同意即告知第三人該次性行為之細節時，也會屬於侵害性自主的荒謬狀況，如此見解將導致刑法保護法益的基礎原理遭到破壞²⁰²。

第三目 本文見解

a. 積極面向的性自主

對此，本文首先認為此觀點在以 deepfake 製成之不實性影像案件中並無適用餘地，蓋此類案件中被害人本即並無真正的拍攝性影像，其也無法透過不親自從事性活動之實踐來避免自己成為受害者，也因此並不會因為將散布不實性影像入罪化，從而對被害人積極面向的性自主有任何助益。退步言之，本文亦根本的質疑將積極面向的性自主設定為刑法保護法益的適當性與合理性。雖自司法院釋

¹⁹⁶ 蕭郁唐（2016），前揭註 94，頁 45-46。

¹⁹⁷ 蕭郁唐（2016），前揭註 94，頁 45-46。

¹⁹⁸ 報導者（2021/11/1），〈姜冠霖、王碩勛／從陌生人到熟人凌辱，韓國 Deepfake 數位性犯罪修法的啟示〉，<https://www.twreporter.org/a/opinion-deepfake-digital-sex-crime-korean-experience>（最後瀏覽日：07/13/2023）

¹⁹⁹ 吳芳毅（2020），前揭註 22，頁 195。

²⁰⁰ ŠEPEC, LANGO (2020), Supra note 147, 127.

²⁰¹ 許恒達（2022），前揭註 4，頁 17-18。

²⁰² 許恒達（2022），前揭註 4，頁 18 頁註 44 部分。

字第 791 號解釋以降，已經承認積極從事性行為之自由亦為憲法第 22 條所保障之概括基本權之一²⁰³，惟如本文前述，並非所有憲法上保障之權利都適合動用刑法加以保護，仍需於個別法體系中審查之。

對於當前我國妨礙性自主罪章體系架構，學者有明確表示所謂性自主係指性自主保障的是對性行為「是否」、「方式」、「樣態」、「時間」、「地點」與「對象」可以自由決定，禁止任何人壓迫他人上述之決定自由，然而其僅保障性自主的消極面向。其理由在於「因為個人主張性發展權利可能是以犧牲他人法益為代價」，此與刑法的意義與目的相悖。刑法在於保護弱者不應與不會成為社會衝突的被害人，至於性行為的積極發展自由，非刑法保護的範圍。」²⁰⁴亦即，法律選擇保障每個人都能「不從事自己不喜歡的性行為」，而並未保障每個人都能「從事自己喜歡的性行為」。

本文贊同前述學者見解結論，亦認為我國刑法所保護者應僅限於消極的性自主，惟並不認同其論理之過程。蓋其所稱的因為「可能是以犧牲他人法益為代價，悖離刑法意義與目的」，所以即不應受刑法保障的邏輯，反過來說亦可以推導出消極性自主不應被保障的結論。亦即，在一場潛在的性互動中，若一方有意進行，而另一方無意進行時，此時其實是前者的積極性自主與後者的消極性自主有所衝突、處於競合之狀態。此際，若要保障積極性自主發展之權利，則確實會如學者所言，必須付出犧牲他人消極性自主之代價。

然而，這句話反過來說亦然，即在保障消極性自主的過程中，也形同國家在這般衝突中做出了「認為另一方的積極性自主較不值得被保障」的價值判斷。因此，以「保護某一法益時會犧牲他人法益」作為判斷該法益不應受刑法保障之理由，並不妥適。本文認為現行刑法體制下之所以選擇僅保障消極性自主，正當性基礎在於以下兩點：

其一，若欲保障積極面向的性自主，將使所欲保障的法益過度抽象，且範圍過於模糊。例如在現行法下，「被老闆臨時要求加班以至於無法與伴侶從事預計性行為的勞工」、「因丈夫被車禍撞傷失去性能力，以至於無法從事所期待之性行為的妻子」其積極面向上的性自主或可皆接被認為受到妨害，然而當今的刑法並無特別針對此類的性自主妨害設計規範加以保護，此可能是因為「有潛在的想行使積極性自主權的意念」是難以使一般人從客觀面上觀察得知的。

換個角度自義務角度加以觀察，現行法下選擇課以每個行為人「不任意侵擾他人性領域」的消極義務，至少能夠透過確認對方外顯意願等方法，使行為人得

²⁰³ 司法院釋字第 603 號解釋理由書。

²⁰⁴ 黃惠婷（2011），〈從性自主權檢視刑法保護兒童及少年之規定與修法建議〉，《司法新聲》，第 97 期，頁 65-66。

以預見、理解義務之範圍，也不至於過度干預行為人基本權。然而若欲對行為人課以「不得妨害他人積極滿足性實踐」的積極義務時，將使行為人必須去確認、甚至配合與滿足他人「想要從事的性行為樣態」，會對其造成過重之負擔。

此外，在消極面向性自主的保護上，行為人每次行為都是對於性自主侵害的實害犯（不想從事特定性行為的自由被侵擾），觀察起來相對清晰。然而在此類案件若欲使積極面向性自主也成為刑法保護的法益，其本質事實上是在處罰「行為人此次的散布行為，可能導致被害人本人，甚至是潛在想要拍攝性影音的一般人，在將來不敢積極實踐性自主」的此種侵害，其在觀察上相對難確定法益侵害的範圍與界限（有誰不敢拍？恐懼程度大小為何？與被告行為的因果關係為何？），因此以此做為刑法保護的法益並不適宜。

其二，更直白素樸的理由可能是，在權利被侵害後對個人造成的傷害程度，具經驗證據上觀之可能亦有差別。例如美國國家合併症調查報告估計，北美一般人均終生患 PTSD 的患病率為 7.8%，然而若是遭受過性侵犯的女性，其終生患有 PTSD 的患病率就會驟升至 50%。此外，性侵犯是女性 PTSD 最常見的原因，益有研究顯示，在所有遭受性侵犯的女性中，有高達 94% 的人會在受害後的前兩週內出現 PTSD 症狀²⁰⁵。上述狀況皆是在單一事件中，若選擇壓抑一方消極面向性自主後可能導致的實際傷害，與之相對的，若在該單一事件中轉而選擇壓抑積極面向的性自主，則可能相對欠缺相關實證支持會對該當事人造成何等程度的傷害。也因此，在保障積極性自主與消極性自主可能發生衝突與競合時，現行刑法體系選擇僅保障消極性自主，自有其正當性。

b. 消極面向的性自主

相對於積極面向之性自主所關照的是被害人能否積極實現自己所欲完成的性行為，所謂消極面向的性自主則著眼於被害人能否拒絕從事自己所不欲進行的性活動。關於「deepfake 條款是否保障消極面向的性自主？」此一問題上，本文較為支持否定見解。

在採取肯定見解的論點當中，細分下來可以觀察出兩條不同的論證取徑。其一較為著重於「性隱私相關資訊被未經同意之揭露」所造成之侵害，其二較為著重在「將被害人置於性客體地位」所造成之侵害。

對於第一條取徑，學者有主張認為目前我國學者對性自主權之定義多指的是「個人有享有免於成為他人性客體的消極自由，而可依其意願，決定在何時與何人從事何種類與方式的性行為」。若認為性自主保障之範圍觸及到性行為後對於資訊的控管，將導致性自主的概念被無端放大，甚至因而導致出現一些較為荒誕

²⁰⁵ Chivers-Wilson (2006), Sexual assault and posttraumatic stress disorder: A review of the biological, psychological and sociological factors and treatments, *Mcgill J Med.* 2006 Jul; 9(2), 112.

的狀況亦須被刑事管制的結論²⁰⁶。

本文認為前述觀點實屬的論，除了如前述學者所舉之性行為後分享細節會被認為屬於侵害性自主外，甚至可能連民事侵害配偶權訴訟中，嘗試證明配偶有通姦行為，或是主張自己「未」與某人發生性行為，都可以被認定是侵害性自主的行為（因為除了「有」從事性行為外，「未」從事性行為本身，亦應屬於一種性相關資訊），的確如學者所說會破壞設定保護法益的基礎原理。況且退步言之，縱使承認性自主保障之範圍觸及到性行為後對於資訊的控管，於散布不實性影像案件中，該資訊亦根本並非被害人實際之性行為資訊，因此難以透過此取徑認為係對於性自主之侵害。

而對於第二條取徑，本文雖同意人應享有「免於成為他人性客體」的消極自由，然而到底到什麼程度的行為可以被認定係「使人成為他人性客體」，必須被細究。此處論者所說的成為性客體，白話一點來講就是前述另一論者所稱的「成為公眾作為性聯想的對象」，然而，刑法是否有要將保護範圍擴大到此程度不無疑問。倘若單純地將他人當作性聯想之對象，即屬於刑法應該處罰、規制的行為，則不僅是使用 deepfake 製造之不實性影音，單純以文字描述將他人當作主角的同人創作成人文學，甚至是單純意淫他人的內心活動經輾轉被本人意外得知，似乎都會落入刑法要處罰的對象²⁰⁷，此無疑是過度擴大刑法的規範範圍。

學者即有指出，我國釋憲實務上自司法院釋字第 567 號解釋、憲法法庭 111 年憲判字第 2 號判決以降，皆承認思想自由應受我國憲法保障，而思想自由之保障內涵，並非僅侷限於抽象無體、無形無貌的思想之理，而係必須透過對於保障人民實質上做成之行為，藉以才能使思想自由受到充分保障。而思想者藉由外在行為或物品為「思想輔助」(support of thinking/for thought)，以及思想者將其內思想以外在行為或物品投射於外在世界之「思想具現」(embodiment of thought)，此等外在行為或物品，亦應受到思想自由保障，不容許國家任意干預。所謂的「思想輔助」指的是「人民藉由外在的文字、圖片、影像、聲音、氣味、空間、裝置、設備或行為等而輔助思想，學者舉例說明如使用計算機輔助數學計算，亦或是作畫時先以炭筆勾勒草稿皆屬之」。作者進而認為製作出此類之不實性影像，是製造者內心思想展露於外的外在行為或物品，此處之內心思想，主要為性幻想，其無論對象為誰、內容為何、於何時何地發生，以及當中情節又是如何煽情、香豔、暴力、低俗或骯髒，均一概受到憲法思想自由之保障²⁰⁸。學者甚至指出，基於其高度私密性，反而因此應該受到憲法提供更高程度的保障²⁰⁹。

²⁰⁶ 許恒達（2022），前揭註 4，頁 18 頁 註 44 部分。

²⁰⁷ 畢竟如果要認為把他人當成性聯想對象就構成「把人當成性客體」，那麼在內心起心動念的那一刻，就已經成立了，是否表述出來似乎只是能否被證明或探知層次的問題。誠然本文承認，同樣是知道自己被當成性聯想對象，只是抽象的知道對方有這念頭、看到對方實際把念頭化為文字書寫出來、看到對方實際把念頭使用 Deepfake 技術呈現出自己參與其中的虛假性相關影片，三者在當事人來看或許所受到的衝擊有量上的區別。但在性質上，三者依照前述學者見解無非都是「把人當成性客體」，本質並無不同。

²⁰⁸ 許昭元（2023），前揭註 23，頁 456-464。

²⁰⁹ 許昭元（2023），前揭註 23，頁 483。

本文不否認在某些情境下，從被意淫的當事人角度出發來觀察，在得知自己成為特定人性聯想的對象時會感到不快。然而反之，從進行性聯想的當事人角度出發，其作出該等聯想抑是自己攸關性實踐的思想自由，甚至是言論自由的體現。面對此兩端的基本權衝突，立法者應該做成適當的調節，並不能徹底單純的偏重一方，認為任何使一方當事人不快的行為都必須透過刑法來保障與解決。此就好比對於 Elon Musk 而言，倘若知道自己被員工用 deepfake 技術做成自己被非人道奴役的不實影像，亦可能也會感到不適，甚至影像中的內容可能也根本違反人性尊嚴、使人淪為工具，然而我們也不會因此就當然認為員工的行為必須被規制。

在現行法下，性騷擾防治法第 2 條規定：「本法所稱性騷擾，係指性侵害犯罪以外，對他人實施違反其意願而與性或性別有關之行為，且有下列情形之一者：二、以展示或播送文字、圖畫、聲音、影像或其他物品之方式，或以歧視、侮辱之言行，或以他法，而有損害他人人格尊嚴，或造成使人心生畏怖、感受敵意或冒犯之情境，或不當影響其工作、教育、訓練、服務、計畫、活動或正常生活之進行。」，同法第 20 條規定：「對他人為性騷擾者，由直轄市、縣（市）主管機關處新臺幣一萬元以上十萬元以下罰鍰。」，亦即，現行法下其實有相關法規在嘗試針對前述兩種基本權之衝突進行評價與調和，例如倘若行為人公開於職場上對被害人口述表達其意淫之內容，而有損害被害人人格尊嚴，或造成使人心生畏怖、感受敵意或冒犯之情境，或不當影響其工作、教育、訓練、服務、計畫、活動或正常生活之進行之狀況時，將可能被認為構成性騷擾。然而即便是如此，現行法下的價值評價系統亦僅認為只成立行政罰鍰而與刑法無涉。

準此，即便是在與「被性聯想之對象」處於緊密關係場域的當事人、在符合比較嚴格的要件而被認為構成性騷擾後，亦僅需受到行政罰鍰上之不利。反過來說，於散布使用 deepfake 技術之性相關影音案件中，行為人所做的「將他人當作性聯想之對象」的行為，甚至可能跟「被性聯想之對象」所處場域相隔甚遠，未必足以擾動到「被性聯想之對象」的生活運行，此際卻須被以刑事制裁，即顯有輕重失衡的狀況。白話的來說，現行法下在性騷擾防治法的價值評價中，一個行為人如果把同事當作性聯想對象，只要不要涉及到該法第 2 條之要件（例如單純跟非被聯想對象分享其意淫的內容，意外被被聯想對象輾轉得知），可能根本不成立任何刑事或行政罰上責任。即便成立該法第 2 條之要件，如此地把同事當成性聯想對象、可能損害他人人格尊嚴、使人心生畏怖、感受敵意或冒犯，甚至不當影響其工作的行為，亦僅涉及行政上罰鍰問題而已。然而，若依照前述將「把人當成性聯想對象」直接化約成「使人成為他人性客體」的見解，卻會使一個遠離某名人生活圈之人，僅因將該名人當成性聯想對象，因此即成立對性自主之侵害而須受刑事制裁。

本文以為若採此見解將導致現行法下對於法益的價值判斷體系被徹底顛覆，且在同時涉及兩方基本權衝突時選擇徹底偏重一方欠缺正當理由，又會導致內心的性幻想只是意外被當事人得知，居然就可能得受到刑事處罰這種幾乎像是思想犯的荒誕結論，因此並不可採。

綜上所述，關於散布不實性影像是否可能涉及性自主法益之侵害，論者有從性自主的積極面向與消極面向出發採取肯定見解，其中消極面向上亦存在兩種不同的論述取徑。惟本文針對以上各論者之主張皆以提出看法，認為其見解並不可採。準此，本文以為散布不實性影像與性自主法益無涉。

第四款 性相關的冒犯性/仇恨性言論

如前所述，比較法上有將不實性影像之使用視為一種性冒犯 (sexual offense) 來對待²¹⁰，或有論者得以藉此主張，此處所稱之冒犯並非著眼於個人法益上之性自主侵害，而係較著眼於超個人法益思維，將其視之為「對於女性此一群體之不特定的攻擊」。在此思維下，某種程度上係將不實性影像之使用，當作一種類似於冒犯性言論 (offensive speech) 或是仇恨性言論 (hate speech) 作為管制。

然而必須注意的是，即便是仇恨性言論，仍可能對於個體乃至於社會而言有其重要價值，因此仍應受法律保護。如即有學者介紹 Dworkin 之主張指出，「仇恨言論與色情或歧視性言論一樣，都是一個人表達他所肯認的基礎性倫理價值的言論，政府如果以法律禁止仇恨言論，無異於以強制侵犯了他和其他人一樣都應被平等關懷(equal concern)的生命客觀重要性，即剝奪了他正視自我生命如何算是活得好的尊嚴自重原則，並且同時侵犯了他和其他人一樣都應被平等尊重 (equal respect) 的倫理獨立性，即剝奪了他為自己選擇的倫理性價值 負起個人責任的尊嚴本真原則。²¹¹」

同時，如前所述，法益概念被期待於立法論上發揮「可罰性限縮機能」，以控制國家刑罰權的範圍。立法者在將某種行為態樣入罪化時，必須能夠合理的解釋該行為侵害了何等利益，而不能僅依憑該行為「不道德」即加以刑事處罰，否則即屬違憲²¹²。單純的冒犯 (mere offense) 亦不足以作為使特定行為被入罪化的正當化理由²¹³。超個人法益需得還原成個人法益，才能確保刑罰作為維護個人良善生活環境而形塑的規範機制所欲達成的目的，同時避免法益喪失批判體系功能²¹⁴。準此，釐清冒犯性言論或是仇恨性言論，究竟除了使人感到冒犯或不道德外，究竟尚侵害了何等利益，以至於得以合理化其刑事上規制之正當性，首應被探究。

就此，我們應該先從學說上如何定義仇恨性言論出發。有學者於綜合整理支持或反對將仇恨言論入罪化的學術論點後指出，不論是哪方的支持者，都往往將仇恨言論界定為：「針對具有某些『人類特徵』之社會群體，以威脅、辱罵、嘲弄、藐視等方式鼓吹憎惡、貶抑或仇視思想，進而挑起或煽動對該群體的仇恨或暴力行為之言論。」學者並認為此定義下仇恨言論之認定必須同時符合「內容」、「表達方式」與「表達後之可能結果」三個要素，並援引了美國聯邦最高法院之

²¹⁰ ŠEPEC, LANGO (2020), *supra* note 147, 127.

²¹¹ 莊世同 (2022)，〈論仇恨言論的概念——一個法哲學觀點的分析〉，《臺北大學法學論叢》，第一二四期，2022 年 6 月，頁 146。

²¹² 黃宗旻 (2019)，前揭註 121，頁 161。

²¹³ Hornle(2016), *supra* note 124, 303-306.

²¹⁴ 黃國瑞 (2014)，前揭註 120，頁 26。

見解作為論證依據²¹⁵。自此可知，仇恨性言論之核心不應僅是對於特定弱勢群體之貶抑，尚須因為該貶抑使該群體陷入不利處境始足當之。

而所謂的不利處境，根據支持將仇恨性言論以法律管制之學者見解，其擔憂仇恨性言論可能導致，並足以正當化其刑事處罰之後果包含「嚴重貶損被仇視群體的每位成員身為平等之人的尊嚴，而且還在社會上形成對這些弱勢群體成員的歧視或敵意環境，深刻傷害了同樣作為有道德自主性之弱勢者的平等尊嚴，……，民主社會便不可能建立良好的公共論述機制，同時也無法……為法律提供充分的民主正當性。」²¹⁶、「仇恨言論不但是一種『群體誹謗』(group libel)，而且是針對弱勢少數群體(vulnerable minorities)的誹謗言論；這類言論固然詆毀被仇恨之弱勢群體的名譽，然而法律禁止的理由最終還是在對個人(individuals)的保障，因為仇恨言論經常以『事實宣稱的形式』(the form of a factual claim)，將某些事實歸罪(imputation)於某個群體的所有成員，藉此貶抑該群體成員的平等地位(equal standing)，羞辱他們身為平等公民的基本尊嚴」²¹⁷、「對 Waldron 而言，以法律禁止仇恨言論，正是要確保弱勢群體參與民主決定的平等尊嚴(地位)不被藐視拒斥」²¹⁸。

自前述整理可知，縱使對於支持以法律管制仇恨性言論的學者而言，僅是「對弱勢群體表達貶抑」都不足以合理化法律介入之正當性，尚須該言論表達「使社會對弱勢群體產生不當的歸罪」、「導致社會對弱勢群體形成敵意環境」、「妨害弱勢群體參與民主決定之可能」，甚至導致「民主正當性因此受妨害」。反面來說，以法律規制仇恨性言論之目的，即是確保以上的結果不發生。然而本文以為，以刑法處罰仇恨性言論，不僅無法達成以上目的，還會導致額外弊害。同時，即便認為仇恨性言論應該被刑法處罰，於散布不實性影像之案件亦與其有所區別，因此 deepfake 條款不宜將保護法益設定為「對於女性此一群體之不特定的攻擊」。

第一目 仇恨性言論未必使弱勢群體處境更為不利

自前述對於仇恨性言論之定義與規制正當性可以觀察得知，其所關照的不是單一事件中對特定個人的名譽或情感受傷²¹⁹，而係必須將範圍擴及到對於該群體的集體性傷害。其所想像的典型因果歷程應係「因為行為人A對弱勢群體表達了貶抑的發言，進而影響了B、C、D、E等第三人也對該群體產生負面觀點，進而形成所謂之『敵意環境』，甚至因此妨害弱勢群體民主參與之可能。」然而本文認為，前述的說法在「行為人對弱勢群體表達了貶抑的發言」到「形成對弱勢

²¹⁵ 莊世同(2022)，前揭註211，頁172-175。

²¹⁶ 莊世同(2022)，前揭註211，頁122-123。

²¹⁷ 莊世同(2022)，前揭註211，頁131。

²¹⁸ 莊世同(2022)，前揭註211，頁169。

²¹⁹ 甚至典型的仇恨性言論事件中，也未必存在一個明確的特定被害人牽涉其中，例如某政黨參選人受訪聲稱「印地安人天生智力較低」等。而，若僅係為處理明確特定人之傷害，訴諸既存之個人法益保障相關條文，例如妨害名譽、恐嚇等罪名，亦足當之，難以說明專門另外設計仇恨性言論相關犯罪之正當性。

群體之敵意環境」兩個客觀事實間因果關係的想像，有些太過理所當然而欠缺客觀論證，甚至事實上也欠缺論證的可能。

質言之，在聽到表達者對某一群體做出負面發言後，對於閱聽人而言會產生什麼影響，實際上應該存在各種可能性。舉例來說，對於高中學生來說，如果是一個平時與之相處和睦、德高望重的老師，對其傳遞「同性戀是邪惡魔鬼的化身」之訊息，可能較有機會使學生信以為真，因此使其對同性戀群體產生負面印象；然而反之，若是一個平時為老不尊、被學生鄙夷的老師，傳遞相同的資訊，反而可能使學生基於逆反心理、覺得跟這樣的老師有同樣的價值觀是很羞恥的事情，因此反倒警醒自己不能對同性戀群體有負面觀感。甚至在政治場域上亦然，自我經驗觀之，可能即存在統派的政治人物，基於自己社會形象的崩壞，反而使公眾對待與中國統一此一政治主張更敬而遠之，因而被戲稱是「台獨教父」。亦曾存在反對同婚之參選人或牧師因為其對同性戀群體荒誕可笑的發言，反而被支持同婚的陣營作為笑柄加以宣傳等等。

本文前述所舉的這些事例，想主張的並非「仇恨性言論不可能使第三人對弱勢群體形成負面觀感」，但本文想提醒的是，「仇恨性言論不一定會使第三人對弱勢群體形成負面觀感」，甚至可能「阻止、妨礙第三人對弱勢群體形成負面觀感」。這些可能狀況，其實都是言論市場自然的消化、因應對弱勢群體負面言論可能自然發生的諸多的可能性之一。站在這樣的基礎上，本文認為僅憑膝反射式的邏輯推理，就欲從「行為人對弱勢群體表達了貶抑的發言」，推導至「第三人因此被影響而對弱勢群體有負面觀感，甚至進而形成敵意環境」是過度去脈絡化的主張。甚至，倘國家機關僅憑此理由就要對行為人加以刑事規制，會是過度魯莽而顛覆的行為。

事實上，前述現象的根源其實在於，「特定個人對某群體的評價，究竟會如何影響集體社會對於該群體的評價」，甚至是「究竟賦予該群體如何的評價方屬合宜」，本來就會隨著時代、場域、脈絡而有所流動，不可一概而論。在時代方面，同樣是「老師在升學輔導時告知全班『女子無才便是德』因此建議女學生別繼續升學，找個人好人家嫁了比較重要」的情境，在 30 年前的台灣或許比較可能真的影響聽眾的價值觀與選擇。但在如今比起過去，卻可能有更高的機會使該老師必須去面對體制內或體制外的投訴與批判，而無法真的使這樣的觀念受到推廣。在場域方面，於 2018 年同婚公投前夕，相較於整體社會的言論市場，對於反對同婚陣營可能帶有歧視性的觀點較有容忍度與接納空間，就筆者自己的經驗而言，在筆者所待的系所社群中，反而是一些基於個人信仰等理由，仍然傾向不樂見同性戀得以直接適用民法結婚的人們，在發言時必須謹小慎微。在這樣的場域當中，若「行為人對弱勢群體表達了貶抑的發言」，比起「使第三人因此被影響而對弱勢群體有負面觀感」，該行為人或許其實更該擔心自己會因此在第三人們心中留下負面觀感。

申言之，本文認為對於仇恨性言論的入罪立法，就性質上宛若是刑法第 190 條之 1 流放毒物罪一般，單一事件中個人的對弱勢群體的負面言論，就好比單一

事件中向環境投入有毒物質，個別觀之可能很難證明到致生弱勢群體陷入不利處境或致生環境遭破壞的結果，但想要立法規制的人心中根本性的擔心是「長期性、累積性、集合性」等特質，害怕破窗效應使邪惡吸引邪惡，最終雪崩後雖知沒有一片雪花是無辜的，卻也處罰不到任何一片雪花。然而這種僅因對於「累積性」的恐懼和想像，就省略了個案中行為人之行為與結果間之因果關係證明，其立法正當性即屬有所疑問，這也讓當初刑法第 190 條之 1 修法移除「致生公共危險者」此一要件時飽受爭議²²⁰。本文認為比起流放毒物罪，仇恨性言論更難論證個案行為與最終有害結果間之因果關係。這是因為如前所述的，雖然邪惡可能吸引邪惡，形成負面的循環，然而邪惡也可能反而激起正道的光。本文認為綜合考量發言者與受眾的身份、發言的時空背景、語意的脈絡這些可能流動的因素，個案中出現的仇恨性言論，究竟會讓整體社會離最終擔心的「弱勢群體陷入不利處境」此一有害結果距離更加接近、距離不變，甚或是更加遠離，皆是有所可能的。在流放毒物罪中至少想像上比較難以出現「行為人投放污染物後，反而使環境更加優化」的結果，極限就是難以證明與環境惡化之因果關係而已，即已經因為可能的過度刑罰前置化而備受批評了，在仇恨性言論的案例當中，行為與最終擔心發生之結果無疑因果關係上更為遙遠、更難證明。

必須澄清的是，本文並非想要主張「放任仇恨性言論反而是保護弱勢群體的適當手段」、「要以毒攻毒」，甚至本文也能夠直言，許多仇恨性言論都是偏狹的、錯誤的、不道德的、令人不適的。然而正如本文前述，在討論刑法規範時，不能僅以一個行為不道德、令人不適，即認為有刑法上之應罰性。倘若連仇恨性言論的發表比起不發表，必然提高使弱勢群體陷於不利處境的風險都無法證明，則本文即認為其欠缺刑法上規制之正當性。也因此，基於處罰仇恨性言論之正當性基礎在於使「弱勢群體陷入更不利之處境」，然而基於言論市場之自然反應與消化之結果，難以證明個案中表達仇恨性言論必然或高度可能導致前述結果，本文基此認為不宜將「對於特定群體的集體性攻擊」納入刑法規制之範圍。也因此，本文不認為 deepfake 條款應將「對於女性此一群体之不特定的攻擊」設定為其保護法益。

第二目 定義仇恨性言論之標準困難，且可能因此導致弊害

²²⁰ 參見：王正嘉（2019），〈刑法第 190 條之 1 修正之問題與商榷〉，《台灣民主法治的經驗與見證—江義雄教授七秩晉五華誕祝壽論文集》，頁 448-450；謝煜偉（2019），〈論排放毒物污染環境媒介罪：與各環境行政罰法規之關連性〉，《國立臺灣大學法學論叢》，第 48 卷 特刊，頁 1423；李聖傑（2019），〈刑法第 190 條之 1 的釋義思考〉，《月旦法學雜誌》，第 290 期，頁 78-81；李茂生（2019），〈新修刑法第 190-1 條之商榷〉，《台灣法學雜誌》，第 368 期，頁 29-32 頁等。

前述學者意見當中，有抽象地批判該次修法「逾越最後手段性原則」者；有從立法論上建議應「整條予以刪除」者；有從解釋論上試圖詮釋縱使係採取抽象結果犯之設計，亦應避免以行為犯之檢驗模式使行為人一有投放行為即成立犯罪者；亦有主張抽象危險犯之設計僅是推定危險，仍可以於偵審階段反推定者。總體而言，認為該立法模式應該以各種形式被限縮甚至刪除之意見者眾多。

我國學者即有指出，基於言論的意義需要詮釋(interpret; deuten)此一特性，某一言論是否真的會造成仇恨或歧視的結果，或該言論是否真的具有仇恨或歧視的意涵(meaning)，皆是事實認定上之難題²²¹。而比較法上亦有學者指出，當代人不曾認為單純異議是一種極端主義或仇恨，然而仇恨與討厭或異議(dislike or disagreement)或是仇恨與批評(critique)，卻都難以區分。甚至，即便是仇恨，亦並非所有的仇恨都為社會所不許而應以法律禁止之。其提出的具體質疑包含「攻擊伊斯蘭基本教義派的反同性戀(homophobia)及壓制女性的主張，是仇恨言論還是批評？攻擊天主教會的戀童癖教士(pedophiliac priests)或者它的墮胎立場，是仇恨言論或批評？和平主義者主張「士兵是兇手」，是仇恨言論或批評？」等²²²。

學者進一步主張，若要區別「尊重爭議合宜性」(respect the decencies of controversy)的言論，以及「令人憤慨而引致仇恨」(outrageous and therefore hate inducing)的言論，唯有藉由「周遭社會規範」(ambient social norm)才能加以判斷。然而，學者近一步認為前述的規範並非固定、單一、靜態、不變，而會隨時間演變發展(evolve)，從而本質上是有爭議的(intrinsically contestable)。基此，其主張仇恨性言論本質上即是難以界定與區分的²²³。

本文同意前述學者觀點，並進一步認為，有一些發言與表達可能並非典型想像中的仇恨性言論，然而其對於弱勢群體之不利處境之惡化影響程度卻可能強、更深，此際應如何定義、認定哪些言論屬於仇恨性言論，即是大問題。舉例而言，我國學者有認為指稱「美國人要團結起來，不要被黑鬼的強暴、搶劫、槍枝、刀械及大麻混同及恫嚇。」比起指稱「美國人要團結起來、不要被黑鬼的惡習影響。」，前者在認定是否屬於仇恨性言論時更容易成立，這是因為前者將具體事實強加於黑人身上，使人更容易將這些具體情狀與黑人連結在一起，也因此更容易判斷是否可能因此導致歧視、仇恨、或暴力後果²²⁴。

然而，此種「批判的越具體，越容易成立仇恨性言論」的思路，在很多狀況下其實會值得再思量。舉例來說，當我們不說「原住民天生就是比較笨」，改說「根據研究結果顯示，在中學生群體中原住民比起漢人的平均智商低 30%」；當我們不說「女人本來就是理科白痴」，改說「根據統計，台大理學院歷年招生比例女性最多不曾超過 10%」；當我們不說「同性戀都是性病的溫床」，改說「男同志雖僅佔總人口 5%，然而每年確診性病的比例卻是異性戀的 8 倍」。在以上三組負面言論的組合中，後者毋寧都是比起前者更具體的敘事。然而，這些披著「理性、科學、統計數據」外衣的負面言論，雖說通常不是文獻上在舉例仇恨性言論時的典型範例，且其言論往往可能是忽略了背景的結構因素，例如智商測驗設計本身可能對於非主流群體不友善、社會風氣自小不鼓勵女性朝物理科學習、同志群體因為病識感較強有著更高的篩檢意願等等，事實上是種偏見或歧視。然而這

²²¹ 蘇慧婕 (2015)，〈從諷刺言論談言論管制的司法審查難題〉，《台灣法學雜誌》，第 264 期，頁 26-28。

²²² 莊世同 (2022)，前揭註 211，頁 125-127。

²²³ 莊世同 (2022)，前揭註 211，頁 132-133。

²²⁴ 莊世同 (2022)，前揭註 211，頁 135-136。

些偏見或歧視卻因為配備著華麗的包裝，因此反而比起前者只是素樸的表現仇恨，在「製造弱勢群體陷於困境」的層面上，可能有更強的效果。對於聽眾而言，這些被包裝過後的負面言論，可能顯得「更講道理、更有邏輯」。此際，倘若認為前者是屬於仇恨性言論的範疇，則後者是否也應屬之？這實在認定上就形成了很兩難的局面。

若要忠於規範的保護意旨，即「確保弱勢群體社會上尊嚴與地位」而從寬認定，則難免陷入過度壓縮理性公眾討論空間的境地。過度滑坡下，甚至可能會形成禁止對弱勢群體進行一切負面評論的狀態，而徹底封殺了某些公共政策討論的空間。且此亦可能使對弱勢群體本就有敵意的人群形成一種「都已經被證明是科學事實了還不准我說」的憤恨感，結論上對於弱勢群體之社會地位維護是否有利即是問題。若要主張仍然容許對於弱勢群體的「合理的負面評論」，則，在「合理的負面評論」與「包裝得體的仇恨性言論」間究竟該以什麼作為判斷標準與依據，亦欠缺可操作的標準。反之，若限縮對於仇恨性言論之認定，只處罰那些「未經包裝的」、素樸的仇恨言論，則亦會面對另外的問題。一來是若其「製造弱勢群體陷於困境」的潛能明明比「包裝得體的仇恨性言論」更低，後者能夠全身而退之餘而前者卻需被刑事處罰，則有輕重失衡、違反平等原則之嫌。再者，若是國家透過判決案例認定了「包裝得體的仇恨性言論」非屬仇恨性言論後，可能反倒形成一種「國家認可此種對於弱勢群體的負面評論是合理合法的」之認證，進而讓對弱勢群體有敵意者可以更恣意的使用那些「包裝得體的仇恨性言論去攻擊弱勢群體，進而形成更不利之環境。此外，此舉亦反而形同鼓勵、促使對弱勢群體有敵意的人們，在攻擊該群體時要記得替負面言論進行包裝，結論上是鼓勵這群人放棄攻擊性比較弱的言論，改為使用那些效果更強的武器，這樣的世界是否真的對該弱勢群體有利即非無疑²²⁵。

綜上所述，若欲以刑事處罰規制仇恨性言論，即勢必需面對處理定義與劃分範圍的問題，而學說上意見已有指出這樣的劃分有其本質上之困難。且本文進一步指出，這樣劃分上之困難不僅是抽象的違反法律明確性原則的問題，而會在各種劃分過寬或過窄的情況下，分別導致一些惡害，結論上都未必對於所欲保護之弱勢群體有利。基此，本文認為不宜將仇恨性言論納入刑事管制，deepfake 條款自然亦不應以「對於女性此一群體之不特定的攻擊」作為保護法益。

第三目 不實性影像之使用難以被認為是仇恨性言論

本文前述已經敘明了不宜將對於特定群體之仇恨性言論納入刑事管制之理由。退步言之，縱使認為應以刑法管制仇恨性言論，則不實性影像之使用亦難以被認為是種仇恨性言論。

如前所述，仇恨性言論是指「針對具有某些『人類特徵』之社會群體，以威

²²⁵ 誠然，可以透過「以威脅、辱罵、嘲弄、藐視等方式」此一表達形式上的要求，將此類道貌岸然的歧視切分出去，主張此非屬仇恨性言論。然而此亦無法回答，若此類道貌岸然的歧視對弱勢群體惡劣處境的固化效果可能有過之而無不及，何以此類行為反而無需被處罰。

脅、辱罵、嘲弄、藐視等方式鼓吹憎惡、貶抑或仇視思想，進而挑起或煽動對該群體的仇恨或暴力行為之言論。」，涵蓋仇恨言論的「內容」、「表達方式」及「可能結果」這三個要素。以此標準回頭檢驗不實性影像之使用，行為人內容上究竟有無「鼓吹憎惡、貶抑或仇視思想」，或是其方式上有無「以威脅、辱罵、嘲弄、藐視等方式」為之，本文其實認為都值得商榷。蓋在私領域將他者作為性幻想對象，與在公領域上對於該他者予以尊重，本質上並非互斥之概念。又如本文前述，性幻想本就是憲法應保障之思想自由之一部分，且應是特別值得被保障的一種基本權利，倘要直接將「把他者當成性幻想對象」就直接等同於「將人視作性客體」、「對人貶抑使他人喪失人性尊嚴」，未免使國家得以干預人民思想自由過深。

然而此部分如同本文前段所述，何種言論屬於「仇恨性言論」、怎樣的內容可以被稱為「鼓吹憎惡、貶抑或仇視思想」，本就在本質上難以被根本性地確定，因此在這層次上若有論者堅持認為不實性影像之使用，在內容上屬於「仇恨性言論」、在表達形式上屬於「鼓吹憎惡、貶抑或仇視思想」，本文似也無法對此給出足以一錘定音的否定論理。惟本文認為自「可能結果」此一要素上，亦應能否定不實性影像之使用屬於應受刑事法規制之仇恨性言論。蓋如前所述，即便是在支持仇恨性言論應被刑事處罰的學者眼中，正當性基礎都一定程度必須奠定在該弱勢群體因為該言論，而面臨了被貶抑到可能影響其在民主公眾對話場域話語、妨礙其平等參與公共決策之可能，甚至因此到達影響民主正當性的程度。

以這般的標準對不實性影像之使用加以檢驗時，本文認為難以得出「女性此一群體，有因為不實性影像之使用，而受貶抑達到妨礙其平等參與公共決策之可能的程度」之結論。蓋在我國著名事件發生後，整體社會言論市場的自然運作結果，並沒有形成純然的惡意佔據上風，讓公眾都理所當然覺得女性應該被貶抑之結論。例如該名嫌犯原先是知名 Youtuber，如今卻直接在該產業市場消聲匿跡。甚至如前所述，2023 年度的刑法修法關於不實性影像之立法，吸引了我國前四大政黨立委共提出 19 個版本的草案，這些現象都反應了在民意輿論下，即便認為不實性影像之使在內容與方式上足以構成仇恨性言論，在結果上亦難以被認為屬於值得被刑法規制的類型。依照我國經驗，光憑自然的言論市場運作即足以平衡掉其不利之影響，而並未存在某一群體被根本否定、噤聲，則即不致於達到需要特別動用刑法規範保護之程度。

綜上所述，本文認為仇恨性言論不宜以刑法加以規制。縱使認為其適合以刑法加以規制，不實性影像之使用亦非值得以刑法規制之仇恨性言論，因此 deepfake 條款不宜以「對於女性群體的集體攻擊」作為其保護法益。

第五款 形象法益

第一目 實務見解

與前述討論性自主法益的脈絡雷同，於 2023 年刑法修法歷程中，綜觀所有文件素材，均未發現任何說明 deepfake 條款係為了保障形象法益的依據。

唯一可能的取徑亦僅有法務部之立法總說明對於刑法第 319 條之 4 部分提到「增訂本條規定，以保護個人之隱私權及人格權。」透過此處所指之人格權，將形象法益涵蓋入內。惟在欠缺其他端倪的狀況下，本文亦難得出立法者有意保護形象法益之結論。

第二目 學說見解

對此，學者有主張此涉及「一般人格權衍生的形象法益」，究其具體論述指出，散布不實性影像，可能造成個人社會形象受損，而個人對自己形象的表達方式與控管，是一般人格權的延伸，並且是實現個人自我認知、自我認同的基本條件，亦是使個人自由與人格有效開展的必要條件。並且學者亦指出，此立法方向與德國刑法第 201a 條第 2 項禁止貶損個人形象之照片或影像相關規定近似。學者並亦特別強調，形象法益應係獨立於名譽權的獨立概念，其中最大區別在於名譽權所涉及者係「公領域生活歷程」，而形象法益所側重者則聚焦在以「私領域性活動」為基礎，構築出形象的公開層面利益²²⁶。

然而在前述見解被提出後，亦有見解對此明確提出質疑。如有論者指出決定屬於名聲或形象的個人公領域與私領域生活歷程之區分標準十分模糊，以形象作為保護法益將使未來實務在操作上面臨極大困難。也因此，其指出德國刑法第 201a 條第 2 項規定在適用的過程中，實務法院因應前述所說之困難，對於何謂「形象」的定義與標準，最終仍傾向向公然侮辱罪的解釋靠攏，以特定個人在社會中應受到的社會評價來決定，因此其認為散布不實性影像相關行為應以名譽權相關規範處理已足，無須另立新法作為規範²²⁷。無獨有偶地，亦有指出個人形象的侵害危險欠缺客觀經驗上可操作的判斷標準，且德國刑法第 201a 條第 2 項規範邏輯本身即已然向公然侮辱靠攏。從而其認為誹謗與德國刑法第 201a 條第 2 項之區別在於，前者係以不實之影像為基礎，反之後者則是透過真實之影像為之。其亦指出欲以刑法來保護形象法益實際上難以發揮成效，這是因為「個人形象的侵害危險欠缺客觀經驗上可操作的判斷標準，例如被害人對拍攝結果感到受辱，有可能只是因為被害人自己過於敏感，但也不排除是客觀第三人無法認同被害人的身著或肢體表現，例如髮型、衣服品味、可笑的姿勢等」進而，其既認為形象法益無法透過刑法加以保護，且認為散布不實性影像相關行為即應以我國既有之誹謗罪處理之，而無須另立新法因應²²⁸。

第三目 本文見解

從前述反對見解觀之，大抵可以分成「形象法益客觀上難以判斷，因此無法藉由刑法加以保護」以及「形象法益與名譽法益之範圍難以區分，因此僅以保護名譽法益之線刑法加以處理即已足」兩個層次。

²²⁶ 許恒達（2022），前揭註 4，頁 22-25。

²²⁷ 前揭註 14，頁 30-32。

²²⁸ 古承宗（2023），前揭註 14，頁 16。

關於第一個層次，本文認為其批評固然有其一定道理，對於個人形象侵害的感受著實因人而異，然而是否得以僅執此理由，即足以否定該法益有被刑法保護之必要性或可能性，本文並不認同。蓋依其論述結構，其實也完全可以適用在名譽權、隱私權等現行法體系中明文保障的法益中。直言之，若依照其使用之句式照樣造句，亦能得出被害人感到名譽被侵害、隱私被侵害，都只是個人過於敏感之結論。然而，顯然現行法體系下並不因為「客觀的判斷標準難以區分」即直接放棄區分，因而認為該法益無須、不能以刑法加以保障。反之，現行的司法體系會嘗試更細緻的透過類型化，去找出至少普遍比較能獲得共識之部分加以保障。例如刑法第 315 條之 1 關於無故竊視竊錄罪中，即是透過法條明文規範「非公開」、「隱私部位」等要件，並輔以實務發展出「合理隱私期待」之判斷基準，在難以區分的標準中盡可能找出其中有共識的公約數加以保護。同理，在形象法益部分，若要根本性的將所有使他人形象被扭曲之行為皆加以刑事處罰，可能有其判斷上為難之處。然而若經立法者先選定「因人而異的感受中，較有共識的公約數」類型的形象侵害加以立法，例如關於性的形象扭曲，則即如前述妨害隱私的形式規範一般，應仍屬於刑事規範中可以容許之範圍。

再者，綜觀前述肯否見解之爭辯，會發現爭議之根源很大程度來自於對於「名譽」與「形象」法益能否區分？應如何區分？有著根本上的認知歧異。其中肯定見解認為兩者能夠以「源自於公領域或私領域」作為標準，明確區分出差別；然而否定見解中有認為上述公私領域之標準難以區分，因此主張散布不實性影像之行為應一概以名譽權被侵害之相關規定處理之。亦有認為兩者之區別僅是「影像係真實與否」，而散布不實性影像之行為應以現行法之誹謗罪處理之。

本文認為前述學者間產生分歧的癥結點，在於其皆將「形象」與「名譽」視為兩個互斥的概念，而因為互斥則必須找出兩者間區別的界線，才因而產生了「有人認為得以明確區別，而有人卻認為區別模糊」的爭端。然而本文認為「形象」與「名譽」本質上並非互斥，而係包含的關係。直言之，所謂形象法益的侵害，本質上是一種「使當事人被以其所不樂見的錯誤面貌呈現於公眾」的狀態，例如當一個黑道老大被不實「指控」其私下其實樂善好施、扶危濟困，其可能因為感到這有損其在道上之凶狠形象，因而認為「自己被以不樂見的錯誤面貌呈現於公眾」。而名譽法益其實也帶有「呈現錯誤面貌於公眾」的特質，然而除了前述特質之外，尚包含了另一要素即「該面貌在社會上有著普遍的負面評價」。例如在一個政府官員被不實指控私下收受賄賂此一典型的誹謗案例中，其實該官員也被以「貪官的錯誤面貌呈現於公眾」，同時這樣的面貌在社會上有著普遍的負面評價。換言之，雖然在討論名譽法益時，學說上總會著眼在被害人「社會評價受影響」此一最終狀態，然而其實無論是誹謗亦或是公然侮辱，要使被害人「社會評價受影響」的過程中，其實都會經歷過「使被害人以錯誤面向被呈現於公眾」此一中間狀態，係透過此途徑作為方法，才最終成功使他人的「社會評價受影響」。

學者有整理實務判決對於名譽法益的理解，其使用的修辭包含「個人在社會

上所保持之人格或地位之評價」、「社群裡之人際關係及聲譽評價」等語²²⁹，其中所提及的「社會地位」、「人際關係」之用字遣詞，事實上都隱約顯示了妨害名譽的本質中必然涉及到了個人於社會與人互動中，呈現面貌的自我控制與決定。依照本文前述的理解，則可以解釋成並非一切的「使當事人被以其所不樂見的錯誤面貌呈現於公眾」的行為都值得被刑法處罰，例如前述的使黑道老大被誤會成樂善好施大善人的案例，即無動用刑法介入之必要。在通常狀況下，形象的扭曲尚需伴隨著「該面貌在社會上有著普遍的負面評價」此一要件始值得刑法介入，如此一來則構成我國刑法體系本有的妨害名譽罪章的相關犯罪。然而，縱使該錯誤面貌在社會上欠缺負面評價，在一些「特別嚴重」的案件類型中，可能仍有加以刑事規範之必要，而「與性相關的形象扭曲」則可能即屬於「特別嚴重」的案件類型。

本文採取此見解之核心目的與底層邏輯在於，此取徑得以一方面不去認定「從事性行為且拍攝影像之行為，在社會上應受負面評價」，從而避免本文前述所擔憂之追認、鞏固標籤化與汙名化之問題，與此同時又能確保對於此類案件之被害人提供充足的法益保護。參照本文前述所介紹諸多此類犯罪下受害人的受訪結果，其幾乎皆提及自己很大的恐懼感來源係擔心被粉絲、家人，乃至於一般社會大眾誤認自己真的曾從事影像中的行為。自其自述觀之，散布不實性影像所帶來的傷害，其實非常符合學者所說的，「個人對自己形象的表達方式與控管」。從而本文以為以「一般人格權衍生的形象法益」，作為立法規制此類散布不實性影像行為的保護法益最為適當。

對此，其實亦有比較法上之討論可相互輝映，如有學者即主張被害人應該享有拒絕被錯誤呈現在公眾下的權利（the victim's right to not be publicized in a false light）²³⁰；又亦有研究者指出，不實性影像會造成一種類似於「扭曲」（distortion）的傷害，操縱（manipulate）他人被公眾認知、評價的方式，並且使他人以不準確的形象暴露（be inaccurately imposed）在公眾前。²³¹如上之學者見解在討論不實性影像可能帶來之傷害時，都是側重在「外顯形象的被扭曲」，而不涉及對於該形象在社會上究竟是褒是貶的價值評價。

此外，法國刑法第 226-8 條規定：「以任何方法，發行對描述他人圖像或言詞進行編輯改造（montage）的出版物，且未經該他人同意者，應處以 1 年監禁和 15,000 歐元的罰款。若該出版物明顯（obvious）係經編輯改造，亦或其係經編輯改造之事實被明確指明者，不在此限。²³²」²³³該條文係規範在法國刑法典第 6 章的「侵害人格罪章」（OFFENCES AGAINST PERSONALITY）其中的第 2

²²⁹ 陳俊偉（2020），〈網路遊戲中侮辱行為的刑法評價難題——以臺灣實務判決的觀點為中心〉，《月旦法學雜誌》，第 305 期，頁 162。

²³⁰ Harris (2019), *supra* note 2, 122.

²³¹ Kugler, Matthew B.; Pace (2021), Carly, *supra* note 46, 611, 657-660.

²³² “A sentence of one year's imprisonment and a fine of €15,000 apply to the publication by any means of any montage made that uses the words or the image of a person without the latter's consent, unless it is obvious that it is a montage, or this fact is expressly indicated.”

²³³ FRENCH PENAL CODE ARTICLE 226-8.

節「侵害個人形象」(OFFENCES AGAINST THE IMAGE OF PERSONS)，此條文之要件上亦不要求該描繪他人之不實圖像必須導致社會對其施以負面評價的結果，而係認為只需通過不實圖像去影響個人對於公眾呈現之面貌，即屬於侵害個人形象此一人格權。

此外，在美國法上隱私權發展歷程中，Prosser 提出的四種分類中包含了「公眾誤導型」(false light in the public eye) 此一類型，其內涵即是因加害人之行為，以致被害人因而受到「公眾之誤解」、「將被害人置於一種公眾誤解狀態下被公開」。其所提出的典型案例，即將劣質的詩作錯誤的以被害人之名公開，事實上也與散布不實性影音的案件得以完全對應。即，「一個人被誤會發表了劣質詩作」的案件結構，完全相當於「一個人被誤會拍攝了性影音」之結構。而學者亦明確提及，本類型之行為與誹謗案件(defamation)有相當之類似性，差異在於誹謗案件中尚要求「使被害人名譽降低此一要件」，但本類案型中只要求「錯誤之訊息」，即便該訊息可能使被害人名譽上升亦可當之，因此此種種類之範圍遠較誹謗案件為廣²³⁴。自前述學者對於「公眾誤導型」隱私權侵害之說明，其主張「公眾之誤解不排除構成誹謗，但也不必然構成誹謗」的兩權利間之互動結構，事實上與本文欲主張的「形象法益與名譽法益係呈現包含關係，而非互斥關係」得以相互輝映，至少說明了本文所主張的兩權利間之互動關係並非純然天馬行空、全然欠缺操作可能性之見解。此外，雖此部分係放在隱私權的民事侵權脈絡下所為之討論，但如本文前述的美國法與大陸法系下對於隱私權的發展脈絡差異，許多在美國法下被納入隱私權範圍內加以保護的權利，在我國事實上都可以被視為人格權而透過憲法第 22 條概括基本權加以承認成為一種獨立的基本權利。並且，學者於討論隱私權背後立論基礎時，亦有指出美國法上之隱私權已隱然形成而具有一般人格權(Right of personhood)之性質，其係以維持個人生活之獨立完整不受侵擾為其核心思考，目的即在於保持個人人格之完整，是以隱私權已形成一類似一般人格權般上位之基本權利。事實上隱私權在美國法上發展出之意義，其實應與大陸法系中一般人格權之概念並無不同²³⁵。且如本文前述的，如此個別處理各基本權的法體系建構，也比較有助於掌握各基本權間之差異與區別，應是可採的取徑。

另外，在民法學說上亦有明確認為形象權屬於獨立之人格權類型者，該學者指出若是一個人格利益無法被其他既存之人格權所涵蓋與概括，並且具有鮮明的特徵和內容，就應當認定其為具體之人格權。其並認為，所謂形象權是指「主體對標表其人格特徵的形象人格利益獨佔享有、使用以及獲取相應利益的具體人格權」，其不應僅被視為財產權之一種，而係兼有維護主體獨立人格完整性與不可侵犯性，並保障主體獲得充分的尊重，同時也為了使自身的價值得到充分的發揮之人格權屬性。其進而認為形象利益具有獨立之屬性與價值、無法被其他人格權所涵蓋，因而應當被視作獨立之人格權類型。其亦明確指出：「形象是民事主體的人格特徵，是固有的人格利益，而不是一般的社會評價，不具有客觀的色彩。」

²³⁴ 林建中(1997)，前揭註 149，頁 58。

²³⁵ 林建中(1997)，前揭註 149，頁 66。

236。誠然，民事上構成侵權未必代表足以作為刑法上所保護之法益，然而自前述觀點至少可見，形象權與涉及社會評價之名譽法益有本質上之差別，且具備成為獨立人格權之可能性。

至此，本文已說明了將形象法益與名譽法益係視為包含關係，而非互斥關係，本文前述透過援引比較法上之立法與學說見解，主張此種區分應該有其正當性與可操作性。而此般的解釋途徑好處有二：

其一是有助於解決前述學者間對於兩者之間得否區分與應如何區分的爭端。簡言之，本文給出的答案是：形象法益保護的是個人不被以錯誤的面貌呈現於公眾，而名譽法益則是在侵害形象法益的基礎上，該錯誤的面貌尚導致被害人因此遭受負面評價。換言之，名譽法益是一種更為嚴重的形象法益侵害，兩者之間以該錯誤的面貌是否有社會普遍性的負面評價作為區別依據；

其二是在於保護了被害人之餘，由於並不認定該面貌係屬負面評價，因此不至於陷入本文前述的將其視為名譽法益可能導致的，存在鞏固汙名化與刻板印象之虞。

至此，本文尚未回答的最後一個問題應係：既然認為通常而言，不導致負面評價的錯誤面貌呈現，則無須動用刑法規制，則何以與性相關的錯誤面貌呈現，縱使不導致負面評價，本文卻仍認為屬於「特別嚴重」的形象扭曲，而應該動用刑法規制？一個人被誤以為有拍攝性影像，與一個黑道被誤以為樂善好施，有何本質上的區別？

面對此問題，本文給出的答案主要有三：

其一是聚焦在此類行為在經驗證據的觀察下，受害者如本文前述所介紹的，有著明顯的傷害反應。其中受害者多自陳自己因為擔心公眾真的誤以為自己是那樣形象的人，因此感到重度焦慮、對人際互動失去信任感，進而影響日常生活運作等狀況。若其他案例類型亦能觀察出類似的現象時，本文立場其實不排斥對於該等形象扭曲亦考慮以刑法加以規制；

其二，也是性相關議題最重要的關鍵，在於澄清的難度與其他類型的形象扭曲有著相當的落差。具體來說，使本文形成此一心證的最關鍵之處係在於，筆者於撰寫本文間分別看了大量 deepfake 不實影像使用之個案，部分與性相關，部分則無關。於此過程中，得以觀察到兩者之間有一個現象上的差異，亦即，在與性相關的案件使用中，有很多受害者都會強調自己會「不敢或難以澄清」。如前述曾提及的華盛頓郵報印度女記者 Rana Ayyub 即提及自己在事發後不敢面對家人、關閉社群軟體等²³⁷，具體理由可能係因擔心澄清反而使該影像散布範圍擴大、被公眾指責受傷是自找的等。反之，在不涉及性相關的案件中，無論是前述烏克蘭 2022 年發生的，駭客利用 deepfake 技術，偽造烏克蘭總統 Vladimir Zelenskyy 出面向烏克蘭士兵喊話，要求他們放下武器停止戰鬥影音的事件²³⁸，亦或是在比利

²³⁶ 楊立新（2009），〈楊立新民法講義（貳）—人格權法〉，自刊，頁 214-218。

²³⁷ Ayyub, *supra* note 135（最後瀏覽日：07/13/2023）

²³⁸ 雷喻翔（2022），前揭註 7，頁 29。

時發生的，有政黨利用 deepfake 技術，偽造時任美國總統 Donald Trump 發言的影像試圖影響政局的事件²³⁹，受影響之人（無論是不是影音中被描繪之人，抑或是間接被影響之人）大都可以沒有顧忌的大力闢謠，而無須擔心可能承受後續的二次傷害等。

此歸根結柢當然與「性」之特殊文化意義，乃至於社會對於女性尤其的結構壓迫有關，茲不贅述。本文想說明的是，正由於本文認為值得被保護的法益是「個人對自己形象的表達方式與控管」，當形象被扭曲時尚有能力自行對公眾澄清者，可能還暫時無刑法介入保護之必要。然而對於形象被扭曲時，試圖澄清卻仍得畏首畏尾、投鼠忌器之被害人來說，透過刑法給與特別之保護應有其正當性與必要性。

其三是如前述介紹名譽法益時本文最後所強調的，雖然「有事實名譽減損，然而卻因考量規範名譽概念後而可能造成汙名化」之情況，不宜被認為係屬名譽法益之侵害。然而此類的「名譽法益未受侵害」，與傳統典型的「因為客觀名譽自始未有減損」類型之「名譽法益未受侵害」，本質上仍有差別。直言之，前者相較於後者，行為人有更強的主觀不法，同時被害人亦受到更強的傷害，且此類行為若因為不能被認為是名譽侵害而自始不受刑事法規範，終局上對於所欲保障之群體也未必有利。本文在前一章指出，應該找出「名譽法益以外的其他法益」作為處理、消化此類「僅事實名譽受減損，而因規範名譽概念不宜被認為侵害名譽」之案件類型的避風港，而本文認為形象法益即可能是個適當的選項。即把「事實名譽受減損」，看作輔助證明「形象扭曲嚴重到具備足以被刑事規制」的依據之一。準此，則若在實然層次上，得以有證據資料證明「被誣指拍攝性影像」真的會傷害到被害人之事實名譽，則可以以此作為依據，認為此屬於「帶有事實名譽受損性質之形象扭曲」，而與前述所舉的「黑道老大被誤認樂善好施」有別，因此屬於值得被刑事保護的形象扭曲。同理，前文所舉的「在種姓社會中被誣指曾與賤民結婚」此類的形象扭曲，亦可能循此途徑改由形象法益來接手、消化，作為校正法秩序之規定。

綜上所述，基於其案件類型所帶來的傷害程度，以及澄清錯誤形象的難度，本文認為散布性影像此類的妨害形象之行為，縱使未如誹謗罪這般會帶來評價上的減損，仍然值得以刑法加以規範。換句話說，既然誹謗罪的結構係「形象被扭曲」+「因此得到負面評價」因而可以被認為相較於普通的形象扭曲（例如前述的，「誹謗」黑道大哥樂善好施）更為嚴重，以至於值得以刑法管制的狀態。則，與性相關的形象扭曲，其結構係「形象被扭曲」+「因此被害人在經驗證據上顯示遭逢了重大傷害」+「被害人難以澄清」+「（若經驗證據得以支持）會使人事實名譽有所減損」等特質，應亦可被認為得以與普通的形象扭曲相較，是更嚴重的類型，而屬於本文前述所說的「因人而異感受中，較有共識的公約數」，進而因受刑事法規制。綜上所述，本文認為，deepfake 條款以形象法益作為其保護法益最為合宜。

²³⁹ Galston, *supra* note 47. (最後瀏覽日：07/13/2023)



第三節 無混淆之虞的不實性影像使用

第一項 實務見解

根據 2023 年度修法增訂刑法第 319 條之 4 deepfake 條款所使用之文字，其構成要件中規範之客體係「以電腦合成或其他科技方法製作關於他人不實之性影像」，於此範圍內並無法得出立法者有意要求該不實性影像需達「致人混淆之虞」之程度始足當之的意圖。

雖立法所使用之文字尚包含「足以生損害於他人者」此一適性犯要件，在論理上似乎得以作為排除「無致人混淆之虞」之不實性影像適用本條款之依據，然而自立法理由以及立法過程中之相關文件觀之，似難以得出立法者有意透過適性犯之規定篩選掉無誤導公眾潛能之不實性影音。如本條規定之立法理由第二項提及：「以電腦合成或其他科技方法而製作關於他人不實之性影像，可能真假難辨，易於流傳...」，第三條提及：「作他人不實性影像之行為手段包括以電腦合成、加工、編輯或其他科技方法，例如以深度偽造技術，將被害人之臉部移接於他人之性影像即屬之。」，依其文義似認為凡是「將被害人之臉部移接於他人之性影像」即當然帶有「『可能』真假難辨」之特質，而光是這個「可能性」，立法者即認為足以該當本條文之適用。

此外，在立法歷程中，有兩處皆討論到「不實性影像」之定義與範圍：

其一是司法院加註意見²⁴⁰提及：「本院認行政院來函會銜刑法第三百十九條之四草案，容有下列疑義，茲說明如下：一、『不實』性影像概念不明，存有違法律明確性原則之疑慮：何謂『不實之性影像』？其概念可能有多種情形，包括(一)將被害人之臉部移接於他人之性影像。(即草案立法說明三所舉例子)、(二)將被害人真正之性影像，透過軟體製作出原影音所無之動作(即臉部、身體係被害人本人，但製作不實之動作內容)或變更被害人之身體(如將胸部變大或變小等)、(三)將被害人真正之性影像，換臉為他人。(如身體、動作係被害人本人之性隱私，但臉部置換為他人)等。則草案所稱『不實性影像』，係規範何種情狀？未予定義、說明，存有違反法律明確性原則之疑慮。此涉法律之解釋與適用，建請於草案或立法理由中予以定義、說明，以資明確。」其中司法院所提之後兩種狀況，應皆無「使公眾誤會、混淆特定人從事性行為」之虞，若法務部對此提出回應則可探求其對於此一要件之態度。惟直到本法公布施行後，法務部皆全然未曾對此明確表態，亦未如司法院建議虞立法理由予以定義、說明；

其二是於黨團協商過程中，鄭運鵬委員與江永昌委員皆有對法務部就特定案件類型提問²⁴¹，該案型簡言之係「以 A 為主角之性影像，B 全然未參與其中，

²⁴⁰ 立法院公報第 112 卷第 14 期委員會紀錄頁 165-166。

²⁴¹ 立法院公報第 111 卷第 84 期委員會紀錄頁 16-19。

惟該影片標題宣稱其是 B 之性影像」或「以 C 之臉，嫁接至 D 身體之不實性影像，惟該影片標題宣稱其是 E 之性影像」，於此類案件中究竟 B 與 E 兩人是否為 deepfake 條款下之被害人？對此，列席之法務部長蔡清祥回應表示：「雖然影像是真的，但是用了不實的文字加註在上面，變成同樣的畫面，那還是不實啊！我認為還是不實，當然個案情節部分還是要由承辦的檢察官來認定。」就此可以觀察出，至少對於提出草案的法務部而言，「不實性影像」之認定範圍極其寬廣，顯然無意排除「無致人混淆之虞」之不實性影像適用本條款。而在此之後，無論是提案的行政機關，亦或是審查之立法部門，亦都未繼續對於「不實性影像」之認定範圍有所表示。

綜上，無論是法條文字本身、立法理由說明，亦或是立法材料中探求實務部門之意旨，應可認為其整體態度係認為，縱使是涉及「無致人混淆之虞」之不實性影像亦有刑事上之可罰性。

茲有附言者，2023 年度刑法修正審議過程中，時代力量黨團有提出自己的修法草案。在其草案版本中，有欲處罰與性無關的普通深偽罪，於第 313 條之 1 規定：「利用工具或設備偽造他人之影像、聲音，且無明確標示或未以他法使閱聽者得以辨識該影像、聲音係屬偽造，足以生損害於公眾或他人者」。而在其欲處罰不實性影像之第 313 條之 2 則規定：「未得同意，利用工具或設備偽造他人從事性行為或裸露性器、肛門或胸部之影像或聲音者」、「未得同意，公開、散布、傳送或以他法使第三人得以視聽或保存前項偽造物者」均予處罰。

兩相對照可知，其在處罰不實性影像之部分，係有意省略「且無明確標示或未以他法使閱聽者得以辨識該影像、聲音係屬偽造」此一明文化之阻卻構成要件要素，亦即，其態度明確表明於不實性影像相關案件中，就是已經使閱聽者得以辨識該影像、聲音係屬偽造，仍屬於應受刑事處罰之範疇。自此亦可端見部分立法者面對「是否要求有混淆公眾之虞」此一要件之態度。

第二項 學說見解

在學說見解方面，應可從各學者認為此類案件涉及之法益類型來分類，以此探求其整體之態度。

有認為「得」以「形象」作為保護法益，而替此行為單獨立法之學者，其表示：「構成要件中必須獨立設計有關深偽色情影音內容的文字，在此深偽影音不宜採用『不實內容』來描述，因為深偽技術的重點在於把兩個無關的真正影像合成，亦即被用以合成的兩部分圖像，均可能為真實，就此以觀，真正重點應是行為人採用深偽技術而『模擬』或『仿真』特定被害人帶有色情意味圖片或性內容之影音。」²⁴²其強調「模擬」或「仿真」，或可藉此解釋成對於該影像對公眾混淆之潛能有一定程度之要求。惟其之所以排斥使用「不實內容」作為構成要件描述之文字之理由，似僅係因為擔心此類案件以「真實素材」加上「真實素材」的

²⁴² 許恒達（2022），前揭註 4，頁 24-25。

合成技術，可能不符合「不實」此一要件。惟其於該文論理過程中，又曾提及：「重要的政治人物頭像，被貼附到人獸性交的影片上，該影音除了滿足特定購買者的性想像之外，沒有任何人會相信該政治人物會作出類似行為，還被拍下影片，要說妨害了被害人的客觀社會名望，似有論理扞格。」²⁴³如此一來，則其是否有意藉由「模擬」或「仿真」之要件，而排除製作、散布「無混淆公眾之虞」之不實性影像之行為適用刑事法規之可能性，真意仍有待研求。

亦有認為不實性影像相關案件係涉及名譽權，而認為無需獨立修法僅以既存妨害名譽罪章處理已足者²⁴⁴。其既然認為應以現行法下之妨害名譽罪章處理，則應即無意排除製作、散布「無混淆公眾之虞」之不實性影像之行為適用刑事法規之可能性。蓋無論認為此類行為涉及刑法第 309 條公然侮辱罪抑或是刑法第 310 條誹謗罪規範中，皆未特別強調所述之言論或所為之意見表達需達「使一般人足以相信其評論、意見表達」之程度。雖說誹謗罪中有要求指摘或傳述之事須「足以毀損他人名譽」，惟實務操作上對此要件之審查似僅聚焦在指摘或傳述之事「是否具貶低人格之性質」，而非審查「是否使一般人足以相信其評論、意見表達」。以前述學者所舉之例，一般理性之人應不至於因為他人隨口之評論即相信被害人真的曾經從事人獸交之行為。惟在我國實務案例中，無論是被告於網路論壇上指摘被害人：「哈哈 有個破麻被[狗臉小圖]幹過呦……蕭啦啊 被狗幹 人獸交。」曾被法院判決涉犯誹謗罪確定²⁴⁵，或是指摘被害人：「聽說鄭吉洪的老婆都玩人獸交，不知道真的還假的」、「你怎麼不跟我分享你老婆人獸交的經過」等語，亦被臺灣桃園地方法院檢察署檢察官起訴²⁴⁶（惟本案因被害人撤回告訴故法院判決公訴不受理）。事實上，綜觀其他誹謗罪判決確定之案例，亦未曾見法官對於「被告之言論如何具備足以取信於公眾之潛能」做論理或說明，顯見此一要件並非實務上審查本罪時決定成罪與否之關鍵。

綜上，若依照前述學者們觀點，認為不實性影像相關案件僅以既存之妨害名譽罪章處理已足，則實際上操作後的結果即會是即便所製作、散布者係「無混淆公眾之虞」之不實性影像，仍可能構成刑事犯罪。

最後，亦有學者主張基於擔心以妨害名譽罪章加以處理可能加劇社會對性與被害人之汙名，同時擔心可能間接影響對於限制適用妨害名譽罪章之檢討，因此認為此類散布性影像之行為，宜論以單一的妨害「性隱私或性人格」犯罪，而不再援引妨害名譽罪相關規定。²⁴⁷惟依照該學者見解，在不適用現行妨害名譽罪章之餘，為求保護其所稱之性隱私或性人格，是否有必要連「無混淆公眾之虞」之不實性影像亦納入規制範圍內，未見其對此有所表態。

²⁴³ 許恒達（2022），前揭註 4，頁 16。

²⁴⁴ 例如：林琬珊（2023），前揭註 14，頁 51；古承宗（2023），前揭註 14，頁 19；前揭註 14，頁 37。

²⁴⁵ 臺灣基隆地方法院 109 年度易字第 259 號判決。

²⁴⁶ 臺灣桃園地方法院檢察署 106 年度偵字第 5815 號。

²⁴⁷ 廖宜寧（2022），前揭註 19，頁 168-169。

第三項 本文見解

自前述對於實務與學說見解之整理可知，目前社會上對於「無混淆公眾之虞」之不實性影像是否有加以刑事上規制的正當性，較為欠缺問題意識以至於相關討論甚少。惟本文認為本爭點涉及到此種類型之不實性影音究竟如何其影響所欲保護之法益的問題，對於應否出動刑法規範有根本的重要性，因此應屬於重要之爭點。

如本文前述，製作出此類之不實性影像，是製造者內心思想展露於外的外在行為或物品，此處之內心思想，主要為性幻想，其無論對象為誰、內容為何、於何時何地發生，以及當中情節又是如何煽情、香豔、暴力、低俗或骯髒，均一概受到憲法思想自由之保障²⁴⁸。甚至，利用他人所製作之不實性影像藉以完成自己的性幻想，亦符合學說上所稱之思想輔助，而亦屬於思想自由之一環。也因此，立法者禁止製作或是散布此類不實性影音之行為即屬於對於人民思想自由之干預，必須要有堅實的理由作為支撐，亦即確實的說明透過刑法的規制確實能保障更為重要的法益。

本文認為 deepfake 條款所欲保護之法益係「形象法益」已如前述，因此，本文勢必得回應與處理的問題即是：製作或散布「無混淆公眾之虞」之不實性影像是否會對影像中被描繪之人的形象法益有所傷害？其傷害之程度是否必須以刑法加以保護？對此，本文回頭檢驗自己認定類型一係屬侵害形象法益，因而得以以刑法規範之論理基礎，結論最終採取否定見解，以下從幾個面向分述之：

第一款 學說基礎

本文前述論證形象法益可以做為一種個別人格權，因此受刑法所保護之論理過程，很大一部分之基礎係以相關行為屬於美國法下關於「公眾誤導型」(false light in the public eye)之隱私權侵害，並說明美國法下之隱私權保護內容於我國法下得以作為個別人格權受刑法保護，作為立論基礎。而就學者介紹，此類侵害類型「著重之部分在於因加害人之行為，以致被害人因而受到『公眾之誤解』亦即加害人之行為必須係『不實的』，方可成立本性質之隱私權侵害。」²⁴⁹。自此可知，不僅需要有加害人「不實之行為」，尚需有「公眾之誤解」的結果，始能成立此類的人格權侵害。

此外，前文亦曾介紹學者有認為此類案件中之被害人被侵害之基本權應係應「拒絕被錯誤呈現在公眾下的權利」(the victim's right to not be publicized in a false light)²⁵⁰，學者於下文即接著強調「被害人可以爭執 deepfake 影音過度逼真 (life-like)，導致彷彿宛若是一起真實發生之事件」²⁵¹可見其本旨上亦高度強調「誤以為真」的風險。又亦有研究者指出，不實性影像會導致的影響是一種類似於「扭

²⁴⁸ 許昭元 (2023)，前揭註 23，頁 456-464。

²⁴⁹ 林建中 (1997)，前揭註 149，頁 56-59。

²⁵⁰ Harris (2019), *supra* note 2, 122.

²⁵¹ Harris (2019), *supra* note 2, 122.

曲」(distortion)的傷害，操縱(manipulate)他人被公眾認知、評價的方式，並且使他人以不準確的形象暴露(be inaccurately imposed)在公眾前。²⁵²自其敘述之形式亦可知，其關心的傷害係以「公眾對被害人之認知與評價，發生被扭曲或被操縱」之結果為前提。此外，亦有學者指出，其清楚知悉學說上有反對見解認為於不實性影像案件中，身體隱私部位並非被害人真正的隱私資訊，因此不屬於隱私權之侵害。然而其仍認為此類案件構成對人格權之侵害，關鍵理由即是因為該性影像的閱聽人可能難以理解該影像是虛假的，進而會相信影像中所描繪之人是真實從事該行為²⁵³。由此可知，可能使公眾誤會影像中被描繪之人確實從事影像中之行為，是此類案件類型的重要可罰性基礎²⁵⁴。

綜上所述，自前述學說基礎可知，形象法益之侵害，尤其特別在此類不實性影像相關案件中，應以公眾確有可能對被害人形成錯誤之印象之風險，作為其正當性基礎，而不能僅以行為人對被害人做出不實之呈現，即認為法益受到侵害。

第二款 比較法對人格法益之立法模式

關於比較法上立法模式部分，本文前述提及了德國法和法國法，作為支持此種保護法益設定之正當性基礎。

在德國法方面，學者介紹：「德國刑法第 201a 條還擴張保護個人對於自我形象的人格權利(第 201a 條第 2 項)，此類保護法益重點在於，不論該個人肖像的取得管得合法或違法，只要使得第三人可以接近或閱覽個人屬最高人格領域的肖像時，即可成立第 2 項罪名，亦即，第 2 項強調個人的自我形象不能因為他人的散布行為，使公眾對個人形象產生誤會或貶低效果。」²⁵⁵自其內容觀之，應可認為其認為「誤會」的結果係重要的，倘若不生誤會之風險則應無刑事處罰之必要。

在法國法方面，前文介紹了法國刑法第 226-8 條規定：「以任何方法，發行對描述他人圖像或言詞進行編輯改造(montage)的出版物，且未經該他人同意者，應處以 1 年監禁和 15,000 歐元的罰款。若該出版物明顯(obvious)係經編輯改造，亦或其係經編輯改造之事實被明確指明者，不在此限。」²⁵⁶²⁵⁷，並且說明該條文係規範在法國刑法典第 6 章的「侵害人格罪章」(OFFENCES AGAINST PERSONALITY)其中的第 2 節「侵害個人形象」(OFFENCES AGAINST THE IMAGE OF PERSONS)，以此說明以「個人形象」作為保護法益應屬可採。自其構成要件可以明顯看出，其有意以但書排除編輯技術不足以使人誤會，或藉由註明經編輯之事實而不誤導人的狀況。準此，則亦應可認為倘若不生誤會之風險，

²⁵² Kugler, Pace (2021), *supra* note 46, 611, 657-660.

²⁵³ ŠEPEC, LANGO (2020), *supra* note 147, 121.

²⁵⁴ 相同見解：廖宜寧(2022)，前揭註 19，頁 168。

²⁵⁵ 許恒達(2022)，前揭註 4，頁 10-11。

²⁵⁶ “A sentence of one year's imprisonment and a fine of €15,000 apply to the publication by any means of any montage made that uses the words or the image of a person without the latter's consent, unless it is obvious that it is a montage, or this fact is expressly indicated.”

²⁵⁷ FRENCH PENAL CODE ARTICLE 226-8.

則應即可認為並無形象被侵害之風險，而無刑事處罰之必要。

第三款 被害人自述之傷害狀況

本文前述尚以援引華盛頓郵報之印度籍記者 Rana Ayyub 之自述，輔以鏡傳媒訪問我國知名案件中之受害人專訪，藉此欲主張散布不實性影像，會對當事人造成很大的侵害，因而證立有刑事規範介入之正當性。然而顯得尷尬的是，在被害人自述或訪談的脈絡中，都並未特別區分討論「有無混淆之虞」對其傷害之程度是否存在影響？若有存在，其程度為何？

然而綜合所有言論觀之，所有受害者都一致提及的點即是「被誤以為真」的**痛苦感**，而此痛苦感也是其導致其不敢面對家人、不敢積極澄清的重要原因，更甚者，在 Rana Ayyub 的例子中群眾在信以為真後對其的後續反應，也是其自述傷害心路歷程的重要構成部分。

回頭看到我國的立法經驗，如同學者所言，應該很難否認我國的立法歷程推動很大程度源自於重大事件的推波助瀾，是一種現象立法²⁵⁸。然而，由此亦可以嘗試觀察究竟立法者們觀察到了怎樣的「現象」，才讓他們起心動念推動如此的立法。在二讀黨團協商中，葉毓蘭委員發言²⁵⁹表示：「被移花接木的被害人們有苦說不出，無法證明影片中的人不是自己，或者應該說，為什麼要這群無辜的受害者去拼命地闢謠，來澄清一件自己從來沒有做過的事？」，而范雲委員則指出²⁶⁰：「最後，deepfake 的被害人也面對同樣的困境，因為現在技術成熟，真偽難以判定，對被害人的傷害，其實很多時候不亞於散布真實性影像的犯罪，目前也無法可罰。」某種程度上亦可以說明，「真假難辨」是目前真正擔心難以處理社會議題的重要核心要素。誠然，前述所有的說明，都不足以反過頭來證偽「可辨真假」的不實性影像對當事人而言就並無傷害。然而，本文知道發動刑事制裁是一件對基本權干預甚深的事，因此必須確認發動後可以防免重大的痛苦再度發生，才有必須動用刑事規範的正當性基礎。也因此，在舉證責任的分配上，作為立法者的國家機關無疑應該擔負起這樣的義務，以更積極的調查或研究去補足資訊上的落差和缺口。

在前文中，筆者確實能從被害人的自述中感受到其擔心被誤會，或根本已經被誤會後，所導致的巨大痛苦，也因此本文認為散布「有混淆之虞」的不實性影像影音，應該值得被刑事懲罰。但就如同前述所說的，基於筆者撰寫本文時所觀察到的資訊細節，實在很難確認對於被害人而言，遇到散布「無混淆之虞」的不實性影像情境時，也會因此感受到巨大痛苦，因而必須動用刑事制裁加以對應。

第四項 結論

²⁵⁸ 古承宗（2023），前揭註 14，頁 6。

²⁵⁹ 立法院公報第 111 卷第 84 期委員會紀錄頁 5。

²⁶⁰ 立法院公報第 111 卷第 84 期委員會紀錄頁 8。

本文一一審視了前文當中，之所以肯認散布「有混淆之虞」的不實性影像時值得被刑事處罰，且保護法益應是「個人形象法益」時，所採取的所有論證途徑與論理基礎。結果發現無論是從學說基礎、比較法上立法模式，或是被害人自述的傷害狀況，都讓本文難以認為散布「無混淆之虞」的不實性影像值得發動刑事懲罰。

事實上，筆者必須承認在撰寫此段落時歷經了相當的內心掙扎，主要是因為第一時間純憑直覺所產生的心證，與最後透過思考、論理過後所得到的結論有著相當的衝突。更甚者，依據筆者自己的生命經驗，也相信本段落所得出之結論會抵觸相當部分民眾的法感情。然而，最終之所以使本文能夠在心證與論理衝突時，堅定的相信論理之結果，主要是基於撰文過程中所閱讀到的一則文獻。

美國有論者透過民調作為手段進行研究，其研究樣本包含 1141 名美國民眾，並且樣本符合美國的年齡、性別、地區、教育水準、種族之分布。其調查得出結論顯示，絕大多數民眾都認為與性相關的 deepfake 使用，無論是否經註明是經編輯的產物，皆是具備很高的傷害性 (harmfulness) 與可責性 (blameworthiness)，並且強烈的認為這類行為應該被刑事規範；而關於與性無關的 deepfake 使用，雖然傷害性與可責性被認為相較與性相關者更低，且經過註明是編輯的產物會再進一步的降低傷害性與可責性，然而，無論在任何狀況下，都仍有較多的人認為這樣的行為仍值得被刑事處罰。研究者甚至發現，就算已經獲得在影片中被描繪者同意，或是其他沒有明顯傷害的 deepfake 使用情境中²⁶¹，受測者仍認為如此行為是可責並且有害的²⁶²。為了解釋此奇怪的現象，研究者援引道德心理學上名曰「道德震撼」(moral dumbfounding) 的理論，其理論內容的核心是指，人類有時會「頑固而困惑的堅持著，欠缺合理論點支撐的判斷」²⁶³。最經典的例子包括面對食人、亂倫、人獸交 (cannibalism, incest, or bestiality) 等議題，人們即便欠缺合理理由，仍會選擇沒來由的對這些行為進行道德上的譴責。研究「道德震撼」的學者指出，人類經常從直覺式的厭惡感，直接躍升到進行道德非難，而於此過程中不曾停下思考關於傷害相關的哲學問題²⁶⁴。學者援引此理論回到 deepfake 立法規制的議題上，認為民眾可能是因為一提到 deepfake 一詞時，直接就優先想到針對名人的非合意不實性影像，也因為受此認知影響，當面對到合意的、與性無關的相關影像這類理應並不製造傷害的影像類型時，仍產生了負面觀感²⁶⁵。同一篇文中學者尚援引了另一起關於「古怪效應」(weirdness effect) 的研究，內容

²⁶¹ 例如製造出一名科學家談起其太太，或是描繪一名政治家唱著荒謬歌曲的 Deepfake 影片。
(such as a deepfake depicting a scientist talking about their life's work or a deepfake depicting a politician singing a silly song.)

²⁶² Kugler, Pace (2021), *supra* note 46, 661.

²⁶³ the stubborn and puzzled maintenance of a judgment without supporting reasons.

²⁶⁴ It shows that people often leap from intuitive feelings of disgust to judgments of moral wrongfulness without stopping to consider coherent philosophical theories of harm.

²⁶⁵ Kugler, Pace (2021), *supra* note 46, 661-663.

發現人們在發覺別人以「古怪但無涉於道德爭議」²⁶⁶的方式 (in weird but morally irrelevant conduct) 從事某項行為時，會傾向認為該人存有道德問題、值得道德非難，並且認為該人是個性較差²⁶⁷的存在²⁶⁸。無論是前述的「道德震撼」亦或是「古怪效應」，都是道德心理學 (moral-psychology) 上，在研究關於人類對於傷害之看法和厭惡感的重要學說之一。並且，在此取徑下學者認為，任何事情只要先使人感到不舒服 (uncomfortable)，就會被視為是錯誤的 (be viewed as wrong)，從而人們會在事後發展出一套關於傷害的理論來正當化他們原始的感受，然而那些關於傷害的理論本身其實都無足輕重，真正的重點還是在於原始的本能反應 (initial gut reaction) ²⁶⁹。

前述的研究結果其實都指向一件事，人類單憑直覺、潛意識的心證所做的道德判斷，很常會是欠缺合理論點支撐的武斷評價。且人類在面對到背離原始生命經驗的、古怪的行為時，會傾向優先認定那些行為應該被道德非難。尤其在面對到 deepfake 如此的新興科技時更是如此，可能正是因為他顛覆了人類社會長時間以來對於「眼見為憑」的認知，讓人類產生一定程度的認知失調，因此很可能對其加諸直覺式的、欠缺理論基礎的道德非難。準此，本文也因此格外謹慎的看待與觀察自己在此議題上的心證擺盪，期盼能確保自己所做的價值判斷能夠有合理的論據進行支撐。

同時，如同本文前述於討論不實性影像之使用與性自主權關係時曾提及的，筆者也開始反思，失去「致人混淆之虞」這項特色的 deepfake 技術，究竟與傳統的各種載體有何區別？有什麼正當性要對這樣的 deepfake 技術差別對待？直言之，本文以為若排除了「致人混淆之虞」的性質，這時的不實性影像使用，本質上跟傳統載體中「散布以某名人作為意淫對象的成人小說」或是「以拙劣的 Photoshop 技術把性感寫真的主角 P 圖成喜歡的藝人」並無二致。或許對於被意淫者來說，知道自己被以上述任何一種狀況對待，都會感到不適。然而就如前文所說的，在被意淫者不想成為幻想中主角的反面，與之存在基本權衝突的也是另一群人的思想自由與言論自由。縱使「想要意淫人的思想自由」聽起來不是特別高尚，然而也並不因此代表其在任何價值衝突的場合都只能無止盡的退讓。甚至有學者根本性的指出，正是因為性幻想的私密性，反而值得更高程度的憲法保障。如本文前面提及的，現行法下如性騷擾防治法就有嘗試劃分、分配兩種腳色之間在權利衝突時所能主張的權利空間，其劃分的結果無論如何都與要把一切的性幻想行為都刑事入罪化有著相當巨大的距離。誠如有論者正確地指出，deepfake 所做成的不實性影像，在屬於一種假資訊之餘，也同時是一種娛樂。一個個體雖然

²⁶⁶ 在本研究中，該古怪行為僅是「把身體塗抹成紅色，並且披上披風」，即容易使受測者認為該人的行為值得被道德非難。

²⁶⁷ more morally wrongful and as indicative of worse character if the person performing them engages in weird but morally irrelevant conduct.

²⁶⁸ Kugler, Pace (2021), *supra* note 46, 663.

²⁶⁹ Kugler, Pace (2021), *supra* note 46, 663.

有權利主張自己不希望被在網路上以錯誤的形象描繪著（not be portrayed online in a false light），然而，即便並不是每一個人都同意自己的照片被以 Photoshop 或其他方式拼貼在他人的身體上，這些行為都仍應該被認為是合法的²⁷⁰。

基於以上種種理由，本文最終仍傾向認為，即便對於此種散布「無混淆之虞」的不實性影項案件中，無論是本文自己，抑或是本文猜測社會輿論之傾向，都會在直覺的心證上傾向覺得此類行為並不可取。但出於認識到人類直覺式的道德判斷在此類議題上容易失之偏狹，為了避免使刑事法成為只為道德恐慌服務的工具，本文仍然忠於前述論理之結果，認為此種散布「無混淆之虞」的不實性影像案件，並無特別侵害到影像中被描繪者任何值得被刑法保護的法益，因此無須特別修法規範。

第四節 有混淆之虞的與性無關不實影像之使用

第一項 實務見解

在 2023 年刑法修正歷程中，連同最終被採納的法務部版本在內，共計有 20 個版本的草案²⁷¹被推出，其中全數都包含了處理不實性影像的條款，而其中有 7 個草案尚且針對非性相關的不實影像做出規範（下稱「普通深偽條款」），且提出版本的立委黨籍分布遍布民進黨、國民黨、時代力量、民眾黨等國會主要政黨，應可認為此部分之立法具跨黨派的共識。

原先法務部提出的版本亦包含普通深偽條款在內，惟據黨團協商時出席之法務部次長表示²⁷²，其係基於「行政院認為當初整個民團及外界覺得有關性影像的部分可能比較急迫，以及犯罪被害人保護法有急需修法的必要。」之理由，才選擇先僅將與性相關的 deepfake 條款送交國會審查。

惟於國會審查之階段，司法院加註意見明確反對普通深偽條款之全數草案，其挑選了委員王定宇等 18 人擬具之「中華民國刑法第三百十五條之一及第三百十五條之二條文修正草案」草案（下稱「王定宇版本」）作為主要評論對象，並於其餘版本之相似條文意見部分，註明其審查意見與對於王定宇版本之意見相同。

王定宇版本之普通深偽條款²⁷³主要內容係於刑法第 315 條之 2 增訂第三款規定：「無故利用工具或設備偽造、變造他人公開或非公開之活動、言論、談話或身體部位之文字、圖片或影像者。」，並於第 315 條之 2 一併納入處罰，同時並加重前述兩個條文之法定刑分別至「處 6 個月以上 3 年以下有期徒刑、拘役或

²⁷⁰ Harris (2019), *supra* note 2, 128.

²⁷¹ 立委提出之部分包含民進黨、國民黨各自有數個由不同黨籍立委提出的版本，而時代力量與民眾黨則是以黨團名義提出版本。

²⁷² 立法院公報第 111 卷第 84 期委員會紀錄頁 117。

²⁷³ 立法院公報第 112 卷第 14 期委員會紀錄頁 119-120。

30 萬元以下罰金」與「1 年以上 5 年以下有期徒刑」。

王定宇版本之普通深偽條款部分，立法理由主張：

「一、深度偽造、變造（deepfake）私密文字、圖片、影像於網路散布外流等案件，是新興的犯罪手法。該類不法行為之究責，因行為人匿名特性，多數被害人無法得知加害者身分資料予以追究，只能經由檢舉請網路平台業者，透過社群規則處理，過程不但難度很高，也曠日廢時，同時因網路散布迅速，不法侵害仍持續擴大，許多受害人在尋求司法救濟同時身心受創甚鉅。

二、以深度偽造影片或圖片等不實訊息在網路流傳之犯行，可能涉及刑法之妨害風化、妨害名譽、恐嚇、詐欺、妨害信用、妨害農工商等罪及違反選舉罷免法、個人資料保護法、著作權法、兒童及少年性剝削防制條例等刑事責任。

三、我國相關法律對於這類 AI 視覺處理技術從事不法行為處罰散見於刑法各章或個別專法、罰則過輕，且無『沒入』規定，導致侵害持續進行擴散，實有必要修正現行條文予以加重罰則並增訂防止侵害擴散機制。」

而司法院加註意見²⁷⁴則指出：

「關於本條之立法目的，依立法說明……可知本條乃以深度偽造技術合成製作他人不實訊息影音之基本條款，惟：

一、草案存有違反明確性原則、比例原則之疑慮：……關於本罪保護之法益究竟為何？規範目的為何？依立法說明，本條行為可能侵害個人隱私、名譽、網路資訊之憑信性、選舉公平、著作權等。惟保護之法益不同，其構成要件、法律效果、是否告訴乃論應有所不同，法官對於法條之解釋與適用，其界線亦有所不同。本罪以概括性之構成要件，規範以一切科技方法偽、變造影片行為，企圖全面保護任何可能受到侵害之法益，並未區別不同保護法益而精準設計構成要件，並異其法定刑，是否符合法律明確性原則、比例原則？建請斟酌。

二、草案存有侵害憲法保障言論自由等基本權之疑慮：網路上以深偽等人工智慧技術製作並散布之不實影片，或基於搞笑、娛樂，或基於時事之諷刺、評論，或基於創作，或基於技術展示，原因不一而足，此涉及憲法保障之言論自由與其他基本權或公益間衝突。草案並未依保護法益、規範目的之不同，訂定對應之主觀或客觀構成要件，反而以概括方式規範，無法特定保護之法益究竟為何，則網路上任何製作、散布關於他人合成影片之行為，均有可能入罪，存有侵害憲法保障言論自由或其他基本權之疑慮，並造成寒蟬效應，是否妥適，建請斟酌。」

綜觀司法院意見可概括整理成三點：1.普通深偽條款保護法益不明，過於紛雜，可能違反明確性原則 2.若認為普通深偽條款係為保護零散於各法規之各式法益，則未異其法定刑可能違反比例原則 3.使用科技技術製作不實之影音可能有其正當目的，普通深偽條款並未區分使用之目的與情形，概括一律禁止將侵害言論自由，可能違反比例原則。

相對於司法院之強烈反對，法務部對於王定宇版本之普通深偽條款僅評論道：

「草案未區分行為客體為性影像與非性影像，刑度均為相同，建請予以區分，以

²⁷⁴ 立法院公報第 112 卷第 14 期委員會紀錄頁 42-43。

符考量罪刑相當原則；又草案均未有『足以生損害於他人者』之要件，對於合法運用 deepfake 技術之創作恐將涉及是否亦屬處罰範圍之疑義。」

觀察法務部之評論，對於有正當理由之應用可能一併被禁止之疑慮與司法院所提之第三點意見相同。此外，法務部尚具焦在其針對性相關與非性相關之不實影像皆課以相同法定刑與以質疑，其並未明言何者應該更重，惟本文推測應該係認為與性相關之不實影像應該比起與性無關之不實影像有更嚴苛之刑度。

特別值得注意的是，雖說司法院對於所有提出普通深偽條款之草案，皆直接認為其評論與對王定宇版本之相同，惟事實上本文認為有其他版本的草案事實上於立法理由部分有著更充分之論述，得以一定程度回應司法院之質疑。

如委員高嘉瑜等 25 人擬具之「中華民國刑法增訂部分條文草案」²⁷⁵（下稱「高嘉瑜」版本）增訂增訂第 15 章之 1「偽造數位圖像、影音、電磁紀錄罪」，同時增訂第 220 條之 1 規定「意圖供行使之用，以電腦技術或其他科技方法，偽造、變造他人之不實照片、影像、語音或電磁紀錄，足以生損害於公眾或他人」及對其行使之行為為處罰。

高嘉瑜版本之立法目的提及：

「互信，是個人自主與相互尊重的表現，是人與人之間可以和平共存，經營社會共同生活，自由發展個人人格的基礎。為維護人際互信，刑法特別禁止施用詐術、傳遞不實訊息，偽造、變造的行為。特別是，為保障公共信用與交易秩序，刑法定有偽造貨幣、有價證券、度量衡以及偽造文書印文罪等罪章。惟上開罪章規範之偽、變造的客體，僅限於文書、準文書、有價證券、可以作為簽帳、提款、轉帳或支付工具之電磁紀錄物、通用貨幣、印章、印文、署押、商標與商號罪以及度量衡，並不包括足以影響公共信用與交易秩序，但不屬於準文書（刑§220）之列的個人圖像、影音或電磁紀錄。然圖象與影音，雖不具有準文書之性質，但於吾人的社會生活，普遍認其具有一定的公信力、對於公共信用或交易秩序具有一定的影響力，為維護社會大眾對於具有社會公信力之數位圖像、影音與其電磁紀錄之信賴，對於偽造、變造或不實製作圖像、影音或其電磁紀錄，足以生損害於公眾或他人者，有特別立法規範之必要。」、「為維護人際互信及網路流通性，行為人意圖供行使之用，以電腦技術或其他科技方法，偽造、變造他人之不實照片、影像、語音或電磁紀錄，足以生損害於公眾或他人之行為，應予處罰，爰增訂本條規定，以維護網路之可信性。」、「有關他人之照片、影像、語音或電磁紀錄，於法律上、政治上或社會、經濟生活上重要事項具有證明之意義，如有不實，除侵害個人法益外，亦影響會大眾使用網路往來之意願，爰規定為本條犯罪行為之客體。」

就高嘉瑜版本增訂章節之位置，與其提出之立法理由觀之，其明確將普通深偽罪定位成侵害社會法益之犯罪，而該法益之具體內涵即是「對於圖像與影音之信賴」，並清楚說明刑法既有條文中對於許多應該被信賴的社會系統，皆設有偽造變造之相關處罰規定，惟對於圖像與影音卻欠缺相關規定。同時說明，在當代

²⁷⁵ 立法院公報第 112 卷第 14 期委員會紀錄頁 86-88。

圖像與影音具公信力、對人際往來之互信有相當影響力，若未被充分保障，將會影響大眾往來之意願等。

依照此部分的論述，應可充分回應司法院之質疑。即，由於保護法益被特定在「對於圖像與影音之信賴」，並不存在保護法益不明的問題，同時亦不生同時保護數法益而無異其刑度之疑慮。同時，其透過適性犯之規定，似亦能確保「正當目的之不實影像使用」不會受刑事處罰。司法院並未對此表示意見甚是可惜。

此外，於時代力量黨團所提出之版本立法理由²⁷⁶亦有提到：「深偽技術升高了辨識資訊真假之成本，又提高了人民形成錯誤認知之可能性。有鑑於此一方面對公共生活之以事實為基礎之理性公開論辯活動、問責制度及共識凝聚機制構成程度不一之威脅，在另一方面，人民為基於可靠資訊規劃社會生活與互動，所仰賴之人際信任關係，亦因深偽之介入而可能遭到濫用和破壞，並對被偽造者經營之名譽和信用影響甚鉅。」「本法之保護法益兼有個人法益與社會法益。」等語，其對於普通深偽條款保護法益之說明，亦殊值傾聽。

綜上所述，目前跨黨派的各黨立委與法務部對於普通深偽條款應該立法之方向有所共識，2023 年度之修法所採之法務部版本僅因擔心立法效率之技術問題而並未一併提出。司法院對於普通深偽條款之立法有提出若干疑慮，惟本文認為某些版本之立法理由已經相當程度地回應了司法院提出之質疑。

第二項 學說見解

在學說見解方面，學者有認為現行法下已經有了充分的規制，例如刑法第 339 條詐欺取財罪、刑法第 313 條詐術妨害信用罪，甚或是詐術妨害考試公正、詐術妨害集會、詐術妨害投票等罪；或以「謠言或不實資訊（或內含此意旨）」作為構成要件之刑法條文，例如刑法第 251 條散布民生必需品不實資訊罪、刑法第 310 條誹謗罪²⁷⁷、選舉罷免法第 104 條意圖使人當選或不當選而散布謠言罪，或陸海空軍刑法第 72 條散布軍事謠言罪等，皆足以對應此種不實影像濫用之社會衝擊，因此並無另行增訂刑法規範的必要²⁷⁸。亦有認為此類與性無關的不實影像對社會帶來之衝擊，可以以有無侵害特定人名譽作為區分，分別以誹謗罪與各種散布不實訊息罪處理之。然而其亦肯認此類不實影像基於其高度逼真與難以辨識的特性，再加上網際網路的傳播作用，會提高了這類行為的法益侵害程度與處罰必要性。

因此，其認為應該在刑法第 310 條第 2 項加重誹謗罪之部分增訂相關規定因應。然而在散布不實訊息相關條文則由於涉及到言論自由干預，因此正當性議題

²⁷⁶ 立法院公報第 112 卷第 14 期委員會紀錄頁 109-111。

²⁷⁷ 技術上來說，本罪之構成要件事實上並未要求所指摘或傳述之資訊必須是謠言或不實資訊，然而由於刑法第 310 條第 3 項規定：「對於所誹謗之事，能證明其為真實者，不罰。但涉於私德而與公共利益無關者，不在此限。」，因此在討論本罪是否成立的層次上，指摘或傳述之資訊是否屬於謠言或不實資訊仍係重要的決定性因素。

²⁷⁸ 許恒達（2022），前揭註 4，頁 15。

更為複雜且極有爭議，因此傾向不特別為了不實影像之社會現象修訂相關法規²⁷⁹。同時，該學者亦質疑「立法者完全沒有交代，為何只有與性有關的不實影像才會損害人格權，其他類型的不實影像卻不會損害人格權或者不具刑法保護的必要性」²⁸⁰，輔以其於該文中認為與性相關之不實影像僅需以現行之妨害名譽罪章處理即可，並不涉及額外之人格權因此無須特別為此立法之態度，其前述對於立法者未對與性相關和與性無關的不實影像「等者等之」之質疑，應可理解成其亦根本性的反對普通深偽罪之訂定。

第三項 本文見解

綜觀前述之整理，能夠發現目前實務界至少就行政端與立法端對於普通深偽罪有著高度的共識，未能趕在 2023 年的刑法修正一併增訂僅是因為擔心立法效率等技術性之問題。然而在學術端討論量能高度集中在與性相關之不實影像的修法議題上，對於普通深偽罪立法正當性之討論則相對付之闕如。前述整理曾發表過見解的學者意見，大抵皆對為此訂定專法處理抱持反對態度，理由則主要集中在認為現行法制即足以因應 deepfake 等新興技術所帶來的社會衝擊。

本文以下將從三個面向來討論對此議題之態度，分別係「deepfake 等新興的技術對社會是否帶來有別於過去之新衝擊？」、「現行法規是否足以充分處理與規制新衝擊？」與「若欲以專法管制衝擊，其得以正當地保護之法益為何？」。

第一款 新興技術確實對社會帶來有別以往之新衝擊

本文認為散布不實影像本身，即可能對於社會的信用機制造成破壞。如學者即有指出，民主對話的根基必須奠基在有「經驗證據」（empirical evidence）做為支撐的，有共識的事實與真理上，才可能讓民主充分的運作²⁸²。而與前述見解相似的，有學者指出，民主制度必須奠基在人民有能力進行「建立在有共同信任的事實基礎之上」（founded on a shared set of truths）的溝通與辯論²⁸³。然而，deepfake 本身即會在很大的層面上侵蝕社會上對於信賴的根基，在選舉機關、法庭、立法者、行政機關等方方面面上都造成影響。舉例而言，在 deepfake 盛行後的世界，新聞記者就會面臨一個兩難（dilemma），也就是當有民眾提供影音資訊且其內容具備高度新聞價值時，那段影音資訊的真實性究竟能否被信任？

雖說這並非嶄新的問題，早在 deepfake 盛行前新聞業者本就需要面對資訊的篩選，然而由於 deepfake 容易製成高擬真影像的特質，且高度挑戰人們傾向認

²⁷⁹ 陳俊偉（2023），前揭註 14，頁 33。

²⁸⁰ 陳俊偉（2023），前揭註 14，頁 37。

²⁸¹ 關於質疑政府「僅針對與性相關之不實影像作刑事規制欠缺正當性」之部分，相似見解：古承宗（2023），前揭註 14，頁 19。

²⁸² Chesney, Citron (2019), *supra* note 5, 1777-1778.

²⁸³ Brown (2020), *supra* note 32, 12.

為有圖有真相（a picture is worth a thousand words）的直覺，因此會使處境更加艱難。學者進一步表示擔憂，在 deepfake 盛行的環境中，可能導致新聞媒體更不願意承擔風險去報導新聞，或至少讓新聞媒體錯過許多報導的即時性。學者即直言，從「fake news」變成「deep-fake news」的世界即將到來²⁸⁴。亦有學者明確提出擔憂，認為當民眾漸漸無法辨別影音是真實或虛偽時，會進一步導致人民對於傳統媒體的信賴程度下降，進而讓真實的新聞報導亦難以發揮影響力²⁸⁵。另有學者指出，據研究顯示如今之美國人已經對於媒體欠缺信任，而 deepfake 的盛行很可能會加劇這種不信任²⁸⁶。

此外，學者亦進一步擔心，deepfake 盛行的環境中將導致「詐欺者紅利」（The Liar's Dividend）的現象，亦即，由於人民對於所見到的影音不再信任，反而開始回頭懷疑起那些並非 deepfake 所製造，貨真價實的影音資訊。這會讓許多有心人士有機可趁，在分明自己做錯事情被影像確實記錄下的場合，更容易透過宣稱該影像是被 deepfake 所偽造，以獲得「詐欺者紅利」。論者鮮明地指出，那樣的社會場景就會像是美國前總統 Donald Trump 面對一切不利於己的醜聞時，會反射式的把那些指控都是「假新聞」當作咒語（mantra）一般的使用著。當 deepfake 散播的更加盛行時，人們將不再能相信自己所見所聞的資訊為真，也因此，deepfake 的廣泛散播將回過頭侵蝕民主政體運作所需的基礎信任。而一旦當不信任持續發生，將是替極權主義的誕生創造極大的空間²⁸⁷。學者即有舉例，在美國過去有許多案例都是藉由影音證據，使政治人物被證實曾有失言或其他不當舉措。好比過去時任總統參選人 Donald Trump 即曾被媒體 Access Hollywood 流出的影像證實，其曾發表對於女性的猥褻言論。學者擔憂在 deepfake 氾濫的時代，Donald Trump 就會利用此優勢主張該影像是經過偽造的，而其支持者也會完全替這般主張買單²⁸⁸。又例如 2012 年時，共和黨籍的總統候選人 Mitt Romney 即被流出影像證實，其曾經在與一群資助者談話時，失言表示有 47% 的美國人都對於僅仰賴政府所提供生活必需品這般的生活感到相當開心；2016 年時，民主黨籍的總統候選人 Hillary Clinton 被爆出批評對手 Donald Trump 的支持者是「一籃子的可悲仔」（a basket of deplorables），這些爆料在當時都是千真萬確、毫無爭議的。然而到了 2020 年，這些在私人場合流出的影像內容真實性都可能被挑戰與質疑。

Google 的首席研究工程師 Nick Dufour 就直言：「deepfake 技術將導致人們在面對地貨真價實影像證據時，仍可以聲稱那是造假的。」來自紐約大學電腦科學與工程系的教授 Nasir Memon 也表示：「deepfake 氾濫的後果是，即便是真相

²⁸⁴ Chesney, Citron (2019), *supra* note 5, 1784-1785.

²⁸⁵ Brown (2020), *supra* note 32, 12.

²⁸⁶ Rini (2020), Deepfakes and the Epistemic Backstop, *PHILOSOPHERS' IMPRINT*, 11.

²⁸⁷ Chesney, Citron (2019), *supra* note 5, 1785-1787; see also: Ray (2021), Disinformation, Deepfakes and Democracies: The Need for Legislative Reform Thematic Issue: Big Technology and the Law, *University of New South Wales Law Journal*, 44, 989.

²⁸⁸ Brown (2020), *supra* note 32, 12.

也不再會被相信。天安門事件中『坦克人』的相片流出，影響了全世界；在水門案（Watergate scandal）中，Richard Nixon 的電話錄音流出導致他失去了總統大位；隨著關於集中營的恐怖影音流出，促進了一些改變發生。當這些『眼見為憑』、『所見即真實』的概念遭到了挑戰，對於人類而言那會是個大麻煩。²⁸⁹也有論者表示擔憂，指出在過去人們就已經會對 911 事件、人類登陸月球的影片產生懷疑的陰謀論了，在 deepfake 盛行的時代這可能只會更嚴重²⁹⁰。前述擔憂也並非孤例，我國亦有學者引用致力於研究資訊生態系統的美國研究者 Aviv Ovadya 之「真相冷漠」（reality apathy）理論，認為當群眾發現自己時常被假資訊給矇騙時，可能會因此傾向直接放棄繼續關注資訊，而這樣的狀況在資訊匱乏的地區會更加顯著。具體的體現是群眾將停止繼續關注新聞，因而導致民主政體運作所需的資訊流通變得極不穩定。而學者明確指出，deepfake 的廣泛散播將導致這樣的資訊冷漠，因此不利社會整體運作²⁹¹。也無怪乎，研究機構 Brookings 做成的報告亦指出，對於影音真實性懸而未決的不確定性，將導致災難性的後果，並且徹底結束人們對於政府的信任²⁹²。

對於前述學者認為 deepfake 技術的發展，將根本性的破壞人們對於經驗證據信賴的擔憂，其實亦有學者提出反對見解。其論理主要是以生物之認知系統運作機制來進行舉例，嘗試說明前述學者之擔憂過於杞人憂天，人民有能力隨著技術發展自然因應去處理新種類的假資訊衝擊。例如其嘗試用以類比的重要生物機制即是：「許多毒蛇會用斑斕的色彩來標誌其有毒進而避免遭獵捕，然而許多其實沒有毒性的蛇類也會為了避免被獵捕而傳遞色彩斑斕的虛假訊號。然而隨著越來越多『假毒蛇』出現，此一『斑斕的色彩即代表有毒』的訊號系統即會失去功用。為了避免訊號系統失效，生物會自然發展採取出幾種因應方式，包含 1.使所有傳遞訊號的個體間有共同利益（common interest）2.使該訊號難以被偽造 3.使偽造該訊號之成本提高 4.懲罰詐欺者。藉由這些機制，生物會很自然而然的避免訊號系統因為詐欺者出現而整體性崩解。」²⁹³針對前述學者見解，本文認為先姑且不論逕自以毒蛇運作生態類比回人類的社會是否合適，究其提出的幾種因應方式來細緻探討之亦會發現，我們現在面臨的 deepfake 不實影像衝擊的社會背景，即是因為欲透過影像來傳遞真實訊息的人與想要傳遞假訊息的人利益本就不一致所致。而目前社會面臨的重大困境亦包含，該虛假之訊號偽造越來越容易、成本越來越低。也因此，是否能以足夠的力量去懲罰詐欺者，似乎就成為維繫人類社會關於「影像是足以眼見為憑的經驗證據」此一訊號系統的最重要途徑。

²⁸⁹ Galston, *supra* note 47.（最後瀏覽日：07/13/2023）；see also: Langa (2021), Deepfakes, *supra* note 34, 773; Brown (2020), *supra* note 32, 11; Thombre (2021), *supra* note 6, 2273;

²⁹⁰ Brown (2020), *supra* note 32, 11.

²⁹¹ 吳芳毅（2020），前揭註 22，頁 189-190。

²⁹² Thombre (2021), *supra* note 6, 2273.

²⁹³ Atencia-Linares, Marc (2022), Deepfakes, shallow epistemic graves On the epistemic robustness of photography and videos in the era of deepfakes, *Synthese* 200 (6):1–22, 518-519.

此外，從目前許多對於 deepfake 實際應用的社會研究觀之，本文亦認為期待人類社會能自然而然僅憑社會自行適應，就得以解決此類新興假訊息的衝擊是癡人說夢。舉例而言，即有學者研究了 2018 至 2021 年於 Reddit 論壇上提及 deepfake 的文章，共計有 6638 篇文章與 86425 則評論，做成觀察分析之研究報告。其研究結果表明，在 2018 年至 2019 年時曾一度興起了討論 deepfake 涉及道德爭議的討論，然而到 2020 年後即銷聲匿跡。學者對此所下的結論係，人們漸漸將其視為常態而忽視了其所帶來的傷害，甚至有更動力創作更多此類作品。且作者發現 deepfake 之適用隨著年代遞嬗，有逐漸朝向營利化的趨勢，涉及的主題範圍亦欲發廣泛，不僅限於與色情相關之內容，皆有消費者願意付費託人製作 deepfake 影像²⁹⁴。此外，亦有人類發展學與經濟學家進行行為決策之研究後發現以下四點結論：1.人類無法可靠的分辨 deepfake 影像 2.無論是提醒人類要提高警覺，亦或是提高其經濟誘因，皆無法提升人類分辨 deepfake 影像的準確度 3.人類傾向誤把 deepfake 製成的不實影像誤以為真（而非將真實影像誤以為其不實）4.人類通常會高估自己的辨別能力。因此，作者指出若人類不再能對影像內容眼見為憑，則會破壞透過影像獲取資訊的能力，而其所作成之研究可以支撐前述假說，並說明人類真的欠缺辨別 deepfake 製成之不實影像的能力，且容易對於自己的辨別能力過分自信²⁹⁵。

本文已經透過前述研究資料表明，「影像是足以眼見為憑的經驗證據」此一認知是當代社會體系運作的重要前提，無論是民主機制的討論，亦或是公民知識的取得，皆必須以此為根基，對於社會有其根本的重要性。然而，deepfake 技術發展後，製成高度擬真的不實影像，將破壞前述認知，而人們無法自然而然透過自力解決這個問題，從過去幾年的社會觀察發現人類不僅傾向將其常態化而忽視風險，且反而會傾向高估自己辨別能力而導致容易上當被騙。若此現象持續的發生，學者擔心將會動搖當代民主社會的根基。前述的擔憂並非危言聳聽，例如位於中非的加彭共和國，在 2019 年時其長年執政的總統 Ali Bongo Ondimba 長期未出席公開活動，因此引發了公眾對於其健康狀況的擔憂與質疑。政府機關為了闢謠，遂釋出了一則其本人錄製的向群眾慶賀新年的祝福影像。然而，該影像卻遭質疑是透過 deepfake 所合成，而專家學者對於該影像之真實性也感到困惑，莫衷一是，難以肯定其是否為真。在這樣使群眾不信任的基礎下，竟釀成了當地的一場軍事政變。學者就直言，這反映了該影片最終就算無論是被證實或證偽，其對社會所造成的信任破壞都已然形成且難以回復²⁹⁶。

又以我國學者於文中所舉之事例為例，其指出若政治人物甲候選人在 A 公園的造勢場合公開發言指責「乙候選人出生自黑金世家」，而行為人透過 deepfake 技術偽造影像，結合甲候選人於 B 舞廳跳舞的畫面，製造出「甲候選人於 B 舞

²⁹⁴ Gamage, Ghasiya, Bonagiri (2022), Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications, Conference: CHI Conference on Human Factors in Computing Systems

²⁹⁵ Köbis, Doležalová, Doležalová (2021), Fooled Twice – People Cannot Detect Deepfakes But Think They Can, iScience, 1, 11.

²⁹⁶ Langa (2021), Deepfakes, *supra* note 34, 767.

廳做出指控乙候選人發言」的虛假影像，於此案件中，學者認為由於其影片中所呈現的雖然並非真實存在的事件，然而「甲候選人確實曾指控乙候選人」一事並非不實，因此並不成立選罷法第 104 條的散布謠言罪²⁹⁷。對於此適用現行法的結論本文敬表贊同，然而值得深思的是，如此事件倘若發生是否真的對社會沒有傷害？倘若前開的事例發生在台灣，可以想見可能發生的後果是，乙候選人陣營的競選團隊或支持者會先持著該虛偽影像指控甲候選人，隨後甲候選人會努力澄清該影像是藉由 deepfake 技術所製的不實影像，甚至出示自己於 B 舞廳跳舞的原始影像自清。然而乙候選人陣營將未必採信甲候選人之抗辯，甚至懷疑其聲稱的原始影像才是被 deepfake 技術所製的偽作。即便甲候選人真的自清成功，乙候選人陣營也將因為曾利用虛偽影像做出錯誤的指控而信譽大減。因此，隨後即便其真的掌握甲候選人有從事不法、違背道德情事的真實事證，在指控時也會被民眾懷疑是否又是 deepfake 技術所製的不實影像，甲候選人也勢必緊咬此事，讓民主政治運作中可以被檢驗的「經驗證據」信用系統遭到巨大破壞。更陰謀論的說，在最先開始甲候選人陣營甚至可以主動釋放出以 deepfake 技術所製，汙蔑甲候選人自己的虛偽影像做為誘餌，吸引乙候選人陣營上鉤執此指控後，再釋出原始影像自清，以癱瘓乙候選人隨後對其的一切攻擊能力。且此影響甚至會溢出單次選舉，而使該候選人陣營，甚至所屬政黨隨後的政治生涯中的信譽都受到干擾。而目前選罷法第 104 條的散布謠言罪僅旨在處理單一事件中的個人名譽或是選舉公正，而非針對散布虛偽影像本身所帶來的社會信任基礎侵蝕，因此對於前開的社會紛擾將無能為力。

綜上所述，本文認為 deepfake 等新興的技術確實會對社會帶來有別於過去之新衝擊，而有積極處理之必要性。

第二款 現行法規不足以充分因應新衝擊

如前所述，目前我國對普通深偽罪發表見解之學者，皆認為現行法即足以因應與處理新技術帶來的衝擊，具體而言指的是透過既有的妨害名譽以及散布謠言等相關刑事規範加以規制。

惟本文認為現行法下規範仍可能存在以下漏洞無法被充分填補：

第一目 偏重於處理「對於被描繪者之負面描述」，難以處理特定案件類型

自論者討論的脈絡觀之，其對於惡意使用 deepfake 的想像多集中在製作者或散布者藉由 deepfake 製造有損於影像中被描繪者形象之事，因此該被描繪之人能夠，且有動機，尋求一些保護個人法益或兼及保護個人法益之刑法規定主張權利，進而達到規制行為人惡意使用之效果。

²⁹⁷ 許恒達（2022），前揭註 4，頁 8 頁 註 17。

如前述我國學者即在文中舉例提及刑法第 313 條詐術妨害信用罪可能得以處裡一部份此類型犯罪時，所舉的例子是「如果行為人利用深偽技術製造銀行家甲自稱破產的影像，並予以散布」²⁹⁸，然而在同樣的例子裡，倘若行為人是利用 deepfake 技術製造該銀行家甲與諸多國際知名企業家於酒宴上推杯換盞、觥籌交錯之影像並予以散布，因而於社會上製造其政通人和之假象時，似乎於現行法下即欠缺足以對應處理之規範。（此類案件若要針對其後續因此博得之商業利益或締約機會等，論及詐欺取財或得利等罪名，可能會在因果關係之判斷上，抑或是是否致生交易對象財產損害等要件上無法通過檢驗）。

又像是美國有學者論及 deepfake 技術之應用時提到，此可能讓人得以在運動領域競爭選秀順位時，亦或是職場上競爭職位時，透過製造對於競爭對手不利的影像流傳於公眾藉以獲利，該學者甚至援引研究指出在美國有超過 9 成的企業主在決定職位歸屬時會嘗試在網路上搜尋關於候選人的相關訊息，因此佐證以此方法去達到陷害的效果會有相當卓越之成效²⁹⁹。事實上，前開學者所提之陷害手段，本身即可能無須達到刑法上足以被評價為「毀損他人名譽」之程度就可以得手，例如在選秀前刻意釋放出作為投手的競爭對手投球球速嚴重失速的影像，即可能使球團認為對方有健康疑慮而卻步，然而此舉卻可能並未符合刑法誹謗罪之要件，因此無法被現行法處理。進一步言，若該行為人釋放、散布的並非競爭對手的負面描述影像，而是利用 deepfake 技術製造自己球速激增之影像，並加以釋出、散布以讓自己取得選秀上競爭優勢時，更是現行法下可能難以處理的類型。

更甚者，若單純要製造他人困擾，在利用 deepfake 技術時也未必須要對對方進行「負面描述」，一些看似中性的假資訊本身也足以製造紛擾。例如針對保守派的反同陣營倡議人士或意見領袖，利用 deepfake 技術製作其出席同志遊行的影像；又像是針對基本教義派的穆斯林宗教領袖，利用 deepfake 技術製作其公開批判可蘭經教義的影像，都足以使其生活領域的信任關係遭到劇烈擾動，卻不足以被認為構成現行法下誹謗等相關刑事規範，因此可能存在著規範上的漏洞。前述舉例亦非筆者空穴來風的假想，在如今 deepfake 的使用脈絡上，即存在許多很難被認為是針對被描繪者進行「負面描述」，然而卻仍可能對社會有所影響之案例。例如 2020 年美國總統大選時，即有有心人士透過 deepfake 技術使美國前總統 Barack Obama 與時任總統 Donald Trump 說出不符其理念的虛偽言論³⁰⁰，然而與個人理念不符卻未必代表著即是「負面描述」。又好比同一場選戰進行中，亦有人使用 deepfake 技術，製造出時任國務卿 Mike Pompeo 高唱「我愛你，中國」的虛假影片，瘋傳全球在一天內獲得破百萬次的點擊數³⁰¹。而在 2018 年的英國，亦有人以 deepfake 製作出兩名候選人替彼此背書（endorsing）地虛偽影片³⁰²，這

²⁹⁸ 許恒達（2022），前揭註 4，頁 8。

²⁹⁹ Chesney, Citron (2019), *supra* note 5, 1774-1775.

³⁰⁰ 許恒達（2022），前揭註 4，頁 3。

³⁰¹ 雷喻翔（2022），前揭註 7，頁 32-33。

³⁰² Thombre (2021), *supra* note 6, 2273.

些例子都同樣難以被以強調處理對於被描繪者進行「負面描述」的相關法規加以對應。也難怪即有學者直言，在適當時機釋放出的，政府官員從事正面行為（engaging in positive behavior）的 deepfake 影音，都仍可能侵蝕人民對於政府機構的信任³⁰³。

總的來說，對於「利用 deepfake 技術製造有利特定人的不實影像」與「利用 deepfake 技術製造他人非負面描述，卻足以干擾其生活圈信任關係的不實影像」的對應處理上，本文認為在既存實定法下，處理這類事件的能量可能仍存在力猶未逮之處，因此有另立新法加以對應的必要。

第二目 僅是正面表列劃定保護範圍，因而有所缺漏

如前所述，我國學者認為考量現行刑事法制上已經存在許多規範足以應對 deepfake 之惡意使用，例如意圖使人當選或不當選而散布謠言罪或散布軍事謠言罪等諸多對於「謠言」或「不實訊息」的規制，因此無須另立規範處理之。然而就以其所舉之條文為例，在我國的脈絡下可能都有適用上之盲區。

例如，在選罷法的散布謠言罪中，以意圖干擾「候選人」之當選與否，作為構成要件之一部分。然而在我國的選戰舉文化上，選戰早在登記成為候選人之前就已開打，各方陣營於彼此尚是擬參選人身分時就開始彼此攻訐的情況也所在多有，此顯然是選罷法系爭規定所無法處理者；而在散布軍事謠言罪之條文中，以散布「軍事上」之謠言或不實訊息作為構成要件。然而在台灣如今的社會脈絡，欲威脅國安所散布之謠言很可能無須直接與「軍事」掛勾，而係可以從外交、經濟等面向上著手，諸如利用 deepfake 技術製造美國總統與中國總書記暢談放棄台灣利益相關規劃，或製造中共國務院總理揚言經濟封鎖台灣等影像，都足以擾動我國之國安狀況。

誠然，本文並未要否定各個專門領域上，各自針對需求設計特別法之構成要件的正當性，甚至也同意前述對於選罷法與散布軍事謠言罪的挑剔，各自可能透過修正構成要件加以解決。然而，對於我國是否真的不需要建構一個一般性的規範，而得以僅單純仰賴各個專門領域上的特別法，各自正面表列劃定保護範圍，就足以如論者所言的「已足對應此類深偽影音，無另行增訂刑法規範必要」，本文仍對此感到高度懷疑。誠如該學者自己也提及「濫用深偽影音侵害他人或社會共同生活秩序的行為，並非沒有處理，只不過處罰必須從既有法益與相應構成要件進行判斷，這導致其制裁範圍相對受限」³⁰⁴，準此，既然亦認為制裁範圍有所

³⁰³ Langa (2021), Deepfakes, *supra* note 34, 767.

³⁰⁴ 許恒達（2022），前揭註4，頁8。

局限，是否即能謂現行法架構已足以徹底應對 deepfake 影像之不當使用，不無疑問。

比較法上即有學者指出，deepfake 所製成之虛假影像所以影響的層面，可能包含了感情機會的喪失、朋友間的支持、失去升遷的機會、商業合作機會被取消等各種方方面面³⁰⁵，本文認為其中許多的影響可能都很幽微，未必能被現行法下正面表列的保護規範所涵蓋。

第三目 偏重於單次事件所生的影響，忽略系統性之傷害

如我國學者所言，其歸納現行法下為數甚多被認為足以應對 deepfake 影像之不當使用的刑事規範後，指出其共通之處在於這些刑事規範的保護對象並非深偽影像的散布本身，而是把深偽影像作為犯罪手法，必須該手法足以影響社會其他利益（被害人名譽或信用／民生秩序／選舉公正），方可成立刑事責任³⁰⁶。

然而這亦啟發本文開始思考，會不會一切問題的根源其實在於，deepfake 相關影像的散布本身其實就足以表彰法益的侵害，而現行法規範正是在前端的 deepfake 相關影像的散布行為欠缺足以截堵之一般性規範，才會顯得在後段欲透過尋求保護其他法益不受侵害的條文，來處理 deepfake 相關影像散布所帶來的問題時，會顯得左支右絀。

誠如本文前一段落提到的，deepfake 技術的濫用雖然可以被用以作為侵害名譽、信用、財產的工具之一，然而除了這些後段的法益侵害外，對於「影像是足以眼見為憑的經驗證據」此一社會原本存在的信任與共識之破壞本身，即應是值得被在意之法益侵害。此影響的不只是單一事件之侵擾，而是整個社會體系對於某種訊息交換系統的信賴感破壞。此並非透過零散於各處的妨害名譽或謠言法規足以充分處理的，反過來說，縱使該行為同時得以以現行法規處罰，也不代表著政府就無須，或不應，對於前述信賴法益之破壞有所因應。舉例來說，學者於介紹偽造文書相關條文時即有提及：「在現實生活中，不乏濫用文書之公共信賴性，以偽造他人名義之文書，而既得不法之財產利益者。惟此際文書之偽造，僅係詐欺、侵占或背信等財產犯罪之準備行為。偽造文書罪之規定，非以防止此等犯罪為目的，而係側重在社會生活中所扮演之任務，藉以確保文書之公共信用，而維護社會生活之安全。因此，偽造文書罪之保護法益，乃以公共信用為對象，要無可疑。」³⁰⁷據此可知，即使某些案件中偽造文書被作為侵害財產權的工具，在該案件中得以對行為人論以詐欺、侵占、背信等罪，亦不妨礙刑法可以，且應該，

³⁰⁵ Chesney, Citron (2019), *supra* note 5, 1774.

³⁰⁶ 許恒達 (2022)，前揭註 4，頁 8。

³⁰⁷ 甘添貴 (1996)，《刑法之重要理念》，頁 270，瑞興圖書。

針對文書之公共信用與以保護。即，不能僅以破壞信用後的單一事件中仍有法規足以處罰，就做為破壞信用本身不應被刑法處罰與規制之理由。

綜上所述，本文認為 deepfake 技術的濫用所造成的傷害，並非僅限於單一事件中的特定法益被干擾，而是在散布的本身，即會破壞社會對於「眼見為憑即為真相」的信任，使人們將來不再能相信有圖有真相，但也知道能更信任些什麼其他足以作為「經驗證據」的事實認定基礎，此一侵害最終可能導致社會運作核心機能的結構性破壞。同時，現行法下有些規範之構成要件偏重於處理「對於被描繪者之負面描述」，難以處理特定案件類型。此外，若僅是仰賴特別法零散的正面表列劃定保護範圍，亦可能因為各法規構成要件之限制，因而有所缺漏。準此，本文認為單以現行法所建構的規範架構中，並不足以充分達到法益保護之機能，存在著規範上之漏洞有需額外立法填補。

第三款 所保護之法益

本文認為此情況下應保護之法益與傳統偽造文書罪章中所欲保護的公共信用法益中，關於無形偽造之部分有著許多相似之處。

直言之，以簡單的例子來說明，當有人於求職時向面試的人資人員出示一封自己假冒的署名為 Donald Trump 的推薦信時，這行為在現行法下毫無疑問的是構成無制作權之人，冒用他人名義，制作內容不實之文書的有形偽造，因而得以成立刑法第 210 條、刑法第 216 條製作與行使偽造文書罪。該條文之立法意旨學說上有認為係保護「公眾對於文書的保證機能的信賴」者，如主張：「文書具有意思傳達機能以及意思表示固定化機能，才會開始於社會中被大眾所使用，並擁有安定性，進而成為一個重要的社會制度。所以這兩個機能應該是刑法對於某種社會制度加以保障的前提要件，而非刑罰規制的基礎。而文書制度的證明機能，其首重文書的內容，是立法例中的實質主義的立論基礎，反之，保證機能則是重視文書成立時的真正機能，其亦是形式主義的立論基礎。」³⁰⁸。另亦有主張係保護「社會往來之公共信用」者，如：「透過刑法對於文書之變造或偽造行為之處罰，用以建立文書之真實性與不可偽造性與不可變造性，以確保文書在經濟交易與社會往來之公共信用與安全可靠。」³⁰⁹、「在現實生活中，不乏濫用文書之公共信賴性，以偽造他人名義之文書，而既得不法之財產利益者。惟此際文書之偽造，僅係詐欺、侵占或背信等財產犯罪之準備行為。偽造文書罪之規定，非以防止此等犯罪為目的，而係側重在社會生活中所扮演之任務，藉以確保文書之公共信用，而維護社會生活之安全。因此，偽造文書罪之保護法益，乃以公共信用為

³⁰⁸ 李茂生（2000），〈再論偽造文書罪中有關有形偽造以及保護法益的問題〉，《刑事思潮之奔騰-韓忠謨教授紀念論文集》，頁 316-317。

³⁰⁹ 林山田（2006），《刑法各罪論(下冊)》，頁 375，自刊。

對象，要無可疑。」³¹⁰另外就實務見解方面亦採取類似見解，如最高法院 110 年度台上字第 3566 號判決指出：「刑法第 210 條之偽造文書，以無制作權之人冒用他人名義而制作該文書為必要，如果行為人對於此種文書本有制作之權，縱令有不應制作而制作，亦無偽造之可言，且刑法上處罰有形的偽造文書，非祇因其虛捏或冒用他人名義，而在於虛偽文書有害於公共信用及社會交往之安全。」最高法院 109 年度台上字第 3575 號判決指出：「又刑法上所謂偽造文書（包括公文書及私文書），係指無制作權人，擅自冒用他人名義，制作內容不實之文書而言。其所侵害者為社會法益或國家法益，旨在保護文書之公共信用。」

自以上學說與實務見解可歸納出幾個結論。首先，當一個系統乘載了重要的社會機能，且該系統之運作高度仰賴、奠基於社會集體的信賴關係時，倘若該系統遭受抽象的攻擊，則即值得以刑法介入規制；再者，前述學者也指出，即便該攻擊系統的行為後續可能額外造成詐欺、侵占或背信等其他法益的侵害，也不妨害在前端的行為本身就足以構成法益的侵害；最後，自偽造文書罪章針對一般人原則上僅處罰形式不實的有形偽造，而不及於內容不實的無形偽造可知，該罪章預設的信賴基礎是文書系統中的「製作名義人」。直言之，我們的社會必須有著共同的信念，共同相信著原則上在文書系統中所傳達之內容即為製作名義人所述，縱使內容不實，只要形式上製作名義人並無錯誤而足以提供追索的依據，原則上即無處罰之必要。

然而，在如今的社會運作中，乘載重要社會機能的信賴基礎未必僅限於文書中的「製作名義人」，在很多場合，「影像中被描繪者」，成為社會信任基礎的重要性，很可能有過之而無不及。舉例而言，同樣以本文前述的例子稍作調整，當有人於求職時向面試的人資人員出示的不是自己假冒的署名為 Donald Trump 的推薦信，而是以 deepfake 製作的 Donald Trump 親口推薦自己入職的虛偽影像，在後者的狀況反而可能欺詐性更高。畢竟在強調有圖有真相的時代，只是出示有知名人士署名的文書很可能被認為容易被偽造，然而就該名人真實開口的影像呈現在眼前時，可信度反而很有可能因此提升。又好比當有業餘棒球選手欲向職業球團毛遂自薦時，出示署名為某退役知名球星替其所做的書面球探報告，相比於出示經 deepfake 編輯過的自己能夠投出超過時速 150 公里速球的不實影像，後者很可能讓球團有著更強的信賴基礎，也因此該基礎被攻擊時社會所受的動盪與影響也更深。既然以「製作名義人」作為基礎的信賴關係遭到攻擊時，現行法即肯認有以刑法積極介入之必要，則在當今社會下以「影像中被描繪者」為基礎的信賴關係遭到 deepfake 所製成的虛偽影像攻擊時，當然也有以刑法進行規制的必要性。而且如前所述，在偽造文書罪章中，並不會因為製作或使用該偽造文書後，後續可能可以成立刑法第 339 條詐欺取財，甚至其他以施加詐術作為構成要件的罪名（例如出示虛假的其他企業家受破產宣告裁定，可能成立前述學者提

³¹⁰ 甘添貴（1996），《刑法之重要理念》，前揭註 309，頁 270。

即亦可透過 deepfake 之行使而成立的刑法第 313 條詐術詐術妨害信用罪)。倘若我們不因為這些偽造文書的後續行為可以成立其他犯罪,就因此認定偽造文書罪章可以因此被取代、廢除的話,則於 deepfake 濫用的脈絡中,亦不應後續有保護其他法益的各罪名可供行使,就因此得以否定針對 deepfake 濫用進行一般性立法規範的正當性。

此見解可能遭受的質疑是,本文前述的各種 deepfake 濫用方式會不會其實透過刑法第 220 條規定:「錄音、錄影或電磁紀錄,藉機器或電腦之處理所顯示之聲音、影像或符號,足以為表示其用意之證明者,亦同。」關於準文書的規定,即可以既有的偽造文書罪章加以處理。本文對此持否定見解,蓋前述各種脈絡中,很可能不存在冒用「製作名義人」的必要,因為在該互動關係中「製作名義人」並不重要,重點在於「影音中被描繪者」。以前述所舉的業餘棒球選手以 deepfake 製作虛偽影片的例子為例,其其實無需假託該影片的「製作名義人」是什麼知名人士,完全可以坦承「製作名義人」即是自己,而事實上,該影片經重製編輯後也確實是由自己製作,「製作名義人」與「實際製作權人」並無不符,因此並未違反偽造文書罪章之相關規定。

以前述公共信用的角度去認定此類型 deepfake 使用所涉及之法益,實質上即是有別於偽造文書基礎條文僅處罰「製作名義人」虛偽不實的有形偽造,而將處罰範圍擴及到「內容不實」之無形偽造類型。事實上,我國刑法體系本即存有若干條文會針對特定情形的無形偽造進行規範,例如刑法第 213 條公務人員登載不實罪,與刑法第 215 條業務登載不實罪即屬之。觀察此類處罰無形偽造之條文,其正當性基礎可能在於公務文書與業務文書對於社會體系之運作有相當之重要性,以及由於前述兩類文書係專業人士所做成使社會對其有著特別之信賴。以同樣的兩個標準檢視此類型 deepfake 使用中,關於「影像中被描繪者」虛偽不實之無形偽造,本文認為亦應等同視之。直言之,如前文所證述的,對於當代民主社會運作中,影像作為經驗證據有著極其重要的意義,同時,基於人類對於影像證據的信任,亦會使人傾向過度樂觀的信賴此類素材。準此,在體系重要性以及社會信賴程度兩個面向觀之,此類型 deepfake 使用都有著與現行法下受刑法規制之無形偽造行為足以相提並論的影響力。

綜上所述,本文認為當今社會往來所信賴的基礎,除了傳統上想像的以「製作名義人」為核心外,其實以「影像中被描繪者」為核心的信賴關係系統也在現在的社會交往中扮演舉足輕重的腳色,然而卻不如前者有受到刑法規範上的一般性保護,此亦與前述所引的,2023 年刑法修法歷程中,高嘉瑜版本與時代力量黨團所提出之版本,對於期待普通深偽罪所欲保護之法益不謀而合。因此,有額外立法擴及規範此類型之無形偽造的必要性。並且,如同傳統偽造文書罪章中,並不因為製造偽作後續仍有保護其他法益的刑法罪名得以行使,就妨害了在前端針對信用法益被破壞的本身立法處罰的可行性或正當性。

併予敘明者，前述本文說明有混淆之虞的不實性影像可能涉及對於形象法益之侵害，其屬於個人法益之範疇。而此類的不實性影像亦可能同時涉及到公共信用法益之侵害，此時兩法益間應該處於法益競合之關係，彼此不影響判斷法益侵害之事實。



第五節 無混淆之虞的與性無關之不實影像之使用

第一項 實務見解

如前所述，在 2023 年刑法修法歷程中，共計有 7 個版本包含了普通深偽條款。其中 6 個於構成要件中，皆未特別規範強調須以該影像有致生公眾混淆之虞始能論罪。其中典型以高嘉瑜版本為例，其規定：「意圖供行使之用，以電腦技術或其他科技方法，偽造、變造他人之不實照片、影像、語音或電磁紀錄，足以生損害於公眾或他人。」應受刑事處罰。究其法條使用之文字，「足以生損害」之適性犯規定，似乎可以透過此一要件篩選掉無致人混淆之虞的不實影像，惟法院實務上對此要件將如何操作其實亦難以預測。

以偽造文書罪中相同之構成要件為例，有大量判決皆援引以廢止之最高法院三十年上字第四六五號判例，以該文書內容是否真實，作為是否「足以生損害」之判斷基準。然而此標準於普通深偽罪中顯然不適用，蓋其本就必然屬於內容不實之影像。另亦有指出：「所謂足以生損害於公眾或他人，固以有損害之虞為已足，不以實際發生損害為必要，惟若已具體認定發生實際之損害，即已明顯涵蓋足以生損害之情形，尚無需再對『足以生損害』一節，贅為多餘之記載及說明。」³¹¹

依此見解係以該文書是否具備實際造成損害之風險作為判斷基準，然而所謂之「損害」究竟範圍限定為何其實仍有疑義。舉例來說，若有人透過 deepfake 技術，製作公眾顯然不會認為係屬真實的，將他人的臉嫁接在豬隻上之影片。其顯然不會破壞到社會對於影像真實性的信賴（蓋常人皆能知道此係偽作），惟其仍可能實際上侵害到被害人之名譽權，準此，依照實務揭櫫之此操作標準仍可能認為係足以生損害於他人之不實影像，而無法徹底排除無致人混淆之虞的不實影像作為規制對象的可能性。惟亦有認為：「且所謂損害，亦不以經濟價值為限，屬抽象危險犯，祇要有令一般人誤認之危險（可能），即可成立。」³¹²，依照此見解確實以是否有令人誤認之風險作為「足生損害」此一要件之判斷基準。若依此，則應能排除無致人混淆之虞的不實影像作為規制對象的可能性。

綜合前述對於實務法院關於「足以生損害」規定之認定基準應可認為，於立法階段僅以此要件就希望達到能排除無致人混淆之虞的不實影像作為規制對象可能性之目的，會存在高度的不確定性。

³¹¹ 最高法院 109 年度台上字第 193 號判決。

³¹² 最高法院 107 年度台上字第 2360 號刑事判決。

此外，特別值得注意的是時代力量黨團所提出之版本，規定：「利用工具或設備偽造他人之影像、聲音，且無明確標示或未以他法使閱聽者得以辨識該影像、聲音係屬偽造，足以生損害於公眾或他人者」應受刑事處罰。其明文規範若使閱聽者得以辨識影音係屬偽造則可以免於刑事處罰，係唯一明確對於「是否需以有混淆公眾之潛力」此一爭點明確表態之草案版本。

第二項 學說見解

如前所述，目前對普通深偽罪表態之學者見解，皆認為無須單獨為此立法，因此亦並未表態若欲立法其規範對象是否應僅以「有混淆公眾之潛力」者為限。

第三項 本文見解

如前一章節討論現行法規範時所述，本文認為此類型之不實影像並不會特別侵害到獨立之法益類型，因此亦無特別為此單獨立法之必要，惟其作為手段若涉犯現行法下既有之犯罪類型時，仍有加以刑事處罰之可能。

第六節 總結

本章依序討論了有混淆之虞的不實性影像（類型一）、無混淆之虞的不實性影像（類型二）、有混淆之虞的與性無關之不實影像（類型三）、無混淆之虞的與性無關之不實影像（類型四）等四種類型的的不實影像，被濫用時可能侵害之法益。

在有混淆之虞的不實性影像方面，本文認為其涉及之法益並非法務部所主張之隱私法益，亦非司法院所主張之名譽法益，也非學說有主張之積極或消極之性自主權，而係攸關個人對公眾展現自我、避免遭扭曲呈現於眾的形象法益。

在無混淆之虞的不實性影像方面，本文從學說討論、比較法立法模式、經驗證據上被害人陳述等方面論述，認為其應該不涉及值得以刑事規制之法益侵害。

在有混淆之虞的與性無關之不實影像方面，本文認為其可能破壞公眾對於影像作為經驗證據之信賴，進而影響、妨礙當代社會理性討論、民主思辨之根基，因此不得僅因其後續可能於單一事件中涉及其他法益之侵害，部分行為得以以現行加以處罰，即忽視此類無形偽造對於公共信用之侵害。

在無混淆之虞的與性無關之不實影像方面，本文認為其在失去了「誤以為真」之特性後，與傳統的溝通媒介並無不同，準此亦未侵害更多法益，由現行既存刑事規範加以因應已足，無須為此修法加以處理。

第四章 不實影像規制之美國州法介紹與我國立法評析



第一節 比較法介紹

第一項 簡介

雖說 deepfake 濫用行為之問題仍屬相對新穎之社會現象，然而在他國比較法上業已出現為了防範此類行為可能造成之惡害，因而針對性的立法規制。他山之石可以為錯，本文期待藉由參照比較法上之經驗，或可提供我國立法管制此類行為時較佳之方案或靈感。

考量筆者本身使用語言之限制以及資料之可觸及性，本文選擇以美國作為比較法之主要研究對象。

直至本文撰寫之際，美國於聯邦法層級尚未出現針對 deepfake 濫用管制的相關立法，僅於 2019 年訂定 Deepfake Report Act，其內容旨在要求國土安全部在 2019 年至 2024 年間，應該對國會定期陳報針對 deepfake 之年度報告，評估外國主體使用這類電子偽造物之狀況，並且評估能夠偵測與緩解這類威脅的可行方法，而並非對於人民使用 deepfake 技術有所規範或限制。惟眾議院議員 Yvette D. Clarke 於 2019 年、2021 年間數次提出立法草案未果，且其於 2023 年的如今仍嘗試推動相關的立法³¹³。而在州法的層級，已有德州、加州、維吉尼亞州、紐約州共 4 州完成 5 部相關法規之制定（其中加州完成 2 部立法）。自規範之影音種類觀之，其中有專門針對與選舉相關之使用，亦有專針對不實性影像之使用作出規範者；自規範效果觀之，亦有民事上責任與刑事上責任之區別。

本文擬於本節依序介紹 5 部法規，並亦會挑選一些雖未完成立法但具參考價值的立法草案，簡要歸納這些法規，並分析要件與效果設計上之重點與異同。下列的這些法規或草案，雖然大多皆非刑事上之規範，然而其於要件設計上之考量與安排，仍可能提供我國立法時作為參照與借鑑。

第二項 Texas Statutes Election Code Sec. 255.004. TRUE SOURCE OF COMMUNICATION. (德州選舉法)

德州於 2019 年修正其原有之《選舉法》(Texas Statutes Election Code) 並於同年 9 月 1 日生效，旨在處理選舉日前夕以 deepfake 技術干擾選舉公正之情形。其相關條文如下³¹⁴：

³¹³ Hsu, As Deepfakes Flourish, Countries Struggle With Response, in <https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html>. (最後瀏覽日：07/13/2023)

³¹⁴

(c) An offense under this section is a Class A misdemeanor.

(d) A person commits an offense if the person, with intent to injure a candidate or influence the result

「(c)本節所規範之犯罪行為屬於 A 級輕罪。」

「(d)任何人意圖傷害候選人或影響選舉結果，而為下列行為，成立犯罪：

(1)製作 deepfake 影像；且

(2)使該 deepfake 影像在選舉日的 30 日內被公開或傳播。」

「(e)本節中所稱的『deepfake 影像』係指帶有欺騙性意圖所製作，而宛若描繪了一個真實人物從事未曾實際發生行為之影像。」

本部法規在時間序列上係美國最早成功針對 deepfake 使用進行管制之立法例，值得注意的特色包含：

1. 違反法規所需承擔的法律效果係刑事處罰。
2. 在構成要件設計方面，客觀上需同時滿足製作加上散布行為始構成違法行為。
3. 針對所欲規範之客體，以「deepfake 影像」(deep fake video)一詞對其進行描述，僅包含了影像(video)，適用範圍明顯小於其他法規所規範之影音作品(audiovisual work、audio or visual media、audiovisual record、audiovisual content)、影像或靜止之圖像(videographic or still image)等，因此將無法處理到僅是針對圖像或是音檔的偽造行為。
4. 針對所欲規範之客體所加之定義相對簡略，僅要求製作時具備欺騙之主觀意圖，而未要求其欺詐性在客觀上達到特定程度。

第三項 California Civil Code Section 1708.86 - Cause of action for depicted individual. (加州民法)

加州於 2019 年於民法典中增修管制 deepfake 濫用之相關規範，其旨在規範製作或散布不實性影像使用的相關行為，並賦予受害人對被告請求損害救濟之民事上權利。由於本部分條文內容繁多，本文將挑選其中具重要性之部分進行介紹：

「(b)被描繪之個體對於做出下列任一行為之行為人，得以提出訴訟上之請求：

(1)製作且故意揭露具有性意涵之素材，且該行為人知悉或可得而知該素材中被描繪之個體並未同意(Consent)製作或揭露該素材。

(2)故意揭露非自己製作之具有性意涵之素材，且該行為人知悉該素材中被

of an election:

(1) creates a deep fake video; and

(2) causes the deep fake video to be published or distributed within 30 days of an election.

(e) In this section, "deep fake video" means a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality.

描繪之個體並未同意 (Consent) 製作該素材。」³¹⁵

「『被描繪之個體』 (Depicted individual) 係指因為數位化，而宛若做出其未曾真實做出之表現，或於『受變造的描繪』 (Altered depiction) 中做出表現的個體。」³¹⁶

「『受變造的描繪』 (Altered depiction) 係指被描繪之個體實際上曾真實作出，然而卻隨後被變造，而違反本節規定之表現。」³¹⁷

「(6)『數位化』係指逼真地描繪下列任一情況：

(A)使另一個人的裸體部位，成為被描繪之個體的裸體部位。

(B)使電腦生成的裸體部位，成為被描繪之個體的裸體部位。

(C)使被描繪之個體從事了其不曾實際從事之性行為。」³¹⁸

「(A)『同意』 (Consent) 係指以通俗淺白的文字寫成的，被描繪之個體明確知情且自願地簽署之書面協議，內容包含對於其參與之具有性意涵之素材與影音作品的大致描述。

(B)除非滿足以下要求之一，否則被描繪之個體得在給予同意之日起三個工作日內，以書面通知撤回同意：

(i)被描繪之個體在簽署協議之前，有至少 72 小時得以檢閱協議條款。

(ii)被描繪之個體所授權之代表提供簽署協議的書面同意。」³¹⁹

315

(b) A depicted individual has a cause of action against a person who does either of the following:
(1) Creates and intentionally discloses sexually explicit material and the person knows or reasonably should have known the depicted individual in that material did not consent to its creation or disclosure.
(2) Intentionally discloses sexually explicit material that the person did not create and the person knows the depicted individual in that material did not consent to the creation of the sexually explicit material.

316

(a) For purposes of this section:

(4) "Depicted individual" means an individual who appears, as a result of digitization, to be giving a performance they did not actually perform or to be performing in an altered depiction.

317

(a) For purposes of this section:

(1) "Altered depiction" means a performance that was actually performed by the depicted individual but was subsequently altered to be in violation of this section.

318

(a) For purposes of this section:

(6) "Digitization" means to realistically depict any of the following:

(A) The nude body parts of another human being as the nude body parts of the depicted individual.

(B) Computer-generated nude body parts as the nude body parts of the depicted individual.

(C) The depicted individual engaging in sexual conduct in which the depicted individual did not engage.

319

(a) For purposes of this section:

(3)

(A) "Consent" means an agreement written in plain language signed knowingly and voluntarily by the depicted individual that includes a general description of the sexually explicit material and the

「『具有性意涵之素材』係指其中之任何部分展演了被描繪之個體裸體，或從事、遭受性行為的影音作品。」³²⁰

「(d)行為人於因本節規定而涉及之訴訟中，即便已在該具有性意涵之素材中註明免責聲明，表明該被描繪之個體於該素材之參與未經授權，或表明該被描繪之個體並未參與該素材的製作與開發，仍不得執此作為訴訟上之抗辯。」³²¹

本部法規具備值得注意的特色如下：

1. 其設計結構之核心條文位於(b)款，規範了兩種違反法規之行為樣態。

然而因為(b)款所用之文字簡明，涉及大量不確定法律概念，因此於(a)款花費大量篇幅進行用詞之定義與解釋，甚至多次出現解釋中亦涉及新的專有名詞因此需要額外再以定義解釋定義的狀況。

在(c)款部分規定有豁免條款，明確定義於特定情況下例如報導不法活動、法庭聽證程序或其他涉及公益之狀況下，無須承擔本條規定下之法律責任。

在(d)款部分特別明文排斥特定類型之訴訟上抗辯，而在(e)款部分規定違反本條文之法律效果，另原告得以向被告請求藉由散布影像所得之獲利³²²、財產上與非財產上之損失，包含情緒上的困擾³²³、懲罰性賠償金(Punitive damages)、合理的律師費(Reasonable attorney's fees and costs)以及包含禁制令在內的其他可行的救濟途徑³²⁴等，並強調本條文規定之責任係屬累積性的，不因此排斥其他法規下之救濟措施。

2. 在(b)款部分，同樣的散布 deepfake 影像之行為會因為行為人是否為製作者，而有相異的主觀要件要求。

audiovisual work in which it will be incorporated.

(B) A depicted individual may rescind consent by delivering written notice within three business days from the date consent was given to the person in whose favor consent was made, unless one of the following requirements is satisfied:

(i) The depicted individual is given at least 72 hours to review the terms of the agreement before signing it.

(ii) The depicted individual's authorized representative provides written approval of the signed agreement.

³²⁰

(a) For purposes of this section:

(14) "Sexually explicit material" means any portion of an audiovisual work that shows the depicted individual performing in the nude or appearing to engage in, or being subjected to, sexual conduct.

³²¹

(d) It shall not be a defense to an action under this section that there is a disclaimer included in the sexually explicit material that communicates that the inclusion of the depicted individual in the sexually explicit material was unauthorized or that the depicted individual did not participate in the creation or development of the material.

³²² An amount equal to the monetary gain made by the defendant from the creation

³²³ Economic and noneconomic damages proximately caused by the disclosure of the sexually explicit material, including damages for emotional distress

³²⁴ Any other available relief, including injunctive relief

對於實際製作者除知悉未得同意外，僅需可得而知未得被描繪之個體的同意而散布，即成立法律上責任；

反之，非實際製作者在散布時，僅有知悉未得被描繪之個體的同意而散布始成立本條文下之法律責任。

3. 針對所欲規範之客體，以「具有性意涵之素材」(sexually explicit material)一詞對其進行描述，在解釋上將「具有性意涵之素材」解釋成「任何部分展演了被描繪之個體裸體，或從事、遭受性行為的影音作品。」，單就這方面而言完全無法看出係針對 deepfake 技術所規範，甚至徹底真實的性相關影音似乎也會落入本法規所欲處理的範圍之下。

然而本法其實是取徑「被描繪之個體」一詞的定義來限縮、特定處理的範圍，將其定義成「係指因為數位化，而宛若做出其未曾真實做出之表現，或於『受變造的描繪』(Altered depiction)中做出表現的個體。」，隨後又分別定義「數位化手段」與「變造的描繪」二詞，總之是透過相當迂迴的途徑才完成對於所欲處理對象的劃定，此部分完全可以看出由於科技的特殊與新穎性，導致化約規範對象時的艱辛與為難。

然而綜合以上諸多定義，可以確定的部分僅有該影像需呈現出被描繪之客體與性相關之「未曾真實做出」或「雖曾做出但經變造」之表現，然而該表現是否要求具備欺詐性？應否要求有足生理性一般人混淆之危險？在法條條文中並未被清晰的呈現。

可能與之有關的用詞可能僅有在定義「數位化手段」一詞時，立法者使用了「逼真地描繪」(realistically depict)一詞去描述「數位化手段」之範圍，或可透過「是否足夠逼真」(realistically)來排除欺騙性不足的影像作品亦落入本條規範之射程範圍內。

4. 本條文對於被描繪之個體所為之同意(Consent)有明確的形式要求規定，除明確要求需書面為之外，亦對於撤回同意之要件有相當之規範。此設計影響最鉅的方面應該會落在於對於被告主觀要件的證明，相較於不對同意做出形式要求，將使散布者更難主張自己不知被描繪之個體並未同意等。
5. (d)款部分明確地限制被告不得主張已在影音作品中加註免責聲明，註明被描繪之個體並未實際參與影音作品之製作，執此作為訴訟上抗辯。此規範可能使一些已經透過文字說明，使觀眾欠缺誤以為真潛力的影音作品，仍落入本法規所欲規制的射程範圍內。從此當中或可觀察出立法者並不在意誤認之可能性（也因此有本文前述的，並未特別以文字要求欺詐性程度的設計出現）。
6. (a)款(9)部分針對惡意(Malice)一詞做出定義：「『惡意』係指被告故意對原告造成傷害(cause harm to the plaintiff)或做出卑鄙行為(despicable conduct)，故意且明知地無視原告的權利。」³²⁵。

而成立「惡意」的法律效果落在(e)款，得以以此提高法定損害賠償（statutory damages）之額度。

反面解釋可知，立法者認為即便散布行為「無意對被告造成傷害」仍然屬於本法所欲規制的射程範圍內，並未對行為人的主觀意圖多有限制。

第四項 Code of Virginia 18.2-386.2. Unlawful dissemination or sale of images of another; penalty.（維吉尼亞州刑法）

德州於 2019 年修正其州法並於同年 7 月生效，旨在處理不實性影像所造成的傷害。其相關條文如下³²⁶：

「任何人意圖脅迫、騷擾、或恐嚇，惡意散布或販賣以一切方式製作的，描繪他人全裸或處於未著衣而暴露生殖器、陰部、臀部或女性乳房之狀態的任何影像或靜止之圖像，而知悉或可得而知其並未有權散布或販賣該影像或靜止之圖像者，構成第一級輕罪。本款所稱之『他人』，包含任何個人其形象被用以製作、改編或修改成為影像或靜止之圖像，而該影像或靜止之圖像意圖描繪一個真實存在之人，或可經面部、相似性或其他可識別特徵，辨識為一個真實存在之人。」

本部法規具備值得注意的特色如下：

1. 違反法規所需承擔的法律效果係刑事處罰。
2. 構成要件設計上，本規範特別強調需具備指定的主觀意圖。此項設計將導致單純分享或營利而散布不實性影像之行為，無法落入本法射程範圍內。
3. 針對所欲規範之客體，文字上僅描述與性相關的影像或靜止之圖像，此部分無法觀察出係專為處理 deepfake 技術製成之影音。

惟其亦在保護對象上，強調該「他人」係圖像被改編藉以製成影音之人，且該影音需「意圖描繪一個真實存在之人，或可經面部、相似性或其他明顯特徵辨識為一個真實存在之人」。

綜合上述文字，立法者似僅要求行為人「主觀上」有意將該影音連結或可得連結到特定真人即可，而不論客觀上是否具相當欺詐性。

(a) For purposes of this section:

(9) "Malice" means that the defendant acted with intent to cause harm to the plaintiff or despicable conduct that was done with a willful and knowing disregard of the rights of the plaintiff.

³²⁶

A. Any person who, with the intent to coerce, harass, or intimidate, maliciously disseminates or sells any videographic or still image created by any means whatsoever that depicts another person who is totally nude, or in a state of undress so as to expose the genitals, pubic area, buttocks, or female breast, where such person knows or has reason to know that he is not licensed or authorized to disseminate or sell such videographic or still image is guilty of a Class 1 misdemeanor. For purposes of this subsection, "another person" includes a person whose image was used in creating, adapting, or modifying a videographic or still image with the intent to depict an actual person and who is recognizable as an actual person by the person's face, likeness, or other distinguishing characteristic.

4. 於D.款特別強調：「本節的規定不應排除根據任何其他法規提起的訴訟。」

327

第五項 New York 52-C. Private right of action for unlawful dissemination or publication of a sexually explicit depiction of an individual. (紐約州民法)

紐約州於2020年於民法典中增修管制 deepfake 濫用之相關規範，其旨在規範製作或散布不實性影像使用的相關行為，並賦予受害人對被告請求損害救濟之民事上權利。

本部法規在文字使用上幾乎與前述的 California Civil Code Section 1708.86 - Cause of action for depicted individual 完全相同，僅在若干處存有不影響文義的字詞修正，故本文對其之分析亦與本節第三項相同，茲不贅述。

第六項 California 2021 California Code Elections Code Section 20010 - Campaign material with superimposed image of person prohibited. (加州選舉法)

加州選舉法經修訂後於2023年生效，旨在處理選舉相關之不實影像濫用問題。其相關條文如下：

「除非符合(b)條款之規定，任何個人、依照 Government Code 第 82013 條所規範之委員會，或其他實體，在候選人將出現在選票上的選舉日之 60 日內，不得意圖傷害候選人之聲譽或欺騙選民投票支持或反對候選人，以真實惡意散布 (e) 條款所規範之，具實質欺騙性的聽覺或視覺素材。³²⁸」

「(b) 任何人、公司、協會、法人、競選委員會或組織，僅有在於競選素材所包含的每張照片或圖像旁，以與競選素材其他部分所使用的最大字體同等大小之字體，註明『本照片所呈現的並非準確的真相。』之警語，且該警語應緊鄰任何被(a)條款所禁止之照片或圖像時，始得生產、散布、公開或傳播包含(a)條款所禁止之照片或圖像的競選素材。³²⁹」

³²⁷ D. The provisions of this section shall not preclude prosecution under any other statute.

³²⁸ Except as provided in subdivision (b), a person, committee, as defined in Section 82013 of the Government Code, or other entity shall not, within 60 days of an election at which a candidate for elective office will appear on the ballot, distribute, with actual malice, materially deceptive audio or visual media, as defined in subdivision (e), of the candidate with the intent to injure the candidate's reputation or to deceive a voter into voting for or against the candidate.

³²⁹ A person, firm, association, corporation, campaign committee, or organization may produce, distribute, publish, or broadcast campaign material that contains a picture or photograph prohibited by subdivision (a) only if each picture or photograph in the campaign material includes the following statement in the same point size type as the largest point size type used elsewhere in the campaign material: "This picture is not an accurate representation of fact." The statement shall be immediately adjacent to each picture or photograph prohibited by subdivision (a).

「公職候選人的聲音或形象出現在違反本節規定而散布的，具實質欺騙性的聽覺或視覺素材中，可尋求禁令或其他公平救濟，禁止散布違反本節規定之聽覺或視覺素材。³³⁰」

「若公職候選人的肖像出現在(a)條款所禁止的圖片或照片中，可以對生產、散布、公開或傳播該圖片的任何個人、公司、協會、法人、競選委員會或組織提起民事訴訟。除合理的律師費和其他開銷外，法院尚得判賠額度相當於違反本節規定之競選素材的生產、散布、公開或傳播費用的損害賠償。³³¹」

「本節法規不適用於構成諷刺或諧擬的具實質欺騙性之聽覺或視覺素材。³³²」

「本節所稱之『具實質欺騙性之聽覺或視覺素材』係指紀錄著候選人外表、言論或行為的圖像、音檔或影像，且被有意操弄而同時符合以下兩個條件：

(1)該圖像、音檔或影像會使理性一般人誤以為真。

(2)該圖像、音檔或影像會使理性一般人，相較於看到或聽到未經變造的原始版本，對其所欲表達的內容產生根本性的不同理解或印象。³³³」

本部法規具備值得注意的特色如下：

1. 違反法規所需承擔的法律效果係遭受民事上禁制令，以及民事上之損害賠償。

2. 本部法規同時透過三個途徑去限縮、排除不具欺詐性之客體。

其一是該影音需具備實質欺騙性（materially deceptive），而所謂實質欺騙性必須「使理性一般人誤以為真」且「使理性一般人因為該客體而產生錯誤的理解與印象」；

其二是明文於(b)規範，倘依照其所規定之格式加註警語，則製作、散布、公

³³⁰ A candidate for elective office whose voice or likeness appears in a materially deceptive audio or visual media distributed in violation of this section may seek injunctive or other equitable relief prohibiting the distribution of audio or visual media in violation of this section. An action under this paragraph shall be entitled to precedence in accordance with Section 35 of the Code of Civil Procedure.

³³¹ A candidate for public office whose likeness appears in a picture or photograph prohibited by subdivision (a) may bring a civil action against any person, firm, association, corporation, campaign committee, or organization that produced, distributed, published, or broadcast the picture or photograph prohibited by subdivision (a). The court may award damages in an amount equal to the cost of producing, distributing, publishing, or broadcasting the campaign material that violated this section, in addition to reasonable attorney's fees and costs.

³³² This section does not apply to materially deceptive audio or visual media that constitutes satire or parody.

³³³ As used in this section, "materially deceptive audio or visual media" means an image or an audio or video recording of a candidate's appearance, speech, or conduct that has been intentionally manipulated in a manner such that both of the following conditions are met:

(1) The image or audio or video recording would falsely appear to a reasonable person to be authentic.

(2) The image or audio or video recording would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.

開或傳播不實影像之行為皆不具不法性；

其三是明文規範了豁免條款，指出本法不適用於構成諷刺或諧擬的具實質欺騙性聽覺或視覺素材。

第七項 Massachusetts petition (accompanied by bill, House, No. 1755). (麻州草案)

麻薩諸塞州於 2019 年經議會議員提出修法草案，試圖擴張既有法規中身分詐欺 (identity fraud) 之定義範圍，藉以規制處理 deepfake 技術濫用之問題，然而最終並未通過修法³³⁴。其預計修法內容如下：

「(b)任何人(1)意圖散布而製作，且意圖透過散布促成刑事犯罪或民事侵權行為而製作 deep fake，或(2)知悉所散布之影音記錄係 deep fake，且意圖透過散布促成刑事犯罪或民事侵權行為而散布該影音記錄，應成立身分詐欺之刑事犯罪……。」³³⁵

「(a)為了本節所使用，以下之字詞除了依其文脈明確另有要求外，定義如下：

影音紀錄 (Audiovisual record) 係指任何電子形式的音檔或影像，包括任何照片、電影、影像記錄、電子圖像或錄音。

deep fake 係指以某種方式創造或變更，而會被理性的觀察者 (a reasonable observer) 錯誤地認為是個人實際言論或行為的影音記錄。」³³⁶

本部法規具備值得注意的特色如下：

1. 違反法規所需承擔的法律效果係刑事處罰。
2. 論者有指出，相較於定義 deepfake 影音之特定用途 (如選舉干擾、非合意色情影音等)，本法規更擴大其適用範圍到一切「促成刑事犯罪或民事侵權行為」的散布行為³³⁷。

惟本文對於這樣的法規設計是否真係擴大適用範圍有所疑義，蓋在如此

³³⁴ Langa (2021), *supra* note 34, 778.

³³⁵

(b) Whoever (1), creates, with the intent to distribute, a deep fake and with the intent that the distribution of the deep fake would facilitate criminal or tortious conduct, or (2) distributes an audiovisual record with actual knowledge that the audiovisual record is a deep fake and with the intent that the distribution of the audiovisual record would facilitate criminal or tortious conduct shall be guilty of the crime of identity fraud

³³⁶

(a)As used in this section, the following words shall have the following meaning unless the context clearly requires otherwise:

“Audiovisual record,” any audio or visual media in an electronic format and includes any photograph, motion-picture film, video recording, electronic image, or sound recording.

“Deep fake”, an audiovisual record created or altered in a manner that the record would falsely appear to a reasonable observer to be an authentic record of the actual speech or conduct of an individual.

³³⁷ Langa (2021), *supra* note 34, 778.

規範下，散布 deepfake 影音是否成立犯罪端繫於該影像之散布本身是否涉及刑事犯罪或民事侵權。

倘若該散布行為本就成立法律上責任，例如散布以 deepfake 技術變造 Donald Trump 辱罵官員之影片，則因本就在民刑事上有相對應之規範得以處理，另立新法也未必能讓行為人之法律上責任進一步加重，未免有疊床架屋之嫌；

反之，若該散布行為本身並不成立法律上責任，例如在我國舊法下散布以 deepfake 技術變造製成之不實性影像，又如本文意見不認為此行為成立誹謗等罪時，則此類的立法亦可能因沒有「促成刑事犯罪或民事侵權行為」而無從動用、形同虛設。

附言者論，參議員 Ben Sasse 於 2018 年在聯邦層級試圖引進立法的 Malicious Deep Fake Prohibition Act 亦採取了相同的立法文字描述。

3. 針對所欲規範之客體，本法明確描述需使「理性的觀察者」誤以為真，始屬於本法射程範圍內所欲處理的對象。

準此，則無論是技術上過於拙劣導致欺詐性不足，亦或是額外附加警語使理性的觀察者不至於誤以為真時，皆無本規範之適用。

第八項 The DEEP FAKES Accountability Act (DEEP FAKES 法草案)

2019 年與 2021 年時，眾議院議員 Yvette D. Clarke 兩度提出「確保剝削受問責，以保護所有人免於以錯誤形象呈現」法案 (Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act, The DEEP FAKES Accountability Act)，惟皆未成功立法。其預計於 2023 年持續嘗試推動修法。其預計修法內容概略介紹如下：

「(a)總論：除非符合(j)款之例外，任何人使用國際或外國通商方式或工具，意圖透過網路散布或知悉該紀錄將會被散布，製作以先進技術偽造之身分紀錄，應確保該紀錄符合下列規定：

- (1) 應符合(b)款規定之浮水印要求 (watermark requirement)，和
- (2) 若該紀錄係屬影音、影像、音檔，分別應符合(c)款、(d)款、(e)款規定之揭露要求 (disclosure requirements)。」³³⁸

338

(a) In General.—Except as provided in subsection (j), any person who, using any means or facility of interstate or foreign commerce, produces an advanced technological false personation record with the intent to distribute such record over the internet or knowledge that such record shall be so distributed, shall ensure such record, complies with—

(1) the watermark requirement under subsection (b); and

(2)

(A) in the case of an audiovisual record, the disclosure requirements under subsection (c);

(B) in the case of a visual record, the disclosure requirements under subsection (d); or

(C) in the case of an audio record, the disclosure requirements under subsection (e).

「(b)浮水印—任何包含動態視覺元素之以先進技術偽造之身分紀錄，應包含嵌入式之數位浮水印，清楚的辨明該紀錄包含經編輯過之聲音或影像元素。」³³⁹

「(c)影音揭露—任何同時包含聲音或影像元素之以先進技術偽造之身分紀錄，應包含：

(1)至少 1 則清晰的口頭聲明，辨明該紀錄包含經編輯過之影音元素，並簡要的敘述編輯之程度；且

(2)在視覺元素的持續期間，於圖像底部以清晰可讀的文字呈現的未受遮蔽之書面聲明，辨明該紀錄包含經編輯過之影音元素，並簡要的敘述編輯之程度。」³⁴⁰

「(d)影像揭露—任何只包含影像元素之以先進技術偽造之身分紀錄，應包含在視覺元素的持續期間，於圖像底部以清晰可讀的文字呈現的未受遮蔽之書面聲明，辨明該紀錄包含經編輯過之影像元素，並簡要的敘述編輯之程度。」³⁴¹

「(e) 語音揭露—任何只包含聲音元素之以先進技術偽造之身分紀錄，應包含在該紀錄開始處，以清晰的口頭聲明，辨明該紀錄包含經編輯過之聲音元素，並簡要的敘述編輯之程度；且若該紀錄長度超過 2 分鐘，應額外於隨後的每 2 分鐘間隔處，增加至少 1 則語音聲明，辨明該紀錄包含經編輯過之聲音元素，並額外簡要的敘述編輯之程度。」³⁴²

339

(b) Digital Watermark.—Any advanced technological false personation record which contains a moving visual element shall contain an embedded digital watermark clearly identifying such record as containing altered audio or visual elements.

340

(c) Audiovisual Disclosure.—Any advanced technological false personation records containing both an audio and a visual element shall include—

(1) not less than 1 clearly articulated verbal statement that identifies the record as containing altered audio and visual elements, and a concise description of the extent of such alteration; and

(2) an unobscured written statement in clearly readable text appearing at the bottom of the image throughout the duration of the visual element that identifies the record as containing altered audio and visual elements, and a concise description of the extent of such alteration.

³⁴¹ (d) Visual Disclosure.—Any advanced technological false personation records exclusively containing a visual element shall include an unobscured written statement in clearly readable text appearing at the bottom of the image throughout the duration of the visual element that identifies the record as containing altered visual elements, and a concise description of the extent of such alteration.

³⁴² (e) Audio Disclosure.—Any advanced technological false personation records exclusively containing an audio element shall include, at the beginning of such record, a clearly articulated verbal statement that identifies the record as containing altered audio elements and a concise description of the extent of such alteration, and in the event such record exceeds two minutes in length, not less than 1 additional clearly articulated verbal statement and additional concise description at some interval during each two-minute period thereafter.



「(f) 罰則—

(1) 刑事責任

(A) 未能揭露—故意違反(a)款者，而：

- (i) 意圖羞辱或騷擾被虛偽展示之人，且該先進科技偽造之身分紀錄包含性相關視覺內容，並顯示該他人參與該性活動或呈現裸體狀態；
 - (ii) 意圖引發暴力或人身傷害，煽動武裝或外交衝突，或干涉包涵選舉在內之公務，且該先進科技偽造之身分紀錄確實構成了引發或促進此類行為的可信威脅；
 - (iii) 涉及詐欺相關的刑事犯罪，包含證券詐欺、電信詐欺、虛偽冒充或身份盜用；或
 - (iv) 外國勢力或其代理人，意圖影響國內公共政策辯論，干涉聯邦、州、地方或地區選舉，或從事該勢力不得合法從事的其他行為；
- 應根據本法處以罰金、五年以下有期徒刑，或兩者兼具。」³⁴³

「(j) 例外—

(1) 揭露—(c)、(d)、(e) 款所規範之要求，不適用於以下種類之以先進科技偽造之身分紀錄：

(A) 含有替代性的揭露措施，依照個案情節觀之足使理性之人(a reasonable person)認為這些揭露相較(c)、(d)或(e)款所要求的揭露效果更為顯著；

(C) 主要包含真實人物(例如表演藝術家)的圖像或錄音，而未經過大幅度的數位修改；

(E) 依照其情境，理性之人不會將偽造的物質活動誤認為是其中出演者所為真正的物質活動，例如模仿表演或出版物、歷史重演或虛構的廣播、電視或電影節目。」³⁴⁴

³⁴³

(f) Penalty.—

(1) CRIMINAL PENALTY.—

(A) FAILURE TO DISCLOSE.—Whoever knowingly violates subsection (a)—

“(i) with the intent to humiliate or otherwise harass the person falsely exhibited, provided the advanced technological false personation record contains sexual content of a visual nature and appears to feature such person engaging in such sexual acts or in a state of nudity;

“(ii) with the intent to cause violence or physical harm, incite armed or diplomatic conflict, or interfere in an official proceeding, including an election, provided the advanced technological false personation record did in fact pose a credible threat of instigating or advancing such;

“(iii) in the course of criminal conduct related to fraud, including securities fraud and wire fraud, false personation, or identity theft; or

“(iv) by a foreign power, or an agent thereof, with the intent of influencing a domestic public policy debate, interfering in a Federal, State, local, or territorial election, or engaging in other acts which such power may not lawfully undertake;

shall be fined under this title, imprisoned for not more than 5 years, or both.

³⁴⁴ (j) Exceptions.—

“(1) DISCLOSURE.—The requirements under subsections (c), (d), and (e) shall not apply with respect to any advanced technological false personation record—



「(n) 定義—

(1)以先進科技偽造之身分紀錄—係指任何 deep fake 同時符合下列條件者：

(A) 理性之人在考慮了記錄的影像或音檔品質以及發行渠道的性質後，會相信該紀錄準確地呈現了—

(i)作品中的在世人士實際上並未從事的任何物質活動，或

(ii)作品中的已逝世人士實際上並未從事的任何物質活動，而這些活動的展演很可能會助長犯罪行為或導致對公務、公共政策辯論或選舉的不當干預。

(B) 未經該在世人士或已逝世人士或其繼承人同意而製作。

(2) 物質活動—係指任何的虛假的言論、行為或描繪，可能導致或理性之人會認識到具有導致可感知的個人或社會損害的傾向，包括誤導、名譽損害、困窘、騷擾、經濟損失、煽動暴力、干預公共政策辯論或選舉，或助長任何違法行為

(3) deep fake 是指任何影像、電影、錄音、電子圖像或照片，或任何展現言論或行為的科技，符合以下標準：

(A) 看似真實地描述了一個人的言論或行為，但實際上該人並未參與該言論或行為；且

(B) 其製作主要取決於技術手段，而非個人在肢體上或言語上冒充該人之能力。」³⁴⁵

“(A) containing alternative disclosures regarding the falsity of the exhibited material activities which a reasonable person would deem to be more prominent than those required under subsection (c), (d), or (e), as the case may be;”

“(C) which primarily contains images or sound recordings of actual persons, such as performing artists, and have not been substantially digitally modified;”

“(E) appearing in a context such that a reasonable person would not mistake the falsified material activity for actual material activity of the exhibited living person, such as parody shows or publications, historical reenactments, or fictionalized radio, television, or motion picture programming.”

³⁴⁵ (n) Definitions.—

“(1) ADVANCED TECHNOLOGICAL FALSE PERSONATION RECORD.—The term ‘advanced technological false personation record’ means any deep fake, which—

“(A) a reasonable person, having considered the visual or audio qualities of the record and the nature of the distribution channel in which the record appears, would believe accurately exhibits—

“(i) any material activity of a living person which such living person did not in fact undertake; or

“(ii) any material activity of a deceased person which such deceased person did not in fact undertake, and the exhibition of which is substantially likely to either further a criminal act or result in improper interference in an official proceeding, public policy debate, or election; and

“(B) was produced without the consent of such living person, or in the case of a deceased person, such person or the heirs thereof.

“(2) MATERIAL ACTIVITY.—The term ‘material activity’ means any falsified speech, conduct, or depiction which causes, or a reasonable person would recognize has a tendency to cause perceptible individual or societal harm, including misrepresentation, reputational damage, embarrassment, harassment, financial losses, the incitement of violence, the alteration of a public policy debate or election, or the furtherance of any unlawful act.

“(3) DEEP FAKE.—The term ‘deep fake’ means any video recording, motion-picture film, sound recording, electronic image, or photograph, or any technological representation of speech or conduct

本部法規具備值得注意的特色如下：

1. 自違反規範的法律效果觀之，本部法規同時課以行為人刑事上與民事上的法律責任。
2. 自所欲打擊的不實影音類型觀之，本部法規同時欲處理了用於數位性暴力、妨礙選舉公正等公務、各類型詐欺、境外勢力干預國家安全等各種類型的 deepfake 影音，打擊範圍甚廣。
3. 自規範的行為類型觀之，本部法規所規制的並非單純的散布或製作行為，準確來說係「帶有特定意圖的，未確實依法嵌入浮水印或揭露資訊而散布」或「帶有特定意圖的，移除或有意掩蓋應揭露資訊且散布」之行為。亦即，在此種規範類型下只要確實嵌入浮水印且揭露此為 deepfake 影音，則無論如何製作、散布該影音皆為法律所容許。同時本部法規亦對於浮水印與揭露的形式規範有比較詳細的規定。
4. 自所欲規範的客體觀之，本部法規定義了以先進科技偽造之身分紀錄（advanced technological false personation record）、物質活動（material activity）、deep fake 三個專有名詞。

其中所稱的 deep fake 係指「看似真實地描述了一個人的言論或行為，但實際上該人並未參與該言論或行為」的一切影音，並未要求該影音需要具備相當的欺詐性，是個較大的集合，但並非本部法規所欲規範之客體。

本部法規真正欲規範之客體係「以先進科技偽造之身分紀錄」，依照本法定義此紀錄必須是 deep fake，但尚須使「理性之人在考慮了記錄的影像或音檔品質以及發行渠道的性質後，會相信該紀錄準確地呈現了作品中的在世人士實際上並未從事的任何物質活動」，其中還區分了出演者若是已逝世之人，則紀錄之成立標準將進一步嚴格，亦即成立「以先進科技偽造之身分紀錄」的可能性將進一步限縮。

同時，其要求該影音讓理性之人誤以為真的言行並非普通的言行已足，而係要求該言行屬於「物質活動」（material activity），亦即「任何的虛假的言論、行為或描繪，可能導致或理性之人會認識到具有導致可感知的個人或社會損害的傾向，包括誤導、名譽損害、困窘、騷擾、經濟損失、煽動暴力、干預公共政策辯論或選舉，或助長任何違法行為」。

綜合以上條件的交集可知，本部法規所欲規範的客體其實是「具備相當欺詐性，且描述之內容將導致傷害之虞的，未得出演者或其繼承人同意，卻描繪出演者從事其未曾真正做出之言行」的影音。

substantially derivative thereof—

“(A) which appears to authentically depict any speech or conduct of a person who did not in fact engage in such speech or conduct; and

“(B) the production of which was substantially dependent upon technical means, rather than the ability of another person to physically or verbally impersonate such person.

5. 此外，在(j)款紀載的，例外得豁免於履行揭露要求要件中，亦可以觀察出「該影音本身有效果更佳的揭露措施」、「該影音經編輯修正的程度不高」、「該影音依照脈絡並不會使理性之人誤以為真」等狀況下，則製作者即不在本法規範的射程範圍內。

從此亦可觀察出立法者將「足以使人混淆」視作本規定處罰的核心正當性基礎。

第九項 比較法整理

自前述已完成立法的法典，亦或是尚未完成立法的草案中，本文得以歸結出幾項於設計構成要件時應該考量的因素，包含了 1.主觀上是否要求行為人具備特殊意圖？是否因客觀上所為行為之不同而有相異的主觀要件？、2.客觀上所欲規範之行為係製作、散布、或其他行為？、3.客觀上所欲規範之客體應以如何的法律文字進行描述？是否要求須具備相當的欺詐性？、4.客觀上所欲規範之客體是否應限定於特定種類之主題，如與性相關、妨礙選舉公正等內容？、5.若所欲規範之客體成立以未得出演者同意為前提，該同意是否需有特定的形式要件？、6.若行為人於影音中主動加註警語、免責聲明、浮水印，會如何影響行為之定性？法規中是否應對前述行為有明確的形式上之要求？、7.有無明文豁免條款等主要差異。本文以下將簡要將各版本之立場加以整理並介紹之。

第一款 主觀要件

觀諸前開立法例與草案，對於主觀要件之設計主要有幾種類型：

第一目 用以限縮所規範之不實影像的種類範圍

a.例如德州選舉法中要求行為人須意圖傷害候選人或影響選舉結果，亦或 DEEP FAKES 草案中，有部分條文以「意圖造成暴力或人身傷害，煽動武裝或外交衝突，或干涉包涵選舉在內之公務」，皆屬於透過主觀意圖要件之設計，將所規範之不實影像種類加以限縮。

b.又如加州選舉法規定於「意圖傷害候選人之聲譽或欺騙選民投票支持或反對候選人」時，行為人始會被追究責任。

c.再者像麻州草案中規定，僅有「意圖透過散布促成刑事犯罪或民事侵權行為」之製作或散布的行為，始有成立不法性。

第二目 對於不實性影像之使用，用以限縮成立不法之範圍

例如維吉尼亞州刑法要求行為人需「意圖脅迫、騷擾、或恐嚇」始能成立犯罪。本規範特別強調需具備指定的主觀意圖，將導致單純分享或營利而散布不實性影像之行為，無法落入本法射程範圍內。

第三目 對於不實性影像之使用，用以調整法律效果

例如加州民法以及紐約州民法皆規定，「『惡意』係指被告故意對原告造成傷害（cause harm to the plaintiff）或做出卑鄙行為（despicable conduct），故意且明知地無視原告的權利。」而成立「惡意」的法律效果落在(e)款，得以以此提高法定損害賠償（statutory damages）之額度。

反面解釋可知，此立法例下並不因行為人欠缺特定主觀要件即認為未成立不法，而僅係作為確立不法後法律效果嚴重程度的依據。

第四目 調整不同種類行為人間，成立主觀不法之要件

例如加州民法以及紐約州民法皆規定，同樣的散布不實性影像之行為，會因為行為人是製作者，亦或是再輾轉之散布者，而有相異的主觀要件要求。

對於實際製作者而言，除知悉未得同意外，僅需可得而知未得被描繪之個體的同意而散布，即成立法律上責任；

反之，非實際製作者在散布時，僅有知悉未得被描繪之個體的同意而散布，始成立本條文下之法律責任。

第二款 所規範之客體之欺詐性

觀諸前開立法例與草案，對於所規範之客體的欺詐性要求，主要有幾種設計類型：

第一目 自製作時之主觀意圖加以限縮

例如德州選舉法規定「本節中所稱的『deepfake 影像』係指帶有欺騙性意圖所製作，而宛若描繪了一個真實人物從事未曾實際發生行為之影像。」，則若該客體並非「帶有欺騙性意圖所製作」，即不屬於該條文所規範之客體。

綜合上述文字，立法者似僅要求行為人「主觀上」有意欺騙即可，而不論客觀上是否具相當欺詐性。

第二目 自對於客體之形容詞描繪加以限縮

a.例如加州民法與紐約州民法中，透過「被描繪之個體」、「受變造的描繪」、「數位化」等詞交互解釋，去建構所欲規範之客體範圍。

自其描述的文字中，可以確定的部分僅有該影像需呈現出被描繪之客體與性相關之「未曾真實做出」或「雖曾做出但經變造」之表現，惟對於欺詐性之要求，似僅能透過「逼真地描繪」（realistically depict）一詞加以限縮，以其不足夠「逼真」（realistically）為由來排除欺騙性不足的影像作品落入本條規範之射程範圍內。

b.加州選舉法規定，本部法規需具備實質欺騙性（materially deceptive），而所謂實質欺騙性必須「使理性一般人誤以為真」且「使理性一般人因為該客體而產生錯誤的理解與印象」，因此若欠缺足以使理性一般人誤會之欺詐性則無本法之適用；

c.而麻州草案規定「deep fake 係指以某種方式創造或變更，而會被理性的觀

察者 (a reasonable observer) 錯誤地認為是個人實際言論或行為的影音記錄。」

針對所欲規範之客體，本法明確描述需使「理性的觀察者」誤以為真，始屬於本法射程範圍內所欲處理的對象。

準此，則無論是技術上過於拙劣導致欺騙性不足，亦或是額外附加警語使理性的觀察者不至於誤以為真時，皆無本規範之適用。

d.此外，DEEP FAKES 草案規定「以先進科技偽造之身分紀錄係指任何 deep fake 同時符合下列條件者：(A) 理性之人在考慮了記錄的影像或音檔品質以及發行渠道的性質後，會相信該紀錄準確地呈現了—(i)作品中的在世人士實際上並未從事的任何物質活動，或(ii)作品中的已逝世人士實際上並未從事的任何物質活動，而這些活動的展演很可能會助長犯罪行為或導致對公務、公共政策辯論或選舉的不當干預。」

綜合所有描述規範客體條件的交集可知，本部法規所欲規範者其實是「具備使理性之人足以誤會之欺詐性，且描述之內容將導致傷害之虞的，未得出演者或其繼承人同意，卻描繪出演者從事其未曾真正做出之言行」的影音。

此外，該草案尚規定「(1) 揭露—(c)、(d)、(e) 款所規範之要求，不適用於以下種類之以先進科技偽造之身分紀錄：(A) 含有替代性的揭露措施，依照個案情節觀之足使理性之人 (a reasonable person) 認為這些揭露相較 (c)、(d)或(e) 款所要求的揭露效果更為顯著；(C) 主要包含真實人物（例如表演藝術家）的圖像或錄音，而未經過大幅度的數位修改；(E) 依照其情境，理性之人不會將偽造的物質活動誤認為是其中出演者所為真正的物質活動，例如模仿表演或出版物、歷史重演或虛構的廣播、電視或電影節目。」

依照其明文例示之豁免規定亦可得知，「以其他替代措施達到不使人誤會之效果」、「本質上不足以使理性之人誤會」等客體，皆非本條文所欲規範的對象。

第三款 所規範之客體之種類限制

觀諸前開立法例與草案，對於所規範之客體之種類限制，主要有以兩個維度之區別，且兩個維度並不彼此影響相容性：

第一目 載體形式之區別

a.有認為僅包含影像 (video) 者或影音作品者 (audiovisual work)，前者如德州選舉法，後者如加州民法、紐約州民法與麻州草案等。

b.有認為包含影像 (video) 與靜止之圖像 (still image) 者，如維吉尼亞州刑法。

c.有認為同時包含了圖像、音檔或影像 (an image or an audio or video recording)，如加州選舉法；同時包含了影音、影像、音檔 (audiovisual record、visual record、audio record)，如 DEEP FAKES 草案。

第二目 載體主題之區別

- a.有限縮在選舉相關者，如德州選舉法、加州選舉法。
- b.有限縮在不實性影像者，如加州民法、維吉尼亞州刑法、紐約州民法。
- c.有限縮在一切「促成刑事犯罪或民事侵權行為」的散布行為，如麻州草案。
- d.有同時欲在同一部法規中處理「性相關視覺內容」、「意圖引發暴力或人身傷害，煽動武裝或外交衝突，或干涉包涵選舉在內之公務」、「涉及詐欺相關的刑事犯罪，包含證券詐欺、電信詐欺、虛偽冒充或身份盜用」、「外國勢力或其代理人，意圖影響國內公共政策辯論，干涉聯邦、州、地方或地區選舉，或從事該勢力不得合法從事的其他行為」者，如 DEEP FAKES 草案。

第四款 客觀上構成要件行為

觀諸前開立法例與草案，對於客觀上構成要件之設計主要有幾種類型：

第一目 僅製作階段即可能成立不法

例如麻州草案規定，「任何人(1)意圖散布而製作，且意圖透過散布促成刑事犯罪或民事侵權行為而製作 deep fake」即可成立刑事不法。

在此立法例下，倘若指定的主觀意圖，則僅在製作階段之行為就可被認為具不法性。

第二目 需達揭露或散布階段始成立不法

a.例如德州選舉法規定，行為人需「製作 deepfake 影像」，且「使該 deepfake 影像在選舉日的 30 日內被公開或傳播。」始成立不法；加州選舉法規定，行為人「在候選人將出現在選票上的選舉日之 60 日內，不得意圖傷害候選人之聲譽或欺騙選民投票支持或反對候選人，以真實惡意散布(e)條款所規範之，具實質上欺騙性的聽覺或視覺素材。」

b.加州民法與紐約州民法規定，行為人需「製作且故意揭露」或「故意揭露非自己製作」之具有性意涵之素材，始成立不法；

c.維吉尼亞州刑法規定，行為人需「惡意散布或販賣」性影像，始成立犯罪。

第三目 除揭露或散布外，尚需符合其他構成要件始成立不法

a.例如加州選舉法規定之構成要件行為為：「除非符合(b)節之規定，……不得意圖傷害候選人之聲譽或欺騙選民投票支持或反對候選人，以真實惡意散布(e)條款所規範之，具實質上欺騙性的聽覺或視覺素材。」，而(b)節則規定：「(b)任何人、公司、協會、法人、競選委員會或組織，僅有在於競選素材所包含的每張照片或圖像旁，以與競選素材其他部分所使用的最大字體同等大小之字體，註明『本照片所呈現的並非準確的真相。』之警語，且該警語應緊鄰任何被(a)條款所禁止之照片或圖像時，始得生產、散布、公開或傳播包含(a)條款所禁止之照片或圖像的競選素材。」

b.而 DEEP FAKES 草案則規定，其主要規制之行為係「(A) 未能揭露」或

「(B) 變更揭露內容」，亦即未加註條文要求之浮水印與警語，或移除他人加註之浮水印與警語，而「為虛假展示」等行為。

在此立法例下，規範不法性的重點不僅是揭露或散布，而係不符合指定規範之揭露或散布。

第五款 警語、免責聲明或浮水印之定性與效果

對此，前開立法例與草案呈現極端對立的狀況，有部分認為若確實加註則可阻卻不法，然而另外亦有部份認為即便加註亦無法阻卻不法：

第一目 明文規定加註後得以阻卻不法

a.如加州選舉法規定，「除非符合(b)條款之規定，任何個人、依照 Government Code 第 82013 條 所規範之委員會，或其他實體，在候選人將出現在選票上的選舉日之 60 日內，不得意圖傷害候選人之聲譽或欺騙選民投票支持或反對候選人，以真實惡意散布(e)條款所規範之，具實質上欺騙性的聽覺或視覺素材。³⁴⁶」、「(b) 任何人、公司、協會、法人、競選委員會或組織，僅有在於競選素材所包含的每張照片或圖像旁，以與競選素材其他部分所使用的最大字體同等大小之字體，註明『本照片所呈現的並非準確的真相。』之警語，且該警語應緊鄰任何被(a)條款所禁止之照片或圖像時，始得生產、散布、公開或傳播包含(a)條款所禁止之照片或圖像的競選素材。³⁴⁷」

b.又如 DEEP FAKES 草案規定，影音、影像、語音須分別符合對應之浮水印要求 (watermark requirement) 與揭露要求 (disclosure requirements)。且對於該浮水印與揭露形式上應如何為之，本草案亦有明確規範。以影音作為舉例，其即指出：「(c) 影音揭露—任何同時包含聲音或影像元素之以先進技術偽造之身分紀錄，應包含：(1)至少 1 則清晰的口頭聲明，辨明該紀錄包含經編輯過之影音元素，並簡要的敘述編輯之程度；且(2)在視覺元素的持續期間，於圖像底部以清晰可讀的文字呈現的未受遮蔽之書面聲明，辨明該紀錄包含經編輯過之影音元素，並簡要的敘述編輯之程度。」

而針對純粹的音檔，其亦指出：「(e) 語音揭露—任何只包含聲音元素之以先進技術偽造之身分紀錄，應包含在該紀錄開始處，以清晰的口頭聲明，辨明該紀錄包含經編輯過之聲音元素，並簡要的敘述編輯之程度；且若該紀錄長度超過

³⁴⁶ Except as provided in subdivision (b), a person, committee, as defined in Section 82013 of the Government Code, or other entity shall not, within 60 days of an election at which a candidate for elective office will appear on the ballot, distribute, with actual malice, materially deceptive audio or visual media, as defined in subdivision (e), of the candidate with the intent to injure the candidate's reputation or to deceive a voter into voting for or against the candidate.

³⁴⁷ A person, firm, association, corporation, campaign committee, or organization may produce, distribute, publish, or broadcast campaign material that contains a picture or photograph prohibited by subdivision (a) only if each picture or photograph in the campaign material includes the following statement in the same point size type as the largest point size type used elsewhere in the campaign material: "This picture is not an accurate representation of fact." The statement shall be immediately adjacent to each picture or photograph prohibited by subdivision (a).

2 分鐘，應額外於隨後的每 2 分鐘間隔處，增加至少 1 則語音聲明，辨明該紀錄包含經編輯過之聲音元素，並額外簡要的敘述編輯之程度。」

第二目 明文規定即便加註亦不能阻卻不法

如加州民法與紐約州民法皆規定：「(d)行為人於因本節規定而涉及之訴訟中，即便已在該具有性意涵之素材中註明免責聲明，表明該被描繪之個體於該素材之參與未經授權，或表明該被描繪之個體並未參與該素材的製作與開發，仍不得執此作為訴訟上之抗辯。」

此規範可能使一些已經透過文字說明，使觀眾欠缺誤以為真潛力的影音作品，仍落入本法規所欲規制的射程範圍內。從此當中或可觀察出立法者並不在意誤認之可能性。

第六款 明文之豁免條款

對此，有部分法規明文規範了豁免例外，使特定種類之不實影像不屬於其欲規範之客體：

(1) 如加州選舉法規定：「本節法規不適用於構成諷刺或諧擬的具實質欺騙性之聽覺或視覺素材。³⁴⁸」

(2) 又如 DEEP FAKES 草案規定：「(A) 含有替代性的揭露措施，依照個案情節觀之足使理性之人 (a reasonable person) 認為這些揭露相較 (c)、(d) 或 (e) 款所要求的揭露效果更為顯著」、「(B) 以最終散布之記錄符合規範為前提，處於生產過程中之中間產物」、「(C) 主要包含真實人物 (例如表演藝術家) 的圖像或錄音，而未經過大幅度的數位修改」、「(D) 基於電影、電視、音樂或類似作品或創作之衍生作品有關的改編創作，其原始內容是在本法頒布之前所完成，且其中被描繪之人於起初時同意出演者」、「(E) 依照其情境，理性之人不會將偽造的物質活動誤認為是其中出演者所為真正的物質活動，例如模仿表演或出版物、歷史重演或虛構的廣播、電視或電影節目」、「(F) 由美國官員或僱員或在其授權下為促進公共安全或國家安全而製作」

綜上所述，本文介紹、整理並略作分析了關於美國目前已經完成之立法例，並介紹了部分的修法草案。本文將以此為基礎，評析我國於 2023 年之修法，並給出本文對其之修法建議。

第二節 2023 年修法評析與建議

於 2023 年，我國完成了針對不實影像之使用，完成了兩次刑法的修法。分

³⁴⁸ This section does not apply to materially deceptive audio or visual media that constitutes satire or parody.

別是 1 月時公布新增刑法第 319 條之 4，處罰製作與散布不實性影像之行為，以及 5 月時三讀通過新增刑法第 339 條之 4 第 4 款規定，將利用不實影像、聲音、電磁紀錄犯詐欺罪之行為，列入加重詐欺罪之處罰事由。

本文將以前一節整理出的，數種比較法上立法設計中有所歧異之考量做為尺規，並輔以第三章中對於保護法益之論述，逐一審視並分析我國立法之優劣，同時給予具體之修法上建議。

第一項 刑法第 319 條之 4 與其相關規定

2023 年 1 月我國立法院三讀通過刑法修正案，其中包含之第 319 條之 4（本項下稱「本條文」）涉及了對於不實性影像之規範。其條文規定：

「意圖散布、播送、交付、公然陳列，或以他法供人觀覽，以電腦合成或其他科技方法製作關於他人不實之性影像，足以生損害於他人者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。

散布、播送、交付、公然陳列，或以他法供人觀覽前項性影像，足以生損害於他人者，亦同。

意圖營利而犯前二項之罪者，處七年以上有期徒刑，得併科七十萬元以下罰金。販賣前二項性影像者，亦同。」

而所謂性影像者，於刑法第 10 條第 8 項新增規定：

「稱性影像者，謂內容有下列各款之一之影像或電磁紀錄：

- 一、第五項第一款或第二款之行為。
- 二、性器或客觀上足以引起性慾或羞恥之身體隱私部位。
- 三、以身體或器物接觸前款部位，而客觀上足以引起性慾或羞恥之行為。
- 四、其他與性相關而客觀上足以引起性慾或羞恥之行為。」

本文分析如下：

第一款 主觀要件

如前所述，在各立法例中可資參考的主觀要件設計包含下列四種：a.用以限縮所規範之不實影像的種類範圍、b.在不實性影像之使用，用以限縮成立不法之範圍、c.在不實性影像之使用，用以調整法律效果、d.調整不同種類行為人間，成立主觀不法之要件。

在本條文中，業已透過客體上之限制將種類限定於與性相關，故 a.之部分之立法體例無須考慮。而其他之立法模式是否有我國得以參照之處，以下析論之：

第一目 以「脅迫、騷擾或恐嚇」作為主觀意圖要件

b.立法模式之內涵在於認為若行為人不具備「脅迫、騷擾或恐嚇」等意圖，

則行為即欠缺不法性。

我國於本條文立法理由中明文主張所參考之「維吉尼亞州刑法道德與風化罪章第十八點二之三八六點二條規定」（即本文前述之維吉尼亞州刑法）即是採取此立法體例，惟本條文最終於立法時卻選擇捨棄主觀意圖之要求，而認為即便不具備此般意圖之製作或散布行為，仍有不法性值得加以刑事處罰。

對於前述行為是否應以主觀意圖要件加以限縮不法成立的範圍，我國過往有類似之討論者多集中在涉及真實性影像之復仇式色情立法討論上。有傾向主張應以主觀意圖要件加以限縮者，如「未另外要求行為人除了一般故意外，尚需具備有其他不法意圖（例如基於傷害被害人、使被害人痛苦或騷擾被害人等之意圖）。……與他國之相關立法相較，似乎過於寬鬆，而容易將許多非基於惡意或報復之散布或傳送影像之行為納入處罰之對象。」³⁴⁹、「未限定行為人應具有傷害或騷擾之故意，且被害人不須確實受有傷害或損害，是否將過於擴大本法之適用，侵害言論自由、新聞自由而有違憲之虞，尚屬可議。」³⁵⁰等。惟比較法上亦有指出「理想的法規設計應禁止於網路上發布虛擬復仇式色情（按：即不實性影像）內容，並且無須要求具備傷害受害者的特定意圖。」³⁵¹。

關於此問題，本文認為應該回歸保護法益的觀點加以審視，直言之，究竟會是否僅因為行為人欠缺這些主觀意圖，即會使被害人的法益不受侵害，或欠缺被保護的必要性？對此，本文持反對見解，蓋如本文前述，此行為所侵害的法益應是形象法益，核心在於使被害人特別私密的面貌被錯誤的呈現在公眾下，且被害人對此在澄清上會有所困難。考量到這些保護法益的內涵，皆無涉於行為人的主觀意圖，縱使其全然欠缺「脅迫、騷擾或恐嚇」被害人的意圖，僅是為了自娛或商業營利而散布不實性影像，仍足以使被害人的形象法益受損。準此，本文肯定本部分之刑法修正，認為成立刑事不法無需以具備特定主觀意圖為必要。

第二目 以「脅迫、騷擾或恐嚇」作為加重處罰之要件

c.立法模式之內涵在於，其認為無論是否具備特定主觀意圖皆足以成立不法性，然而在具備前開主觀意圖時加重行為人應該負擔之法律責任。而本條文在設計上並未要求有此般主觀意圖要件。

本文認為在思考上述問題時，首應注意的是若帶有「脅迫、騷擾或恐嚇」之意圖而為散布不實性影像之行為，本身即可能成立犯罪，如刑法第 305 條之恐嚇罪，亦或是跟蹤騷擾防治法第 18 條、第 19 條之罪。在此情形下，屬於一行為侵害到形象法益與其他法益，理應依照想像競合論處。若是以此意圖作為散布不實影像罪之加重處罰事由，某種程度上即屬於實質上之結合犯。然而，結合犯之立法正當性於學說上本即遭受批評。學說即有認為，此類犯罪類型係特別排除刑法

³⁴⁹ 法思齊（2021），前揭註 183，頁 221。

³⁵⁰ 蕭郁唐（2017），〈日本情色報復受害防治法之評析〉，《科技法律透析》，第 29 卷第 2 期，頁 27。

³⁵¹ ŠEPEC, LANGO (2020), Supra note 147, 128.

總則競合規則之適用，且往往是提高應有的法定刑。然而，究竟此提高法定刑的正當性基礎為何？何以選擇性的挑選特定種類犯罪？往往欠缺正當理由。支持者往往將正當性訴諸於同時侵害數法益「惡性更重」，然而個別的惡性在各自的法定刑中皆業已經過評價，且兩犯罪的惡性評價本來就是競合論欲處理的問題，因此以「惡性更重」作為正當化結合犯之基礎顯無理由。而此悖於刑法總則競合論的選擇性立法，本身即是違反平等原則與罪責原則³⁵²。

本文完全同意前述見解，蓋以與侵害形象法益近似之妨害名譽罪章為例，也並未因行為人具備「脅迫、騷擾或恐嚇」之意圖而作為加重法定刑之理由。思考上之所以容易將前述種類的主觀意圖與散布不實性影像做連結，本質上可能僅是因為其伴隨而生的頻率較高，是兩種容易同時成立的犯罪類型。然而兩種犯罪類型容易伴隨相生，也不足以作為正當化將其作為結合犯加重法定刑之理由。例如行使偽造文書與詐欺罪亦在實務上時常伴隨而生，如出示假的買賣契約誘使人交付財物等，然而實然上沒有、應然上也不該因此就「意圖供詐欺之用」作為加重行使偽造文書罪之加重事由。準此，本文亦認為不應因行為人具有「脅迫、騷擾或恐嚇」等主觀意圖要件而調整法定刑。

第三目 針對不同種類行為人，設計相異之主觀不法要件

d.立法模式之內涵在於，其對於「製作不實性影像而散布之行為人」與「非製作不實影像之散布行為人」規範有不同的主觀不法構成要件。對於前者，「知悉」或「可得而知」(the person knows or reasonably should have known)未得被描繪之個體同意而散布皆可成立主觀不法，惟對於後者，僅有「知悉」(knows)未得被描繪之個體同意而散布始可成立主觀不法。而本條文在設計上，並未對於兩種行為人間之主觀不法要件有歧異之規定，所有的散布行為皆以故意为限始成立主觀上不法而得以構成本條文之犯罪。

對此，本文認為此處所稱之「知悉」，得以對應到我國刑法下所稱之故意應較無疑義。而相較於我國刑法第13條第2項、第14條第2項分別規定之間接故意、有認識之過失，本質上仍需對結果之發生「實際上有預見」，前述比較法所謂「可得而知」與之相較，使行為人縱使並未實際上有預見仍得成立主觀不法，其主觀要件應屬更為寬鬆，在光譜上可能更接近我國刑法第14條第1項所稱之過失。如此一來，則對於「非製作不實影像之散布行為人」限於主觀上須具備故意始成立不法並無疑義，剩下必須討論的則是，是否有必要使「製作不實性影像而散布之行為人」有著更低的主觀不法成立門檻，令其僅需具備類似過失之主觀不法程度即成立犯罪。

本文對此持否定看法，蓋觀察我國刑法體例中與此類型行為侵害法益、程度相似之妨害名譽罪章，亦未特別設有處罰過失犯之規定。又，本文亦難想像散布不實影像之行為有「容易出現過失未得同意之狀況」，以至於必須擴大處罰範圍

³⁵² 黃榮堅(2012)，《基礎刑法學(下)》，4版，頁891-893，台北：元照。

加以預防之必要。綜上，本文認為本條文無須以「脅迫、騷擾或恐嚇」之意圖作為犯罪成立與否，或罪刑加重之事由。並且，亦無需令「製作不實性影像而散布之行為人」有著更低的主觀不法成立門檻，以故意作為其主觀不法之要件較為妥適。準此，則本文認為本條文於主觀要件之設計上並無必須更動之處。

第二款 所規範之客體之欺詐性

如前所述，在各立法例中可資參考的客體之欺詐性設計包含下列兩種：a.自製作時之主觀意圖加以限縮、b.自對於客體之形容詞描繪加以限縮。前列立法模式是否有我國得以參照之處，以下析論之：

第一目 自製作時之主觀意圖加以限縮，不足以排除無混淆之虞的散布不實性影像行為成立犯罪

a.立法模式之內涵在於，透過行為人在製作不實影像時須帶有特殊主觀意圖作為限縮。例如德州選舉法即以「帶有欺騙性意圖所製作」(created with the intent to deceive)形容所欲規範之客體，限縮了法規得以適用之範圍。而本條文對於客體之規範並未有如此之設計。

而如本文前述，本文主張僅有存在使人混淆之虞之不實性影像，始有侵害法益之可能，進而存在對其加以刑事上規範之正當性。準此，則得否於構成要件中排除無混淆之虞之散布行為即至關重要。然而，本文認為以 a.立法模式，無法充分達到前述目標。一來單從邏輯論之，事實上，是否會導致形象法益受侵害之關鍵，並非在於行為人製作時之主觀意圖，而係在於客觀上之混淆潛能。即便行為人主觀上有意欺騙，其客觀上仍可能基於技術上缺陷等理由，使其製作出之作品自始欠缺足以使人混淆的能量。反之，行為人即便主觀上並無欺騙之意圖，只是一時失察忘記加上本欲為之的浮水印，此時該作品仍可能存在使人混淆之虞。二來自實務運作上論之，由於法官無法站在上帝視角去探明行為人製作時之主觀意圖，因此勢必需要透過客觀的事證輔助推論，進而「有可能」使法官因為該不實性影像欠缺足以使人混淆之潛能，反推得出行為人欠缺欺騙之主觀意圖的結論。然而此僅是歪打正著的結果，將正確適用法律之可能，僅繫於射倖成分，且於個案中給予法官過度模糊裁量之空間，難謂適當。

準此，本文認為是否有混淆之虞自本質上就與行為人主觀意圖並無必然關係，自製作時之主觀意圖加以限縮，並不足以排除無混淆之虞的散布不實性影像行為成立犯罪，因此本文認為不宜採取此立法模式。

第二目 對於客體之描述

b.立法模式之內涵在於，透過形容詞去描繪所欲規範之客體，限定法規之適用範圍，進而排除無混淆之虞的散布不實性影像行為成立犯罪。而前述立法模式中所使用之文字（「足以使...」），很大程度上可以對應到我國立法體系中適性犯之相關規定，故本文亦將適性犯之相關規定較廣義地納入此環節進行討論。

本條文對於客體之規範，係以主要「不實」一詞，用以描繪所欲規範之客體即「性影像」。惟本文認為本條文此般之規範，並不足以排除無混淆之虞的散布不實性影像行為成立犯罪。蓋就文義解釋上，本文前述舉例的各種「雖經過 deepfake 技術編輯，然而因為技術拙劣、經明確註記係偽作，或依使用情境使公眾不會混淆影像中被描繪之人確實從事影像中之言行」性影像，皆仍可以落入所謂「不實」性影像之範疇。又，按照歷史解釋探尋立法時歷史背景資料，亦如本文前述本次修法提出草案之主責機關即法務部，於黨團協商階段明確表示即便僅是「影像標題表明係 A 所出演，惟影像內容係與 A 完全不相似之 B 出演」的狀況，對其而言亦屬於其所認定的「對於 A 之不實性影像」³⁵³。準此亦可知，循此解釋方法亦仍無從排除無混淆之虞的散布不實性影像行為成立犯罪。

此外，本條文規範有「足以生損害於他人者」之適性犯規範，「理論上」可以提供司法部門藉此排除掉無混淆之虞的散布不實性影像行為成立犯罪。惟所謂「足以生損害於他人者」所指稱之「損害」為何，其實仍具備高度不確定性，而並未被限縮在「形象可能受損害」此一可能性。舉例而言，即便是無混淆之虞之不實性影像，仍可能對於影像中被描繪之人產生「未經同意使用其肖像之民事上權利侵害」等。同時，此處所稱之「損害」「理論上」可以藉由立法理由、歷史解釋，去輔助司法部門將其限縮，然而如前所述，無論是自立法理由或立法歷史背景資料觀之，都無從得出立法者有限縮無混淆之虞的散布不實性影像行為成立犯罪之意圖，自亦無從期待能藉此對司法部門加以限縮。故，本文認為仍有參考比較法上立法模式，進一步調整構成要件文字之必要。

具體而言，立法例上有透過「實質欺騙性」(materially deceptive)一詞來形容所欲規範之客體，而所謂的「實質欺騙性」指的是「該圖像、音檔或影像會使理性一般人誤以為真 (would falsely appear to a reasonable person to be authentic)」且「該圖像、音檔或影像會使理性一般人，相較於看到或聽到未經變造的原始版本，對其所欲表達的內容產生根本性的不同理解或印象」³⁵⁴。亦有以「會被理性的觀察者錯誤地認為是個人實際言論或行為的影音記錄」³⁵⁵作為描繪，去對所規範之客體加以限縮。或有規定「以先進科技偽造之身分紀錄一係指任何 deep fake 同時符合下列條件者：(A) 理性之人在考慮了記錄的影像或音檔品質以及發行渠道的性質後，會相信該紀錄準確地呈現了—(i)作品中的在世人士實際上並未從事的任何物質活動，或(ii)作品中的已逝世人士實際上並未從事的任何物質活動，而這些活動的展演很可能會助長犯罪行為或導致對公務、公共政策辯論或選

³⁵³ 立法院公報第 111 卷第 84 期委員會紀錄頁 16-19。

³⁵⁴ would cause a reasonable person to have a fundamentally different understanding or impression of the expressive content of the image or audio or video recording than that person would have if the person were hearing or seeing the unaltered, original version of the image or audio or video recording.

³⁵⁵ would falsely appear to a reasonable observer to be an authentic record of the actual speech or conduct of an individual.

舉的不當干預。³⁵⁶」綜合上述立法例可知，其共同提及之元素即是強調須以「理性一般人」之觀點，而對於何謂「混淆」，應是「足使人相信，客體中被描繪之人做出與其實際所為不相符之言論或行為」。

本文認為前述之規範係屬合理，蓋將判斷標準設定為理性一般人觀點，足以強調使「公眾」混淆之潛能。又，所謂形象法益受損如本文前述，核心內涵應係「以不準確的形象暴露（be inaccurately imposed）在公眾前」等，故將「混淆」之定義明確限定在「使人相信被描繪之人做出與實際不符之言行」，應能最貼合保護法益之核心。準此，本文認為無論係以主觀意圖加以限縮，亦或是如本條文原先以「不實」此一形容詞，或適性犯之規範加以限縮，皆不足以排除無混淆之虞的散布不實性影像行為成立犯罪。本文認為本條文宜將使用之文字調整，在客體部分加上形容詞子句的後綴，強調該客體需「足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為」，藉以排除無混淆之虞的散布不實性影像行為成立犯罪。

第三款 所規範之客體之種類限制

所謂的客體之種類限制，如前所述包含了載體形式之區別與載體主題之區別兩種。惟本條文適用之主題已明確限定在性相關之影像，故本段落僅針對載體形式之區別進行討論。

如前所述，在各立法例中可資參考的載體形式之區別包含下列三種：a.

僅包含影像（video）者或影音作品者（audiovisual work）、b.同時包含影像（video）與靜止之圖像（still image）者、c.同時包含了圖像、音檔、影像（an image or an audio or video recording）或影音、影像、音檔（audiovisual record、visual record、audio record）者。前列立法模式是否有我國得以參照之處，以下析論之：

參照前述各種立法模式，皆有交集而沒有疑義者應係不實之「影像」必須被規制，而本條文亦僅以「影像」作為規範客體。惟對於純粹靜止之「圖像」，甚至是欠缺視覺元素而僅有聲音之「音檔」是否應受規制，各種立法體例上價值取

356

(A) a reasonable person, having considered the visual or audio qualities of the record and the nature of the distribution channel in which the record appears, would believe accurately exhibits—

“(i) any material activity of a living person which such living person did not in fact undertake; or

“(ii) any material activity of a deceased person which such deceased person did not in fact undertake, and the exhibition of which is substantially likely to either further a criminal act or result in improper interference in an official proceeding, public policy debate, or election

(A) a reasonable person, having considered the visual or audio qualities of the record and the nature of the distribution channel in which the record appears, would believe accurately exhibits—

“(i) any material activity of a living person which such living person did not in fact undertake; or

“(ii) any material activity of a deceased person which such deceased person did not in fact undertake, and the exhibition of which is substantially likely to either further a criminal act or result in improper interference in an official proceeding, public policy debate, or election

捨上顯然有所落差。

而我國立法者於本條文中亦無意將「圖像」或「音檔」納入規制範圍。此與我國於 2023 年 5 月三讀通過增訂刑法第 339 條之 4 第 4 款加重詐欺罪之規定，將客體描述成「不實影像、聲音或電磁紀錄」相較有明顯之區別，應是立法者有意之選擇。

第一目 應包含圖像

在圖像部分，我國於 2023 年修法時所提出之 20 個版本中，確實有部分版本認為應包含圖像，例如葉毓蘭等委員所提出之版本明文表態也欲納入規範。其指出：「為明確性影音之規定，爰增訂第八項。無論錄影、聲音、靜態圖片……等，均有可能足以辨識被害人身分，而應納入保護範圍。行政院版之對照條文，僅以『性影像』為範圍，有所不足。³⁵⁷」惟由於立法效率之考量，最終委員會決議僅將法務部版本交付院會審查，故究竟保護之客體是否應擴及圖像，並未有實際上之討論。

對於此議題，本文的思考脈絡是：在已經有所共識認為「影像」是有使人混淆、侵害形象法益之虞的載體之前提下，「圖像」在使人混淆之潛能、是否足以識別特定被害人的層面上，是否與「影像」有別？

本文必須承認，此議題上由於欠缺實證證據，故本文在判斷時僅能高度仰賴個人生命經驗去臆測、推論，故本文於此部分結論亦僅能以較保留之態度表達最終之看法。首先，本文認為在過往，「圖像」在使人混淆之潛能上可能與「影像」有別，蓋在過往已經有較多的精細修圖技術，技術也有相當的普及，因此當出現了圖像作品，而其中被描繪之人主張那是經修圖「換臉」後的假照片，公眾普遍應該是相對可以理解的。反之，由於在過往的技術環境中，較難看到全程精細的「換臉」影像作品，因此對公眾而言，當看到該影像的時候應該會比較傾向相信所見即是真實。然而在如今，由於 deepfake 技術的普及，公眾應會逐漸能夠理解即便是影像中見到被描繪之人做了特定言行，也不能 100%保證被描繪之人實際上有做出前述作為。在公眾意識到「影像」與「圖像」都同樣可能被編輯到很精緻的狀況下，其使人混淆的潛能應無明顯區別。又，由於「影像」與「圖像」皆是以視覺元素去形塑被描繪之人，故在其他條件都相同的前提下，兩種載體在是否足以識別特定被害人層面上，亦應無顯著差別。準此，本文認為並無理由對「圖像」施以差別對待，將其排除在規制客體之範圍外。

第二目 不應包含音檔

在音檔部分，我國於 2023 年修法時同樣亦有部分版本表明欲將其納入規範範圍，惟亦因立法效率之考量，最終委員會決議僅將法務部版本交付院會審查。

然而在音檔部分，於黨團協商環節中，劉建國委員曾詢問法務部長蔡清祥為

³⁵⁷ 立法院公報第 112 卷第 14 期委員會紀錄頁 190。

何法務部提出之版本並未將聲音納入，對此，蔡清祥部長之回應係：「如果大家認為聲音的部分也需要納入相關條文中，法務部並沒有一定反對，主要是因為聲音的嚴重性並非如同影像這麼嚴重。」，後續劉建國委員有持續追問：「請問這兩者之間到底要怎麼去界定影像比較嚴重、聲音比較不嚴重？」，然而法務部仍堅持原先之立場，這段討論於此就不了了之。

此部分之爭執，其實也很忠實的反映了本文在思考上的困境，即，由於欠缺實證證據的支持，因此相關討論僅能訴諸生命經驗的主觀感受，然而每個人的主觀感受本就無法互相驗證、說服，因此在爭執上很容易流於空轉。然若非得表態，本文傾向認為規範之客體不應包含音檔，理由在於本文認為音檔相較以視覺為基礎的影像、圖像，在識別特定被害人的潛力上較差，進而也會影響「使公眾誤信該被描繪之人有從事其並未真正從事言行」之可能。

直言之，本文認為倘若有人以某名人之臉譜做成不實圖像、影像，宣稱該圖像、影像中所描繪之言行即為該名人真正之言行，由於該臉譜比較能確實使公眾特定到該名人，也因此較有可能令公眾誤信。反之，倘若有人以某名人之聲音做成不實音檔，同樣宣稱該音檔中所描繪之言行即為該名人真正之言行，由於聲音並不如視覺呈現容易特定到該名人，可能比較容易使公眾懷疑只是找到聲音近似的人進行模仿，因此未必懷疑或甚至相信該名人確實有音檔中的言行。基此，則不應將本條文規範客體涵蓋進純粹之音檔。

綜上，本文認為基於「圖像」在使人混淆之潛能、是否足以識別特定被害人的層面上，皆與「影像」無明顯區別，故本條文規範客體宜涵蓋進「圖像」。惟在「音檔」部分，基於其識別特定被害人的潛力上較差，進而也會影響使人混淆之潛能，故不應將「音檔」納入規範範圍。惟本文仍需提醒，此部分之結論如前所述，很大程度僅仰賴本文基於生命經驗的臆測，故本文對此部分之結論持較保留之態度。

第四款 客觀上構成要件行為

如前所述，在各立法例中可資參考的客觀上構成要件行為設計包含下列三種：
a.僅製作階段即可能成立不法、b.需達揭露或散布階段始成立不法、c.除揭露或散布外，尚需符合其他構成要件始成立不法。前列立法模式是否有我國得以參照之處，以下析論之：

第一目 不應於製作階段即予以處罰

a.與 b.兩種立法模式間，最大的分歧即是，對於「國家機關應在風險輸出的哪個階段加以介入」有著不同的想像。而本條文之規範係認為除散布行為外，僅需達「意圖散布而製作」之階段，即應受本條文處罰。

對此，我國文獻上業已出現批評之聲浪，例如有認為「意圖散布而『製作』關於他人不實之性影像的行為，既然尚未達到散布於外的行為階段，到底可以對

於人格權造成何種的侵害³⁵⁸」，或有認為「但是既尚未散布，對於臉孔被盜用者是否真的造成剝削、恐懼及精神上傷害，乃至於具有妨害名譽之危險，值得法律在既有處罰規範外再加強介入，尚待實證支持。³⁵⁹」

對於前述問題，本文認為宜限於散布階段後始得處罰。蓋僅有到達散布階段，始開啟被害人於公眾前形象可能遭損害之因果歷程，因此意圖散布而製作之行為，僅是散布行為之實質預備犯。對於此類過度前置之實質預備犯立法，即有學者提醒在違憲審查的層次上，倘若獨立處罰過度前置且無適足引發損害可能性的預備行為，可能無法通過比例原則的審查³⁶⁰。此外亦有學者表明，必須要在要件中要求具有經驗上更高度的發生蓋然性，以及影響範圍的廣泛性，例如將行為客體限定在不特定多數人³⁶¹。按此標準，本條文所規範下之行為客體，在個案中皆應得以被特定至個別被害人上，而非影響不特定多數人，因此似無以實質預備犯之立法形式加以管制之正當性。

此外，如同妨礙名譽罪章基於是否確實導致對被害人社會評價有所貶損實難證明之，因而多數學說認為其性質係屬抽象危險犯³⁶²，本文認為本條文所欲防免之對於形象法益之侵害，亦僅是具備扭曲之潛能而無法證明實害，屬於抽象危險犯之性質。若散布不實性影像之行為屬於抽象危險犯，本身已屬於將刑罰一定程度前置化之規範類型，則意圖製作而散布之行為即是「抽象危險犯之實質預備犯」。對於此種前置又前置之立法模式，學說上多有質疑，如針對持有毒品罪，即有學者指出：「作為販賣毒品行為的前階段行為或實質預備犯之持有行為，可能會變成抽象危險犯的抽象危險犯、抽象危險犯的實質預備犯、抽象危險犯的截堵式構成要件。在學說將持有行為解釋為截堵式構成要件的見解之下，甚至會是截堵式構成要件之截堵式構成要件，產生雙重截堵的效果，從刑法謙抑性、節制刑罰的觀點來看，顯然有很大的問題。」又好比針對刑法第 362 條製作犯罪電腦程式罪，即有學者表示其屬於刑法第 358 條入侵電腦或其相關設備罪之實質預備犯，而刑法第 358 條之罪又是刑法第 359 條破壞電磁紀錄罪之實質預備犯，準此，刑法第 362 條之罪則亦形成實質預備犯之實質預備犯之結構。對此，學者認為其處罰的界線在實際上不斷地被提前之狀況，「其正當性應被慎重檢視」³⁶³。

綜上，若欲將處罰範圍觸及「意圖散布而製作不實性影像」此種實質預備之行為，本文認為一來在影響範圍的廣泛性上，其並未影響到不特定多數人。而在風險發生之蓋然性上，在其預備所犯之核心條文本身即屬抽象危險犯之前提下，實難認為其有足夠正當性去證立其將處罰時點提前至預備行為之正當性。準此，本文認為本條文應刪除第一項關於處罰意圖散布而製作行為之規定，僅保留第二

³⁵⁸ 陳俊偉（2023），前揭註 14，第 333 期，頁 37。

³⁵⁹ 許昭元（2023），前揭註 23，頁 486。

³⁶⁰ 許恒達（2022），〈交付人頭帳戶的獨立刑事制裁？—評洗錢防制法第 15 條之 1 修正草案〉，《當代法律》，第 8 期，頁 26。

³⁶¹ 謝煜偉（2016），〈風險社會中的抽象危險犯與食安管制—「攙偽假冒罪」的限定解釋〉，《月旦刑事法評論》，第 1 期，頁 81-82。

³⁶² 曾思堯（2017），前揭註 194，頁 699。

³⁶³ 徐育安（2014），〈資訊風險與刑事立法〉，《北大學法學論叢》，第 91 期，頁 26-27。

項關於散布不實性影像之規定。

第二目 除散布行為外之其他客觀構成要件

c.立法模式之內涵在於，直接透過構成要件內容之設計去調整不法成立之範圍，所處罰者並非單純之散布行為，而係「未能依照規定揭露浮水印與警語，而為散布不實作品行為」與「針對他人已散布之不實作品，移除其依照規定揭露之浮水印與警語而虛偽展示」等行為。其設計之目的應該在於，其認為充分的揭露即可以防免使人混淆之虞，進而弭平不實作品可能造成之傷害。

值得注意的是，此立法模式之案例中，兩部法規對於所欲規範的客體，分別要求必須「該圖像、音檔或影像會使理性一般人誤以為真。」與使「理性之人在考慮了記錄的影像或音檔品質以及發行渠道的性質後，會準確地相信影音中的出演者實際上並未進行的任何物質活動」，換句話說，兩個採取此立法模式之案例中，皆係從「所規範之客體之欺詐性」與「客觀上構成要件行為」兩個要件雙管齊下，避免將「無使人混淆之虞」之客體落入打擊範圍內。

本文基本同意亦心儀此立法模式。之所以希望雙管齊下的理由，係因為單從「所規範之客體之欺詐性」與單從「客觀上構成要件行為」加以限縮，各自可能面臨些許疑慮。若單從「所規範之客體之欺詐性」加以限縮，則行為人仍可能因為難以捉摸、參透司法部門對於「足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為」認定之標準，因而陷入兩難。可能在選擇積極發表的狀況下誤觸法網，或者直接選擇自我審查消極的不發表，而對言論自由、思想自由有不必要的箝制。前述「所規範之客體之欺詐性」終究是不確定法律概念，在認定與適用上必然存在曖昧而難以判斷之空間，舉例來說，可能行為人認為自己之創作情境中臉譜所貼附之我國名人，顯然與影像中主角所使用之外國語言不符，因此並無使人混淆之虞。然而此一案件或許在司法部門看來會認為，未必每個觀眾在觀看該影片的時候都會在意到腳色所使用之語言，因此仍然認為其有使人混淆的疑慮。

平心而論，本文認為以上兩種思路單就邏輯跟論理上都很難根本性的否定其正當性，然而這樣的模糊空間就可能使行為人在積極作為與消極不作為時各自承擔風險與損失。因此，本文希望透過一個相對清晰的「客觀上構成要件行為」，使創作者能夠有把握「至少自己只要做到指定行為，就一定會在法律上是安全的」，進而敢於積極創作。

前述這種提供行為人「雙保險」的立法模式，本文認為在釋字第 617 號解釋公布後的刑法第 235 條散布猥褻物品罪中，亦會出現類似的結構，即便此目的應非司法部門之本意。舉例來說，對於某出版商來說，在此號解釋公布前，縱使其認為正考量將出版之作品係藝術創作，並不足以引起普通一般人羞恥，因此並非猥褻物。然而其仍可能基於沒有把握實務法院會如何認定「猥褻」之定義，而直接選擇放棄出版。然而在此號解釋公布後，其就可以較明確的知悉，無論實務法院如何認定「猥褻」之定義，只要其出版之作品無涉於有暴力、性虐待或人獸性

交等情節，而又經適當隔絕措施之行為後，就必定不會有刑事上之責任，進而敢於出版作品。誠然在認定是否係屬硬蕊之判斷上仍然存在模糊空間，但在釋字第617號解釋提供了「適當隔絕措施」此一選項後，本文認為應係更擴大了行為人敢於表意之空間。而本文對於本條文之期待抑是希望能夠提供行為人較明確之行為指引，無論其認為自己散布之作品究竟有無「混淆之虞」，都可以僅需透過履行「合理揭露警語與浮水印」之要求，來免於使自己落入法網，進而平衡保障行為人表意自由與影像中被描繪之人之法益。

反之，若單從「客觀上構成要件行為」加以限縮亦會有所不當，蓋縱使「未合理揭露警語與浮水印」，亦不必然代表該作品有使人混淆之虞，仍可能基於作品本身技術過於拙劣等理由，而使作品本身不具備侵害法益之潛力。若僅因行為人未滿足形式上之行為規範即對其施以刑罰，則違反了刑罰保護法益之目的而與罪責原則有違。

而具體立法模式上，本文認為可以本條文之第2項規定中之構成要件行為調整成「散布、播送、交付、公然陳列，或以他法供人觀覽...，而未合理揭露警語與浮水印者」，且同時如本文前述透過對於客體之描述，排除掉不具備侵害法益之潛力的作品。亦即，即便未依規定揭露而散布，但只要該客體不具備足夠之欺詐性，該行為仍不應受本條文處罰。

至於具體細節上應如何「合理」揭露警語與浮水印，於本文所參考之加州選舉法亦是在(a)節中規定其構成要件行為，內容為「除非符合(b)節之規定」之散布行為，而在(b)節中規範警語之位置(「該警語應緊鄰任何被(a)節所禁止之照片或圖像」)、內容(「註明『本照片所呈現的並非實際的真相。』」)、字體大小(「以與競選素材其他部分所使用的最大字體同等大小之字體」)等。而 DEEP FAKES 法草案中，係將警語與浮水印之所有細節皆規範在同一條文之其他款項中，具體包含到要求聲明之數量(「至少1則口頭聲明」)、內容(「註明該紀錄包含經編輯過之影音元素，並簡要的敘述編輯之程度」)、間隔(「若該紀錄長度大於2分鐘，應額外於隨後的每2分鐘間隔處，包含至少1則口頭聲明」)、清晰程度(「以明確可讀的文字」)等。

本文較支持前述之立法模式，即在刑法規範中即細緻的規範得以阻卻構成要件之行為內容。準此，則應於刑法第10條中明確規範：「本法所稱之合理揭露警語與浮水印，謂下列措施：.....」，而措施細節應至少包含聲明之數量、內容、位置、清晰程度等。此般規定於刑法中其他涉及不實影像之條文中(例如2023年新修法之第339條之4，亦或是尚未修法之普通深偽罪等)，亦可能有適用餘地。惟此類以涉及高度細節的內容做為阻卻構成要件的立法模式，於我國刑法本文中確實較為欠缺立法先例。此等模式較多出現在行政刑法中，且時常仰賴空白刑法授權主管機關以行政命令詳加規範。例如112年憲判字第5號判決所涉之標的，即證券交易法第43條之1第3項規定：「任何人單獨或與他人共同預定取得公開發行公司已發行股份總額達一定比例者，除符合一定條件外，應採公開收購方式為之。」、第43條之1第4項後段規定：「.....前項之一定比例及條件，由

主管機關定之。」。

此處所稱之「除符合一定條件外」要件之定位，即有如本文欲在本條文中規範之「未合理揭露警語與浮水印者」。在證券交易法之規範中，選擇以空白刑法之方式授權行政機關填充規範細節，經大法官憲法判決認定並未違反法律明確性原則與授權明確性原則，且即便有少數大法官提出不同意見書，其中所主要針對者亦非此部分之規範³⁶⁴。準此，本文亦未反對選擇此種立法模式，惟在授權時應明確表明授權目的係由於不實影像之製作技術可能日新月異，為確保警語與浮水印之相關規定能夠時刻符合排除掉「無使人混淆之虞」作品落入刑法規範之意旨，故授權行政機關於此範圍內對於警語與浮水印之細節加以規定，以符合授權明確性原則之要求。

最後必須釐清的是，前述這樣的立法模式中，除最終的散布行為外，行為人「未合理揭露或移除浮水印與警語」此一要件究竟應如何定性？對此，本文認為可能性包含「客觀構成要件」與「客觀處罰條件」兩者，而結論上應以前者為宜。蓋所謂「客觀處罰條件」係指立法者對於已經具有應刑罰性的不法且有責行為，基於需刑罰性的角度所增加之犯罪成立要件。其中，具備正當性而不至於違反罪責原則者係「純正客觀處罰條件」，其必須以該條件與行為之不法內涵無關，功能上純屬限制犯罪成立之事由，始屬當之。反之，若將與行為之不法內涵有關之條件設定做為客觀處罰條件，僅為了規避證明行為人主觀上對該條件之認識，此則屬於學說上所稱之「不純正客觀處罰條件」，有違反罪責原則之虞³⁶⁵。

前者較無爭議者應如刑法第 233 條詐術結婚罪中，「以詐術締結無效或得撤銷之婚姻」本身即以足夠支撐刑事處罰之不法性，然而由於避免刑罰過度發動，而以「因而致婚姻無效之裁判或撤銷婚姻之裁判確定」作為限縮刑罰發動之客觀處罰條件；而後者則如於討論刑法第 185 條之 4 肇事逃逸罪中之「致人傷害」、「致人於死或重傷」，或是於討論刑法第 227 條與幼年男女性交罪中之「未滿十六歲」等要件時，若基於避免行為人抗辯主觀上對該要件欠缺認識而難以證明其故意，而將渠等要件解釋成客觀處罰條件時，則基於該要件本身即與不法內涵有關，因此屬於「不純正客觀處罰條件」之典型範例。依此理解，「未揭露或移除浮水印與警語」此一要件是否可以正當的定性為「客觀處罰條件」，即取決於其是否與行為之不法內涵有關。

³⁶⁴ 例如由蔡宗珍大法官提出，張瓊文大法官加入之部分不同意見書明確指出其不同意之部分係：「其所稱『預定取得』及『預定於 50 日內取得』之文句部分，不符刑罰明確性原則」；而蔡明誠大法官提出之不同意見書則指出：「此外，關於『一定條件』之用語，若干法規內容使用此用語，不乏其例。至於使用『一定比例』之用語，在法律上明定此要件者較少，……，惟於刑事處罰相關法律規定，同時使用一定比例與一定條件之用語者，並非常見。……尤其是使用雙重不確定概念，是否容許主管機關隨其管制之目的，而可不經由立法機關參與，逕為訂定所謂一定之比例與條件。」其所針對的亦是系爭法規中，將足以成立犯罪之要件（一定比例）與足以排除犯罪成立之要件（一定條件）同時授權予行政機關決定。惟若如本文所正討論者，僅將足以排除犯罪成立之要件授權予行政機關決定，其是否亦會持同樣立場即屬未知。

³⁶⁵ 薛智仁（2022），〈刑法解釋與適用之憲法界限——近期刑事判決之綜合評析〉，《興大法學》，第 32 期，頁 48。

如前所述，本文認為浮水印與警語之目的即在確保該作品無使人混淆之虞，而此本身即與保護法益與行為之不法性密切相關。直言之，倘若有行為人主觀上對於自己「未揭露浮水印與警語」欠缺認識，客觀上卻真的「未揭露或移除浮水印與警語」，例如有事實足認其確實認知到自己上傳了有加上浮水印的影像版本，然而卻因為技術失誤而上傳了尚未加上浮水印的影像初稿等，此際，行為人於主觀上並未有攻擊法益之意思，自不能將該要件強加解釋成「客觀處罰條件」而認為行為人仍然具備主觀不法，令行為人必須負擔刑事責任。

本文認為「意圖散布而製作不實性影像」係散布不實性影像之實質預備行為，在影響範圍的廣泛性及風險發生之蓋然性上，都欠缺足夠正當性去證立將處罰時點提前至預備行為之正當性，因此應刪除本條文第一項關於處罰意圖散布而製作行為之規定，僅保留第二項關於處罰散布不實性影像之規定。

而基於為了確保行為人有所依循，能夠知悉自己在滿足特定狀況下即無受刑事制裁之風險，調和表意自由與法益保護，本文認為宜參照比較法中的「雙保險」立法模式，除在「所規範之客體之欺詐性」要件中排除不具混淆潛能之作品外，於「客觀上構成要件行為」中，應將所處罰之行為限縮在「散布、播送、交付、公然陳列，或以他法供人觀覽……，而未合理揭露警語與浮水印者」。

而所謂的「合理揭露警語與浮水印」，本文認為最適當的立法模式應是在刑法第 10 條明確規範：「本法所稱之合理揭露警語與浮水印，謂下列措施：……」，而措施細節應至少包含聲明之數量、內容、位置、清晰程度等。另外若考量刑法本文並無此立法模式之先例，本文亦不排斥透過空白刑法之方式授權行政機關規範此部分細節，惟在授權時應明確表明授權目的，以符合授權明確性原則之要求。

而前述之立法模式於刑法體系之定位，基於其本身即與不法內涵有關，應被認為構成要件之一部分，而不能認為其屬於客觀處罰條件，因此行為人主觀上必須對「未合理揭露警語與浮水印」有所故意始能成立不法。

第五款 明文之豁免條款

此類立法模式在比較法上主要出現在加州選舉法與 DEEP FAKES 草案中，而此二立法例亦是惟二採前述「雙保險」立法模式之案例，直言之，其立法例上在設計了雙保險後仍然擔心某些個案中之行為人可能會被不當的入罪，因此為其設計了更多保險措施，可見其立法者在價值取捨上高度希望兼顧表意自由之決心。

而此類立法模式中，豁免條款之設計相當紛雜，以下本文將整理成幾種類型析論之：

第一目 因欠缺混淆潛能而豁免

此類型如 DEEP FAKES 草案規定中之「(A) 含有替代性的揭露措施，依照個案情節觀之足使理性之人 (a reasonable person) 認為這些揭露相較 (c)、(d) 或 (e) 款所要求的揭露效果更為顯著」、「(C) 主要包含真實人物 (例如表演藝術家) 的圖像或錄音，而未經過大幅度的數位修改」、「依照其情境，理性之人不會將偽

造的物質活動誤認為是其中出演者所為真正的物質活動，例如模仿表演或出版物、歷史重演或虛構的廣播、電視或電影節目。」等。

對於此種類型之豁免條款，本文認為較欠缺必要性，蓋其在判斷上完全可以為「所規範之客體之欺詐性」所涵蓋，進一步規範豁免條款有疊床架屋之嫌，不僅未能更適當的排除掉無混淆之虞之作品落入打擊範圍，反而存在使原先之客體判斷上陷入混亂的危險。直言之，在判斷「替代性揭露措施是否效果足夠『顯著』」、「是否經『大幅度』數位修改」、「理性之人『不會誤認』」時，其本質上之判斷標準即理應是從「所規範之客體之欺詐性」出發。在邏輯上，本文實難想像可能存在某種情況是在客體的判斷時認為具備欺詐性，卻能透過上述此種類之豁免條款排除犯罪成立者。

反而，若設計有這種豁免條款，在解釋論上可能使司法部門有採取法學方法論中反面解釋之空間，即主張「立法者若認為『依照脈絡是否會使理性人誤認』能夠成為『所規範之客體之欺詐性』之判斷依據之一，則其就沒有必要在豁免條款額外對此規定。故反面推論下可知，立法者有意排除『依照脈絡是否會使理性人誤認』能夠成為『所規範之客體之欺詐性』之判斷依據之一」的論述。從而，本文認為本條文不宜規定此種類之豁免條款。若符合前述情況，則應在「所規範之客體之欺詐性」之判斷時，認為該客體欠缺欺詐性而並非本條文所欲規範之打擊範圍。

第二目 因有正當目的而豁免

此類型如加州選舉法規定：「本節法規不適用於構成諷刺或諧擬（constitutes satire or parody）的具實質欺騙性影音。」，或 DEEP FAKES 草案規定：「(F) 由美國官員或僱員或在其授權下為促進公共安全或國家安全而製作」等。本文認為此類規範之正當性基礎在於其行為在侵害法益之餘，亦同時保障了若干利益條件下之衡量，其立法概念如同我國刑法第 311 條規定在若干條件下以善意發表言論，得以免於公然侮辱或誹謗之刑事處罰。

本文對此之立場係，凡是涉及利益衡量而阻卻行為不法性之判斷，在現行法下本就有正當防衛、緊急避難、依法令之行為等明文之阻卻違法要件可資應用，且前述種類之明文阻卻違法事由各自已經在學說與實務上發展出相對完整、穩定、具可操作性之判斷流程與基準。也因此，立法者無須也不宜在個別刑法條文中，增設以利益衡量為考量之阻卻不法事由，蓋其可能會破壞或架空適當的利益衡量之結果，徒增適用法規上的困難。

以加州選舉法之規定為例，本文能夠理解有些時候諷刺或諧擬的作品，有其針砭時事的公益性，但該公益性的本就能夠且應該依個案而認定、判斷，不可一概論之。否則難以說明，何以該作品凡是只要帶有絲毫諷刺或諧擬之性質，其所帶來之利益就必然能夠凌駕於行為侵害之法益？若真的要在特定條文中加入獨有之阻卻不法要件，也至少應該保留給司法部門足夠之彈性去做個案判斷，例如前述提及之刑法第 311 條規定中即包含了「善意」、「合法之利益」、「可受公

評之事」、「適當之評論」等不確定法律概念，可供司法部門於個案中適當地去衡量所侵害之法益與保護之利益之間的取舍。

惟本文仍然不偏好此種立法模式，蓋本文實在難以想像有任何情況是其在既有的明文化阻卻違法事由審查下，皆認為不得阻卻違法後，然而卻又令人在價值判斷上認為其實理應阻卻不法，進而有需要為其特別設立獨有之阻卻不法事由。這類為特定犯罪專設之阻卻不法事由，一來難以說明其挑選特定條文設立之標準為何，二來也可能架空既有明文化審查事由中已經形成之考量因素，反而在個案判斷中流於恣意。

綜上所述，參考比較法上立法例中之豁免條款，本文整理出兩大類型。在因欠缺混淆潛能而豁免此一類型中，本文基於其規範可能與客體欺詐性之判斷形成疊床架屋，甚至反而干擾實務上對於客體欺詐性之判斷，因此認為本條文並無增設之必要。在因有正當目的而豁免此一類型中，本文基於其正當性基礎在於利益衡量，而我國現行法下已有眾多明文化之阻卻違法事由可供操作。這些明文化之事由皆有發展較完整之體系可加以操作，足以確保個案中利益衡量之適當，因此較不傾向專為此類犯罪特別設立獨有之阻卻不法事由。

第六款 不應以「意圖營利」作為加重刑罰事由

本條文第3項規定：「意圖營利而犯前二項之罪者，處七年以下有期徒刑，得併科七十萬元以下罰金。販賣前二項性影像者，亦同。」，其立法理由則以：「至於行為人意圖營利而犯前二項之罪，或販賣前二項性影像者，惡性更重，應予加重處罰，爰為第三項規定，以符罪刑相當原則。」此般規定在比較法立法例上從未出現，為我國所獨創。

所謂的「惡性更重」的營利意圖究竟所指為何，在學理上可能是特殊的違法要素，若依此解釋，則須以後續將可能產生更高度的法益侵害危險性為前提。亦可能是指獨立的主觀責任要素，標誌著更高的主觀責任非難程度。惟對此區分我國立法者向來欠缺問題意識，自本條文之立法歷程觀之亦無法清楚之探明。惟無論是採取何種解釋定位，本文皆認為其欠缺正當性。若要認為是「特殊的違法要素」，則即必須說明「意圖營利」與法益侵害之關聯，即為何行為人帶著這樣的意圖，即當然可能使被害人形象法益被攻擊之風險增加？

對此可能的思路應係認為當帶有營利意圖時，則可能使該作品之擴散範圍更大等。惟此思路在實證經驗上可能欠缺依據，蓋真正在營利意圖下之經營者往往為了獲利而有動機將作品限縮在小範圍之私人群組³⁶⁶，惟若該作品被無營利意圖之人輾轉上傳至免費、對所有人皆公開之成人網站上，則其擴散範圍反而可能更大。若要在判決階段，以散布之範圍與程度決定量刑範圍，本文並無疑義。惟若在立法階段即在欠缺實證基礎之前提下，妄加猜測具備營利意圖即必然代表著更大的擴散範圍，因此象徵更高的不法性而需要提升刑度，本文難以支持。

³⁶⁶ 例如我國知名案件即臺灣新北地方法院 111 年度審訴字第 356 號判決中所載之犯罪事實即是如此。

而若認為此意圖僅標誌更高的主觀責任非難程度，問題則會聚焦在「何以帶著營利意圖所為之散布行為，在主觀上就更應被非難？」對於此問題，由於涉及每個個人之道德價值判斷，確實難以透過正面論述來證實或證偽，然而本文認為從一些座標來橫向比較，可能會有較清楚之面貌。如本文前述，本文基於平等原則、罪責原則，反對因「脅迫、騷擾或恐嚇」之意圖而作為加重刑罰事由，而事實上我國立法者原先在本條文之立法選擇上亦未認為帶有前述三種意圖之行為「惡性更重大」。如此一來，則此時問題就可以化約成「意圖營利之散布行為，是否比意圖脅迫、騷擾或恐嚇而散布之行為，有著更大的惡性？」本文對此的判斷是否定的，蓋基於營利意圖而散布，雖然知道「可能造成被害人傷害」卻仍執意為之，因此值得非難，惟其動機終究不是直接奔著傷害被害人而去；反之，若是意圖脅迫、騷擾或恐嚇而為之散布行為，其動機自始即是為了傷害被害人而做。兩種動機跟思維之關係，本文認為有些近似於間接故意與直接故意之關係³⁶⁷，而學說上則有認為直接故意係比起間接故意，更「情節重大」之狀況³⁶⁸。準此，在判斷「惡性重大與否」之時，若認為「意圖脅迫、騷擾或恐嚇」無須且不得作為加重事由，則「意圖營利」自亦更無加重刑責之正當性。

準此，本文認為本條文第 3 項關於「意圖營利而犯前二項之罪者，處七年以上有期徒刑，得併科七十萬元以下罰金。販賣前二項性影像者，亦同。」之規定應直接刪除。

第七款 結論

本文在參考了比較法中各立法例下，針對主觀要件、所規範之客體之欺詐性、所規範客體載體形式之區別、客觀上構成要件行為、明文之豁免條款等要素進行之立法模式後，認為本條文應修正如下：

「刑法第 319 條之 4

散布、播送、交付、公然陳列，或以他法供人觀覽不實性影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」

此外亦應於刑法第 10 條第 9 項新增規定如下：

³⁶⁷ 會說是近似，而不能全然畫上等號之理由在於，所謂「故意」之判斷應係行為人對於構成要件事實之認識與意欲，基於兩類不同意圖所為之行為，皆可能充分認識到自己正在散布不實性影像而有意使其發生，因此皆屬直接故意。惟本文仍認為基於兩類意圖所為之行為，其背後動機在「是否積極樂見被害人感受到傷害與痛苦」此一事上，仍有細微之區別，而此區別即是類似於直接故意與間接故意區別之所在。

³⁶⁸ 林慈偉（2019），〈首件適用《公政公約》第 36 號一般性意見之死刑裁判例：評臺灣彰化地方法院 107 年重訴字 第 13 號刑事判決〉，《台灣人權學刊》，第 5 卷第 2 期，頁 186。

「本法所稱之合理揭露警語與浮水印，謂下列措施：……」，而措施細節應至少包含聲明之數量、內容、位置、清晰程度等。

第二項 刑法第 339 條之 4

2023 年 5 月我國立法院三讀通過刑法修正案，其中包含之第 339 條之 4 第 4 款（本項下稱「本款」）涉及了對於利用不實影像為詐欺行為之加重刑責規範。其條文規定：

「犯第三百三十九條詐欺罪而有下列情形之一者，處一年以上七年以下有期徒刑，得併科一百萬元以下罰金：

四、以電腦合成或其他科技方法製作關於他人不實影像、聲音或電磁紀錄之方法犯之。」

本文認為本款規定應直接刪除，理由如下：

對於加重詐欺罪之法律性質，我國學說多認為本規定中之變體構成要件，係屬於刑法第 339 條詐欺罪之加重構成要件³⁶⁹。而作為加重構成要件之正當性，學說明確指出應考量罪責相當原則，而必須「找到特別規範類型中涉及不法或罪責的加重或減輕原因。³⁷⁰」此行為之所以應加重之原因，根據立法理由係因：「易使被害人陷於錯誤，且可能造成廣大民眾受騙，其侵害社會法益程度及影響層面均較普通詐欺行為嚴重，有加重處罰之必要。」其中所稱之「侵害社會法益」究竟所指為何，實令人難以理解。蓋詐欺罪本質上係屬財產法益之犯罪類型，何以利用不實影像完成詐欺行為，卻會侵害到「社會法益」，立法理由說理並不充分。若所指的是「廣大民眾受騙」之部分，則那也只說明了「許多人的財產法益」會因此受害，何以即因此會量變引發質變，進而成為「社會法益」之侵害？論理上仍屬難以理解。

又，若立法者在意的點是受害者的數量，則加重事由選擇從「詐欺手段」著手，正當性亦有高度瑕疵。蓋利用不實影像為手段行詐欺行為之人，未必皆造成許多受害者，令其必需因為「其他行為人也用相同手段詐欺了許多人」而因此必須加重刑責，根本性的違反了罪責原則的核心精神。此外，若其認為加重刑罰之正當性基礎在於「易使被害人陷於錯誤」，其正當性基礎亦相當薄弱。於加重詐欺罪前一次修法時，即有學者明確指出是否容易使被害人陷於錯誤，其不法性本身已經表現在成立既遂或是未遂的罪名之上，因此以既遂可能性作為加重處罰的理由有高度疑問。學者舉例指出，就如趁被害人離開現場竊取其財物，相較當被

³⁶⁹ 游明得（2021），〈刑法加重詐欺罪之檢視——以跨國電信詐欺案為核心〉，《開南法學》，第 12 期，頁 15。

³⁷⁰ 李聖傑（2015），〈三人以上共同犯詐欺罪的規範適用檢視〉，《月旦裁判時報》，第 39 期，頁 54-55；李聖傑（2010），〈刑法第二二二條第一項第一款「二人以上共同犯之」之適用思考〉，《政大法學評論》，第 115 期，頁 30。相似見解：謝煜偉（2022），〈2021 年刑事立法與實務發展回顧：以治安與治療為名〉，《臺大法學論叢》，第 51 卷特刊，頁 1225。

害人之面竊取，確實更容易提高既遂之可能性，惟立法者亦未對兩種竊取行為易其法定刑一般，在加重詐欺罪中僅因使被害人較高可能陷於錯誤即作為加重理由並不正當³⁷¹。此外於討論刑法第 321 條加重竊盜罪時，學者亦有類似看法，其提出質疑道：「為什麼竊盜行為人利用乘客上下交通工具前後疏於防範時的竊盜必需被加重處罰？」，並因此指出此般立法只會使行為人轉向使用其他易於得手之犯罪手段，因此欠缺刑事政策上之效益，而應該廢除該加重事由³⁷²。

對於前述學者意見本文完全持肯定看法，蓋詐欺罪作為財產法益之犯罪，其侵害之不法內涵應該以侵害財產之程度作為最重要之指標，輔以是否侵害到其他法益做為合併觀察。以不實影像為手段行詐欺行為，難以被認為有侵害其他法益已如本文前述，則此際判斷不法內涵之基準則應回歸到財產法益侵害程度本身。同樣是詐欺得手 100 萬元，若只是因為手段上使用了欺詐程度較高的手段，就因此必須比使用傳統手段進行詐欺之人有著較苛之刑事責任，本文認為毫無正當性基礎。綜上，本文認為本款規定欠缺加重刑罰之正當性，應直接刪除。

第三節 其他修法建議

如本文前述，於 2023 年刑法修正之討論階段，針對與性無關之普通深偽罪修法，事實上已經有相當程度之跨黨派共識，僅因立法效率等問題而並未一併通過。

其中較值得參考的草案有如高嘉瑜等委員提出之版本³⁷³，其認為應增訂第 15 章之 1「偽造數位圖像、影音、電磁紀錄罪」，並增訂第 220 條之 1 規定「意圖供行使之用，以電腦技術或其他科技方法，偽造、變造他人之不實照片、影像、語音或電磁紀錄，足以生損害於公眾或他人」及行使行為之處罰。

其於提案之立法理由中提及，係因為「為保障公共信用與交易秩序，刑法定有偽造貨幣、有價證券、度量衡以及偽造文書印文罪等罪章。惟……，並不包括足以影響公共信用與交易秩序，但不屬於準文書（刑§220）之列的個人圖像、影音或電磁紀錄。然圖象與影音，雖不具有準文書之性質，但於吾人的社會生活，普遍認其具有一定的公信力、對於公共信用或交易秩序具有一定的影響力。」或如時代力量黨團提案³⁷⁴，新增第 313 條之 1 規定：「利用工具或設備偽造他人之影像、聲音，足以生損害於公眾或他人者」應受刑事處罰。其於提案之立法理由中提及，係因為「用深偽技術升高了辨識資訊真假之成本，又提高了人民形成錯

³⁷¹ 蔡聖偉（2019），〈參與犯罪組織罪與加重詐欺罪的競合——評最高法院 107 年度台上字第 1066 號刑事判決〉，《月旦法學雜誌》，第 292 期，頁 184。相同意旨：游明得（2021），前揭註 368，頁 30-31。

³⁷² 黃榮堅（2012），〈2011 年刑事法發展回顧：法律說詞與說詞之外〉，《臺大法學論叢》，第 41 卷特刊，頁 1551-1552。

³⁷³ 立法院公報第 112 卷第 14 期委員會紀錄頁 86-87。

³⁷⁴ 立法院公報第 112 卷第 14 期委員會紀錄頁 110-111。

誤認知之可能性。有鑑於此一方面對公共生活之以事實為基礎之理性公開論辯活動、問責制度及共識凝聚機制構成程度不一之威脅，在另一方面，人民為基於可靠資訊規劃社會生活與互動，所仰賴之人際信任關係，亦因深偽之介入而可能遭到濫用和破壞，並對被偽造者經營之名譽和信用影響甚鉅，實有管制之必要。」

以上兩個版本之結構與立法理由之說明與本文之理解較為相近，分別強調了影像、圖像對於公共信用之影響力，以及強調了對於事實證據之檢驗之作為民主社會根基之影響力。且時代力量版本中，尚強調構成要件行為中需「且無明確標示或未以他法使閱聽者得以辨識該影像、聲音係屬偽造」與本文傾向之立法模式更是不謀而合，惟其版本將法益侵害理解為對於個人信用之個人法益侵害，與本文之理解有所歧異。而本文亦認為，基於新科技下，有混淆之虞的不實影像對於公共信用法益之侵害，亦有以刑法規制之正當性與必要性。

在條文設計上，與高嘉瑜版本相似，本文傾向以偽造文書罪章之架構為核心基礎，而將規範客體替換成不實影像。而關於「所規範之客體之欺詐性」要件部分，則於本章前述對於不實性影像之要求相同。惟本文仍認為有些許部分之修法建議與既存之偽造文書罪章有所不同，析論如下：

第一項 客觀構成要件行為，須達「行使」階段始有應罰性

於刑法偽造文書罪章中，第 210 條規定偽造文書本身即有應罰性，同時於第 216 條規定行使偽造文書亦成立犯罪。偽造文書所保護之法益在於免於「參與法律交往之當事人無法根據該文書而找到實際上應該擔保負責之人」³⁷⁵，依此，則事實上僅有在該文書已經經行使後，始開啟此「無法追索實際製作人」之風險歷程，而偽造文書之行為僅是製造危險源之預備行為³⁷⁶。

而為何有必要將處罰時點如此提前，學說上有指出³⁷⁷「如果該危險源無法為我們的生活帶來好處，或者其所帶來的好處小於其所引發的危險時，這是一種不容許的危險，此等危險源原則上便應加以禁止。」、「至於偽造文書之行為到底為我們現實生活帶來什麼樣的好處？除了研究、好玩或無聊打發時間（或練習偽造技術？）之外，我們實在想不出來有什麼好處，……，這樣的自由相對於其所帶來的危險已經超過容許的界限，是一種不容許的危險，刑法於此等危險階段即予以介入，係為了避免具體利益侵害發生，提早化解危險於無形，具有一般預防之作用，這是可以接受的。」

先不論前述論點是否足以正當化處罰偽造文書之時點，在製作不實影像的情境中，事實上存在本文前述的各種教育上、娛樂上、自我實現機能上的社會利益，絕不像是假冒他人製作名義一般，無法得出對社會之利益。也因此，前述試圖正當化處罰偽造文書時點之論述，在製作不實影像的情況中是無法適用的。而如同

³⁷⁵ 吳耀宗（2005），〈偽造文書罪保護法益之研究〉，《月旦法學雜誌》，第 128 期，頁 139。

³⁷⁶ 舉例來說，若有人偽造了知名企業家名義，寫了替自己的推薦信，在其實際對企業人資行使該推薦信之前，皆不存在「破壞公眾對文書系統之信賴」的可能。

³⁷⁷ 吳耀宗（2005），前揭註 374，頁 139-140。

本文於本章曾經提及的，對於抽象危險犯設計針對其之實質預備犯，在正當性上可能很難通過檢驗。而本文亦必須承認，在單一事件中行使不實影像，對於社會可能會因此無法信賴影像系統，進而產生對於公共信用法益之破壞，其應係屬難以證明實害之抽象危險犯。在這樣的前提下，本文認為在尚未行使的製作階段即以刑法加以規制，實有將刑罰過度擴張之虞。準此，本文認為與偽造文書罪章不同，於不實影像之使用上應待行使之行為階段始有納入刑法管制之必要性與正當性。

第二項 仍以「有使人混淆之虞」作為要件

單就條文文義上，偽造文書罪章中假冒製作名義人之行為，並未要求其假冒之欺騙性須達特定程度始足當之。直言之，即便在以相當拙劣之技術，冒名知名大學之在學證明，按一般理性之人應不會相信該文書的情形下，似仍有偽造文書相關罪名之適用。惟在實務上已有透過偽造文書罪章適性犯之規定，排除無混淆之虞的客體成立該罪名之判決。如有認為：「且所謂損害，亦不以經濟價值為限，屬抽象危險犯，祇要有令一般人誤認之危險（可能），即可成立。」³⁷⁸其將「是否有令人誤認之風險」作為「足生損害」此一要件之判斷基準。若依此，則應能排除無致人混淆之虞的不實影像作為規制對象的可能性。

而在不實影像之使用情境中，本文亦認為應堅守「有使人混淆之虞」作為要件。蓋一欠缺混淆潛能之不實影像，無論係因本身技術過度拙劣，或係有明確加上警語註記等，本身就使觀眾不會產生「影像中被描繪之人有從事特定言行」之信任，而自始沒有信任，則當然無所謂信任被破壞之問題。直言之，本文所擔心之法益侵害歷程具象化的例子會是，一個公司的人資看了某企業家推薦求職者之影像後信以為真，認為該影片中所描繪之人真的做了特定言行，事後卻發現了該影片其實是偽作。此後，其可能皆因此不敢再相信一切影像，進而使行使真正影像之求職者也無法透過此一系統作為證明。反之，若自始就沒有誤信，即不產生後續信用系統被破壞之因果歷程。準此，本文認為於不實影像之行使行為方面，仍應以「有使人混淆之虞」作為要件，始能加以刑事處罰。

第三項 「足以生損害於公眾或他人者」之認定

於偽造文書罪章中，於「文書」一詞之定義學說上即認為應包含「證明性」，即其足以證明「法律關係或社會活動中之重要事實」。而在不實影像之使用上，由於在客體之定義上欠缺此一限制，在適用上似可能涵蓋進與「法律關係或社會活動中之重要事實」無關之客體，使之亦納入刑法打擊範圍。

惟本文認為與「法律關係或社會活動中之重要事實」無關之客體，其破壞公共信用之潛能本質上較差。舉例而言，在本文前述所舉之例中，之所以該公司人資會對影像系統失去信任，一定程度即是基於其該影像係為證明重要事實之用。

³⁷⁸ 最高法院 107 年度台上字第 2360 號刑事判決。

反之，若有學生於群組中上傳某我國藝人滑倒之搞笑影片，事後才得知該影片是 deepfake 所為之不實影像，實際影像中之人只是外國的路人。此際對於看過該不實影片之人來說，確實仍然會在心中留下「看似真實之影像仍可能是虛假的」之印象，然而此對於其對於影像系統信賴感的破壞程度應該相對有限，主要是因為這樣情境下的「被騙」無足輕重完全可以一笑置之。準此，本文認為與「法律關係或社會活動中之重要事實」無關之客體不應在規範之打擊範圍之內，而在法規操作上得以排除此類客體之方法即是透過「足以生損害於公眾或他人者」此一要件之解釋。事實上，在偽造文書罪章中之「足以生損害於公眾或他人者」要件，實務之有力見解向來認為係以該文書內容是否真實，作為是否「足以生損害」之判斷基準³⁷⁹，惟此標準於不實影像之使用中顯然不適用，蓋其本就必然屬於內容不實之影像，故本文認為以「足以生損害於公眾或他人者」要件去篩選掉與「法律關係或社會活動中之重要事實」無關之客體應屬適當。

第四項 結論

綜上所述，本文參考了比較法中之立法例與偽造文書罪章之架構，並針對不實影像之特質做出些許調整後，認為應新增條文如下：

「刑法第 220 條之 1

行使不實影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為，而足以生損害於公眾或他人者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」

前述條文中所稱之「合理揭露警語與浮水印」，具體細節如本章前述，宜於刑法第 10 條中增設定義性規範。

第四節 總結

第一項 針對刑法第 319 條之 4 與其相關規定

本文認為應刪除原條文中第一項關於意圖散布而製作之規定，以及第三項關於意圖營利之加重規定。

而原條文應修正如下：

「刑法第 319 條之 4

散布、播送、交付、公然陳列，或以他法供人觀覽不實性影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相

³⁷⁹ 最高法院 30 年上字第 465 號判決。

符之言論或行為者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」

此外亦應於刑法第 10 條第 9 項新增規定如下：

「本法所稱之合理揭露警語與浮水印，謂下列措施：……」，而措施細節應至少包含聲明之數量、內容、位置、清晰程度等。

第二項 針對刑法第 339 條之 4 第 4 款規定

本文認為本款規定應全文刪除。

第三項 新增刑法第 220 條之 1 規定

本文認為應新增對於與性無關的不實影像濫用之刑事規範如下：

「刑法第 220 條之 1

行使不實影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為，而足以生損害於公眾或他人者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」



第五章 結論

第一節 研究成果

近年來由於 deepfake 技術之興起，高度擬真之不實影像遂進入了大眾的視野，進而對於社會既存之秩序與認知系統帶來了莫大的衝擊。為了因應此般衝擊，我國政府機關看似果決的做出了修法的應對措施，然而在倉促的立法過程中，正因為各黨派間毫無疑義的一致認為應該通過修法以安撫民心，反而在看似一片祥和之氣的遮掩下，使人容易忽視過程中出現的各種異音。

對此，本文提出問題意識如下：「對於各種類之不實影像，若不採特別訂立專法之模式，單以既存之實定法是否得以充分應對 deepfake 技術所帶來之衝擊？」、「對於各種類之不實影像之使用，可能涉及之保護法益為何？」、「對於各種類之不實影像，若需訂立專法加以因應，其條文應如何設計較為適當？」。

於第一章，本文對於近年來針對不實影像議題發表看法之文獻進行回顧，發現研究量能高度集中在與性相關的不實影像議題上，因此在針對與性無關之不實影像方面討論高度匱乏，連帶的使其是否真的得以僅透過既存之實定法即可充分保護？可能涉及之法益侵害為何？若欲特別立法規制，應如何設計方屬適當？等問題，皆欠缺較為系統化的分析與研究。

而在針對與性相關之不實影像方面，本文回顧發現對於是否「無需另訂專法，僅依照原先既有之刑法系統即足以充分對應」，學說上即出現了紛呈的意見。究其根本，爭議其實即源自於學者間對於其可能涉及之保護法益認知有所歧異，然而在各學者之論述中，極大比例的篇幅都聚焦在正面論述自己認為合宜的保護法益，而相對缺少對於彼此觀點的檢驗與交鋒。最後，由於比較多數的意見都認為以既有之刑法系統即足以因應，故對於現行法下所選擇的，以專法加以規制的立法模式，究竟應如何設計其構成要件方屬適當，討論上亦僅見零星意見，而缺乏系統性的整理與論述。

於第二章，本文首先以「是否與性相關」與「是否有致使理性一般人混淆之虞」兩條軸線，將不實影像分成 4 種類型，並在隨後依序加以檢討對其之使用在原先既有之刑法體系中，可能成立何種犯罪。

在有混淆之虞且與性相關（類型一）方面，本文指出其無法成立刑法第 315 條之 2 散布竊錄影音罪以及刑法第 310 條誹謗罪、很大一部分情形無法成立刑法第 235 條散布猥褻物品罪以及個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪，且縱使得以成立上述罪名之案例，其亦未充分評價侵害之法益，且未提供被害人充足之保護；

在無混淆之虞且與性相關（類型二）方面，本文指出其在既存刑法體系中成立犯罪之情形大抵與類型一相同，惟基於其「不使公眾誤認行為人真有拍攝性影像」此一特質，使其更難成立刑法第 310 條誹謗罪以及個人資料保護法第 41 條非公務機關未於蒐集之特定目的必要範圍內利用個人資料罪；

在有混淆之虞且與性無關(類型三)方面,本文同意學者見解,認為既有的刑法體系中,關於施用詐術以及散布謠言之犯罪類型,得以部分處理與規制此類不實影音濫用之情形,惟本文認為基於若干理由其保護仍不充分。然而由於行文之順暢,本文於此章節並未深究前述所稱之理由,詳細之論理將留待第三章討論保護法益時一併處理;

在無混淆之虞且與性無關(類型四)方面,本文認為其本質上抹平了 deepfake 技術之所以使社會受到衝擊之特性,而與過去之影音編輯技術無異,因此總體而言將其視為普通言論即可。

於第三章,本文依序檢討 4 種類型之不實影像可能涉及之保護法益。(問題二)

在有混淆之虞且與性相關(類型一)方面,本文首先援引比較法上學者見解,以及被害人自陳之意見,指出了此種類型之影像確實可能使被害人蒙受巨大傷害。而此傷害背後所象徵之法益,本文整理實務與學說見解,共計討論了隱私權、名譽權、性自主權、性相關的冒犯性/仇恨性言論、形象法益等 5 種法益類型。

在隱私權部分,本文指出我國多數學者皆認為由於裸露之隱私部位,並非臉譜所有人真實之隱私資訊,故非屬隱私權之侵害,惟比較法上亦有學者考量過前述論點後,仍認為其係屬隱私之侵害。細究其原因,本文認為根本因素在於美國法上隱私權之發展脈絡,很大程度上乘載了引進新興人格權時,作為概括媒介條款之機能,而此機能在我國法的發展上則通常無須藉由隱私權作為媒介,而得以直接作為獨立的人格權被繼受進我國之法律系統當中。基於如上之法學文化背景差異,又未避免隱私權概念過於混雜以至於難以與其他人格權做為區分,本文結論上同意我國多數學者觀點,認為隱私權並非此種類不實影像所涉之法益;

在名譽權部分,本文並不同意多數學者認為「傳遞關於他人曾拍攝性影像,即會使該他人名譽受減損」之意見。在實然層次上,本文援引了美國最新之研究,指出社會一般通念對於此類行為的評價可能會隨著時間、環境而有所流轉,因此對於此類行為之價值判斷應與時俱進,不可妄加論斷。在應然層次上,本文亦援引了學者提出的規範名譽觀點,指出考量名譽概念時不得僅在事實層次上考量「名譽是否減損」,而應一併從應然層次上考量「名譽應否減損」。直言之,即本文認為若司法系統,乃至於整個政府機關,以公權力的姿態宣示了「使人知悉拍攝性影像是一種名譽減損」,則可能固化甚至加劇社會對於性之汙名化,此與我國學界當前提倡將性之概念與「猥褻」所象徵之「羞恥」等意涵脫鉤的思潮有所乖離,同時,當企圖保護法益之舉措,反倒成為傷害被害人的武器之一,其與刑法法益保護之機能亦有所違背。是故,本文認為不應認為此類型不實影像與名譽法益有涉。惟本文仍同意,雖同屬「未侵害名譽法益」,此種「事實名譽有受侵害」之類型,仍有被法律系統特別處理之必要性,最理想之作法應係找到其他法益吸納對於此部分侵害之評價;

在性自主權方面,本文首先指出由於此類影像所揭露之並非當事人真實之性資訊,故既不會侵害積極面向之性自主,亦無「性資訊控制權」被侵害之問題。

至於在此類型之不實影像，是否會形同「將人視為性客體」，因此侵害性自主權之問題上，本文亦持否定見解。理由在於本文說明了性幻想本身亦受憲法思想自由保障，有其正面價值，且若將「把人作為性聯想對象」直接與「將人視為性客體」畫上等號，此將使單純意淫之行為都會落入侵害性自主之範疇，有過度擴張刑罰之虞。最後，本文透過現行法下性騷擾防治法關於敵意環境性騷擾設計有行政罰規定作為座標，指出在某些前提下，「把人作為性聯想對象」且將此聯想付諸於若干行動，可能形成性騷擾而會受處罰，然而此種處罰亦僅係行政罰而已，而此類型之不實影像使用所造成之侵擾很多時候相較前述情形會更低，自然不應認為是刑法上所欲保障之性自主遭侵害；

在「性相關的冒犯性/仇恨性言論」此一法益方面，論者將此類型不實影像之使用，視為針對女性此一弱勢群體之集體性攻擊。對此，本文持否定看法。理由主要基於本文認為仇恨性言論本身，未必會使弱勢群體於社會上所受處境更為惡劣。此外，若欲處罰仇恨性言論，則勢必會陷入定義與區分上之困難，此一困難不僅抽象上可能違反法律明確性原則、平等原則、罪責相當原則，具體上亦可能使弱勢群體反而陷於更不利之地位，與其伊始之規範目的有間。最後，縱使認為仇恨性言論得以以刑事處罰，此類不實影像之使用亦非仇恨性言論所能涵蓋，蓋此類不實影像之使用，在我國經驗下從社會之反饋觀之，其並未根本性的使女性群體於社會上失去平等公眾參與之能動性，而得以透過言論市場之自我糾錯功能與以消化；

最後在形象法益方面，本文指出學說上肯否見解之最核心歧異，在於對於形象法益得否與名譽法益加以區分有不同的認知。本文認為形象法益與名譽法益之互動並非互斥關係，而應係包含關係，蓋名譽法益之侵害，本身即帶有形象扭曲之特質。並非所有的形象扭曲都值得以刑罰加以規制，例如使人誤認「某黑道老大背地裡其實樂善好施」，此種情形雖屬形象扭曲，惟刑法並無介入之必要。僅有「特別重大」的形象扭曲始有令刑罰介入之正當性，例如名譽法益即是一種「使人獲得負面評價的形象扭曲」，其因為除了形象扭曲外，尚帶有「使人獲得負面評價」此一特性，因此得以有別於其他形象扭曲，而成為值得受刑法保護之情況。而本文認為與性相關之形象扭曲，亦得以被認為是一種「特別重大」的形象扭曲，本文證立其「特別重大」之特殊性，主要基於兩點，其一是實證經驗上顯示被害人確實遭逢特別嚴重之傷害，其二是此類影像使用後，被害人會出現難以透過自力積極澄清，將形象導正之現象。此外，本文尚認為若實證上得以支持「遭散布不實性影像會傷害被描繪之人事實名譽」之結論，則「事實名譽受減損」之評價，亦可為形象法益所侵害，作為論證其形象扭曲「特別重大」之標準之一。準此，本文因此主張，使用有混淆之虞且與性相關（類型一）之不實影像，可能涉及之法益應為形象法益。

在無混淆之虞且與性相關（類型二）方面，本文則反向檢驗本文於證成類型一之影像係屬涉及形象法益而應受刑罰保護之理由，論證脈絡中包含了學說基礎、比較法對人格法益之立法模式，以及被害人自述之傷害狀況，三者都不足以支撐

在此類型影像之使用上，可以得出其具有應受刑法保障法益之結論，因此結論上認為此類型之不實影像使用並未涉及法益侵害。

於有混淆之虞且與性無關（類型三）方面，本文論證指出 deepfake 作為新興的技術，確實替社會帶來前所未有之新衝擊，主要是針對作為民主社會根基的，對於「眼見為憑」此一認知事實之根本性顛覆。而如此之顛覆在既存之法規範下，基於「法規皆僅偏重於處理『對於被描繪者之負面描述』，難以處理特定案件類型」、「僅是正面表列劃定保護範圍，因而有所缺漏」，以及「偏重於單次事件所生的影響，忽略系統性之傷害」等理由，而無法充分的保護到行為所侵害之法益。為此，本文認為此類型之不實影像，可能涉及目前未受刑法規範的無形偽造類型，其影響力在於對社會體系運作之重要性以及社會對其的信賴程度方面，與現行法下已納入規範之無形偽造類型如公務或業務文書相較，都足以相提並論，而存在獨立立法保障之必要。

於無混淆之虞且與性無關（類型四）方面，本文則承接第二章所述，認為其與過去之影音編輯技術無異，並未額外侵害法益。

於第四章，本文整理了美國法上共計 7 部與不實影像有關之法規/草案後，指出在「主觀要件」、「所規範之客體之欺詐性」、「所規範之客體之種類限制」、「客觀上構成要件行為」、「警語、免責聲明或浮水印之定性與效果」，以及「明文之豁免條款」等要件之部位上，可能可以存在設計上之差異。

以此為基礎，本文隨後檢視刑法第 319 條之 4 之立法，建議修正條文如下：

「刑法第 319 條之 4

散布、播送、交付、公然陳列，或以他法供人觀覽不實性影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為者，處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」

此外亦應於刑法第 10 條第 9 項新增規定如下：

「本法所稱之合理揭露警語與浮水印，謂下列措施：……」，而措施細節應至少包含聲明之數量、內容、位置、清晰程度等。

此外，針對刑法第 339 條之 4 之立法，本文檢驗其立法理由後認為其作為詐欺罪之加重處罰事由並無正當性，應全文刪除。

最後，針對與性無關之不實影像立法部分，本文認為應新增條文如下：

「刑法第 220 條之 1

行使不實影像與圖像，而未合理揭露警語與浮水印，足使人相信其中被描繪之人做出與其實際所為不相符之言論或行為，而足以生損害於公眾或他人者，

處五年以下有期徒刑、拘役或科或併科五十萬元以下罰金。」



第二節 研究限制、反思與展望

第一項 關於有混淆之虞且與性無關之不實影像（類型三）

至此，本文針對最初提出之問題意識，皆已盡力地完成了作答。然而若捫心自問，這份答卷中確實存在某些區塊是本文回答起來相對沒把握的地方。

首先是對於有混淆之虞且與性無關之不實影像（類型三）部分，本文最終最後得出了現行法下之規範不足以充分保護公共信用法益，而應受獨立之刑事法規制的結論。

惟在形成心證的過程中，本文亦曾陷入些許掙扎，主要是基於「人民相較於偽造之文書更容易相信不實影像」是一回事，但未必代表「影像就確實乘載了與文書相同之社會信賴機能」。簡言之，以本文最常拿出之推薦信與推薦影片的例子來說，目前大多商務實作、求職入學等情境中，皆仍都會要求提出文書以證明特定事實，然而卻未必要求提出影像據以證明。本文最終形成心證之理由是在於本文認為持平而論，文書與影像在社會實務上不同場合中，對於「證明特定事實」之效果與泛用性應是各佔勝場，因此仍然得出若偽造文書應受刑法規制，則散布此種類型不實影像亦應如此之結論。然而此論理是否合理，仍不排除有再思量之空間。

此外，在於「民主社會運作究竟有多仰賴影像證據？」以及「不實影像究竟會多大程度傷害社會對於影像證據之信賴？」此兩關鍵問題上，基於國內目前文獻上之討論量能相對匱乏，而本文所大量仰賴之比較法文獻中則是較為欠缺我國法體系中在意的「法益」概念，因此本文論證過程中仰賴大量個案的堆疊，輔以各種專家學者之斷言，去使人「感受」、「判斷」其嚴重性。不諱言，本文的立場能夠從論述中「感受」到嚴重性，因此也最終做出應以刑罰規制之結論。然而主觀感受畢竟會隨著個體生命經驗之落差而有著劇烈之波動，若僅能以此種路徑作為論證之基礎，終究是較為使人不安的推導過程。

綜合上述本文對於論理過程之自我反省，終局上本文無意改變原先給出之答案，然而亦必須承認在論理過程中存在諸多必須被填補之空白，需仰賴將來之研究者進一步開展。

第二項 關於無混淆之虞且與性相關之不實影像（類型二）

對此，本文最終給出其並未對法益有所侵害，因此不應受刑法規制之結論。然而細譯本文論理的過程，其實並非「積極的證明未有法益侵害，因此主張不應以刑法處罰」，毋寧更是有些被動的，選擇檢驗自己證立類型一應受刑法處罰之論理基礎，發現其皆難以支撐此種類型亦應受刑法規制，因此「消極的認為難以證明有法益侵害，因此主張不應以刑法處罰」。

誠然，在刑法的基本原則上，當找不出處罰一個行為的正當性時，不應予以處罰絕對是必然的正確答案，因此本文亦無意更改所給出之答案。然而在選擇

論證之取徑上，以消極的路線去說理，終究還是讓人較難以心安。

而深究此現象之根本原因，本文認為應係目前國內外文獻於討論不實影像時，皆未存在「有無使人混淆之虞乃一重要因子」的問題意識，因此無論是在實證證據上討論被害人實際受害情形，抑或是法理上討論可能涉及之法益侵害，皆很難探明究竟在失去此一重要因子後，是否仍實質上可能導致被害人傷害或在法理上涉及法益侵害。惟無論在何種論述取徑下，似乎都想當然耳的把「公眾可能誤以為真」當作論證的重要環節，準此，本文仍認為欠缺使人誤以為真潛能之不實影像，即應不具可罰性。

自此，雖然礙於本文能力限制，未能在此議題下產出更多積極面向上之論述，僅採取比較消極的姿態去檢驗與挑剔既存文獻中論理過程之缺憾，惟仍希望能透過本文作為楔子，使將來之研究方向對於此議題有所著墨，完成本文未竟之論理空白。

參考文獻



一、中文文獻

(一) 專書

- 甘添貴 (1996),《刑法之重要理念》,瑞興圖書
林山田 (2006),《刑法各罪論(下冊)》,自刊
黃榮堅 (2012),《基礎刑法學(下)》,4 版,元照

(二) 期刊論文

- 許恒達 (2020),〈行為規範、保護法益與通姦罪的違憲審查——評釋字第 791 號解釋〉,《月旦法學雜誌》,第 305 期,頁 15-32
許恒達 (2022),〈深度偽造影音及其刑法規制〉,《法學叢刊》,第 265 期,頁 1-29
許恒達 (2022),〈交付人頭帳戶的獨立刑事制裁?—評洗錢防制法第 15 條之 1 修正草案〉,《當代法律》,第 8 期,頁 15-27
雷喻翔 (2022),〈臉,為何會被挖走?深度偽造(Deepfake)介紹〉,《清流雙月刊》,第 42 卷,頁 29-35
林建中 (1997),〈隱私權概念初探—從美國法之觀點切入〉,《憲政時代》,第 23 卷 1 期,頁 53-78
林政佑 (2020),〈刑法第一百四十四條侮辱公務員及公署罪之保護法益之探討〉,《國立中正大學法學集刊》,第 74 期,頁 65-125
林琬珊 (2018),〈妨害名譽罪與負面標籤〉,《月旦刑事法評論》,第 9 期,頁 63-78
林琬珊 (2023),〈「性影像」與隱私之刑法保護——新修刑法「性影像」規定之評析〉,《月旦法學雜誌》,第 333 期,頁 40-58
古承宗 (2023),〈刑法作為抗制「數位暴力」之手段——以惡意影音為例〉,《月旦法學雜誌》,第 333 期,頁 6-20
陳俊偉 (2020),〈網路遊戲中侮辱行為的刑法評價難題 —以臺灣實務判決的觀點為中心〉,《月旦法學雜誌》,第 305 期,頁 159-184
陳俊偉 (2023),〈重思深度偽造影音的入罪化——以人工智慧的科技風險規制為思考方向〉,《月旦法學雜誌》,第 333 期,頁 21-39
廖宜寧 (2022),〈從「妨害風化的猥褻物品」到「妨害性隱私的性影像」——散布性關聯影音圖像之不法性質〉,《月旦法學雜誌》,第 327 期,頁 158-171
吳芳毅 (2020),〈深度偽造為色情報復之侵害與規制〉,《檢察新論》,第 28 期,頁 185-200
許昭元 (2023),〈論以深度偽造 (Deepfake) 技術製造性私密影像是否受憲法

- 思想自由之保障——以我國近期刑法修正草案為對象》，《憲政時代》，46 卷 4 期，頁 449-494
- 張天一（2019），〈散布私密影像之刑事責任〉，《月旦法學教室》第 198 期，頁 26-29
- 黃國瑞（2014），〈法益論之解構〉，《輔仁法學》，第 48 期，頁 257-299
- 黃宗旻（2019），〈法益論的侷限與困境：無法發展立法論機能的歷史因素解明〉，《臺大法學論叢》，第 48 卷第 1 期，頁 159-210
- 葉啟洲（2012），〈罵人「共匪」或「胡錦濤」與名譽權侵害——臺灣板橋地方法院一百年度訴字第一三五一號民事判決〉，《月旦裁判時報》，第 13 期，頁 87-95
- 葉啟洲（2020），〈通姦罪違憲論的初始觀點——大法官釋字第 554 號解釋的聲請典故〉，《台灣法學雜誌》，第 392 期，頁 43-47
- 薛智仁（2021），〈通姦罪之憲法審查——評司法院釋字第七九一號解釋〉，《成大法學》，第 42 期，頁 1-69
- 薛智仁（2022），〈刑法解釋與適用之憲法界限——近期刑事判決之綜合評析〉，《興大法學》，第 32 期，頁 1-105
- 廖緯民（1996），〈論資訊時代的隱私權保護——以「資訊隱私權」為中心〉，《資訊法務透析》，第 8 卷第 11 期，頁 20-27
- 法思齊（2021），〈論以刑事法規範復仇式色情（Revenge Porn）之可能——以日本情色報復受害防治法為中心〉，《月旦法學雜誌》，第 314 期，頁 210-223
- 張麗卿（2016），〈刑法對誹謗言論的合理規制〉，《臺大法學論叢》，第 45 卷第 4 期，頁 1999-2029
- 蕭郁唐（2016），〈性隱私內容外流風波——從美國立法例論我國違反本人意願散布性隱私內容之入罪化〉，《科技法律透析》，28 卷 10 期，頁 27-55
- 蕭郁唐（2017），〈日本情色報復受害防治法之評析〉，《科技法律透析》，第 29 卷第 2 期，頁 21-28
- 簡至鴻（2022），〈個人資料保護法第 41 條「違法侵害個資罪」之基本課題——以最高法院 109 年度台上大字第 1869 號刑事裁定的檢討為契機〉，《刑事政策與犯罪防治研究》，第 32 期，頁 131-184
- 劉定基（2020），〈個人資料保護法〉，《月旦法學教室》，第 218 期，頁 59-67
- 曾思堯（2017），〈從公然侮辱罪的存廢談保護法益的實態〉，《國際刑事法學之新脈動：余振華教授六秩晉五壽誕祝賀論文集》，第 1 期，頁 677-706
- 黃惠婷（2011），〈從性自主權檢視刑法保護兒童及少年之規定與修法建議〉，《司法新聲》，第 97 期，頁 62-77
- 莊世同（2022），〈論仇恨言論的概念——一個法哲學觀點的分析〉，《臺北大學法學論叢》，第一二四期，頁 113-179
- 王正嘉（2019），〈刑法第 190 條之 1 修正之問題與商榷〉，《台灣民主法治的經驗與見證——江義雄教授七秩晉五華誕祝壽論文集》，頁 437-455

- 謝煜偉 (2016),〈風險社會中的抽象危險犯與食安管制—「攙偽假冒罪」的限定解釋〉,《月旦刑事法評論》,第 1 期,頁 70-90
- 謝煜偉 (2019),〈論排放毒物污染環境媒介罪:與各環境行政刑罰法規之關連性〉,《國立臺灣大學法學論叢》,第 48 卷 特刊,頁 1375-1434
- 謝煜偉 (2022),〈2021 年刑事立法與實務發展回顧:以治安與治療為名〉,《臺大法學論叢》,第 51 卷特刊,頁 1223-1255
- 李聖傑 (2010),〈刑法第二二二條第一項第一款「二人以上共同犯之」之適用思考〉,《政大法學評論》,第 115 期,頁 1-54
- 李聖傑 (2015),〈三人以上共同犯詐欺罪的規範適用檢視〉,《月旦裁判時報》,第 39 期,頁 52-58
- 李聖傑 (2019),〈刑法第 190 條之 1 的釋義思考〉,《月旦法學雜誌》,第 290 期,頁 69-82
- 李茂生 (2000),〈再論偽造文書罪中有關有形偽造以及保護法益的問題〉,《刑事思潮之奔騰—韓忠謨教授紀念論文集》,頁 299-317
- 李茂生 (2019),〈新修刑法第 190-1 條之商榷〉,《台灣法學雜誌》,第 368 期,頁 29-32
- 蘇慧婕 (2015),〈從諷刺言論談言論管制的司法審查難題〉,《台灣法學雜誌》,第 264 期,頁 26-28
- 徐育安 (2014),〈資訊風險與刑事立法〉,《北大學法學論叢》,第 91 期,頁 113-167
- 林慈偉 (2019),〈首件適用《公政公約》第 36 號一般性意見之死刑裁判例:評臺灣彰化地方法院 107 年重訴字 第 13 號刑事判決〉,《台灣人權學刊》,第 5 卷第 2 期,頁 183-194
- 游明得 (2021),〈刑法加重詐欺罪之檢視—以跨國電信詐欺案為核心〉,《開南法學》,第 12 期,頁 1-45
- 蔡聖偉 (2019),〈參與犯罪組織罪與加重詐欺罪的競合—評最高法院 107 年度台上字第 1066 號刑事判決〉,《月旦法學雜誌》,第 292 期,頁 181-191
- 黃榮堅 (2012),〈2011 年刑事法發展回顧:法律說詞與說詞之外〉,《臺大法學論叢》,第 41 卷特刊,頁 1537-1574
- 吳耀宗 (2005),〈偽造文書罪保護法益之研究〉,《月旦法學雜誌》,第 128 期,頁 120-141
- 余和謙 (2019),〈人工智慧之治理—以深度偽造為例〉,《科技法律透析》,第 31 卷第 8 期,頁 52-72
- 楊立新 (2009),〈楊立新民法講義(貳)—人格權法〉,自刊,頁 211-219

(三) 政府文書

- 立法院公報第 111 卷第 84 期委員會紀錄
- 立法院公報第 112 卷第 14 期委員會紀錄



(四) 網路文獻

鏡傳媒 (05/06/2021)，〈臉被偷走之後—無法可管的數位性暴力？台灣 Deepfake 事件獨家調查〉，

<https://www.mirrormedia.mg/projects/deepfaketaiwan/>

中國時報 (10/19/2021)，〈法務部祭「小玉條款」1 個月內提修法重罰科技犯罪〉，<https://www.chinatimes.com/realtimenews/20211019002132-260407?chdtv>

三立新聞網 (11/17/2021)，〈防再有小玉「深偽」換臉 法務部公布修法 5 大重點〉，<https://www.setn.com/News.aspx?NewsID=1028157>

鏡傳媒 (05/04/2022)，〈【換臉挨打怒提告 2】遭換臉不雅片曝光 陳沂坦言「身心留下陰影」〉，

<https://www.mirrormedia.mg/story/20220503ent012/>

weprol80 科技新聞 (05/26/2022)，〈【Video Hunt】Elon Musk 投資平台保證 3 成回報率 三個破綻識破 Deepfake 騙局〉，

https://www.weprol80.com/deepfakescam_220526/

婦女救援基金會，〈有錢就能散布別人裸照，誰讓加害者有恃無恐？〉，

<https://twrf.org.tw/info/title/151>

自由時報 (05/05/2019)，〈注意言辭！法院認證罵人「張淑晶」屬公然侮辱〉，<https://news.ltn.com.tw/news/society/breakingnews/2780500>

華視新聞 (03/12/2021)，〈納豆身高遭法院認證侮辱？本人傻眼... 大咖藝人相挺〉，

<https://news.cts.com.tw/cts/entertain/202103/202103122034442.html>

CTWANT (09/01/2021)，〈嗆「長得像黃安」法院認證公然侮辱！黃安怒飆髒話回應了〉，<https://www.ctwant.com/article/137187>

報導者 (2021/11/1)，〈姜冠霖、王碩勛／從陌生人到熟人凌辱，韓國 Deepfake 數位性犯罪修法的啟示〉，

<https://www.twreporter.org/a/opinion-deepfake-digital-sex-crime-korean-experience>

二、英文文獻

(一) 期刊論文

Harris (2019), Deepfakes: False Pornography Is Here and the Law Cannot Protect You, *Duke Law & Technology Review* 17, pp. 99-128

Chesney, Citron (2019), Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review*, Vol. 107, Issue 6, pp. 1753-1820

- Thombre (2021), Deconstructing Deepfake: Tracking Legal Implications and Challenges, *International Journal of Law Management & Humanities*, Vol. 4 Issue 4, pp. 2267-2274
- Brown (2020), Deepfakes and the Weaponization of Disinformation, 23 VA. J.L. & TECH, pp. 1-59
- Langa (2021), Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes, *Boston University Law Review*, Vol. 101, Issue 2, pp. 761-802
- Olson (2022), The Double-Side of Deepfakes: Obstacles and Assets in the Fight against Child Pornography Notes, *Georgia Law Review*, pp. 865-[i]
- Kugler, Pace (2021), Deepfake Privacy: Attitudes and Regulation, *Northwestern University Law Review* 116, Issue 3, pp. 611-680
- Nagumotu (2022), Deepfakes Are Taking over Social Media: Can the Law Keep up?, *IDEA: The Law Review of the Franklin Pierce Center for Intel* Issue 2, pp. 102-146
- Bendor, Dancig-Rosenberg (2016), Unconstitutional Criminalization, *New Criminal Law Review* 19 *New Crim. L. Rev.*, pp. 171-207
- Hornle (2016), Theories of Criminalization, *Criminal Law and Philosophy* 10 *Crim. L. & Phil.* , pp. 301-314
- Šepec, Lango (2020), Virtual Revenge Pornography as a New Online Threat to Sexual Integrity, *Balkan Social Science Review*, pp 117-134
- Parent (1983), A New Definition of Privacy for the Law, *Law and Philosophy* 2, pp. 305-338
- Chivers-Wilson (2006), Sexual assault and posttraumatic stress disorder: A review of the biological, psychological and sociological factors and treatments, *Mcgill J Med.* , pp.111-118
- Rini (2020), Deepfakes and the Epistemic Backstop, *PHILOSOPHERS' IMPRINT*, pp. 1-16
- Ray (2021), Disinformation, Deepfakes and Democracies: The Need for Legislative Reform Thematic Issue: Big Technology and the Law, *University of New South Wales Law Journal*, 44, pp. 983-1013
- Atencia-Linares, Marc (2022), Deepfakes, shallow epistemic graves: On the epistemic robustness of photography and videos in the era of Deepfakes, *Synthese* 200 (6), pp.1-22
- Gamage, Ghasiya, Bonagiri (2022), Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications, Conference: CHI Conference on Human Factors in Computing Systems
- Köbis, Doležalová, Doležalová (2021), Fooled Twice – People Cannot Detect Deepfakes But Think They Can, *iScience*, 1, pp. 1-17



(二) 網路文獻

Karen, Deepfake porn is ruining women's lives. Now the law may finally ban it, in_ <https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/>.

Ray, Bot Generated Fake Nudes Of Over 100,000 Women Without Their Knowledge, Says Report, in: <https://www.forbes.com/sites/siladityaray/2020/10/20/bot-generated-fake-nudes-of-over-100000-women-without-their-knowledge-says-report/?sh=7507150b7f6b>.

Mihalcik, California laws seek to crack down on deepfakes in politics and porn, in_ <https://www.cnet.com/news/politics/california-laws-seek-to-crack-down-on-deepfakes-in-politics-and-porn/>.

Galston, Is seeing still believing? The deepfake challenge to the truth in politics, BROOKINGS, in: <https://www.brookings.edu/research/is-seeing-still-believing-the-deepfake-challenge-to-truth-in-politics/>.

Abrams, Elon Musk deep fakes promote new BitVex cryptocurrency scam, bleepingcomputer, in: <https://www.bleepingcomputer.com/news/security/elon-musk-deep-fakes-promote-new-bitvex-cryptocurrency-scam/>

Dator, The NBA logo's complicated history, explained, SBNATION, in: <https://www.sbnation.com/nba/22300942/nba-logo-history-jerry-west-kobe-bryant>.

Ayyub, I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me, HUFFPOST UK, in: https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316.

SWNS, Shocking number of people have sent nude photos, poll finds, New York Post, in: <https://nypost.com/2022/09/13/shocking-number-of-people-have-sent-nude-photos-poll-finds/>.

Hsu, As Deepfakes Flourish, Countries Struggle With Response, in_ <https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html>.