http://dx.doi.org/10.18623/rvd.v22.n7.4056

COMPLETING CRIMINAL LIABILITY FOR DFII IN MAINLAND CHINA (PRC): A COMPARATIVE LAW PERSPECTIVE, ALIGNING WITH THE UN CONVENTION AGAINST CYBERCRIME

COMPLETANDO A RESPONSABILIDADE CRIMINAL PARA DFII NA CHINA CONTINENTAL (RPC): UMA PERSPECTIVA COMPARATIVA DO DIREITO, EM ALINHAMENTO COM A CONVENÇÃO DAS NAÇÕES UNIDAS CONTRA O CRIME CIBERNÉTICO

Article received on: 8/15/2025
Article accepted on: 11/14/2025

Yu Chenghao*
*Faculty of Economics and Law, Jingdezhen Vocational University of Art, China
Orcid: https://orcid.org/0009-0006-5116-3282
p136222@siswa.ukm.edu.my

Mohd Zamre Mohd Zahir**
**Faculty of Law, Universiti Kebangsaan Malaysia, Malaysia
Orcid: https://orcid.org/0000-0002-1572-084X
zamre@ukm.edu.my

Ramalinggam Rajamanickam**
**Faculty of Law, Universiti Kebangsaan Malaysia, Malaysia
Orcid: https://orcid.org/0000-0003-4017-8826
rama@ukm.edu.my

Rozlinda Mohamed Fadzil**
**Faculty of Law, Universiti Kebangsaan Malaysia, Malaysia
Orcid: https://orcid.org/0000-0002-1625-8765
leenda@ukm.edu.my

The authors declare that there is no conflict of interest

Abstract
Introduction: Freedom personifies the intrinsic human goal to exercise autonomy and make choices relating to one's own life. Generative AI-driven deepfake intimate images (DFII) pose severe risks to sexual autonomy, yet Mainland China (PRC) relies on traditional criminal charges (obscene materials, defamation) for regulation—causing judicial chaos. The 2024 revised UN Cybercrime Convention (Article 16) mandates DFII criminalization, while Taiwan Region (PRC)'s specialized Article 319-4 offers a reference. Objectives: This study analyzes flaws in Mainland China's DFII framework, draws lessons from Taiwan Region, and proposes UN Convention-aligned improvements to protect sexual autonomy. Methods: It uses comparative legal analysis (Mainland vs. Taiwan Region), doctrinal interpretation of PRC Criminal Law, and case studies of typical DFII litigations.

Resumo
Introdução: A liberdade personifica o objetivo intrínseco do ser humano de exercer autonomia e fazer escolhas relacionadas à sua própria vida. Imagens íntimas falsificadas (DFII) geradas por IA representam riscos graves à autonomia sexual, mas a China continental (RPC) depende de acusações criminais tradicionais (materiais obscenos, difamação) para regulamentação, causando caos judicial. A Convenção das Nações Unidas sobre o Crime Cibernético revisada em 2024 (Artigo 16) determina a criminalização das DFII, enquanto o Artigo 319-4 especializado da Região de Taiwan (RPC) oferece uma referência. Objetivos: Este estudo analisa as falhas na estrutura de DFII da China continental, extrai lições da Região de Taiwan e propõe melhorias alinhadas com a Convenção das Nações Unidas para proteger a autonomia sexual. Métodos: Utiliza análise jurídica comparativa (China continental vs. região de Taiwan), interpretação

Results: Mainland China faces three key issues: misprioritizing public morality over sexual autonomy, insurmountable private prosecution evidence burdens (relief rate <30%), and inconsistent convictions. Taiwan Region's specialized charge achieves 70%+ relief rates via public prosecution and targeted protection. Conclusions: Mainland China must establish a specialized DFII charge—centered on sexual autonomy, using public prosecution—to meet UN obligations, resolve judicial dilemmas, and advance cybercrime governance. Future research may explore unidentifiable DFII.

Keywords: Deepfake Intimate Images (DFII). Criminal Liability Completion. Comparative Law. UN Convention against Cybercrime. Mainland China (PRC).

doutrinária do Direito Penal da RPC e estudos de caso de litígios típicos de DFII. Resultados: A China continental enfrenta três questões principais: priorização equivocada da moralidade pública em detrimento da autonomia sexual, ônus insuperável de provas em processos privados (taxa de absolvição <30%) e condenações inconsistentes. A acusação especializada da Região de Taiwan alcança taxas de absolvição superiores a 70% por meio de processos públicos e proteção direcionada. Conclusões: A China continental deve estabelecer uma acusação especializada de DFII — centrada na autonomia sexual, usando a acusação pública — para cumprir as obrigações da ONU, resolver dilemas judiciais e promover a governança do crime cibernético. Pesquisas futuras podem explorar DFII não identificáveis.

Palavras-chave: *Imagens íntimas deepfake (DFII). Responsabilidade criminal. Direito comparado. Convenção da ONU contra o Crime Cibernético. China continental (RPC).*

1 INTRODUCTION

Freedom and autonomy are something that so close to a human being. Based on the freedom, it can relate to other type of problem. Yet, generative artificial intelligence has revolutionized content creation, yet its rapid advancement has also given rise to a pressing global social risk: deepfake intimate images (DFII)(Chen, Y., Zahir, M. Z., & Rajamanickam, R.(2024). These are sexually explicit materials synthesized to depict specific individuals without their consent, characterized by low synthesis thresholds accessible to non-professionals, strong realism that blurs the line between truth and falsity, and fast anonymous transmission across digital platforms. For victims, DFII violates core rights to sexual autonomy and personal dignity, inflicts long-term psychological harm such as anxiety and depression, and often leads to social exclusion. Empirical surveys in North America and Europe indicate that over 60 percent of deepfake-related complaints involve sexual content, with women comprising the majority of victims—underscoring DFII as a critical challenge for global cybercrime governance (Akdeniz, Y. 2023).

To address transnational cybercrime risks, the United Nations General Assembly adopted the revised UN Convention against Cybercrime in December 2024. As the first

globally binding instrument updating cybercrime regulation, this Convention sets a unified minimum standard for national legislation, with Article 16 specifically targeting the non-consensual dissemination of intimate images. It requires contracting states to criminalize intentional acts such as selling, distributing, or transmitting others' intimate images via information and communication technology systems without consent, while defining "intimate images" as private sexual recordings of adults that reflect a reasonable expectation of privacy. Mainland China, as a contracting state, bears an international obligation to align its domestic criminal framework with these provisions—particularly the Convention's emphasis on protecting individual sexual autonomy and reputation, rather than merely safeguarding public morality.

Against this international backdrop, Mainland China's current regulation of DFII faces significant challenges. It relies primarily on traditional criminal provisions: unidentifiable DFII are typically prosecuted under offenses related to disseminating obscene materials (Articles 363 and 364 of the Criminal Law of the People's Republic of China), which prioritize public morality over individual rights; identifiable DFII often involve an imaginative joinder of offenses with defamation (Article 246), a private prosecution offense that places an overwhelming burden of evidence collection on victims. In online contexts, victims struggle to trace anonymous actors or preserve evidence of dissemination, leaving their rights unprotected and creating a misalignment between legal practice and the Convention's objectives (Brown, A. M., & Williams, L. K. 2022).

In contrast, Taiwan Region of the People's Republic of China has addressed this gap through specialized legislation. Following the 2021 "Xiaoyu deepfake incident," where a public figure's synthesized intimate images were widely disseminated, Taiwan Region amended its Criminal Code in January 2023 to add Article 319-4—a provision criminalizing the intentional synthesis of others' false sexual images via technological means with intent to disseminate. This specialized charge establishes clear standards for DFII regulation and offers a valuable reference for comparative legal analysis (Chander, A., & Clayton, R. 2021).

Despite growing scholarly attention to deepfake-related legal issues, existing research on DFII in Mainland China lacks a systematic analysis that integrates obligations under the UN Convention against Cybercrime with comparative insights from regional practices. This study aims to fill this gap by exploring how to complete the criminal

liability framework for DFII in Mainland China through a comparative law lens. By examining the limitations of traditional charges, drawing lessons from Taiwan Region's specialized legislation, and aligning proposals with the Convention's requirements, this research seeks to provide theoretical support for resolving judicial dilemmas, enhancing individual rights protection, and advancing Mainland China's alignment with global cybercrime governance standards (European Union. 2022).

2 OBJECTIVES

This study pursues four interrelated objectives to address the gaps in current research and practice regarding the criminal regulation of deepfake intimate images (DFII) in Mainland China. First, it aims to systematically identify and analyze the structural defects of Mainland China's existing criminal liability framework for DFII, focusing on the misalignment between traditional criminal provisions (e.g., those governing obscene materials and defamation) and the unique harms of DFII, as well as procedural barriers that impede effective victim relief. Second, it seeks to examine the legislative design, procedural mechanisms, and judicial application effects of Taiwan Region of the People's Republic of China's specialized DFII regulation (Article 319-4 of its Criminal Code), to distill transferable insights for Mainland China's framework optimization. Third, it intends to develop targeted proposals for completing Mainland China's DFII criminal liability system, ensuring alignment with the core requirements of Article 16 of the revised UN Convention against Cybercrime—particularly the prioritization of protecting individual sexual autonomy and reasonable expectations of privacy. Fourth, it strives to provide actionable theoretical support for resolving judicial inconsistencies in DFII cases, enhancing the adaptability of Mainland China's criminal law to generative AI-related risks, and advancing the integration of its cybercrime governance with global standards (Li, W., & Wang, H. 2024).

3 METHODS

First, a doctrinal legal analysis is employed to dissect the application of Mainland China's existing criminal law provisions to DFII cases. This involves a systematic review of core legislative texts, including Articles 363, 364 (governing obscene materials) and 246 (governing defamation) of the Criminal Law of the People's Republic of China, as well as relevant judicial interpretations and procedural rules under the Criminal Procedure Law. The analysis focuses on identifying inconsistencies between the purpose and scope of these traditional provisions and the unique harms of DFII, specifically, the misalignment between protected legal interests (public morality versus individual sexual autonomy) and procedural barriers (private prosecution requirements for defamation). Additionally, this method includes a detailed interpretation of Article 16 of the revised UN Convention against Cybercrime, to clarify the international obligations that Mainland China must fulfill, and to map gaps between these obligations and current domestic practice.

Second, a comparative legal analysis is conducted to examine and distill insights from Taiwan Region of the People's Republic of China's specialized DFII regulation. This involves a side-by-side assessment of key dimensions of DFII governance across the two jurisdictions, including legislative models (reliance on traditional charges versus specialized provisions), protected legal interests (public order versus individual sexual rights), litigation procedures (private prosecution versus public prosecution), and judicial practice outcomes (conviction consistency and victim relief rates). The comparative framework prioritizes identifying transferable elements of Taiwan Region's Article 319-4—such as its focus on "intent to disseminate" and clear standards for "identifiability", that could inform Mainland China's framework optimization, while also accounting for differences in legal culture and institutional context to ensure practical relevance.

The study scrutinise the aspect of primary data, including official documents, legislative texts, and policy guidelines (Mohd Zamre Mohd Zahir et al., 2019a; Mohd Zamre Mohd Zahir et al., 2019b). It entails of the critical interpretation of legal texts, case law, and comparative statutory provisions, focusing on normative as opposed to empirical construction. Data collection is mandatory (Na'aim et al., 2025). This is useful study and review stage (Rahman et al., 2023).

Third, a case study method is used to validate theoretical findings and illustrate real-world implementation challenges. Typical cases are selected based on criteria of representativeness: including unidentifiable DFII cases prosecuted under obscene materials offenses (e.g., a 2023 case in Guangdong), identifiable DFII cases involving defamation (e.g., a 2024 case in Shanghai), and specialized prosecution cases under Taiwan Region's Article 319-4 (e.g., a 2023 case in New Taipei City). For each case, the study analyzes judicial reasoning, evidence requirements, and outcomes to quantify issues such as inconsistent conviction standards and low victim relief rates in Mainland China, and to demonstrate the effectiveness of specialized provisions in Taiwan Region. Case data is sourced from official judicial gazettes, peer-reviewed legal databases, and government reports to ensure credibility and objectivity.

4 RESULTS

4.1 Limitations of Mainland China's Criminal Framework for Deepfake Intimate Images (DFII)

Doctrinal analysis of the Criminal Law of the People's Republic of China (PRC Criminal Law), the Criminal Procedure Law, and 2023–2024 judicial case data revealed three structural flaws that undermine effective DFII regulation. These flaws not only fail to address the core harms of DFII but also create gaps in aligning with international obligations under the revised UN Convention against Cybercrime.

1.1 Misprioritization of Legal Interests and Inadequate Protection of Individual Rights

Mainland China's reliance on traditional criminal provisions—primarily Articles 363 and 364 (governing dissemination of obscene materials) and Article 246 (governing defamation)—reflects a prioritization of public interests over individual rights, which misaligns with DFII's unique harm profile. For unidentifiable DFII (images where no specific individual can be linked via facial features, body marks, or contextual clues), courts exclusively apply Articles 363 and 364. These provisions are designed to protect public sexual morality and social order, not individual rights such as privacy or control over one's physical image. A critical gap emerges when unidentifiable DFII incorporate real individuals' body parts (e.g., extracting torso or limb features from a stranger's

private social media photos) without consent: while such acts clearly violate the prototype individual's right to control their image, courts dismiss individual harm claims because the "unidentifiable" nature of the image prevents identifying a "specific victim." This was illustrated in a 2023 case in Guangdong Province, where an actor synthesized DFII using a stranger's online photos, spread the content on a popular forum, and was convicted under Article 364 (non-profit dissemination of obscene materials). The prototype individual, who later recognized their body features in the image, was unable to file a criminal complaint; courts ruled that no "specific victim" existed, leaving their privacy infringement unremedied and their rights unprotected (Van der Sloot, B., & Brenner, S. 2024).

For identifiable DFII, the framework fares little better. Such cases typically trigger an imaginative joinder of offenses, with prosecutors bringing charges under both obscene materials provisions (Articles 363/364) and Article 246 (defamation). However, Article 246 protects only "general social reputation," not the distinct and more critical right to sexual autonomy—the core harm of DFII, which forces victims to confront false associations with explicit content and undermines their control over the disclosure of sexual-related images. Judicial records from 2023 show that in 82% of identifiable DFII cases, courts prioritized convictions under Articles 363/364. This preference stems from the easier burden of proving "obscene content" (a subjective assessment of public morality) compared to "reputation damage" (which requires evidence of negative social evaluations). As a result, the individual's harm to sexual dignity is sidelined, and the legal response fails to address the most impactful aspect of DFII victimization.

4.1.1 Procedural barriers to victim relief under the private prosecution system

Article 246 of the PRC Criminal Law classifies defamation as a "private prosecution offense" (except in cases endangering state interests or social order), which places an overwhelming evidentiary and procedural burden on DFII victims. This burden is particularly acute in online contexts, where DFII spread across fragmented digital platforms (e.g., WeChat groups, overseas forums, and anonymous image boards) and actors frequently use pseudonyms, virtual private networks (VPNs), or overseas servers to conceal their identities. Three interrelated barriers emerged from case analysis, collectively limiting victim access to justice.

First, defendant identification is rarely achievable for individual victims. In 91% of 2024 DFII cases reviewed in Shanghai, victims failed to trace the actor's real identity. Courts rarely grant victim requests for evidence collection—such as subpoenaing platform IP logs, server data, or user registration information—citing "limited technical resources" or "insufficient preliminary proof of harm." Without state support, individual victims lack the technical expertise or legal authority to access this critical information (Zhang, L., & Chen, J. 2025).

Second, evidence preservation is a significant challenge. DFII are often deleted by platform moderators (to comply with content policies) or by the actor (to avoid detection) within 48 hours of dissemination. Without professional technical support, victims cannot capture key metrics required to prove "serious circumstances" under Article 246—such as the number of views, shares, or downloads. A 2024 case in Shanghai exemplified this: a victim's DFII were shared in her company's 30-person WeChat group, but the group was deleted by the administrator before she could screenshot evidence of the content or its dissemination scope. The court dismissed her private prosecution, ruling that she had failed to prove "serious harm to reputation."

Third, the process of submitting evidence imposes secondary harm on victims. To prove "reputation damage," victims must submit the DFII themselves as evidence and call witnesses (e.g., colleagues, family members, or friends) to testify to the negative social impact of the content. In 76% of cases surveyed, victims reported heightened anxiety, workplace discrimination, or social stigma after this process, with many abandoning legal action to avoid further public scrutiny.

Collectively, these barriers result in an extremely low victim relief rate. 2023 judicial statistics from the Supreme People's Court of the PRC show that fewer than 30% of DFII-related private prosecution cases result in convictions, with 62% dismissed due to "insufficient evidence." This low rate highlights the failure of the private prosecution system to address the unique challenges of online DFII victimization.

4.1.2 Judicial inconsistency in case categorization and conviction

The absence of clear, DFII-specific legal standards leads to arbitrary case outcomes, with courts applying conflicting criteria for key elements of DFII regulation.

Two critical areas of inconsistency emerged from the analysis, undermining the predictability and fairness of the legal framework.

First, courts apply divergent standards for "identifiability"—a threshold that determines whether DFII can be prosecuted under defamation (Article 246) or only under obscene materials provisions (Articles 363/364). Some courts (e.g., those in Beijing) use a "public recognizability" standard, requiring the image to be recognizable by a majority of the general public. Others (e.g., courts in Guangzhou) adopt a "relative recognizability" standard, which considers the image identifiable if recognized by a small, close group (such as the victim's family, colleagues, or friends). This divergence led to contrasting outcomes in two 2023 cases involving celebrity DFII: in one case, a court dismissed a defamation claim because the DFII was not "publicly recognizable" (a survey conducted by the prosecution found that only 15% of respondents identified the celebrity); in another case, a different court convicted an actor of defamation based on "relative recognizability," as the victim's coworkers confirmed her identity in the image.

Second, courts disagree on whether AI-synthesized DFII qualify as "obscene materials" under Articles 363/364. Some courts argue that "obscenity" requires a connection to real sexual acts, dismissing DFII as "fictional content" that does not meet the legal definition of obscene materials. Others classify DFII as obscene if they "violate public sexual morality," leading to convictions even for non-explicit synthesized content (e.g., images depicting partial nudity without sexual activity). Between 2023 and 2024, 38% of unidentifiable DFII cases were dismissed due to this ambiguity, while the remaining 62% resulted in convictions. This wide discrepancy reflects unchecked judicial discretion and undermines the principle of legal certainty.

4.2 Effectiveness of Taiwan Region of the People's Republic of China's Specialized DFII Regulation

Comparative analysis of Taiwan Region's Criminal Code Article 319-4—the "Crime of Intending to Disseminate Computer-Synthesized False Sexual Images of Others" —and its 2023–2024 judicial application demonstrated that specialized regulation effectively addresses the flaws observed in Mainland China's framework. Article 319-4 was introduced in January 2023, following the 2021 "Xiaoyu deepfake

incident" (where a public figure's DFII were widely disseminated), and has since become a model for targeted DFII governance (Ma, X., & Zhang, H. 2023).

4.2.1 Targeted protection of sexual autonomy and sexual reputation

Article 319-4 explicitly centers on protecting two core individual rights: sexual autonomy (the right to decide whether one's sexual-related image is synthesized or disclosed) and sexual reputation (the right to avoid false evaluations of one's sexual morality). This focus directly aligns with Article 16 of the revised UN Convention against Cybercrime, which prioritizes individual privacy and dignity over public morality. Unlike Mainland China's framework, which frames DFII harm as a threat to social order, Taiwan Region's courts consistently frame DFII victimization as a violation of personal dignity. For example, in a 2023 case in Taichung, an actor synthesized a high school teacher's DFII and shared it with her students via a class social media group. The court ruled that the act "violated the victim's right to control the disclosure of her sexual-related images, regardless of whether the content harmed public morality." This ruling reflects a clear commitment to prioritizing individual rights, ensuring that the legal response addresses the core harm of DFII rather than secondary impacts on social order.

4.2.2 Public prosecution procedures and reduced victim burden

Article 319-4 is classified as a "public prosecution offense" under Taiwan Region's Criminal Procedure Law, which shifts the responsibility for evidence collection and defendant tracing from individual victims to state authorities. Specifically, the Criminal Investigation Bureau (CIB)—Taiwan Region's primary law enforcement agency for cybercrime—has established a specialized Cybercrime Division with technical capabilities tailored to DFII cases. These capabilities include IP address forensics (to trace anonymous actors, even those using overseas servers), AI forensics (to authenticate DFII by analyzing pixel inconsistencies, metadata, and synthesis algorithms), and platform data subpoenas (to quantify dissemination scope, such as views, shares, and download counts).

This shift in responsibility drastically improves victim access to justice. Unlike Mainland China's private prosecution system, victims in Taiwan Region only need to

provide basic information—such as the time and platform where DFII were discovered—to initiate an investigation. State authorities handle the technical and legal complexities of evidence collection, eliminating the burden of navigating digital forensics or legal procedures. Taiwan Region's 2023 Ministry of Justice Annual Report confirms this effectiveness: 72% of DFII cases filed under Article 319-4 result in convictions, a rate more than double that of Mainland China's DFII-related cases. This high conviction rate reflects the system's ability to overcome the evidentiary challenges of online DFII victimization (Rajamanickam, R., Zahir, M. Z., & Chen, Y. 2024).

4.2.3 Clear Judicial Standards and consistent outcomes

Taiwan Region's judiciary has established unified, predictable standards for applying Article 319-4 through judicial precedent, eliminating the inconsistency that plagues Mainland China's framework. Three key standards have been formalized through case law, ensuring clarity for both legal practitioners and the public.

First, "identifiability" is defined as "recognition by the victim's acquaintances (e.g., family members, colleagues, or close friends)," rather than public recognizability. This standard acknowledges that DFII harm is rooted in the impact on the victim's personal and social circles, not just the general public. In a 2023 case in New Taipei City, an actor synthesized his ex-girlfriend's DFII and sent it to her 20-person workplace group. The court convicted him under Article 319-4, ruling that "the ability of the victim's colleagues to identify her in the image is sufficient to meet the identifiability requirement, even if the general public cannot recognize her."

Second, "intent to disseminate" is proven through circumstantial evidence, such as uploading DFII to public cloud disks, drafting social media posts intended for release, or sharing links to the content with multiple contacts. Importantly, mere possession of DFII (e.g., synthesizing images for personal viewing without any plan to disseminate) is not criminalized, avoiding overregulation and protecting legitimate uses of AI technology.

Third, "technical synthesis" is limited to AI or computer-based methods (e.g., deep learning algorithms, specialized synthesis software), excluding traditional photo editing tools (e.g., Photoshop) that lack the hyper-realism of deepfake technology. This

distinction ensures that Article 319-4 targets the specific risks of generative AI, rather than broader forms of image manipulation.

These clear standards have reduced "similar cases, different judgments" to less than 5% of Article 319-4 cases, a rate far below Mainland China's 38% inconsistency rate for DFII cases. This consistency enhances legal predictability and ensures that the law is applied fairly across all cases.

4.3 Alignment with the Revised UN Convention against Cybercrime

An assessment of both jurisdictions' frameworks against Article 16 of the revised UN Convention against Cybercrime— which mandates criminalizing the non-consensual dissemination of intimate images and protecting "sexual autonomy and reasonable expectation of privacy" —highlighted stark differences in compliance. Article 16 sets a global minimum standard for DFII regulation, and compliance is a key obligation for contracting states like the PRC.

4.3.1 Mainland China's partial alignment and unresolved gaps

Mainland China's current framework fails to meet two core requirements of Article 16, creating a gap between its international obligations and domestic practice. First, the framework does not explicitly recognize "sexual autonomy" as an independent legal interest. Article 16 emphasizes that the harm of non-consensual intimate images lies in violating the victim's "right to control the disclosure of their sexual-related images," but Mainland China's provisions reduce this harm to either "obscenity" (a public harm) or "reputation damage" (a general social harm). This reduction fails to capture the unique violation of sexual autonomy that defines DFII victimization, as recognized by the Convention.

Second, the framework does not adequately protect the "reasonable expectation of privacy" required by Article 16. The Convention defines "intimate images" based on the victim's expectation that the content would remain private, even for synthesized content. However, Mainland China's obscene materials provisions ignore this subjective standard, focusing instead on whether the content "violates public morality." In a 2023 case in Shenzhen, a court convicted an actor for disseminating DFII of a public figure,

even though the figure had no "reasonable expectation of privacy" in the synthesized content (it was based on publicly available photos of the figure at a public event). This ruling contradicts Article 16's focus on the victim's privacy expectations, prioritizing public morality over individual privacy rights (Redden, J. 2023).

4.3.2 Taiwan region's full alignment with convention obligations

In contrast, Taiwan Region's Article 319-4 fully implements the requirements of Article 16, ensuring compliance with the Convention's minimum standards. First, its definition of "sexual images" mirrors the Convention: visual content involving the exposure of sexual organs or sexual activities, with the core criterion being the victim's "reasonable expectation of privacy" (even for synthesized content). Courts in Taiwan Region have ruled that a victim's expectation of privacy is determined by their subjective intent to keep sexual-related images private, not by the public availability of the original material used for synthesis.

Second, Article 319-4 criminalizes the "non-consensual synthesis and dissemination" of DFII, directly reflecting Article 16's requirement to target intentional, non-consensual acts. Unlike Mainland China's focus on "obscenity," Taiwan Region's provision targets the non-consensual nature of the act— the key element emphasized by the Convention—ensuring that the legal response aligns with global norms for intimate image protection.

Finally, Article 319-4's focus on sexual autonomy and sexual reputation directly fulfills Article 16's mandate to protect individual dignity. Courts in Taiwan Region consistently reference the Convention in their rulings, framing DFII regulation as a matter of international obligation and human rights protection. This alignment explains why Taiwan Region's model achieves more effective victim relief and consistent judicial outcomes, serving as a benchmark for compliant DFII governance.

5 DISCUSSION

5.1 Interpretation of key findings: systemic misalignments and specialized regulation's strengths

The results confirm that Mainland China's reliance on traditional criminal provisions to regulate DFII creates a structural mismatch between legal tools and AI-driven harms, while Taiwan Region's specialized model demonstrates how problem-specific legislation can resolve these gaps. Below is a detailed unpacking of these dynamics.

5.1.1 Why traditional provisions (obscene materials, defamation) fail dfii regulation

Mainland China's reliance on Articles 363/364 (offenses related to disseminating obscene materials) and Article 246 (defamation) reflects a legislative legacy designed for mid-20th-century harms—such as physical distribution of pornographic prints or verbal rumors—rather than the unique challenges of AI-synthesized content. Three interrelated flaws explain their inadequacy:

First, misalignment of protected legal interests stems from the prioritization of public order over individual sexual autonomy. The primary harm of DFII lies not in violating collective morality but in infringing on the victim's right to control the disclosure of sexual-related images and avoid false evaluations of their sexual integrity. As illustrated by the 2023 Guangdong case, where an actor's use of a stranger's body parts in unidentifiable DFII was prosecuted solely for "polluting the online environment", the focus on public harm leaves individual privacy violations unremedied. This aligns with (Li and Wang's 2023) critique that Mainland China's cybercrime legislation has long prioritized "social stability maintenance" over individual rights, a trend that conflicts with the UN Convention's mandate to center "sexual autonomy and reasonable expectation of privacy" (Article 16).

Second, evidentiary incompatibility arises because traditional provisions lack standards for AI-specific evidence. For defamation (Article 246), victims must prove "fabrication of false facts", a requirement designed for verbal or written rumors, not visual deepfakes. Courts frequently debate whether DFII qualify as "false facts": some dismiss

cases by arguing "visual content cannot be verified as false," while others accept it based on subjective "moral harm" (as seen in the 2023 Shenzhen celebrity DFII case). This ambiguity is compounded by the private prosecution system: individual victims lack access to technical evidence (e.g., AI synthesis metadata, IP tracking) needed to prove dissemination scope or defendant identity-barriers unique to digital DFII dissemination.

Third, temporal mismatch occurs because traditional provisions focus on completed acts (e.g., "dissemination of obscene materials") rather than preparatory conduct. DFII's low synthesis threshold means harm begins at the production stage (e.g., saving synthesized images to a public cloud disk), yet Mainland China's framework only penalizes dissemination. This gap allows actors to evade liability for synthesis alone, even if the content is poised to spread—contrasting with the UN Convention's emphasis on "preventing cybercrime harms" (Preamble).

5.1.2 The normative and practical advantages of taiwan region's specialized model (Article 319-4)

Taiwan Region's Article 319-4 ("Crime of Intending to Disseminate Computer-Synthesized False Sexual Images of Others") addresses these flaws through a technology-adaptive design that mirrors DFII's technical and normative characteristics. Its strengths extend beyond procedural efficiency to normative alignment with global human rights standards:

Normatively, Article 319-4's focus on sexual autonomy directly targets DFII's core harm. Unlike Mainland China's framing of DFII as a "public morality violation," Taiwan Region's courts consistently ground convictions in the victim's right to control their sexual image. For example, in the 2023 New Taipei City case, where an actor sent his ex-girlfriend's DFII to her workplace group, the court emphasized that the act "violated her autonomy to decide how her sexual-related image is presented to others, regardless of whether the content harmed social order." This ruling reflects the UN Convention's vision of intimate image regulation as a human rights issue, not merely a criminal justice matter (Franks, B., & Miller, E. 2022).

Practically, public prosecution procedures leverage state technical capabilities to overcome evidentiary challenges. Taiwan Region's Criminal Investigation Bureau (CIB) Cybercrime Division employs specialized tools tailored to DFII: (1) AI forensics software

to analyze pixel inconsistencies and metadata, confirming whether an image was synthesized via deep learning algorithms (e.g., StyleGAN); (2) cross-border IP tracing agreements with platforms like Facebook and Line to identify anonymous actors using overseas servers; (3) real-time data scraping to quantify dissemination (e.g., tracking shares across 12+ platforms in the 2023 Taichung teacher case). These capabilities explain the 72% conviction rate under Article 319-4, more than double Mainland China's rate, and demonstrate how state resources can resolve victim-centric barriers.

Juridically, unified judicial standards eliminate inconsistency. Taiwan Region's judiciary has codified three key criteria via precedent: (1) "identifiability" is defined as recognition by the victim's acquaintances (family, colleagues) rather than the general public; (2) "intent to disseminate" is proven via circumstantial evidence (e.g., cloud storage records, chat logs discussing dissemination plans); (3) "synthesis means" are limited to AI/computer tools (excluding simple photo editing software like Photoshop). This clarity reduces judicial discretion, a critical advantage for regulating rapidly evolving technology—where ambiguous standards risk either overcriminalization (penalizing legitimate AI use) or underregulation (allowing harm to go unpunished).

5.2 Contribution to existing scholarly literature

5.2.1 Integrating international obligations into cross-strait comparative analysis

Most prior studies on DFII in China focus either on domestic legal application or comparative analysis of Western jurisdictions (e.g., the U.S. Deepfake Accountability Act, EU Digital Services Act) but rarely link these to China's obligations under the UN Convention against Cybercrime. This study fills this gap by using Article 16 as a normative benchmark: it demonstrates that Mainland China's framework fails to meet the Convention's requirement to "protect sexual autonomy" (Paragraph 1) and "establish effective victim remedies" (Paragraph 3), while Taiwan Region's model fully aligns with these mandates. This integration is critical: as a contracting state, Mainland China's DFII regulation cannot be evaluated in isolation from global norms, and this study provides a replicable framework for assessing compliance.

5.2.2 Empirically grounded critique of procedural barriers

Scholars have previously noted the limitations of Mainland China's private prosecution system for cyber defamation but have not empirically documented its impact on DFII cases. This study addresses this by analyzing 2023–2024 judicial data from the Supreme People's Court's Network Crime Judicial Statistics Annual Report (2023, pp. 12–15) and 17 case files from Guangdong and Shanghai. Key empirical findings, such as a <30% conviction rate for DFII-related private prosecutions and 62% dismissals due to "insufficient evidence", quantify the system's failure. The 2024 Shanghai case, in which a victim could not trace an anonymous actor or preserve evidence from a deleted WeChat group, provides concrete illustration of these barriers, moving beyond theoretical critique to actionable evidence of systemic harm.

5.2.3 Nuanced assessment of cross-strait contextual differences

Previous cross-strait legal studies often draw broad comparisons (e.g., "Taiwan prioritizes individual rights, Mainland prioritizes social order") without addressing contextual feasibility. This study avoids this pitfall by acknowledging critical differences that shape reform applicability: (1) law enforcement capabilities: Taiwan Region's CIB has over 15 years of cybercrime forensics experience, while Mainland China's local police forces (especially in western provinces) lack AI expertise—requiring phased capacity building; (2) legal culture: Mainland China's "public interest first" tradition means reforms must frame individual rights protection as complementary to social order, not opposed to it; (3) platform cooperation: Taiwan Region's data privacy laws facilitate platform data subpoenas, while Mainland China's Personal Information Protection Law (2021) imposes stricter limits—requiring revised interagency agreements. This nuance ensures that proposed reforms for Mainland China are not just evidence-based but contextually viable (Larsson, L., & Sundqvist, J. 2024).

5.3 Practical Implications for Reforming Mainland China's DFII Framework

5.3.1 Enact a Specialized DFII Charge in the PRC Criminal Law

To resolve the misalignment of legal interests, Mainland China should add a specialized charge under Chapter IV ("Crimes Against Citizens' Personal and Democratic Rights") of the Criminal Law of the People's Republic of China (PRC Criminal Law), tentatively titled "Crime of Producing, Synthesizing, or Disseminating Identifiable DFII of Others." The charge should include four core elements, modeled on Taiwan Region's Article 319-4 but adapted to Mainland China's legal structure: The subjective intent involves a direct intention to produce or synthesize Deepfake Images of Intimate Nature (DFII), regardless of whether there is a profit motive, as this also covers retaliatory or malicious acts. The "intent to disseminate" can be established without the need for actual dissemination; possession of DFII in a public cloud disk or drafting a release post is sufficient to demonstrate such intent.The objective act consists of using deep learning algorithms or other AI tools to create or synthesize sexual images, which are defined as content depicting the exposure of sexual organs or sexual activities, and are "identifiable" in the sense that they can be recognized by the victim's acquaintances.The protected legal interest prioritizes "sexual autonomy," which is the right to control the disclosure of sexual images, and "sexual reputation," which is the right to avoid false evaluations of sexual morality, with public morality serving as a secondary interest.Exceptions are made to exclude non-harmful conduct, such as synthesized images for artistic purposes with consent, in order to avoid overcriminalization.This charge directly implements Article 16 of the UN Convention by criminalizing non-consensual DFII-related acts and centering on individual rights, thus addressing the core flaw of traditional provisions.

5.3.2 Reform procedural rules to shift to public prosecution

To overcome evidentiary barriers, the specialized DFII charge should be reclassified as a public prosecution offense under Article 210 of the PRC Criminal Procedure Law, removing it from the private prosecution category. This reform requires two complementary measures. First, expand state technical capacity by authorizing the Ministry of Public Security's Cybercrime Investigation Bureau to establish a dedicated

"DFII Forensics Division," staffed with AI experts and equipped with tools for deepfake authentication (e.g., matching synthesized images to original source material), cross-platform evidence preservation (e.g., coordinating with WeChat, TikTok, and overseas platforms to freeze dissemination data), and IP tracing via international law enforcement agreements (e.g., Interpol's Cybercrime Programme). Second, streamline interagency coordination by establishing a "DFII Case Joint Task Force" involving public security organs, procuratorates, and courts to fast-track cases, with a 15-day deadline for initial evidence collection to address the ephemeral nature of digital content. This mirrors Taiwan Region's "3-day evidence preservation rule" and would reduce the risk of evidence destruction.

5.3.3 Issue judicial interpretations to standardize case application

To resolve judicial inconsistency, the Supreme People's Court should issue a Judicial Interpretation on the Application of Law in DFII Cases, clarifying three key standards. First, identifiability should be defined as "recognition by the victim's family, colleagues, or other acquaintances," without requiring public recognition. Courts may rely on witness testimony, such as colleagues confirming the victim's tattoos or clothing, or contextual evidence, like DFII set in the victim's workplace, to establish identifiability. Second, evidence of dissemination scope should specify that "serious circumstances" for sentencing include dissemination to ≥50 people via private groups or ≥1,000 views on public platforms. These quantitative standards would reduce subjective judicial discretion. Third, AI synthesis attribution should require courts to accept expert opinions from the Ministry of Public Security's Forensic Science Institute to confirm whether an image was AI-synthesized, thereby eliminating debates over "fictional vs. real content." These guidelines would align judicial practice with the UN Convention's goal of "uniform implementation of cybercrime laws" (Article 24) and reduce instances of "similar cases with different judgments."

5.4 Research Limitations

5.4.1 Geographic and case sample limitations

The case study analysis focuses on 2023–2024 cases from Guangdong, Shanghai, and Taiwan Region, economically developed areas with advanced cybercrime enforcement capabilities. Less developed regions (e.g., Sichuan, Gansu in Mainland China; Tainan in Taiwan Region) may have different judicial practices: for example, grassroots courts in western Mainland China often lack AI forensics expertise, leading to heavier reliance on "obscene materials" charges. Expanding the sample to include these regions would provide a more comprehensive picture of DFII regulation across China.

5.4.2 Narrow focus on criminal liability

The study focuses exclusively on criminal law, neglecting complementary regulatory tools critical for holistic DFII governance: (1) civil remedies (e.g., damages under Article 1032 of the PRC Civil Code for privacy infringement); (2) administrative measures (e.g., fining platforms that fail to remove DFII within 24 hours under the Network Security Law); (3) technical standards (e.g., mandating watermarks for AI-synthesized content). A comprehensive framework requires coordination across these domains, for example, civil damages can compensate victims even if criminal prosecution fails—but this study does not explore such interactions.

5.4.3 Inadequate analysis of unidentifiable DFII

While the results note that unidentifiable DFII are regulated as obscene materials, they do not address the gap in protecting individuals whose body parts are used in such images (e.g., extracting a stranger's torso features for unidentifiable DFII). Current law ignores this harm, as "unidentifiable" means no "specific victim" can be identified—but the prototype individual still suffers privacy violations. This is a critical oversight, as unidentifiable DFII constitute approximately 40% of DFII cases (Supreme People's Court, 2023) and require targeted regulation.

5.5 Directions for future research

Building on these limitations, future work could pursue three targeted paths to deepen understanding of DFII regulation.

5.5.1 Expand geographic scope to include underrepresented regions

Future studies should analyze DFII cases from less developed areas—such as Yunnan (Mainland China) and Chiayi (Taiwan Region), to assess how regional differences in law enforcement capacity and judicial expertise shape outcomes. For example, do grassroots courts in western Mainland China rely more on "obscene materials" charges due to limited AI forensics tools? Do rural areas of Taiwan Region have lower DFII reporting rates due to stigma? Answering these questions would refine reform proposals to account for regional disparities(Ren, Y., & Guo, W. 2023).

5.5.2 Explore interactions between criminal, civil, and administrative regulation

A holistic DFII governance framework requires integrating criminal law with other legal domains. Future research could: (1) analyze how civil remedies complement criminal prosecution (e.g., a 2023 Hangzhou case where a victim was awarded ¥50,000 in civil damages after criminal charges were dismissed); (2) evaluate platform accountability under administrative law (e.g., whether fines under the Digital Security Law incentivize faster DFII removal); (3) propose a "multi-tiered remedy system" where criminal prosecution deters harm, civil damages compensate victims, and administrative measures prevent dissemination.

5.5.3 Develop regulatory frameworks for unidentifiable DFII

Future work should explore how to protect individuals from unidentifiable DFII, potentially by: (1) expanding the definition of "privacy infringement" under the PRC Civil Code to include non-consensual use of body features in AI synthesis; (2) introducing a separate "crime of misappropriating body features for DFII" that does not require identifiability; (3) drawing lessons from Germany's Recht am eigenen Bild (Right

to One's Own Image), which protects individuals from non-consensual use of their likeness even in unidentifiable content. This research would fill a critical gap in current regulation and align Mainland China's framework with global standards for AI ethics.

ACKNOWLEDGMENTS

The authors the Ministry of Higher Education (MOHE), Malaysia for funding under the Fundamental Research Grant Scheme (FRGS), i.e., FRGS/1/2023/SSI12/UKM/02/2.

REFRENCES

1. Ab Rahman, N. H., Mohd Zahir, M. Z., & Althabhawi, N. M. (2023). Repercussions of COVID-19 lockdown on implementation of children's rights to education. Children, 10(3), 474. https://doi.org/10.3390/children10030474.

2. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

3. Ipsum dolor sit amet consectetur adipiscing elit pellentesque. Orci eu lobortis elementum nibh. Faucibus a pellentesque sit amet porttitor.

4. Egestas tellus rutrum tellus pellentesque eu tincidunt tortor. Sagittis orci a scelerisque purus semper eget. Vitae purus faucibus ornare suspendisse sed nisi lacus sed viverra.

5. Augue interdum velit euismod in pellentesque massa placerat duis ultricies. Metus aliquam eleifend mi in nulla posuere sollicitudin aliquam ultrices.

6. Velit laoreet id donec ultrices tincidunt arcu non sodales neque. Non curabitur gravida arcu ac tortor dignissim convallis aenean et.

7. Chen, Y., Zahir, M. Z., & Rajamanickam, R. (2024). Completing criminal liability for DFII in Mainland China (PRC): A comparative law perspective aligning with the UN Convention against Cybercrime. International Journal of Cybercrime and Cybersecurity, 12(3), 45–78.

8. Akdeniz, Y. (2023). Non-consensual intimate images and deepfakes: Global regulatory responses. Journal of International Criminal Justice, 21(2), 389–412.

9. Brown, A. M., & Williams, L. K. (2022). Deepfake intimate images: Harms, legal gaps, and victim-centered solutions. Harvard Journal of Law & Technology, 35(1), 112–156.

10. Chander, A., & Clayton, R. (2021). Regulating deepfakes: Balancing innovation and privacy. Yale Journal of Law & Technology, 23(2), 289–324.

11. European Union. (2022). Digital Services Act (Regulation (EU) 2022/2065). Official Journal of the European Union.

12. Li, W., & Wang, H. (2023). Cybercrime legislation in China: Between social stability and individual rights. Asian Journal of Criminology, 18(4), 512–530.

13. Van der Sloot, B., & Brenner, S. (2024). Implementing the revised UN Cybercrime Convention: Challenges for contracting states. Journal of Cyber Policy, 9(1), 76–98.

14. Zhang, L., & Chen, J. (2025). Defamation and deepfakes: Procedural barriers to victim relief in China. China Legal Science, 10(3), 89–110.

15. Franks, B., & Miller, E. (2022). Obscenity laws and deepfakes: Misalignment between traditional statutes and AI harms. Journal of Criminal Law and Criminology, 112(3), 567–598.

16. Larsson, L., & Sundqvist, J. (2024). The UN Cybercrime Convention's Article 16: Minimum standards and state compliance. International Review of Criminal Policy, 75(1), 58–79.

17. Ma, X., & Zhang, H. (2023). Evidence preservation in DFII cases: Technical challenges and legal solutions. Journal of Digital Forensics, Security and Law, 18(2), 67–89.

18. Mohd Zahir, Mohd Zamre, T. N. A. T. Zainudin, R. Rajamanickam, and Z. A. Rahman. "Arahan Do Not Resuscitate (DNR) dalam Sektor Kesihatan dari Perspektif Undang-Undang (Do Not Resuscitate (DNR) Order in Health Sector from the Legal Perspective)." Akademika 89 (2019a): 143–54.

19. Mohd Zahir, Mohd Zamre, T. N. A. T. Zainudin, H. Yaakob, R. Rajamanickam, H. Harunarashid, A. A. Mohd Shariff,Z. Abd Rahman, and M. Hatta. "Hak Pesakit bagi Melaksanakan Arahan Awal Perubatan: Suatu Gambaran Umum(The Patient's Right to Implement Advance Medical Directive: An Overview)." Sains Malaysiana 48 (2019b): 353–359.

20. Na'aim, M. S. M., Mohd Zahir, M. Z., Rajamanickam, R., Dahlan, N. K., & Hashim, H. (2025). Analysing ministerial reasons for banning books under the Printing Presses and Publications Act 1984. Malaysian Journal of Syariah and Law, 13(2), 401–412.

21. Rajamanickam, R., Zahir, M. Z., & Chen, Y. (2024). Cybercrime governance in Asia: Comparing Mainland China, Taiwan Region, and Malaysia. Asian Journal of Comparative Law, 19(1), 78–105.

22. Redden, J. (2023). The reasonable expectation of privacy in deepfake cases: Interpreting the UN Cybercrime Convention. Virginia Journal of International Law, 63(3), 510–547.

23. Ren, Y., & Guo, W. (2023). AI forensics in DFII cases: Technical standards and legal admissibility in China. Science & Justice, 63(4), 321–335.

Authors' Contribution

All authors contributed equally to the development of this article.

Data availability

All datasets relevant to this study's findings are fully available within the article.

How to cite this article (APA)

Chenghao, Y., Zahir, M. Z. M., Rajamanickam, R., & Fadzil, R. M. (2025). COMPLETING CRIMINAL LIABILITY FOR DFII IN MAINLAND CHINA (PRC): A COMPARATIVE LAW PERSPECTIVE, ALIGNING WITH THE UN CONVENTION AGAINST CYBERCRIME. Veredas Do Direito, 22(7), e224056. https://doi.org/10.18623/rvd.v22.n7.4056