

Artificial Intelligence and Sexual Offences: An Analysis of Deepfake Pornography in Light of Criminal Law

Túlio Felipe Xavier Januário

ORCID ID: <http://orcid.org/0000-0003-0400-1273>
Faculty of Law, University of Coimbra
Pátio da Universidade
Faculdade de Direito da Universidade de Coimbra
3004528 Coimbra, Portugal
<https://ror.org/04z8k9a98>
E-mail: tuliofxj@gmail.com

Artificial Intelligence and Sexual Offences: An Analysis of Deepfake Pornography in Light of Criminal Law

Túlio Felipe Xavier Januário

(University of Coimbra (Portugal))

This paper explores the legal implications of non-consensual sexual deepfakes, specifically analyzing whether the creation, distribution, exhibition, or possession of such content involving adults or minors could be considered sexual offenses under the Portuguese law. By applying a deductive methodology, the study reviews Portuguese, European, North American, and Brazilian legal frameworks, doctrines, and case law related to sexual crimes and *Artificial Intelligence* (AI), applying them to the issue of deepfakes.

The research begins by discussing the increasing prevalence and impact of non-consensual sexual deepfakes, a form of digital manipulation where AI is used to superimpose a person's face onto someone else's body in fabricated explicit content. With the rise of easily accessible deepfake technology, the realism of such videos has made it increasingly difficult to distinguish them from authentic material, resulting in significant harm to victims, both adults and minors. This includes emotional and reputational damage, as well as severe psychological consequences such as PTSD, anxiety, and depression, particularly among adults, and developmental harm in the case of minors. The research also highlights the alarming potential for deepfakes to depict child sexual abuse, exacerbating concerns about the exploitation of minors.

The study investigates whether non-consensual sexual deepfakes could be classified as criminal offenses under the Portuguese law. It finds that, while the current legal provisions effectively address deepfakes involving minors, a distinction must be made between two types of virtual child pornography: fully virtual representations, and those involving partially real images of minors. Only the latter, where real children's images are used, should be classified as a crime, as it violates the minor's development conditions.

For deepfakes involving adults, however, the research identifies a legal gap. Existing offenses, such as 'computer fraud', 'illicit recordings or photographs', or 'aggravated defamation', may cover some aspects of deepfake cases, but they do not adequately address the unique harm caused by non-consensual sexual deepfakes. The paper argues that a new criminal offense should be introduced specifically to protect sexual privacy and prevent the creation or distribution of non-consensual sexual deepfakes involving adults. Drawing on the Brazilian legal framework as a model, the study suggests the implementation of tailored provisions that criminalize these acts in order to safeguard the victims' sexual privacy.

Received: 09/01/2025. **Accepted:** 31/03/2025

Copyright © 2025 Túlio Felipe Xavier Januário. Published by Vilnius University Press

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The study concludes by emphasizing the need for a legal reform that would balance technological advancements with individual rights. While the paper cautions against excessive expansion of criminal law or a moralistic approach to sexual crimes, it advocates for the creation of specific legal protections against non-consensual sexual deepfakes, while acknowledging the enduring societal impact of this issue. These solutions would help address the harmful effects of deepfakes and provide better legal safeguards for victims of this rapidly evolving digital phenomenon.

Keywords: Criminal law, Portuguese law, artificial intelligence, sexual offences, deepfake.

Dirbtinis intelektas ir seksualiniai nusikaltimai: giliųjų klastočių (*deepfake*) pornografijos analizė baudžiamosios teisės kontekste

Túlio Felipe Xavier Januário

(Koimbros universitetas (Portugalija))

Straipsnyje nagrinėjami be asmens sutikimo sukurtų seksualinio pobūdžio giliųjų klastočių (angl. *deepfake*) naudojimo padariniai, vertinama, ar tokio turinio – suaugusiųjų ar nepilnamečių atvaizdų – kūrimas, platinimas, demonstravimas ar laikymas gali būti pripažįstamas seksualiniu nusikaltimu pagal Portugalijos teisę. Pasitelkus dedukcijos metodą, tyrime peržiūrimos Portugalijos, Europos, Šiaurės Amerikos ir Brazilijos teisės sistemos, teisės doktrinos ir teismų praktika, susijusi su seksualiniais nusikaltimais ir dirbtiniu intelektu (DI), ir jų taikymas giliųjų klastočių problematikai.

Tyrime visų pirma aptariamas be asmens sutikimo sukurtų seksualinių giliųjų klastočių – vaizdų, kuriuose, pasitelkiant dirbtinį intelektą, asmens veidas įkomponuojamas į kito asmens kūną – paplitimas ir poveikis. Dėl lengvos prieigos prie giliųjų klastočių technologijos ir šių vaizdų tikroviškumo tampa sudėtinga juos atskirti nuo autentiškos medijos, o tai daro didelę žalą aukoms – tiek suaugusiems asmenims, tiek nepilnamečiams. Tokia žala – tai emociniai ir reputaciniai nuostoliai, sunkūs psichologiniai padariniai, įskaitant potrauminio streso sutrikimą, nerimą ir depresiją suaugusiems, o nepilnamečiams gali sukelti raidos sutrikimų. Tyrime taip pat atkreipiamas dėmesys į pavojus, kylančius dėl giliųjų klastočių naudojimo vaizduojant vaikų seksualinį išnaudojimą, kas dar labiau didina susirūpinimą nepilnamečių apsauga.

Straipsnyje vertinama, ar be asmens leidimo sukurtų seksualinių giliųjų klastočių turinys gali būti laikomas nusikalstamu pagal Portugalijos teisę. Nustatyta, kad galiojantis teisinis reglamentavimas pakankamai veiksmingai apima nepilnamečius vaizduojančias giliąsias klastotes, tačiau būtina diferencijuoti dvi virtualios vaikų pornografijos kategorijas: visiškai dirbtinius vaizdus ir iš dalies tikrus nepilnamečių atvaizdus. Nusikalstama veika turėtų būti laikoma tik pastaroji kategorija, kadangi ji tiesiogiai pažeidžia nepilnamečių teises ir daro neigiamą poveikį jų psichologinei raidai. Straipsnyje teigiama, kad būtina įtvirtinti naują baudžiamąją veiką, skirtą seksualinio privatumo apsaugai, draudžiančią nepageidaujamo seksualinio giliųjų klastočių turinio kūrimą ar platinimą suaugusiems asmenims.

Kalbant apie suaugusiuosius, tyrime atskleidžiamas teisinis trūkumas. Nors tam tikrais atvejais tokios nusikalstamos veikos kaip „kompiuterinis sukčiavimas“, „neteisėtas įrašų ar nuotraukų darymas“ ar „šmeižtas sunkinančiomis aplinkybėmis“ gali apimti giliųjų klastočių kūrimą ar platinimą, jos nepakankamai atspindi specifinę žalą, kurią sukelia nepageidaujamas seksualinis giliųjų klastočių turinys. Straipsnyje teigiama, kad būtina įtvirtinti naują baudžiamąją veiką, skirtą seksualinio privatumo apsaugai ir užkertančią kelią be sutikimo sukurtų seksualinių giliųjų klastočių kūrimui ar platinimui, susijusiam su suaugusiaisiais. Remiantis Brazilijos teisės sistema kaip pavyzdžiu, siūloma įtvirtinti specialias teises nuostatas, kurios kriminalizuotų šias veikas ir užtikrintų aukų seksualinio privatumo apsaugą.

Tyrime daroma išvada, kad būtina teisės reforma, kuri užtikrintų technologijų pažangos ir asmens teisių apsaugos pusiausvyrą. Nors straipsnyje įspėjama dėl pernelyg plataus baudžiamosios teisės taikymo ar moralizuojančio požiūrio į seksualinius nusikaltimus, jame pasisakoma už konkrečias teises apsaugos priemones, skirtas kovoti su nepageidaujamu seksualiniu giliųjų klastočių turiniu ir geriau apsaugoti nukentėjusiuosius nuo šio sparčiai besivystančio skaitmeninio reiškinio.

Pagrindiniai žodžiai: baudžiamoji teisė, Portugalijos teisė, dirbtinis intelektas, seksualiniai nusikaltimai, gilioji klastotė.

Introduction

The impact of artificial intelligence across various sectors of society is undeniable. Its ability to process vast amounts of data rapidly and make highly accurate autonomous decisions has made it highly promising in fields such as transportation, healthcare, justice administration, and law enforcement, among others. However, alongside its potential, significant concerns and uncertainties arise regarding

its limitations and associated risks. Notable issues include privacy violations, the potential for discriminatory biases in decision-making, and challenges in human comprehension and contestation of AI-generated outcomes.

This investigation¹ specifically focuses on one of these dangers: the use of AI in criminal activities². More specifically, we explore the growing issue of non-consensual sexual deepfakes, which involve hyper-realistic manipulation of a person's face onto someone else's body to create fabricated pornographic content (Spivak, 2019, p. 339). Although this technology has existed for some time, it became more prominent in 2017, when a *Reddit* user released a program that allowed even those with minimal technical expertise to create deepfake videos (Spivak, 2019, p. 345). The widespread use of this technology has led to deepfake pornography targeting both celebrities and private individuals worldwide, causing significant reputational and emotional harm. Moreover, advancements in AI have made it increasingly difficult to distinguish between real and fake videos, thereby intensifying the problem (Rodrigues, 2023). The situation becomes even more alarming when considering its potential use in creating virtual depictions of sexual activities involving minors.

This research aims to explore the role of criminal law in addressing these issues. Specifically, it will investigate whether the production, distribution, exhibition, or possession of non-consensual sexual deepfakes, involving either adults or minors, could be considered sexual offences under Portuguese law. The investigation will also address the obstacles to classifying these actions as crimes and propose possible solutions.

To address these questions, the study will first analyze the technology behind deepfake videos and how AI has advanced their creation, distribution, and associated risks. It will then examine whether non-consensual sexual deepfakes could be considered crimes, focusing on both minor and adult victims. The hypothesis is that while these cases are significant, there are ongoing debates regarding whether the damage caused justifies criminal protection, especially when viewed through the lens of the 'legal good' (*Rechtsgut*) principle. In cases involving minors, it will be necessary to differentiate between situations in which realistic images of non-existent underaged individuals are represented, from those in which there is, at least partially, the use of real minors' images. Only in the latter case, can we speak of a violation of the minor's development conditions. As for deepfakes involving adults, there is currently a significant legal gap to be observed. Given the harm caused by these cases, the paper argues for the creation of a new criminal offence to protect sexual privacy from non-consensual sexual deepfakes.

To achieve these objectives, the study will use a deductive approach, analyzing Portuguese, European, North American, and Brazilian legislation, doctrine, and case law regarding sexual crimes and AI, applying them to the issue of non-consensual sexual deepfakes. The research will primarily draw from the European legal doctrine, particularly Portuguese sources, but will also reference comparative law wherever helpful. Given the advanced state of research in AI and deepfakes, the study will also incorporate insights from North American academic work on these topics.

¹ This investigation was carried out within the scope of the project entitled "Autoria e responsabilidade em crimes cometidos através de sistemas de inteligência artificial", funded by the *Fundação para a Ciência e a Tecnologia – FCT* (2020.08615.BD). For a broader and more in-depth analysis of the topic in question, see: Janeiro, 2024.

² As it can be seen, the focus of the paper is on situations in which AI is deliberately used to commit crimes. For a detailed analysis of criminal liability for situations in which the production or use of this technology causes unintentional harm, see: Janeiro, 2020, 2021, 2022, 2023.

1. Comprehending Non-Consensual Sexual Deepfakes

Artificial Intelligence (AI) systems are designed to make inferences and generate outputs – such as predictions, content, or decisions – based on data inputs (OECD, 2024, p. 4). These systems, which can function with varying degrees of autonomy, are capable of perceiving their environment, interpreting data, reasoning, and deciding on the best course of action based on predefined parameters. Additionally, AI can be programmed to learn from its past actions and adapt its behavior (Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission, 2019, p. 6).

The technology's ability to process large amounts of data rapidly and generate precise outputs makes it highly applicable across multiple sectors, aiding human activities in transportation, healthcare, law enforcement, and more. However, AI also presents significant risks, including its opaque nature (Burrell, 2016, p. 1; Price II, 2017, p. 10; Rodrigues, 2020, p. 25), concerns about the security, accuracy, and legality of the input data (Miró Llinares, 2018, p. 122 ff.), and the unpredictability of the resulting outputs, which can frequently turn out to be unlawful (Sousa, 2020, p. 64).

Generally, deepfake videos can be understood as those that utilize artificial intelligence to place one person's face onto another person's body. These videos can serve both legal and illegal purposes, including the creation of non-consensual sexual content (Mania, 2024, p. 117). Deepfake technology relies on a method called *Generative Adversarial Networks* (GANs)³, which employs both generative and discriminative algorithms. In this process, the generative algorithm produces new frames, while the discriminative algorithm attempts to determine if the frames, which contain the face overlay, are genuine or fake. If the discriminative model fails to distinguish between real and fake images, humans will also be unable to tell the difference (Spivak, 2019, p. 344–345).

The spread of applications like *FakeApp* has made it easy for amateurs to create deepfake videos, contributing to a rise in non-consensual sexual content. Such videos can cause severe harm, both reputational and emotional, to their victims. Moreover, as technology improves, these videos are becoming more difficult to distinguish from real ones (Öhman, 2020, p. 133).

The creation of non-consensual sexual deepfakes can be classified within the broader category of image-based sexual abuse⁴, that is, as one of the possible forms of non-consensual creation and/or distribution of private sexual images (McGlynn and Rackley, 2017, p. 535–537). The creation of this content can be driven by various motives, extending those well beyond revenge, such as greed, voyeurism, and the need for personal validation (Franks, 2021, p. 2).

Although the generated images are not authentic, the harm they cause is very much real, and it is of little importance to the victims whether the content was originally sexual or altered (McGlynn *et al.*, 2017, p. 33). Research conducted between 2014 and 2015 with 18 victims revealed significant mental health issues, including PTSD, suicidality, anxiety, and depression (Bates, 2017, p. 30–34; Toparlak,

³ In that sense: “In the proposed adversarial nets framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles” (Goodfellow *et al.*, 2014, p. 1).

⁴ On that matter: “<...> the notion of the continuum underpins the development of our concept of image-based sexual abuse which encompasses all forms of the non-consensual creation and/or distribution of private sexual images. It includes, therefore, a range of abusive behaviours beyond the familiar example of ‘revenge porn’, such as ‘sexualised photoshopping’, sexual extortion (often labelled as ‘sextortion’), ‘upskirting’, voyeurism and many other similar forms of sexualised abuse. We developed the concept in order to bring into focus the overlapping nature of various forms of abuse, including modes and motives of perpetration, and effects on victim-survivors” (McGlynn *et al.*, 2017, p. 28).

2023, p. 5). Additionally, the culture of ‘victim-blaming’, which is commonly seen in other forms of image-based sexual abuse, is also prevalent in cases of non-consensual pornography (Jurasz, Barker, 2021, p. 795; Uhl *et al.*, 2018, p. 62; Toparlak, 2023, p. 5).

It is common for such content to be shared alongside personal information about the victim, including their identity, contacts, social media profiles, and even baseless accusations about their character, thus making them targets for online harassment (McGlynn and Rackley, 2017, p. 545). Moreover, the consequences of these actions can also include financial harm, such as damage to the victim’s career or when the crime is perpetrated for extortion (Chesney and Citron, 2019, p. 1772–1773; Toparlak, 2023, p. 5). Deepfake content is rarely created just for personal gratification, but it also has a high potential for being used to blackmail, intimidate, and degrade its victims. This issue will likely worsen as the technology advances, making it harder to distinguish fake videos from real ones (Harris, 2019, p. 102).

In this context, the issue of virtual representation of sexually explicit behaviours involving minors becomes particularly significant. Although the creation of virtual child pornography may not directly involve the abuse of children, it holds the potential to lead to such abuse, which is why it is criminalized in many countries, particularly following the adoption of the Convention on Cybercrime and Directive 2011/92/EU (Custers, 2022, p. 205ff.).

Without a targeted legal response, there is no indication that the issue of pornographic deepfakes will subside. On the contrary, if technological advancements are any indication, we can expect these types of content to become more easily accessible and to be of higher technical quality (Gieseke, 2020, p. 1486).

2. The Approach of Portuguese Criminal Law to Deepfake Child Pornography

Sexual offenses are a good indicator on whether the modern criminal law is achieving its goals. On the one hand, the Social and Democratic State demands a clear identification of the specific legal interests protected by these offenses, while, on the other hand, the *Rechtsgut* must never be reduced to the scope of morals⁵, ethics, or religion. This is made clear by the evolution of the Portuguese legislation. While the sexual offences defined in the 1982 Penal Code were aimed at protecting ‘sexual morality’, the 1995 reforms redefined these crimes as offences against individuals, thus emphasizing that the legal interest being protected is of a personal and individual nature (Leite, 2016, p. 61–62).

Since the early 21st century, two opposing trends in sexual criminal law reforms have been evident. The first one seeks to separate sexual offences from sexual morality, viewing them instead as crimes against supra-individual legal goods (*Rechtsgüter*). Meanwhile, the other trend aims to extend criminal protection beyond the individual *Rechtsgut* of sexual freedom and self-determination. A notable example of the latter is the recent European and international directives on child pornography, which now include not only real children but also individuals who resemble children and realistic depictions of non-existent minors. As Maria João Antunes and Sónia Fidalgo warn, this shift may lead to a return to a sexual criminal law that emphasizes the protection of morals and ‘good customs’ (Antunes and Fidalgo, 2008, p. 121–123). In fact, many of the recent amendments to the Portuguese Penal Code regarding child pornography are the result of Portugal’s commitments to the European and international community, rather than arising from deeper doctrinal reflection on the matter (Antunes and Sousa, 2019, p. 239–240).

⁵ According to Luís Greco, the moral or sexual order should not be safeguarded by criminal law, as they are aspects of what the author refers to as ‘a good life’, which falls within the personal domain of the individual (Greco, 2023, p. 193).

The crime of child pornography was made autonomous in the Portuguese Penal Code by Law 59/2007 of September 4, and it is currently provided for by Article 176, under the section addressing crimes against sexual self-determination. Paragraph 4 specifies a prison sentence of up to two years for anyone who “engages in the acts described in items c) and d) of Paragraph 1⁶, using pornographic material that portrays a realistic representation of a minor” (Fidalgo, 2022, p. 56–57). This provision is based on Framework Decision 2004/68/JHA, adopted by the Council on December 22, 2003, which focuses on combating the sexual exploitation of children and child pornography. It mandates the implementation of criminal laws to punish activities related to the production, distribution, possession, or display of realistic images of non-existent children (Lopes and Milheiro, 2023, p. 275).

The aforementioned number 1, item d), can be considered a typical possession offense in some of its provisions, a legislative choice that has sparked significant controversy. As Susana Aires de Sousa explains, in such cases, the legislator regards the dangerousness of the conduct as the basis for criminalization, by adopting a political-criminal stance of preemptively safeguarding the legal interest. Therefore, these legal provisions typically take the form of crimes of abstract danger, not requiring a concrete result (Sousa, 2014, p. 493; Antunes and Sousa, 2019, p. 247).

Manuel da Costa Andrade, for example, criticizes the extremity of the legislative decision to criminalize actions like the possession of pornographic material involving minors, by arguing that such behaviors are only remotely connected to the protection of minors’ sexual freedom and self-determination, rendering them lacking in criminal legitimacy (Andrade, 2009, p. 34).

Similarly, Maria João Antunes and Susana Aires de Sousa stress that the criminal validity of the actions outlined in paragraphs c) and d), as well as Sections 4, 5, and 6 of Article 176, is not founded on protecting the individual legal interest of “child sexual self-determination”. If that were the case, these provisions would always face constitutional challenges based on the principle of the *Rechtsgut* and the requirements it imposes on crimes involving abstract danger (Antunes and Sousa, 2019, p. 254).

Highlighting the controversies surrounding this criminalization in many countries, Lopes and Milheiro refer to the *U.S. Child Pornography Act*, passed in 1996, which expanded the definition of prohibited child pornography to include any visible representation, even computer-generated images, that depicted or appeared to depict a child involved in sexually explicit conduct. However, the U.S. Supreme Court declared this law unconstitutional in *Ashcroft v. Free Speech Coalition* (April 16, 2002), by citing concerns over violations of legitimate freedom of expression and refuting the argument that such images promoted deviant behavior (Lopes and Milheiro, 2023, p. 275–276).

Regarding the Portuguese criminal law, the authors argue that the legal interest being protected by such criminalization is not clearly identifiable, though it is certain that it does not relate to sexual freedom or self-determination, especially since the minors targeted by the crime might not even exist. They assert that the offense reinforces moral censorship, which could conflict with the right to artistic freedom, even if the expression of that freedom is shocking in some cases (Lopes and Milheiro, 2023, p. 276–277). Similarly, André Lamas Leite raises significant concerns about the material constitutionality of the crime of possessing pornographic material that realistically depicts a minor (Leite, 2016, p. 69).

In our opinion, it is essential to make a distinction between ‘total virtual child pornography’ and ‘partial virtual child pornography’. The first category refers to realistic depictions of non-existent children, while the second involves pornographic materials that feature realistic representations of minors

⁶ “c) Produce, distribute, import, export, disseminate, display, transfer or make available in any capacity or by any means, the materials provided for in the previous paragraph; Acquire, hold or host materials referred to in paragraph b) for the purpose of distributing, importing, exporting, disseminating, displaying or giving them;” (Portugal, 1982).

which, although proceeding from graphic technology and the imagination of the creator, they are at least partially based on real images of minors (Antunes and Santos, 2012, p. 883–884).

In the case of ‘total virtual child pornography’, there is no risk to the developmental conditions of any minor, which is why we argue that these cases would not fall within the scope of the norm, as doing so would violate the principle of the *Rechtsgut*. Conversely, when elements of real minors are used, such representations can be damaging to the psychological and relational development of underage individuals, and there is even a possibility that the minors involved could be identified through specific characteristics. Therefore, in these cases, criminalization is warranted (Albergaria and Lima, 2010, p. 214–219; Januário, 2024, p. 493–494).

In any case, there are important points to consider specifically regarding the possession of child pornography. As Albergaria and Lima highlight, while there is no question about the legitimacy of criminal protection for minors’ sexual self-determination, there are difficulties in having it as a reference for protection in the specific case of the detention of child pornography. Arguments that go towards legitimising protection in the need to overcome, at a substantive level, possible procedural difficulties in proving that the holder was the one who harmed the legal interest (that is, punishing those who possess pornography, as if they were actual abusers or distributors of these materials) do not deserve to be prospered. Similarly, the argument that possessing such content might encourage other harmful behaviors or be used to lure minors into illegal acts is unconvincing. The authors note that there is currently no scientific evidence showing a strong link between the consumption of pornography and the commission of sexual violence. Some even suggest that pornography consumption could have a cathartic effect. Moreover, it cannot be assumed that the mere possession of such material suggests an increased likelihood of its use in committing other criminal acts (Albergaria and Lima, 2010, p. 200–204).

For these reasons, we align with Albergaria and Lima’s view that there are difficulties in linking the mere possession of child pornography to the sexual self-determination of minors. From a retrospective perspective, the harm to the *Rechtsgut* has already occurred, regardless of the possessor’s involvement. From a forward-looking perspective, it seems exaggerated to suggest that possession alone poses a future threat to the sexual freedom and self-determination of minors. The *Rechtsgut* being protected here is broader, concerning the minor’s overall development, including both their internal (psychological and moral) and external (social and relational) aspects (Albergaria and Lima, 2010, p. 206–207).

In a similar vein, Maria João Antunes and Susana Aires de Sousa explain that, in the crime of child pornography, there are behaviours in which there is a direct relationship with a minor who is a victim of the crime, while in others this relationship is not observed. The criminal dignity of these latter situations is not based on the protection of the individual *Rechtsgut* ‘sexual self-determination of the minor’. Since the object of these actions is no longer the minor, but, rather, pornographic materials with a minor (or realistic representation thereof), they would refer to the protection of a supra-individual *Rechtsgut*, regardless of whether they are systematically included in the title of crimes against the sexual self-determination (Antunes and Sousa, 2019, p. 248–256).

Given the above, and in response to one of the key issues raised in this paper, we argue that, under the Portuguese law, the production, distribution, importation, exportation, dissemination, exhibition, transfer, or provision of sexual deepfakes involving minors, as well as the acquisition, possession, or storage of such materials with the intent to distribute them, is punishable by up to two years in prison. This crime is based on protecting the *Rechtsgut* related to the developmental conditions of the minor (Januário, 2024, p. 495–496).

3. On the Possible Responsibility Gaps Regarding Non-Consensual Sexual Deepfakes with Adults

A more controversial issue concerns the point whether acts such as the production, distribution, exhibition, or possession of non-consensual sexual deepfakes involving adults would fall under criminal provisions in Portugal. Given that there is currently no specific crime addressing these situations, we must question whether they might be covered by other existing offenses, even those intended for different types of cases (Januário, 2024, p. 496).

Some scholars argue, for instance, that such acts could potentially be classified under crimes like ‘computer fraud’⁷, ‘illegal recordings or photographs’⁸, or ‘aggravated defamation with publicity and slander’⁹ (see, for example, the interview given by Ana Raquel Conceição, in: Costa, 2023).

However, we contend that classifying non-consensual sexual deepfakes as computer fraud is problematic. While the creation of this content does involve the alteration of digital data to make it appear genuine, the requirement that the offender must have the ‘intention of causing deception in legal relations’ limits the applicability of this crime. As we have previously noted, there are various reasons why deepfake videos might be created, and the intent behind their creation could range from fulfilling the creator’s own sexual desires, to satisfying the desires of third parties, or even to mocking the victim. Furthermore, the legal focus of computer fraud crimes is primarily on safeguarding computer security, which contrasts with the more immediate harm caused by non-consensual sexual deepfakes, including physical, social, and psychological harm to the victims. These forms of harm, in our view, should be the central concern in these cases (Januário, 2024, p. 496).

A slightly less contentious issue arises when considering whether these cases might fall within the framework of ‘illegal recordings or photographs’. In these situations, although the content may have originally been created legally and with consent, it is later repurposed for uses beyond what was agreed to by the subject of the images. However, the provision itself imposes limitations that significantly narrow its applicability regarding photographic and film content. According to the article, while the recording of words (and their use) is already illegal when carried out without consent, the taking of

⁷ “Art. 3. Computer fraud. 1 – Whoever, with the intention of causing deception in legal relations, introduces, modifies, erases or suppresses computer data or in any other way interferes in the computer processing of data, producing non-genuine data or documents, with the intention that these be considered or used for legally relevant purposes as if they were, is punished with a prison sentence of up to 5 years or a fine of 120 to 600 days. [...] 4 – Whoever produces, acquires, imports, distributes, sells or holds any device, program or other computer data intended for carrying out the actions set out in paragraph 2, is punished with a prison sentence of 1 to 5 years” (Portugal, 2009).

⁸ “Art. 199. Illegal recordings and photographs. 1 – Whoever without consent: a) Records words spoken by another person and not intended for the public, even if they are addressed to them; or b) Uses or allows the recordings referred to in the previous paragraph to be used, even if lawfully produced; is punished with a prison sentence of up to 1 year or a fine of up to 240 days. 2 – The same penalty applies to anyone who, against the victim’s will: a) Photographs or films another person, even in events in which they have legitimately participated; or b) Uses or allows the use of photographs or films referred to in the previous paragraph, even if lawfully obtained” (Portugal, 1982).

⁹ “Art. 180. Defamation. 1 – Whoever, addressing a third party, imputes to another person, even in the form of suspicion, a fact, or formulates a judgment about him/her, offensive to his/her honor or consideration, or reproduces such an imputation or judgement, is punished with a prison sentence of up to 6 months or a fine of up to 240 days”. “Art. 182. Equivalence. Verbal defamation and insults are equated with those made in writing, gestures, images or any other means of expression”. “Art. 183. Publicity and slander. 1 – If in the case of crimes provided for in Art. 180, 181 and 182: a) The offense is committed through means or in circumstances that facilitate its disclosure; or, b) In the case of the imputation of facts, it is found that the agent knew the imputation was false; The penalties for defamation or insult are increased by one third in their minimum and maximum limits. 2 – If the crime is committed through social media, the offender is punished with a prison sentence of up to 2 years or a fine of no less than 120 days” (Portugal, 1982).

photographs and films (and their use) will only be illegal when they are against the will of the holder (Andrade, 2012, p. 1185).

This distinction, though subtle, is significant. In cases where real images or videos are used to create sexual deepfakes, this requirement could potentially allow the defense that the victim had not explicitly objected to the creation of the content, and might not even have been aware of it until later¹⁰. Moreover, the provision also specifies that the recording must involve another person. That means that self-recorded images or videos do not fall under this offense, even if such content is later used improperly without consent (Andrade, 2012, p. 1212–1213).

Additionally, the criminal protection of one's image under this offense is tied to specific technical processes, such as photography and filming. This is why, for instance, Manuel da Costa Andrade suggests that other forms of visual representation – such as paintings, drawings, caricatures, or live streaming – are not covered. Furthermore, the offense does not extend to protecting the integrity or truth of the image, which means that alterations such as distortions or montages do not trigger this provision (Andrade, 2012, p. 1202–1203).

Given these strict limitations, we have doubts about whether the creation of sexual deepfakes by using data or images from big data could be protected under this offense, even when these images do not reflect the victim's real physical attributes or the context in which they were originally taken (Januário, 2024, p. 497–498).

Lastly, we can consider whether the cases being examined could fall under the category of 'aggravated defamation with publicity and slander'. In fact, the creation or distribution of non-consensual sexual deepfakes to third parties or to the victims themselves involves attributing facts or judgments that are potentially damaging to their honor, conveyed through images or similar forms of expression, and using means that facilitate their spread (including the media). However, the mere possession of such material would likely not fall under this offense (Januário, 2024, p. 498).

The main issue here is whether these actions can be considered offenses against an individual's honor. On the one hand, it is evident that non-consensual sexual deepfakes can damage a victim's honor, which encompasses both their personal worth and external reputation. On the other hand, framing these cases as violations of honor could reinforce harmful stereotypes, particularly the notion that pornographic content undermines a woman's reputation, thus contributing to the perpetuation of gender inequality (Toparlak, 2023, p. 10; Citron, 2019, p. 1898).

Considering the points above, while certain situations could indeed be classified as crimes of 'computer fraud', 'illicit recordings or photographs', or 'aggravated defamation with publicity and slander', we argue that a more fitting and coherent solution to address the issue of non-consensual sexual deepfakes involving adults would be the introduction of a criminal provision specifically tailored to these cases (Januário, 2024, p. 498).

A relevant example can be found in Article 216-B, sole paragraph, of the Brazilian Penal Code, introduced by Law 13.772/2018. This provision stipulates that anyone who creates a montage in photography, video, audio, or any other form of record placing with the intent to place an individual in a nude scene or in a sexual or lewd act of an intimate nature is subject to a penalty ranging from six months to one year, along with a fine (Brasil, 1940). This law aims to protect the victim's privacy and intimacy (Bitencourt, 2021, p. 107), focusing solely on the creation of the montage itself, without requiring its public dissemination (Gueiros and Japiassú, 2023, p. 486).

¹⁰ In the opposite sense, however, Manuel da Costa Andrade understands that going against the presumed will of the holder of the right to the image will suffice. See: Andrade, 2012, p. 1214.

However, if such content is made public, the offense falls under Article 218-C, which imposes a sentence of one to five years for anyone who offers, exchanges, makes available, transmits, sells, exposes for sale, distributes, publishes, or disseminates a sex scene, nudity, or pornography without the victim's consent. In this case, the victim's lack of consent regards specifically to the public disclosure of the content, irrespective of whether the initial creation of the image or voice was consented to (Sousa, 2021, n.p.).

Although we are not proponents of expanding the criminal law in an immediate response to societal concerns, we believe that, in the case of non-consensual sexual deepfakes, such an expansion would not be unreasonable or disproportionate. As we have discussed, we are no longer dealing with crude, easily recognizable fake montages, nor are we in a scenario where the technology for creating these deepfakes is restricted to a select few with the resources to access it. The current ease with which deepfake content can be produced stands in stark contrast to the extent of the harm it causes to the victims, who suffer significant psychological, social, and reputational damage (Januário, 2024, p. 500).

While civil liability remains a necessary recourse, it may not be enough to fully address the issue. Many victims may be reluctant to pursue legal action for fear of further exposure and publicity (Citron and Franks, 2014, p. 346). Moreover, identifying the perpetrators behind the deepfake content is often difficult, and victims may struggle to find support from the websites that host this material. Even efforts to remove such content from the internet are frequently futile, given the speed and ease with which it can be disseminated online.

Conclusions

1. Non-consensual sexual deepfakes have become an increasingly serious issue. Advances in artificial intelligence have not only made the technologies for creating these types of content widely available, but have also enhanced their realism, thus making it progressively harder to distinguish them from authentic material. As a result, the harm caused to the victims of these deepfakes – whether adults or minors – is undeniable. Examples of the devastating effects include PTSD, suicidality, anxiety, and depression in adults, as well as damage to the developmental well-being of children.
2. In relation to deepfake child pornography, despite some doctrinal disagreements on specific aspects, it is clear that the Portuguese legal system addresses this issue effectively. The creation, distribution, import, export, dissemination, exhibition, transfer, or provision of sexual deepfakes involving underaged – where the minors are, at least in part, real – along with the acquisition, possession, or storage of such materials with the intent to distribute, import, export, disseminate, display, or provide them, is punishable by up to two years in prison under Article 176, Paragraph 4. This offense is grounded in the protection of the legal interest related to the development conditions of minors.
3. However, when it comes to non-consensual sexual deepfakes involving adults, although there are situations that may fall under the offenses of 'computer fraud', 'illicit recordings or photographs', or 'aggravated defamation with publicity and slander', we identify potential liability gaps. Following the example of the Brazilian legal framework, we argue that a more appropriate solution would be the introduction of legal provisions specifically tailored to these cases. Such provisions would criminalize the creation or distribution of non-consensual sexual deepfakes, with the aim of protecting the sexual privacy of the victims.
4. While remaining mindful of the demands posed by new technologies, yet cautious not to allow an excessive expansion of the criminal law – much less a regression to a 'moralistic' approach to

the sexual criminal law – we believe that the solutions suggested here could effectively address this phenomenon, which has already become deeply ingrained in contemporary society, but whose impact is likely to endure for a long time.

Bibliography

Legal acts

- Brasil. (1940). *Decreto-Lei n.º 2.848, de 7 de dezembro de 1940: Código Penal* [online]. Available at: https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848.htm [Accessed 06 November 2024].
- Portugal. (1982). *DL n.º 48/95, de 15 de Março: Código Penal de 1982: Versão consolidada posterior a 1995* [online]. Available at: https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=109&tabela=leis [Accessed 06 November 2024].
- Portugal. (2009). *Lei n.º 109/2009, de 15 de Setembro: Lei do Cibercrime* [online]. Available at: https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=1137&tabela=leis. [Accessed 06 November 2024].

Special literature

- Albergaria, P. S., Lima, P. M. (2010). O crime de detenção de pseudopornografia infantil – evolução ou involução? *Julgar*, 12(esp.), 195–220 [online]. Available at: <https://julgar.pt/o-crime-de-detencao-de-pseudopornografia-infantil-evolucao-ou-involucao/>
- Andrade, M. C. (2009). “*Bruscamente no verão passado*”, a reforma do Código de Processo Penal: observações críticas sobre uma lei que podia e devia ter sido diferente. Coimbra: Coimbra Editora.
- Andrade, M. C. (2012). Artigo 199.º (Gravações e fotografias ilícitas). In: Dias, J. F. (ed.) *Comentário Conimbricense do Código Penal: Parte Especial: Tomo I: Artigos 131.º a 201.º*. 2. ed. Coimbra: Coimbra Editora.
- Antunes, M. J.; Fidalgo, S. (2008). Os crimes contra a liberdade e a autodeterminação sexual. In: Franco, A. S. et al. (ed.) *Justiça penal portuguesa e brasileira: tendências de reforma: Colóquio em Homenagem ao Instituto Brasileiro de Ciências Criminais: 13 e 14 de abril de 2007: Faculdade de Direito da Universidade de Coimbra*. São Paulo: IBCCRIM.
- Antunes, M. J., Santos, C. C. (2012). Artigo 176º (Pornografia de menores). In: Dias, J. F. (ed.) *Comentário Conimbricense do Código Penal: Parte Especial: Tomo I: Artigos 131.º a 201.º*. 2.ed. Coimbra: Coimbra Editora.
- Antunes, M. J., Sousa, S. A. (2019). Da relevância da identificação do bem jurídico protegido no crime de pornografia de menores, *Revista Portuguesa de Ciência Criminal*, 29(2), 239–264.
- Bates, S. (2017). Revenge Porn and Mental Health: A Qualitative Analysis of the Mental Health Effects of Revenge Porn on Female Survivors, *Feminist Criminology*, 12(1), 22–42, <https://doi.org/10.1177/1557085116654565>.
- Bitencourt, C. R. (2021). *Tratado de direito penal*, vol. 4. 15. ed. São Paulo: Saraiva.
- Burrell, J. (2016). How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms, *Big Data & Society*. <https://doi.org/10.1177/2053951715622512>
- Chesney, B., Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, *California Law Review*, 107, 1753–1820. <https://dx.doi.org/10.2139/ssrn.3213954>
- Citron, D. K., Franks, M. (2014). Criminalizing Revenge Porn, *Wake Forest Law Review*, 49, 345–391, https://digitalcommons.law.umaryland.edu/fac_pubs/1420/
- Citron, D. K. (2019). Sexual Privacy. *The Yale Law Journal*, 128(7), 1870–1960 [online]. Available at: <https://www.yalelawjournal.org/article/sexual-privacy>
- Custers, B. (2022). AI in Criminal Law: An Overview of AI Applications in Substantive and Procedural Criminal Law. In: Custers, B., Fosch-Villaronga, E. (ed.) *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*. The Hague: Springer.
- Fidalgo, S. (2022). Crimes no ciberespaço: os desafios do metaverso. In: Rodrigues, A. M. (ed.). *A inteligência artificial no direito penal*, vol. II. Coimbra: Almedina.
- Franks, M. A. (2021). *Drafting and Effective “Revenge Porn” Law: A Guide for Legislators*. Cyber Civil Rights Initiative. <http://dx.doi.org/10.2139/ssrn.2468823>
- Gieseke, A. P. (2020). “The New Weapon of Choice”: Law’s Current Inability to Properly Address Deepfake Pornography, *Vanderbilt Law Review*, 73(5), 1479–1515 [online]. Available at: <https://scholarship.law.vanderbilt.edu/vlr/vol73/iss5/4>
- Goodfellow, I. et al. (2014). Generative Adversarial Nets, *Advances in Neural Information Processing Systems*, 27, 1–9. <https://doi.org/10.48550/arXiv.1406.2661>

- Greco, L. (2023). Pornografia punível no estado liberal: fundamentos e limites dos §§ 184, 184A- D do Código Penal Alemão (STGB). In: Camargo, B. C. *et al.* (ed.) *Questões fundamentais do direito penal sexual*. São Paulo: Marcial Pons.
- Gueiros, A., Japiassú, C. E. (2023). *Direito Penal: Parte Especial*. 2.ed. São Paulo: Tirant lo Blanch.
- Harris, D. (2019). Deepfakes: False Pornography is Here and the Law Cannot Protect You, *Duke Law & Technology Review*, 17(1), 99–127 [online]. Available at: <https://scholarship.law.duke.edu/dltr/vol17/iss1/4>
- Januário, T. F. X. (2021). Inteligência artificial e manipulação do mercado de capitais: uma análise das negociações algorítmicas de alta frequência (high-frequency trading) à luz do ordenamento jurídico brasileiro, *Revista Brasileira de Ciências Criminais*, 29(186), 127–183.
- Januário, T. F. X. (2023). Inteligencia artificial y responsabilidad penal de personas jurídicas: un análisis de sus aspectos materiales y procesales, *Estudios Penales y Criminológicos*, 44(ext.), 1–39, <https://doi.org/10.15304/epc.44.8902>
- Januário, T. F. X. (2024). Artificial Intelligence and Image-Based Sexual Abuse: an analysis of non-consensual deepfakes in light of Portuguese criminal law, *RIDP*, 95(2), 483–505 [online]. Available at: <https://www.maklu-online.eu/nl/tijdschrift/ridp/2024/researching-boundaries-sexual-integrity-gender-vio/artificial-intelligence-and-image-based-sexual-abu/>
- Januário, T. X. (2020). Veículos autónomos e imputação de responsabilidades criminais por acidentes. In: Rodrigues, A. M. (ed.). *A Inteligência Artificial no Direito Penal*, vol. I. Coimbra: Almedina.
- Januário, T. X. (2022). Inteligência artificial e direito penal da medicina. In: Rodrigues, A. M. (ed.) *A inteligência artificial no direito penal*, vol. II. Coimbra: Almedina.
- Jurasz, O. & Barker, K. (2021). Sexual Violence in the Digital Age: A Criminal Law Conundrum? *German Law Journal*, 22(5), 784–799, <https://doi.org/10.1017/glj.2021.39>
- Leite, A. L. (2016). As alterações de 2015 ao Código Penal em matéria de crimes contra a liberdade e autodeterminação sexuais – nótulas esparsas, *Julgar*, 28, 61–74 [online]. Available at: <https://julgar.pt/as-alteracoes-de-2015-ao-codigo-penal-em-materia-de-crimes-contra-a-liberdade-e-autodeterminacao-sexuais-notulas-esparsas/>
- Lopes, J. M., Milheiro, T. C. (2023). *Crimes sexuais: análise substantiva e processual*, Coimbra: Almedina.
- Mania, K. (2024) Legal Protection of Revenge and Deepfake Porn Victims in the European Union: Findings from a Comparative Legal Study, *Trauma, Violence & Abuse*, 25(1), 117–129. <https://doi.org/10.1177/15248380221143772>
- McGlynn, C., Rackley, E. (2017). Image-Based Sexual Abuse, *Oxford Journal of Legal Studies*, 37(3), 534–561. <https://doi.org/10.1093/ojls/gqw033>
- McGlynn, C. *et al.* (2017). Beyond ‘Revenge Porn’: The Continuum of Image-Based Sexual Abuse, *Feminist Legal Studies*, 25, 25–46. <https://doi.org/10.1007/s10691-017-9343-2>
- Miró Llinares, F. (2018). Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots, *Revista de Derecho Penal y Criminología*, 3(20), 87–130. <https://doi.org/10.5944/rdpc.20.2018.26446>
- Öhman, C. (2020). Introducing the pervert’s dilemma: a contribution to the critique of Deepfake Pornography, *Ethics and Information Technology*, 22, 133–140. <https://doi.org/10.1007/s10676-019-09522-1>
- Price II, W. N. (2017). Artificial Intelligence in Health Care: Applications and Legal Issues, *The SciTech Lawyer*, 14(1), 10–13 [online]. Available at: <https://repository.law.umich.edu/articles/1932/>
- Rodrigues, A. M. (2020). Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e a Europeização. In: Rodrigues, A. M. (ed.) *A Inteligência Artificial no Direito Penal*, vol. I. Coimbra: Almedina.
- Rodrigues, P. G. L. S. (2023). Deepfakes pornográficas não-consensuais: a busca por um modelo de criminalização. *Revista Brasileira de Ciências Criminais*, 31(199), 277–311.
- Sousa, L. A. (2021). *Direito Penal: Parte Especial: Art. 155 a 234-B do CP*, vol. 3. 2. ed. São Paulo: Thomson Reuters Brasil.
- Sousa, S. A. (2014). *A responsabilidade criminal pelo produto e o topos causal em direito penal: contributo para uma protecção penal de interesses do consumidor*. Coimbra: Coimbra Editora.
- Sousa, S. A. (2020). “Não fui eu, foi a máquina”: teoria do crime, responsabilidade e inteligência artificial”. In: Rodrigues, A. M. (ed.) *A inteligência artificial no direito penal*, vol. 1. Coimbra: Almedina, 59–94.
- Spivak, R. (2019). “Deepfakes”: The Newest Way to Commit One of the Oldest Crimes, *Georgetown Law Technology Review*, 3(2), 339–400 [online]. Available at: <https://georgetownlawtechreview.org/deepfakes-the-newest-way-to-commit-one-of-the-oldest-crimes/GLTR-05-2019/>

- Toparlak, R. T. (2023). Criminalising Deep Fake Pornography: A Gender-Specific Analysis of Image-Based Sexual Abuse, *Cognitio*, 1, 1–16. <https://doi.org/10.5281/zenodo.7791799>
- Uhl, C. A. *et al.* (2018). An examination of nonconsensual pornography websites, *Feminism & Psychology*, 28(1), 50–68, <https://doi.org/10.1177/0959353517720225>

Other sources

- Costa, F. (2023). Deep fakes são uma ameaça crescente e ainda falta resposta legal adequada, *Público*, 11 November [online]. Available at: <https://www.publico.pt/2023/11/11/tecnologia/noticia/deep-fakes-sao-ameaca-crescente-falta-resposta-legal-adequada-2069760> [Accessed 06 November 2024].
- Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission. (2019). *A Definition of AI: Main Capabilities and Disciplines* [online]. Available at: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> [Accessed 06 November 2024].
- OECD. (2024). Explanatory memorandum on the updated OECD definition of an AI system, *OECD Artificial Intelligence Papers*, 8. <https://doi.org/10.1787/623da898-en> [Accessed 06 November 2024].

Túlio Felipe Xavier Januário is a PhD Candidate in Law at the University of Coimbra (Portugal), with a fellowship from the Fundação para a Ciência e a Tecnologia – FCT. M.Sc. in Law by the University of Coimbra (Portugal), with a research internship of the *ERASMUS+* Program at the Georg-August-Universität Göttingen (Germany). He completed Graduate Studies in International Criminal Law at Siracusa International Institute for Criminal Justice and Human Rights (Italy), Graduate Studies in Economic Criminal Law and Crime Theory at the University of Castilla-La Mancha (Spain), Graduate Studies in Compliance and Criminal Law at IDPEE (Portugal) and Graduate Studies in Criminal Law – General Part at IBCCRIM/IDPEE (Brazil/Portugal). He holds a Bachelor's Degree in Law by the Universidade Estadual Paulista – UNESP (Brazil).

Túlio Felipe Xavier Januário – Koimbros universiteto (Portugalija) teisės doktorantas, gavęs Fundação para a Ciência e a Tecnologia (FCT) stipendiją. Įgijo teisės magistro laipsnį Koimbros universitete (Portugalija), atlikdamas mokslinių tyrimų stažuotę pagal „ERASMUS+“ programą Georgo Augusto universitete Getingene (Vokietija). Baigė tarptautinės baudžiamosios teisės podiplomines studijas Sirakūzų tarptautiniame baudžiamosios justicijos ir žmogaus teisių institute (Italija), ekonominės baudžiamosios teisės ir nusikaltimų teorijos podiplomines studijas Kastilijos-La Mančos universitete (Ispanija), atitiktis ir baudžiamosios teisės podiplomines studijas IDPEE institute (Portugalija) bei bendrosios baudžiamosios teisės dalies podiplomines studijas IBCCRIM/IDPEE (Brazilija / Portugalija). Įgijo teisės bakalauro laipsnį São Paulo valstybiniame universitete – UNESP (Brazilija).