# CRIMINAL LIABILITY FOR AI-GENERATED CONTENT: EMERGING LEGAL CHALLENGES AND THE NEED FOR A GLOBAL REGULATORY FRAMEWORK

*by*

Ms. Megha Pant[1]

## ABSTRACT

Generative artificial intelligence has changed how people prepare and share digital content. A user can now create or edit text, images, voices and videos within a very short time. These tools are useful in education, business, entertainment and many other areas. However, they can also be misused. For example, a person may use AI to imitate another person's voice, prepare a fake intimate image, spread a false video or support an online fraud. Such incidents raise an important question: who should be held criminally responsible when AI-generated content is used to harm someone?

The answer depends on the facts of each case. The person who gives the prompt, the developer who designs the tool, the organisation that uses it and the platform that shares the content may have different roles. This paper examines whether traditional principles of criminal law, including actus reus, mens rea, causation and individual responsibility, can deal with these situations. It follows a doctrinal and comparative method and considers selected legal developments in the European Union, the United States, the United Kingdom and the Council of Europe. A qualitative comparative matrix is used to examine the legal materials. The study finds that existing criminal laws can cover many forms of deliberate misuse, but practical difficulties remain in identifying offenders, proving intention, collecting digital evidence and investigating cases across borders. The paper argues that a balanced legal response should focus on the conduct and responsibility of each actor. It should protect victims without treating every error made by an AI system as a crime.

**Keywords:** Artificial Intelligence; AI-Generated Content; Criminal Liability; Deepfakes; Cybercrime; Digital Accountability; Comparative Criminal Law; Global Regulation.

[1] Graduate in Law, Asian Law College, India.



Indian Journal of Law and Jurisprudence | Licensed under Creative Commons — https://creativecommons.org/licenses/by/4.0/ | DOI: 10.5281/zenodo.24112066 | Title: Criminal Liability for AI-Generated Content: Emerging Legal Challenges and the Need for a Global Regulatory Framework

## 1. INTRODUCTION

Artificial intelligence is now part of everyday life. It is no longer used only by research laboratories or large technology companies. Students, professionals, businesses, content creators and ordinary internet users can access generative AI tools. With a simple instruction, a person can prepare written material, create an image, imitate a voice or produce a video. This has made content creation easier. But the same development has also created new ways to misuse digital technology.

Imagine receiving a voice message that sounds like a family member asking for money because of an emergency. The voice may have been created with AI, even though that family member never sent the message. In another case, someone's photograph may be altered to create an intimate image that looks real. A fake video may also be shared to damage a person's reputation or influence public opinion. These examples show that digital content does not need to be genuine to cause genuine harm.[2]

The problem is not that false information or impersonation has appeared for the first time. Fraud, forged documents, threats and misleading statements existed long before generative AI. What has changed is the speed, scale and realism with which such material can be produced and circulated. A person who has little technical knowledge can now create content that may be difficult for an ordinary viewer to identify as false. Once uploaded, the material can quickly move between platforms and countries.

Criminal law usually requires proof of a prohibited act and the mental element prescribed for that offence.[3] Depending on the law, this may mean intention, knowledge, dishonesty or recklessness. AI creates difficulties because the final output may not always be fully predictable. A user may deliberately misuse a system, while a developer may have created a general-purpose tool for lawful purposes. A platform may also host content without knowing that it is unlawful. It would be unfair to treat all these actors in the same manner without examining their actual role.

The central issue of this paper is therefore how criminal responsibility should be decided when AI-generated content is used to commit or assist an offence.[4] The paper does not suggest that an AI system itself should be punished. The focus is on the conduct of people and, where the law allows, organisations. The relevant question is whether an actor has committed, assisted, encouraged or knowingly facilitated conduct that is already prohibited by law.

[2] https://www.gov.uk/government/publications/deepfake-detection-technology/deepfake-detection-technology?
[3] Jeremy Horder & Andrew Ashworth, Ashworth's Principles of Criminal Law (9th ed. 2019).
[4] Horder & Ashworth, supra note 2.

This issue is not limited to one country.[5] The person creating the content, the victim, the platform and the computer server may all be located in different places. Content may be generated in one country, uploaded through a service based in another and viewed by a victim somewhere else. Differences in criminal laws, evidence rules and investigation procedures can make a case difficult to pursue. Even where a country has a relevant offence, authorities may face problems if the suspect or important digital records are outside its territory.

This paper considers selected approaches in the European Union, the United States and the United Kingdom, along with the wider human-rights framework of the Council of Europe.[6] These approaches are not identical. The EU AI Act mainly creates regulatory duties, including certain transparency requirements. The United States has adopted a targeted federal law dealing with the disclosure of non-consensual intimate visual depictions, including certain digital forgeries. The United Kingdom has introduced specific offences relating to the creation and requesting of purported intimate images of adults. The Council of Europe Convention provides a broader framework based on human rights, democracy and the rule of law. A comparison of these approaches can help identify the strengths and limits of the present legal response.

The paper first explains AI-generated content and the ways in which it may be misused. It then discusses traditional principles of criminal liability and the possible roles of users, developers, deployers and platforms. After that, it compares the selected legal frameworks. The comparative analysis is followed by a discussion of practical challenges, findings and recommendations.

### 1.1 Research Problem

The main problem examined in this study is the difficulty of deciding who should be held criminally responsible when AI-generated content causes harm. Existing laws may already cover fraud, harassment, impersonation, extortion or non-consensual intimate imagery. However, the involvement of AI raises further questions about intention, knowledge, causation, participation and proof. There is also no single international criminal-law

[5] Convention on Cybercrime arts. 29–35, Nov. 23, 2001, E.T.S. No. 185; Second Additional Protocol to the Convention on Cybercrime on Enhanced Co-operation and Disclosure of Electronic Evidence, May 12, 2022, C.E.T.S. No. 224.

[6] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689) 1; TAKE IT DOWN Act, Pub. L. No. 119-12, §§ 2-3, 139 Stat. 55 (2025); Data (Use and Access) Act 2025, c. 18, § 138 (UK); Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, May 17, 2024, C.E.T.S. No. 225.

framework that deals with every form of harmful AI-generated content. The challenge is to apply existing law fairly while identifying where clearer legal rules may be needed.

### 1.2 Research Questions

- Are traditional principles of criminal law sufficient to determine liability in cases involving AI-generated content?
- In what circumstances may users, developers, deployers and digital platforms be held criminally responsible?
- What can be learned from the legal approaches adopted by the European Union, the United States, the United Kingdom and the Council of Europe?
- What legal and institutional steps can improve accountability while protecting freedom of expression and innovation?

### 1.3 Objectives of the Study

- To explain the main forms of AI-generated content and the risks arising from its misuse.
- To examine how actus reus, mens rea, causation and participation apply to offences involving AI-generated content.
- To compare selected legal and regulatory responses.
- To identify difficulties relating to evidence, offender identification, jurisdiction and platform responsibility.
- To suggest a proportionate approach to criminal accountability that also protects individual rights.

### 1.4 Research Methodology

This paper uses doctrinal and comparative legal research. The primary materials include legislation, treaty texts and official publications issued by governments or international institutions. Secondary sources are used to understand the legal and policy background. The comparative study focuses on four approaches: the European Union, the United States, the United Kingdom and the Council of Europe. They have been selected because they represent different legal methods: broad AI regulation, a targeted federal criminal law, specific domestic criminal offences and a human-rights treaty framework.

The data analysis in this paper is qualitative and document-based. It does not claim to measure the number of AI-related offences or the success rate of prosecutions. Instead, the selected

legal materials are compared through common criteria, such as the conduct covered, the mental element, the actors addressed, enforcement duties, cross-border cooperation and safeguards for rights. The comparative matrix represents the author's assessment of the legal texts and is not an official rating of any jurisdiction.

## 2. UNDERSTANDING AI-GENERATED CONTENT AND ITS MISUSE

AI-generated content means material produced wholly or partly with the help of an artificial intelligence system.[7] It may take the form of text, images, audio, video or other digital material. Generative systems use patterns learned from data to produce an output in response to a prompt or instruction. Some systems can also change existing material, such as replacing a face, changing a voice or editing parts of a video.

A deepfake is synthetic or manipulated content that makes a person appear to say or do something that did not actually happen.[8] Deepfakes are not always created for harmful reasons. They may be used in films, satire, education, art and entertainment. However, the same methods can be used to create fake endorsements, impersonate public officials, make non-consensual sexual images or deceive people into sending money.

For legal purposes, the context and conduct are more important than the technology alone.[9] A fictional image clearly presented as fiction is different from an altered image circulated to threaten or exploit a real person. In the same way, an unexpected output from an AI system is not automatically a criminal offence. The prosecution must establish the elements of a recognised offence. Possible forms of misuse include online fraud, identity-related offences, forged evidence, extortion, harassment, impersonation and the creation or circulation of non-consensual intimate imagery.

It is also necessary to distinguish harmful content from criminal content. Some material may be offensive, misleading or unethical without satisfying the legal requirements of a criminal offence.[10] Criminal law should not become a general method of controlling every false or unpleasant statement online. It should be used where the conduct meets the threshold established by law.

## 3. TRADITIONAL PRINCIPLES OF CRIMINAL LIABILITY

[7] Regulation (EU) 2024/1689, art. 3(60).
[8] Artificial Intelligence Act, supra note 6, art. 3(60); Dep't for Sci., Innovation & Tech., supra note 1.
[9] Horder & Ashworth, supra note 2.
[10] Horder & Ashworth, supra note 2

### 3.1 Actus Reus

Actus reus refers to the prohibited act or external element of an offence.[11] In cases involving AI-generated content, the relevant conduct may include creating, possessing, publishing, transmitting or using synthetic material. It may also include asking another person to create such material or helping to distribute it as part of a fraudulent plan. The precise conduct must be identified under the offence being considered. Simply having access to an AI tool, or using it to create lawful material, should not be treated as a criminal act.

### 3.2 Mens Rea

Mens rea refers to the mental element required for an offence.[12] Depending on the applicable law, it may involve intention, knowledge, dishonesty, recklessness or another specified state of mind. A person who deliberately creates a fake intimate image to humiliate someone is in a different position from a person who receives an unexpected output without knowing how it was produced. The prosecution must prove the mental element required by the particular offence. It should not be assumed only because AI technology was used.

### 3.3 Causation and Foreseeability

Causation can become difficult where several people and technical steps are involved.[13] A developer may create a model, a user may enter a prompt, a service may generate the output and a platform may later distribute it. The court may need to consider whether the accused person's conduct was sufficiently connected to the prohibited result and whether another act affected that connection. The fact that a tool made an offence easier does not, by itself, prove that its developer caused the offence in the criminal-law sense.

### 3.4 Participation and Organisational Responsibility

Rules relating to aiding, abetting, conspiracy, counselling or procuring may apply when a person intentionally assists an offence, subject to the requirements of the relevant jurisdiction.[14] Corporate criminal liability may also arise where the law provides for it and the necessary connection between the offence and the organisation is established. These existing principles provide a more careful starting point than a general rule that would make every person involved in the AI supply chain automatically responsible.

[11] Horder & Ashworth, supra note 2.
[12] Id.
[13] Horder & Ashworth, supra note 2
[14] Horder & Ashworth, supra note 2.

## 4. RESPONSIBILITY OF USERS, DEVELOPERS, DEPLOYERS AND PLATFORMS

A user who knowingly creates or distributes harmful content is often the most direct person to examine for criminal liability.[15] For example, a user may deliberately employ AI to impersonate another person, commit fraud, threaten someone or create prohibited intimate imagery. Depending on the facts, an existing offence or a specific offence dealing with synthetic material may apply. The use of AI may increase the scale or impact of the conduct, but the legal elements of the offence must still be proved.

Developers should not be held criminally responsible only because someone later misuses a general-purpose tool.[16] A system may be designed for lawful uses and made available to a large number of users. The developer may have no knowledge of a particular offence. Liability may be considered where there is evidence of intentional assistance, deliberate facilitation or another legally recognised basis of responsibility. If a legislature wants to create an offence based on negligence, it should define that offence clearly. Courts should not create a broad negligence offence simply by extending existing provisions beyond their proper meaning.

Deployers and organisations may have a different position because they decide how an AI system is used in a particular setting.[17] An organisation might use synthetic media in advertising, political communication or customer services. Depending on the applicable law, it may have duties relating to risk assessment, internal controls, record keeping and transparency. However, a breach of a regulatory duty does not automatically become a criminal offence. The law must clearly state when a breach attracts criminal punishment and what must be proved.

Digital platforms also matter because they may host and distribute content to a large audience.[18] Reporting systems, notice-and-action procedures, evidence preservation and timely responses can help reduce harm. At the same time, criminal liability should not arise merely because unlawful content appeared on a platform. The relevant law must identify the platform's duty, the required knowledge or notice, the time allowed for action and the legal consequence of non-compliance. A platform should not automatically be treated as the original creator of content uploaded by a user.

[15] TAKE IT DOWN Act, supra note 5, § 2; Sexual Offences Act 2003, c. 42, §§ 66E-66F (UK), as inserted by Data (Use and Access) Act 2025, c. 18, § 138 (UK).
[16] Horder & Ashworth, supra note 2
[17] Artificial Intelligence Act, supra note 6, arts. 9, 17, 50.
[18] TAKE IT DOWN Act, supra note 5, § 3; Online Safety Act 2023, c. 50, § 10 (UK).

## 5. COMPARATIVE ANALYSIS OF SELECTED LEGAL FRAMEWORKS

### 5.1 European Union

The European Union Artificial Intelligence Act, Regulation (EU) 2024/1689, establishes a risk-based system for regulating artificial intelligence. Article 50 deals with transparency obligations for certain AI systems. As at 30 October 2026, Article 50 is applicable from 2 August 2026. It requires providers of certain systems that generate synthetic audio, image, video or text to ensure that outputs are marked in a machine-readable format and can be detected as artificially generated or manipulated, subject to the qualifications in the provision. Deployers also have disclosure duties in relation to certain deepfakes. There is a specific rule for AI-generated text published to inform the public on matters of public interest, subject to the conditions and exceptions in Article 50.[19]

The AI Act is not a general criminal code. It mainly establishes regulatory duties and provides for administrative enforcement. It does not replace national criminal offences dealing with fraud, threats, harassment or other harmful conduct. It is also important to mention the implementation dates accurately. Certain obligations have different application dates, and the limited transition for some systems placed on the market before 2 August 2026 runs until 2 December 2026. Therefore, the relevant provision and date should be checked before making a statement about a particular duty.

### 5.2 United States

The United States adopted the TAKE IT DOWN Act, Public Law 119-12, on 19 May 2025. The Act creates a federal criminal prohibition on the intentional disclosure of non-consensual intimate visual depictions, including certain digital forgeries. It also requires covered platforms to establish a process for receiving valid requests and removing the material within the period set by the Act.[20]

The law is important because it recognises that a digitally created or manipulated intimate depiction can cause serious harm even when the image does not show an event that actually took place. Its strength is its clear focus on a particular form of abuse. However, it does not create a complete criminal-law framework for every form of AI-generated fraud,

[19] Artificial Intelligence Act, supra note 6, arts. 50, 113.
[20] TAKE IT DOWN Act, supra note 5, §§ 2-3.

impersonation, harassment or disinformation. It should therefore be understood as one part of the wider legal response, not as a solution to every problem caused by generative AI.

### 5.3 United Kingdom

The United Kingdom has introduced specific legal provisions dealing with intimate images. The Online Safety Act 2023 amended the law relating to the sharing or threatened sharing of intimate images.[21] In addition, section 138 of the Data (Use and Access) Act 2025 inserted new sections 66E and 66F into the Sexual Offences Act 2003. Section 66E concerns intentionally creating a purported intimate image of an adult without that person's consent and without a reasonable belief in consent. Section 66F concerns intentionally requesting the creation of such an image in the circumstances set out in the legislation. Section 138 came into force on 6 February 2026.

This development is significant because the law addresses the creation and requesting of purported intimate images, rather than focusing only on their later publication. At the same time, these are specific offences. They do not resolve every question about AI-generated content, nor do they automatically make developers or platforms criminally liable. Their application depends on the facts and the statutory requirements.

### 5.4 Council of Europe

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225) was opened for signature in Vilnius on 5 September 2024.[22] It aims to ensure that activities within the life cycle of AI systems are consistent with human rights, democracy and the rule of law. This is relevant because AI-related harm may affect privacy, equality, dignity and democratic participation.

The Convention is not a uniform criminal code and does not itself provide a complete list of offences relating to AI-generated content. The Council of Europe's Convention on Cybercrime, commonly known as the Budapest Convention, and its Second Additional Protocol are also relevant. They provide a framework for cooperation on cybercrime and electronic evidence. Together, these instruments show why international cooperation is

[21] Online Safety Act 2023, c. 50, § 188 (UK); Data (Use and Access) Act 2025, c. 18, § 138 (UK); Data (Use and Access) Act 2025 (Commencement No. 5) Regulations 2026, S.I. 2026/31, reg. 2 (UK).

[22] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, May 17, 2024, C.E.T.S. No. 225; Convention on Cybercrime, Nov. 23, 2001, E.T.S. No. 185; Second Additional Protocol to the Convention on Cybercrime, May 12, 2022, C.E.T.S. No. 224.

important, even though the details of criminal offences continue to depend largely on domestic law.

Table 1. Comparative overview of selected legal approaches

European Union - AI Act: Uses risk-based regulation and transparency duties. Its AI-related focus includes machine-readable marking and disclosure requirements for certain synthetic content. Its main limitation is that it is not a general criminal code; criminal offences remain mainly within national law.

United States - TAKE IT DOWN Act: Creates a targeted federal criminal prohibition and a platform removal process. It addresses intentional disclosure of non-consensual intimate visual depictions, including certain digital forgeries. Its main limitation is its narrow subject matter.

United Kingdom - Data (Use and Access) Act 2025 and Sexual Offences Act 2003: Creates specific criminal offences relating to intentionally creating or requesting the creation of purported intimate images of adults without consent. Its main limitation is that it addresses a particular form of abuse and does not settle wider questions of attribution.

Council of Europe - AI Framework Convention: Provides a human-rights and rule-of-law framework for AI activities. It supports lifecycle safeguards and accountability, but it does not harmonise all AI-related criminal offences.

## 6. DATA ANALYSIS AND DISCUSSION

This section compares the selected legal materials through a qualitative matrix. The documents themselves are treated as the study material. The purpose is not to count prosecutions or estimate how common deepfake offences are, because this study has not collected a verified dataset of reported cases from the countries discussed. Instead, each legal framework is considered through six indicators: (1) whether it directly addresses AI-generated or manipulated content; (2) whether it creates or supports criminal liability; (3) whether it identifies the actors with legal duties; (4) whether it imposes transparency or platform obligations; (5) whether it supports cross-border cooperation; and (6) whether it contains safeguards for rights and lawful expression. The assessments below are the author's qualitative interpretation of the selected legal materials. They are not official ratings issued by the jurisdictions concerned.

### 6.1 Comparative coding of the legal materials

EU AI Act: Direct coverage of synthetic content is high because Article 50 contains express transparency duties. Its direct criminal-offence element is low because the Act is primarily regulatory. It places duties on providers and deployers. Article 50 itself is not a general platform criminal-liability provision. The EU framework supports coordination between Member States, and its obligations must be read alongside applicable rights and exceptions.

US TAKE IT DOWN Act: Direct coverage is high for the specific category of non-consensual intimate visual depictions, including digital forgeries. It creates a targeted federal offence and an express notice-and-removal process for covered platforms. The Act focuses on the person who intentionally discloses the depiction and the duties of covered platforms. Cross-border enforcement may still raise practical issues.

UK DUAA 2025 and Sexual Offences Act 2003: Direct coverage is high for purported intimate images. Sections 66E and 66F create offences relating to intentional creation and requesting of creation in the circumstances defined by law. These provisions focus on the creator and requester. Wider platform duties arise under other parts of the online-safety framework and should not be confused with the two new offences.

Council of Europe AI Framework Convention: Its coverage is broad because it addresses AI activities through human-rights and rule-of-law principles. It does not create a specific criminal offence for every harmful use of synthetic content. It provides a general framework for accountability and safeguards, while cybercrime cooperation is supported by separate instruments.[23]

### 6.2 What the comparison shows

The comparison shows that the four approaches are different, but they can work alongside one another.[24] The EU model has a wider regulatory scope, although it does not operate as a complete criminal code. The United States and the United Kingdom have introduced targeted offences dealing with non-consensual intimate imagery. The Council of Europe Convention sets out broader principles but does not harmonise the details of criminal liability. As a result, the international response is developing through several legal methods rather than one common system.

[23] Artificial Intelligence Act, supra note 6, art. 50; TAKE IT DOWN Act, supra note 5, §§ 2-3; Data (Use and Access) Act 2025, supra note 20, § 138.

[24] Artificial Intelligence Act, supra note 6, art. 50; TAKE IT DOWN Act, supra note 5, §§ 2-3; Data (Use and Access) Act 2025, supra note 20, § 138; Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, supra note 21.

There is also a difference between regulating technology and punishing harmful conduct. Transparency requirements can help people recognise synthetic material, but a label alone cannot stop someone who intends to commit fraud or harassment. On the other hand, a criminal offence may punish an individual after harm has occurred, but it may not improve the design or use of the systems involved. A useful response therefore needs preventive regulation as well as carefully defined criminal offences.

The selected laws also approach responsibility in different ways. The UK provisions focus on the person who intentionally creates or requests the creation of a purported intimate image.[25] The US Act focuses on intentional disclosure and also places removal duties on covered platforms. The EU AI Act places certain regulatory obligations on providers and deployers. These duties should not be treated as proof of criminal guilt. The comparison supports a role-based approach in which responsibility depends on the conduct, knowledge, control and contribution of the particular actor.

The mental element remains important. The UK provisions refer to intention, lack of consent and the absence of a reasonable belief in consent. The US law also focuses on intentional disclosure. These examples show why a person should not face criminal punishment merely because an AI output caused harm. A user who deliberately misuses a system is not in the same position as a developer who supplied a general-purpose tool without knowledge of the offence.

Enforcement is another concern. A legal prohibition may have limited effect if investigators cannot identify the person responsible, preserve the original file, obtain account information lawfully or secure evidence from another country. The Budapest Convention and its Second Additional Protocol are relevant because they support cooperation in cybercrime investigations and access to electronic evidence. However, cooperation still depends on treaty participation, domestic procedures and respect for fundamental rights.[26]

Finally, safeguards for rights and lawful expression are necessary. Synthetic content may be used for satire, parody, political criticism, art, journalism and research.[27] A criminal law drafted too broadly could restrict legitimate expression and create uncertainty for creators. The law should distinguish harmful deception, exploitation and threats from protected forms

[25] Data (Use and Access) Act 2025, supra note 20, § 138; TAKE IT DOWN Act, supra note 5, §§ 2-3; Artificial Intelligence Act, supra note 6, arts. 9, 17, 50.

[26] Convention on Cybercrime, supra note 4, arts. 29-35; Second Additional Protocol to the Convention on Cybercrime, supra note 4.

[27] European Convention on Human Rights art. 10, Nov. 4, 1950, 213 U.N.T.S. 221; Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, supra note 21.

of communication. Consent, intention, reasonable belief, public interest and available defences should be considered carefully when offences are drafted.

### 6.3 Illustrative application of the matrix

Example A: A person uses a freely available tool to create a fake intimate image of a former partner and shares it with the intention of humiliating that person. The user's conduct is direct and deliberate. The UK provisions may be relevant to the creation or requesting of the image, depending on the facts. The US TAKE IT DOWN Act may be relevant to intentional disclosure if its territorial and other statutory requirements are satisfied. The EU AI Act's transparency framework may also be relevant to providers or deployers falling within its scope, but it does not replace the criminal offence.[28]

Example B: A developer releases a general-purpose image tool for lawful creative work. A user independently uses it to commit an offence. If there is no evidence of intentional assistance, knowledge or another recognised legal basis for liability, automatically prosecuting the developer would be difficult to justify. The developer may still have regulatory duties under the law that applies to its role.[29]

Example C: A platform receives a credible report that an intimate image is non-consensual but does not act promptly. Its position must be examined under the relevant notice-and-removal law and any other applicable duties. Liability should not be assumed simply because the content was hosted. The law must identify the duty, the effect of the notice, the time allowed for action and the consequence of failing to comply.[30]

These examples show why one general rule cannot fairly decide every case. The legal analysis must consider the conduct, the role of each actor, the required mental element, the jurisdiction and the available evidence.

## 7. MAJOR CHALLENGES IN ESTABLISHING CRIMINAL LIABILITY

— IJLJ —
ISSN: 2583-2794

### 7.1 Identification of the Offender

AI-generated content may be created through anonymous accounts, virtual private networks, temporary email addresses or accounts registered with false details. A person may also use someone else's device or account. Identifying the actual user requires lawful access to digital

[28] Data (Use and Access) Act 2025, supra note 20, § 138; TAKE IT DOWN Act, supra note 5, § 2; Artificial Intelligence Act, supra note 6, art. 50.
[29] Horder & Ashworth, supra note 2
[30] TAKE IT DOWN Act, supra note 5, § 3; Online Safety Act 2023, supra note 17, § 10.

records and careful investigation. An account name or profile photograph, by itself, may not prove who operated the account at the relevant time.

### 7.2 Proving Intention and Knowledge

A prompt history, chat record or payment record may help investigators understand what happened. However, such material must be authenticated and considered in its full context. AI systems can sometimes produce unexpected results, and a person may claim that the output was accidental. The court must distinguish a genuine mistake from deliberate misuse and apply the mental element required by the particular offence.[31]

### 7.3 Technical Evidence and Authenticity

Digital files may be edited, compressed, copied or stripped of metadata. AI-detection tools can also produce false results, and their reliability may change as technology develops. Expert evidence should explain the method used, its limitations and the level of uncertainty. A detection result should not be treated as conclusive proof of who created the content or whether that person had criminal intent.[32]

### 7.4 Cross-Border Jurisdiction

The victim, suspect, platform and server may be located in different countries. States may have different rules on jurisdiction, criminal offences, privacy and disclosure of digital information. Mutual legal assistance procedures can take time, while electronic evidence may be deleted or lost. International cooperation should therefore be timely, but it must also follow clear legal procedures and safeguards.[33]

### 7.5 Freedom of Expression and Overcriminalisation

Not every manipulated image or AI-generated statement should be treated as a crime. Satire, parody, fiction, political commentary and artistic expression may all involve synthetic material. An offence based only on falsity or offensiveness could discourage lawful speech. Criminalisation should be limited to clearly defined conduct and harm, and should respect necessity, proportionality and due process.[34]

[31] Horder & Ashworth, supra note 2
[32] Dep't for Sci., Innovation & Tech., supra note 1.
[33] Convention on Cybercrime, supra note 4, arts. 29-35; Second Additional Protocol to the Convention on Cybercrime, supra note 4.
[34] European Convention on Human Rights, supra note 26, art. 10; Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, supra note 21.

## 8. FINDINGS AND RECOMMENDATIONS

### 8.1 Principal Findings

- Traditional criminal-law principles remain relevant, but applying them becomes more difficult when AI systems, several actors and cross-border platforms are involved.[35]
- Deliberate misuse by a user is generally easier to connect with criminal responsibility than the conduct of a general-purpose developer or a platform that had no relevant knowledge.
- The EU, US and UK approaches address different parts of the problem. None of them, by itself, creates a complete global framework for all forms of harmful AI-generated content.[36]
- Transparency and content labelling can help with prevention and assessment, but they cannot replace criminal offences or effective remedies for victims.
- Digital evidence, identification of offenders and international cooperation remain important practical challenges.
- Any new legal framework should protect privacy, dignity and safety while avoiding unnecessary restrictions on lawful expression, research and innovation.

### 8.2 Recommendations

1. Define offences clearly. Legislatures should identify the prohibited conduct, the required mental element, the relevant harm or risk and any available defences. Terms such as deepfake and synthetic content should be explained clearly enough for the public, investigators and courts to understand.
2. Use a role-based model of responsibility. The law should distinguish between the user who intentionally commits an offence, the person who requests or assists it, a developer who knowingly facilitates it, an organisation that deploys the system and a platform that distributes the material. Liability should depend on proof of the relevant conduct and mental element.
3. Improve remedies for victims. Victims should have accessible reporting channels, timely procedures for removing unlawful intimate imagery, support services and

[35] Horder & Ashworth, supra note 2.
[36] Artificial Intelligence Act, supra note 6, art. 50; TAKE IT DOWN Act, supra note 5, §§ 2-3; Data (Use and Access) Act 2025, supra note 20, § 138.

protection against repeated circulation. Where possible, relevant evidence should be preserved lawfully before the content is removed.

4. Strengthen digital investigation. Law-enforcement agencies need training in synthetic-media analysis, chain of custody, metadata, platform records and the limits of AI-detection tools. Expert reports should explain the methods used and any uncertainty in the results.
5. Improve international cooperation. States should make effective use of existing cybercrime and mutual legal assistance frameworks. They may also develop contact points for urgent evidence requests and compatible procedures for preserving electronic records. Such cooperation must remain subject to judicial oversight, privacy and human-rights safeguards.
6. Encourage technical transparency. Providers should develop reliable content-provenance and marking systems where technically possible. Platforms should make reporting and appeal procedures easy to understand. These systems should be treated as useful safeguards, not as perfect proof that content is genuine or fake.
7. Protect lawful expression. Legal exceptions and safeguards should take account of satire, art, journalism, research and public-interest communication. Criminal law should focus on clearly defined conduct such as deception, exploitation and threats, rather than on the mere use of generative AI.
8. Review the law regularly. AI technology changes quickly. Legislatures should periodically review legal definitions, enforcement practices and safeguards. Such reviews should be based on evidence and should include consultation with victims, civil society, researchers and the technology sector.

## 9. CONCLUSION

Generative artificial intelligence has made it easier to create and share digital content. It has also increased the speed and realism of some forms of online abuse. The main legal task is not to punish technology itself. It is to identify the conduct of the person or organisation responsible and decide whether that conduct satisfies the requirements of a criminal offence.

The comparison in this paper shows that different jurisdictions have chosen different approaches. The European Union has focused on risk-based regulation and transparency. The United States has adopted a targeted federal offence and removal duties for covered platforms in relation to non-consensual intimate visual depictions. The United Kingdom has introduced offences relating to the intentional creation and requesting of purported intimate images

without consent. The Council of Europe provides a wider human-rights framework, while international cybercrime instruments support cooperation on investigations and electronic evidence. These measures are useful, but they do not yet form one complete global criminal-law system.[37]

Traditional principles such as actus reus, mens rea, causation and participation continue to provide the foundation for criminal liability. They must be applied carefully. A person who deliberately uses AI to commit fraud, threaten someone or create non-consensual intimate material should not be able to avoid responsibility merely by saying that technology was involved. At the same time, a developer or platform should not be held criminally responsible without proof of the conduct and mental element required by law.[38]

A balanced global response should therefore combine clearly drafted offences, responsibility based on each actor's role, effective investigation, international cooperation, victim protection and safeguards for freedom of expression. This approach can make it more difficult for offenders to hide behind technology while ensuring that legitimate innovation and communication are not restricted without proper reason.[39]

— I J L J —
ISSN: 2583-2794

[37] Artificial Intelligence Act, supra note 6, art. 50; TAKE IT DOWN Act, supra note 5, §§ 2-3; Data (Use and Access) Act 2025, supra note 20, § 138; Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, supra note 21.

[38] Horder & Ashworth, supra note 2.

[39] European Convention on Human Rights, supra note 26, art. 10.

## REFERENCES


### A. Legislation and Treaty Materials

1. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689) 1.
2. TAKE IT DOWN Act, Pub. L. No. 119-12, 139 Stat. 55 (2025).
3. Online Safety Act 2023, c. 50 (UK).
4. Data (Use and Access) Act 2025, c. 18 (UK).
5. Sexual Offences Act 2003, c. 42 (UK).
6. Data (Use and Access) Act 2025 (Commencement No. 5) Regulations 2026, S.I. 2026/31 (UK).
7. Convention on Cybercrime, Nov. 23, 2001, E.T.S. No. 185.
8. Second Additional Protocol to the Convention on Cybercrime on Enhanced Co-operation and Disclosure of Electronic Evidence, May 12, 2022, C.E.T.S. No. 224.
9. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, May 17, 2024, C.E.T.S. No. 225.
10. European Convention on Human Rights, Nov. 4, 1950, 213 U.N.T.S. 221.

### B. Books and Secondary Authorities

1. Jeremy Horder & Andrew Ashworth, Ashworth's Principles of Criminal Law (9th ed. 2019).

### C. Official Institutional Materials

1. Department for Science, Innovation & Technology, Deepfake Detection Technology (Mar. 26, 2026).
2. European Commission, AI Act Service Desk, "Article 50: Transparency Obligations for Providers and Deployers of Certain AI Systems."
3. European Commission, AI Act Service Desk, "When Does Enforcement Start?"


### Methodological Note

The comparative matrix and examples used in this paper are the author's qualitative analysis of the legal materials listed above. They are not empirical statistics, survey results or official assessments issued by the jurisdictions discussed. Before submission, the author should check the journal's required citation style, add pinpoint references to the relevant statutory provisions

and official materials, and confirm whether a separate declaration on originality, funding or conflicts of interest is required.



— I J L J —
ISSN: 2583-2794