

# The Deepfake Dilemma: Where the Digital Services Act Fails

Olivia Morse

International Policy Review - Policy of the Month

Date of Publication

## Introduction

In March 2026, German television presenter Collien Fernandes stood in front of thousands of protesters in Hamburg wearing a bulletproof vest. For nearly a decade, hundreds of AI-generated pornographic images of her had circulated on the internet, stealing her likeness and dignity. She alleged that it was her husband, Christian Ulmen, who had been the creator of the false images of her.<sup>1</sup> Upon filing her criminal complaint in Germany, she came to discover there was no law that explicitly criminalised the creation of such material. Instead, she was forced to pursue legal action in Spain, stating that Germany was “an absolute haven for perpetrators”, and that Spain had stronger protection of gender-based digital violence.<sup>2</sup> Fernandes’s case was not an isolated error, but a symptom of a larger structural problem within the European Union’s regulatory framework. While the EU has positioned itself as the global leader in technological governance, victims of abuse from AI-generated images continue to suffer without proper avenues to seek justice.

As such, this article will explore the European Union’s Digital Services Act (DSA), which is the main mechanism for combating the uncontrollable spread of AI-generated non-consensual pornographic material, and whether its platform-based regulatory framework is adequate.<sup>3</sup> While ambitious in its design, the DSA is severely limited by its heavy reliance on Member States definitions of illegality and enforcement systems, leaving victims such as Collien Fernandes unprotected from online abuse.

### 1.1 The Digital Services Act: Provisions and Objectives

The Digital Services Act (Regulation 2022/2065) entered into force in 2022 and became fully applicable in February 2024, establishing a legal framework that addresses content moderation and platform transparency across the European Union.<sup>4</sup> It established a series of mechanisms to attempt to stop the dissemination of illegal content, including, in principle, AI-generated non-consensual intimate imagery.<sup>5</sup> Crucially, however, the DSA does not define what illegal content is, instead relying on Union or national

---

<sup>1</sup> Zoledziowski, Anya. 2026. “German celebrity Collien Fernandes says husband spread sexual deepfakes of her for years.” CBC, April 4, 2026.

<https://www.cbc.ca/news/entertainment/collien-fernandes-christian-ulmen-sexual-deepfake-allegations-9.7136770>.

<sup>2</sup> Bethany Bell, “German Outcry Over Deep Fake Porn Targeting TV Presenter Prompts Plans for Change in Law”, BBC, March 24, 2026

<sup>3</sup> European Parliament and Council of the European Union, Regulation (EU) 2022/2065 of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act), October 19, 2022, <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>

<sup>4</sup> Wikipedia contributors. 2026. “Digital Services Act.” Wikipedia. April 10, 2026. [https://en.wikipedia.org/wiki/Digital\\_Services\\_Act](https://en.wikipedia.org/wiki/Digital_Services_Act).

<sup>5</sup> “The Digital Services Act.” n.d. Shaping Europe’s Digital Future. <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act>.

law, meaning that whether a particular category of content falls within the scope of the DSA is dependent entirely on whether it has already been criminalised in the Member State.<sup>6</sup>

In Article 16 of the Act, the notice and action mechanism is established. This mechanism acts as the primary content removal tool and it requires the online platform, such as Facebook, to establish a mechanism allowing individual users to notify it of illegal content. The provider must process the notice “in a timely, diligent, non-arbitrary and objective manner.” This places most of the responsibility for content moderation on the individual platforms themselves, which they are then obliged to provide a report on available to the public once a year, as per Article 15.<sup>7</sup>

Additionally, in Article 22, the system for “trusted flaggers” is established, which are entities that have been officially certified by a national Digital Services Coordinator who are experts at detecting certain types of illegal content online, including AI-generated intimate imagery.<sup>8</sup> When the flagger submits notices to online platforms identifying content they consider illegal, it is required that their submissions are treated with priority compared to ordinary users by the platforms. In order to become a trusted flagger, the entity must fulfil three conditions: (1) expertise and competence, (2) independence and (3) diligence, accuracy and objectivity.<sup>9</sup> If these three qualities are fulfilled, the entity is awarded the “trusted flagger” status that is valid across the entire EU. To ensure full transparency, the flaggers are obligated to publish one annual report detailing their activities.<sup>10</sup>

## **1.2: Analysis: Can the DSA Protect Victims?**

While these provisions seemingly provide a robust framework for addressing illegal online content, the central question this article addresses is whether the DSA’s platform-based regulatory model is adequate to protect victims of AI-generated non-consensual intimate imagery. From the Act’s structure, it becomes clear that there are major enforcement gaps, most notably, the illegality dependency of the notice-and-action mechanism, which is the main content removal mechanism.<sup>11</sup> If a Member State has not criminalized certain content at the time of the crime, as Germany did not formally criminalise AI-generated non-consensual

---

<sup>6</sup> Ibid.

<sup>7</sup> *Digital Services Act*, Regulation (EU) 2022/2065.

<sup>8</sup> Ibid, Article 22

<sup>9</sup> Ibid.

<sup>10</sup> Ibid.

<sup>11</sup> *Digital Services Act*, Regulation (EU) 2022/2065, art. 16; see also European Commission, “The Digital Services Act,” <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act>.

pornographic imagery during Collien Fernandes's case, then the victim is awarded no protection.<sup>12</sup> Inherently, this leads to the DSA being held captive by the weakest jurisdiction. If illegal AI-generated images of a Spanish citizen are distributed across Spain, Germany and France, but only Spain and France have criminalised such material under national law, platforms operating in Germany are not obliged to remove the content under Article 16. This leaves the injured without recourse in an entire jurisdiction and completely vulnerable.

Even in jurisdictions where national law criminalises AI-generated non-consensual intimate imagery, the Digital Services Act's enforcement procedure meets another limitation: the speed at which this form of content spreads. Due to the accessibility of the tools necessary to create such content, the lack of skill required and the ease of sharing such content across the world, these intimate images can spread at a speed and scale that the DSA's structure is unable to handle.<sup>13</sup> A single person can rapidly disseminate dozens of fabricated images, which can be reuploaded and shared across platforms in minutes.<sup>14</sup> In order for a single photo to be taken down from a platform, the victim or a third party must first submit a notice to the hosting service under Article 16.<sup>15</sup> The provider must then process this notice in an indeterminate period of time, since no timeframe is required by the DSA. While this entire process transpires, the fabricated image may be shared thousands of times across dozens of platforms, each requiring its own separate notice. Ultimately, this process leaves the victim in the same position as if no protection was provided at all. Even if the platform removes the photo, for victims like Fernandes, the result is a system where the harm is immediate but the remedy, in practice, is largely symbolic.

### **1.3: Broader Implications of the DSA:**

Collectively, these enforcement gaps exemplify that the DSA was exclusively designed as a platform-based moderation framework, not as an instrument for victim protection. It presupposes that the content has been uploaded, identified and reported, offering no protection during the creation or initial dissemination of the material. The burden of the DSA's inadequacies falls disproportionately on women and girls, for whom AI-generated non-consensual intimate imagery is not just something to be moderated but a form of online gender-based violence. According to the United Nations, "90%-95% of all online

---

<sup>12</sup> *Digital Services Act*, Regulation (EU) 2022/2065, art. 16; see also Bethany Bell, "German Outcry Over Deep Fake Porn Targeting TV Presenter Prompts Plans for Change in Law", BBC.

<sup>13</sup> UNWomen, "AI-powered online abuse: How AI is amplifying violence against women and what can stop it," November 18, 2025

<https://knowledge.unwomen.org/en/articles/faqs/ai-powered-online-abuse-how-ai-is-amplifying-violence-against-women-and-what-can-stop-it>.

<sup>14</sup> Ibid.

<sup>15</sup> *Digital Services Act*, Regulation (EU) 2022/2065., Art. 16

deepfakes are non-consensual pornographic images” with around “90% depicting women”.<sup>16</sup> Yet within the DSA’s own regulatory framework, platforms have failed to treat this as a priority, letting hours go by without taking down initial images or responding to notices from users.<sup>17</sup> The result is a regulatory architecture that, despite its ambition, offers women and girls protection in principle but not in practice.

#### **1.4: Policy Recommendation:**

If the DSA is to fulfil its promise as a global regulatory standard, Article 16 must be reformed to act independently from national legislation.<sup>18</sup> An essential first step would be introducing a harmonised EU-wide definition of AI-generated non-consensual intimate imagery in order to ensure that content removal obligations are uniform across Member States, regardless of domestic law. Alongside this definitional reform, Article 16 should enforce takedown timelines. For example, a 24-hour removal window for content flagged as “high risk” and a 12-hour takedown window for confirmed illegal content under the new harmonised definition. These two changes, while seemingly simple, introduce a system where the notice-and-action mechanism is a meaningful remedy rather than a procedural formality.

#### **1.5: Conclusion:**

Despite the Digital Service Act’s ambitious structural design, ultimately, it fails to protect victims from online criminal behaviour, including AI-generated non-consensual intimate images. For individuals such as Collien Fernandes they are often left completely vulnerable to the false images generated of them, potentially in countries that do not even define this conduct as illegal.<sup>19</sup> Even if the illegality threshold is met, the process to take down such an image may take hours or days, at which point the damage has already occurred, and the photo will likely never be fully erased from the internet.<sup>20</sup> These risks are inherently not gender neutral because this form of crime is nearly always targeted and against women or young girls.<sup>21</sup> As such, victims, who are almost always women, are left in a situation where they cannot stop others from accessing the materials necessary or producing AI-generated pornographic images of themselves, and without a way to protect themselves if it does happen. The DSA acts as a form of reassurance rather than an actual remedy, and when the time comes that a victim needs platform accountability, no actual remedies are

---

<sup>16</sup> UNWomen, “AI-powered online abuse: How AI is amplifying violence against women and what can stop it”, 18 November 2025

<sup>17</sup> Ibid.

<sup>18</sup> “The Digital Services Act and the EU as the Global Regulator of the Internet | Chicago Journal of International Law.” n.d. <https://cjl.uchicago.edu/print-archive/digital-services-act-and-eu-global-regulator-internet>.

<sup>19</sup> Bell, “German Outcry Over Deep Fake Porn Targeting TV Presenter Prompts Plans for Change in Law”

<sup>20</sup> *Digital Services Act*, Regulation (EU) 2022/2065., Art. 16

<sup>21</sup> UNWomen, “AI-powered online abuse: How AI is amplifying violence against women and what can stop it”, 18 November 2025

available. The EU has positioned itself as a global leader in technology regulation, but leadership demands more than ambition; it demands a framework that protects the people it was written for.

## Bibliography

“AI-powered online abuse: How AI is amplifying violence against women and what can stop it.” 2025. UNWomen.Org. November 18, 2025. Accessed April 13, 2026.

[https://knowledge.unwomen.org/en/articles/faqs/ai-powered-online-abuse-how-ai-is-amplify  
Ng-violence-against-women-and-what-can-stop-it](https://knowledge.unwomen.org/en/articles/faqs/ai-powered-online-abuse-how-ai-is-amplify-<br/>Ng-violence-against-women-and-what-can-stop-it).

Bell, Bethany. 2026. “German Outcry Over Deep Fake Porn Targeting TV Presenter Prompts Plans for Change in Law.” March 24, 2026. <https://www.bbc.com/news/articles/ckg9xvrxz8jo>.

“The Digital Services Act.” n.d. Shaping Europe’s Digital Future.  
<https://digital-strategy.ec.europa.eu/en/policies/digital-services-act>.

“The Digital Services Act and the EU as the Global Regulator of the Internet | Chicago Journal of International Law.” n.d.  
<https://cjl.uchicago.edu/print-archive/digital-services-act-and-eu-global-regulator-internet>.

Wikipedia contributors. 2026. “Digital Services Act.” Wikipedia. April 10, 2026.  
[https://en.wikipedia.org/wiki/Digital\\_Services\\_Act](https://en.wikipedia.org/wiki/Digital_Services_Act).

Zoledziowski, Anya. 2026. “German celebrity Collien Fernandes says husband spread sexual deepfakes of her for years.” CBC, April 4, 2026.  
[https://www.cbc.ca/news/entertainment/collien-fernandes-christian-ulmen-sexual-deepfake-allegatio-  
ns-9.7136770](https://www.cbc.ca/news/entertainment/collien-fernandes-christian-ulmen-sexual-deepfake-allegatio-<br/>ns-9.7136770).