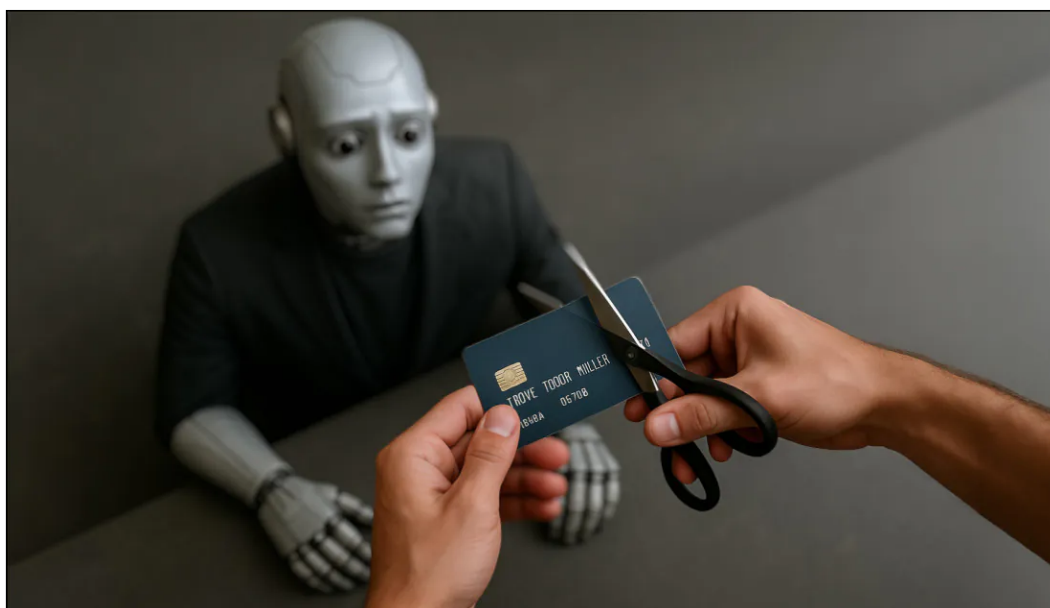


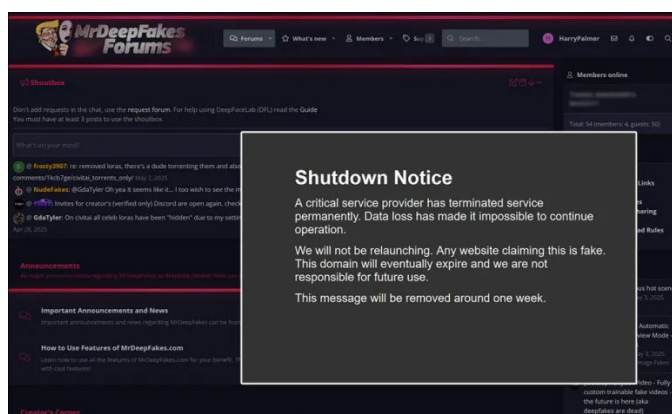
CivitAI in New Payment Provider Crisis, as Trump Signs Anti-Deepfake Act

Originally published at <https://www.unite.ai/civitai-in-new-payment-provider-crisis-as-trump-signs-anti-deepfake-act/>



President Trump has now signed the Take It Down Act, criminalizing sexual deepfakes at a federal level in the US. At the same time, the CivitAI community's bid to 'clean up its act' regarding NSFW AI and celeb output has ultimately failed to appease payment processors, leading the site to seek alternatives or face shutdown. All this in the mere two weeks since the oldest and biggest deepfake porn site in the world went offline...

It has been a momentous few weeks for the state of unregulated image and video deepfaking. Just over two weeks ago, the number #1 domain for the community sharing of celebrity deepfake porn. Mr. Deepfakes, suddenly took itself offline after more than seven years in a dominant and [much-studied](#) position as the global locus for sexualized AI celebrity content. By the time it went down, the site was receiving an [average of more than five million visits](#) a month.



Background, the Mr. Deepfakes domain in early May; inset, the suspension notice, now replaced by a 404 error, since the domain was apparently purchased by an unknown buyer on the 4th of May, 2025 (<https://www.whois.com/whois/mrdeepfakes.com>). Source: mrdeepfakes.com

The cessation of services for Mr. Deepfakes was officially [attributed](#) to the withdrawal of a 'critical service provider' (see inset image above, which was replaced by domain failure within a week). However, a collaborative journalistic investigation had [de-anonymized a key figure](#) behind Mr.Deepfakes directly prior to the shutdown, allowing for the possibility that the site was shuttered for that individual's personal and/or legal reasons.

Around the same time, CivitAI, the commercial platform widely used for celebrity and NSFW [LoRAs](#), imposed a set of unusual and controversial [self-censorship measures](#). These affected deepfake generation, model hosting, and a broader slate of new rules and restrictions, including full bans on [certain](#) marginal NSFW fetishes. and what it termed 'extremist ideologies'.

These measures were prompted by payment providers apparently threatening to withdraw services from the domain unless changes regarding NSFW content and celebrity AI depictions were made.

CivitAI Cut Off

As of today, it appears that the measures taken by CivitAI have not appeased VISA and Mastercard : a [new post](#)[†] at the site, from Community Engagement Manager Alasdair Nicoll, reveals that card payments for CivitAI (whose 'buzz' virtual money system is mostly powered by real-world credit and debit cards) will be halted from this Friday (May 23rd, 2025).

This will prevent users from renewing monthly memberships or buying new buzz. Though Nicoll advises that users can maintain current membership privileges by switching to an annual membership ([costing](#)^{††} \$100-\$550 USD) before Friday, clearly the future is somewhat uncertain for the domain at this time (It should be noted that annual memberships [went live](#) at the same time that the announcement about the loss of payment processors was made).

Regarding the lack of a payment processor, Nicoll says 'We're talking to every provider comfortable with AI innovation'.

As to the failure of recent efforts to adequately rethink the site's [oft-criticized](#) policies around celeb AI and NSFW content, Nicoll states in the post:

'Some payment companies label generative-AI platforms high risk, especially when we allow user-generated mature content, even when it's legal and moderated. That policy choice, not anything users did, forced the cutoff.'

A comment from user 'Faeia', designated as the company's chief of staff in their CivitAI [profile](#)^{*}, adds context to this announcement:

'Just to clarify, we're being removed from the payment processor because we chose not to remove NSFW and adult content from the platform. We remain committed to supporting all kinds of creators and are working on alternative solutions.'

As a [traditional driver of new technologies](#), it's not uncommon for NSFW content to be used to [kick-start interest](#) in a domain, technology or platform – only for the initial adherents to be rejected once enough 'legitimate' capital and/or a user-base is established (i.e., enough users for the entity to survive, when shorn of a NSFW context).

It seemed for a while that CivitAI would follow [Tumblr](#) and diverse other initiatives down this route towards a 'sanitized' product ready to forget its roots. However, the additional and growing [controversy/stigma](#) around AI-generated content of *any kind* represents a cumulative weight that seems set to prevent a last-minute rescue, in this case. In the meantime, the official announcement advises users to adopt crypto as an alternative payment method.

Fake Out

The advent of President Donald Trump enthusiastically [signing](#) the Federal [TAKE IT DOWN](#) Act is likely to have influenced some of these events. The new law criminalizes the distribution of non-consensual intimate imagery, including AI-generated deepfakes.

The legislation mandates that platforms remove flagged content within 48 hours, with enforcement overseen by the Federal Trade Commission. The criminal provisions of the law take effect immediately, allowing for the prosecution of individuals who knowingly publish or threaten to publish non-consensual intimate images (including AI-generated deepfakes) within the purview of the United States.

While the law received rare bipartisan support, as well as backing from tech companies and advocacy groups, critics argue it may suppress legitimate content and threaten privacy tools like encryption. Last month the Electronic Frontier Foundation (EFF) declared [opposition to the bill](#), asserting that the takedown mechanisms it mandates target a broader swathe of material than the narrower definition of non-consensual intimate imagery found elsewhere in the legislation.

‘The takedown provision in TAKE IT DOWN applies to a much broader category of content—potentially any images involving intimate or sexual content—than the narrower NCII definitions found elsewhere in the bill. The takedown provision also lacks critical safeguards against frivolous or bad-faith takedown requests.

‘Services will rely on automated filters, which are infamously blunt tools. They frequently flag legal content, from fair-use commentary to news reporting. The law’s tight time frame requires that apps and websites remove speech within 48 hours, rarely enough time to verify whether the speech is actually illegal.

‘As a result, online service providers, particularly smaller ones, will likely choose to avoid the onerous legal risk by simply depublishing the speech rather than even attempting to verify it.’

Platforms now have up to one year from the law’s enactment to establish a formal notice-and-takedown process, enabling affected individuals or their representatives to invoke the statute in seeking content removal.

This means that although the criminal provisions are immediately in effect, platforms are not legally obligated to comply with the takedown infrastructure (such as receiving and processing requests) until that one-year window has elapsed.

Does the TAKE IT DOWN Act Cover AI-Generated Celebrity Content?

Though the TAKE IT DOWN Act crosses all state borders, it does not necessarily outlaw all AI-driven media of celebrities. The act criminalizes the distribution of non-consensual intimate images, *including* AI-generated deepfakes, only when the depicted individual had a *reasonable expectation of privacy*:

The act states:

“(2) OFFENSE INVOLVING AUTHENTIC INTIMATE VISUAL DEPICTIONS.—

“(A) INVOLVING ADULTS.—Except [for evidentiary, reporting purposes, etc.], it shall be unlawful for any person, in interstate or foreign commerce, to use an interactive computer service to knowingly publish an intimate visual depiction of an identifiable individual who is not a minor if—

“(i) the intimate visual depiction was obtained or created under circumstances in which the person knew or reasonably should have known the identifiable individual had a reasonable expectation of privacy;

“(ii) what is depicted was not voluntarily exposed by the identifiable individual in a public or commercial setting [i.e., self-published porn];

“(iii) what is depicted is not a matter of public concern; and

“(iv) publication of the intimate visual depiction—

“(I) is intended to cause harm; or

“(II) causes harm, including psychological, financial, or reputational harm, to the identifiable individual.

The ‘reasonable expectation of privacy’ contingency applied here has [not traditionally favored](#) the rights of celebrities. Depending on the case law that eventually emerges, it’s possible that even *explicit* AI-generated content involving public figures in public or commercial settings may not fall under the Act’s prohibitions.

The final clause about determining the *extent* of harm is [famously elastic](#) in legal terms, and in this sense adds nothing particularly novel to the legislative burden. However, the *intent* to cause harm would seem to limit the scope of the Act to the context of ‘revenge porn’, where an (unknown) ex-partner publishes real or fake media content of an (equally unknown) other ex-partner.

While the law's 'harm' requirement may seem ill-suited to cases where anonymous users post AI-generated depictions of celebrities, it could prove more relevant in stalking scenarios, where a broader pattern of harassment supports the conclusion that an individual has deliberately and maliciously targeted a public figure across multiple fronts.

Though the Act's reference to 'covered platforms' excludes private channels such as Signal or email from its takedown provisions, this exclusion applies only to the obligation to implement a formal removal mechanism by May 2026. It does not mean that non-consensual AI or real depictions shared through private communications fall outside the scope of the law's criminal prohibitions.

Obviously, a lack of on-site reporting mechanisms does not hinder affected parties from reporting what is now illegal content to the police; neither are such parties precluded from using whatever conventional contact methods a site may make available to make a complaint and request the removal of offending material.

The Rights Left Behind

More than seven years of mounting public and media criticism over deepfake content appear to have culminated within an unusually short span of time. However, while the TAKE IT DOWN Act offers sweeping federal prohibitions, it may not apply in every case involving AI-generated simulations, leaving certain scenarios to be addressed under the growing patchwork of state-level deepfake legislation, where the laws passed often reflect 'local interest'.

For instance, in California, the [California Celebrities Rights Act](#) limits the exclusive use of a celebrity's identity to themselves and their estate, even after their death; conversely, Tennessee's [ELVIS Act](#) focuses on safeguarding musicians from unauthorized AI-generated voice and image reproductions, with each case reflecting a targeted approach to interest groups that are prominent at state level.

Most states [now have laws](#) targeting sexual deepfakes, though many stop short of clarifying whether those protections extend equally to private individuals and public figures. Meanwhile, the political deepfakes that reportedly helped [spur](#) Donald Trump's support for the new federal law may, in practice, [run up against constitutional barriers](#) in certain contexts.

[†] Archived version: <https://web.archive.org/web/20250520024834/https://civitai.com/articles/14945>

^{††} Archived version (does not feature monthly prices):
<https://web.archive.org/web/20250425020325/https://civitai.green/pricing>

* The actual 'chief of staff' to the CEO at CivitAI is [listed](#) at LinkedIn under an unrelated name, while the similar-sounding 'Faiona' is an [official CivitAI staff moderator](#) at the domain's [subreddit](#).

First published Tuesday, May 20, 2025

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More