

Meta 'supreme court' takes on cases of deepfake porn

April 16 2024



Meta's independent oversight board can make recommendations regarding the social media giant's deepfake porn policies but it is up to the tech firm to actually make any changes.

Meta's oversight board said Tuesday it is scrutinizing the social media titan's deepfake porn policies, through the lens of two cases.

The move by what is referred to as a Meta "supreme court" for content moderation disputes comes just months after the widespread sharing of lewd AI-generated images of megastar Taylor Swift on X, formerly Twitter.

The Meta board picked its two cases, regarding images shared on Instagram and Facebook, to "assess whether Meta's policies and its enforcement practices are effective at addressing explicit AI-generated imagery," it said in the release.

The board can make recommendations regarding the social media giant's deepfake porn policies but it is up to the tech firm to actually make any changes.

The first case taken up by the Meta Oversight Board involves an AI-generated image of a nude woman posted on Instagram.

The woman pictured resembled a public figure in India, sparking complaints from users in that country.

Meta left the image up, later saying it did so in error, the board said.

The second case involves a picture posted to a Facebook group devoted to AI creations.

That image depicted a nude woman resembling "an American public figure" with a man groping one of her breasts, the board said in a release.

The board did not name the woman, who it said was identified in a caption on the synthetic image at issue.

Meta removed the image for violating its harassment policy, and the user

who posted the content appealed the decision, according to the board.

People were invited to submit comment, particularly on the gravity of harms posed by deepfake pornography and the harm it does to women who are public figures.

Deepfake porn images of celebrities are not new, but activists and regulators are worried that easy-to-use tools employing generative AI will create an uncontrollable flood of toxic or harmful content.

The targeting of Swift, one of the world's top-streamed artists whose latest concert tour propelled her to the top of American fame, shined a spotlight on the phenomenon, with her legions of fans outraged at the development.

"It is alarming," said White House Press Secretary Karine Jean-Pierre, when asked about the images at the time.

"Sadly we know that lack of enforcement (by the tech platforms) disproportionately impacts women and they also impact girls who are the overwhelming targets of online harassment," Jean-Pierre added.

© 2024 AFP

Citation: Meta 'supreme court' takes on cases of deepfake porn (2024, April 16) retrieved 2 May 2024 from <https://techxplore.com/news/2024-04-meta-supreme-court-cases-deepfake.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.