ALM | LAW.COM

New York Law Journal

The 'Deepfake Era': How To Navigate AI-Generated Content

By Sheila Swaroop and Sara Witty

October 7, 2024

Deepfakes, a high-profile category of AI-generated content, have received widespread attention in connection with the upcoming U.S. presidential election. Two recent examples involving celebrity Taylor Swift demonstrate the intersection between AI-generated content, politics, and intellectual property rights:

- An AI-generated image of Taylor Swift dressed as Uncle Sam, accompanied by the message "Taylor wants YOU to VOTE for DONALD TRUMP."
- AI-generated images of women wearing "Swifties for Trump" T-shirts.

After the Trump campaign and others shared these images on social media, Swift entered her "Deepfake Era." She responded to the images by articulating her political views and expressing support for Vice President Kamala Harris with an authorized social media posting.

These examples of AI-generated images linking a public figure and her fan base to a political endorsement highlight the need for appropriate legal protection to protect against deepfakes. This article examines a few of those options, the remedies that can be achieved if a claim is successful, and new laws targeting AI-generated content.

By freshidea/Adobe Stock

Protect Through Copyright Applications

When AI platforms generate lifelike images based on real people, such as the image of Swift dressed as Uncle Sam, one potential avenue to address these deepfakes is through copyright law. Copyrights protect artists from unauthorized use or distribution of their work, and also protect against the creation of derivative works based on a copyrighted image. If the AI system generating the image of Swift was

trained using copyrighted images, that activity may support a claim for copyright infringement. Artists have begun asserting these types of claims against AI image platforms based on the platforms' use of copyrighted images without consent, and these cases are starting to work their way through the judicial system. Copyright claims require an artist to register their work with the Copyright Office and to show similarity between the copyrighted work and the unauthorized image. A successful claim for copyright infringement may not be, to quote a Taylor Swift song, "a great war," but it can result in a court order stopping further infringement, monetary damages, and recovery of attorneys' fees and costs.

Secure Trademark Registrations

AI-generated images of fans wearing "Swifties for Trump" T-shirts could be addressed through trademark law. Trademark law protects against unauthorized use of a mark when it is likely to cause confusion as to the source of the goods or services at issue. The "Swifties for Trump" t-shirts may seem to have a familiar "style" as Swift herself has several federal trademark registrations for the mark "SWIFTIE," including one for clothing. Those trademark registrations could form the basis for a trademark infringement claim if there is a likelihood of confusion as to the source of the "Swifties for Trump" T-shirts. This could be shown with evidence that consumers associate the T-shirts with Swift. The remedies for a successful trademark infringement claim include a court order stopping use of the infringing mark, destruction of the products at issue, monetary damages, and recovery of attorneys' fees and costs.

Assert Right of Publicity Claims

Another potential avenue to address the "bad blood" caused by deepfakes is to assert a right of publicity claim. The requirements for these claims vary by state, but they generally involve an unauthorized use of a name, image, or likeness for commercial benefit, and the likelihood of causing injury from that use. In these examples, a right of publicity claim could be available for Swift, but could be difficult for the unidentified fans unless the images are based on actual individuals. Right of publicity claims can be brought against those who create the unauthorized images, and those who distribute them. In the context of AI-generated deepfakes, that could include the creators of the images, as well as those who benefit from distribution of the images. The remedies for this claim can include an order stopping further use of the unauthorized images and monetary damages.

Consider Defamation Claims

A claim for defamation is another option. These claims typically involve the publication or communication of a false statement to someone else, fault by the person who publishes or communicates the statement, and harm to the person who is the subject of the statement. If the subject of the defaming statement chooses not to "tolerate it," the available remedies may include an injunction, monetary damages, and a retraction of the false statement.

Stay Current On New Laws

Some states have created new laws to address the increasing prevalence of AI-generated images. In September 2024, California enacted several new laws in this area. This includes SB 942, known as the California AI Transparency

Act, which requires certain content providers to provide an AI detection tool at no cost to users that allows for an assessment of whether content was created or altered by the provider's generative AI system.

The law also requires content providers to include disclosures when AI-generated content is used in images, videos, or audio materials. The content providers subject to this law are those who create, code, or produce a generative AI system that is both publicly accessible within California and has over one million monthly visitors or users.

In addition, the law creates a civil penalty of $5,000 per violation for content providers who do not comply. This law will go into effect on January 1, 2026, so there will be some "blank space" before the public can evaluate the disclosures and the tools that will be used to identify AI-generated contact.

California has also implemented bills to address the use of AI-generated content during elections. AB 2655, known as the "Defending Democracy from Deepfake Deception Act of 2024," requires that large online platforms block the posting of "materially deceptive content" related to elections in California, during specified periods before and after an election. The platforms subject to this law are websites and applications with at least one million California users during the previous year. Materially deceptive content is audio or visual media that is digitally created or modified such that it would appear to be an authentic record of the content. This does not include media that contains minor changes that do not significantly change the meaning of the original content.

As part of this bill, large online platforms must develop procedures to identify and remove materially deceptive content, and to label certain additional content as inauthentic, fake, or false during the same time periods. The platforms must provide an easily accessible reporting mechanism for California residents to report the "hoax" posts. Candidates who are the subject of the materially deceptive images can seek injunctive relief if the online platform does not take action within 72 hours of a submitted report.

The rise of AI-generated images has created new legal challenges, particularly for the unauthorized use of an individual's name, content, and images. The recent examples involving images of Taylor Swift and her fan base demonstrate the risks of deepfakes when used in the political context. Instead of just "shaking it off," those concerned with AI-generated content can proactively address this issue by securing copyright and trademark protection and by relying on additional state laws.

Sheila Swaroop *is a partner at Knobbe Martens and the firm's litigation practice chair.* **Sara Witty** *is an associate at the firm whose practice includes litigation and prosecution work for trademarks, copyrights, and trade dress.*