# AI-generated videos offer creative leaps but demand urgent ethical safeguards



# The Promise and Perils of AI-Generated Video Content

The emergence of artificial intelligence (AI) has profoundly transformed various aspects of our lives, notably in the realm of content creation. While AI has gained notoriety for generating deepfakes of public figures such as Taylor Swift and President Barack Obama, it also presents immense potential for enhancing video content in both enterprise and consumer sectors. This duality was at the forefront of discussions during Fortune’s Brainstorm AI conference held in London, where industry experts emphasised the pressing need to balance innovation with ethical considerations in AI's implementation.

Daniel Hulme, CEO of Satalia—WPP's enterprise AI arm—articulated this need during a panel discussion. He noted, “AI allows us to create content incredibly rapidly, but you have to have the right guardrails and structures in place to mitigate the risks.” This perspective resonates particularly in light of the recent proliferation of harmful deepfake content, notably in the realm of nonconsensual imagery which has raised alarm bells throughout society. Legal experts like Carrie Goldberg have highlighted the urgency for enhanced legal protections, especially as deepfake technology grows more accessible.

The vulnerabilities exposed by AI misuse are not limited to celebrities; they affect everyday individuals, including minors. Legislation such as the proposed Preventing Deepfakes of Intimate Images Act is part of a broader initiative to address these issues. Yet, many of these bills have historically struggled to gain traction in Congress. Without substantial federal legislation, the challenge of regulating AI-generated content remains formidable, allowing harmful deepfakes to proliferate across platforms.

Peter Hill, the Chief Technology Officer of Synthesia, outlined a hopeful vision for AI in the creative domain. He advocated for a redefinition of AI, moving away from the idea that it merely replicates human abilities. Hill proposed that future AI should focus on being goal-directed and adaptive, thus reflecting the innate resilience and creativity of human behaviour. He stated, “We tend not to see AI systems that are very adaptive. I think that’s the new and next opportunity.”

To illustrate the advancements in AI technology, Hill presented an AI-generated video of himself that was strikingly lifelike, showcasing its potential in training and corporate communications. However, the ethical responsibilities associated with this technology cannot be understated. Hulme cautioned that brands must navigate the risks inherent in AI-generated content, citing an AI avatar created for a commercial that bore an uncanny resemblance to Jennifer Lopez. He remarked, “It’s our responsibility to make sure that their brand is put in the absolute best light,” indicating that corporate usage of AI necessitates rigorous oversight to avoid misrepresentation or unintended consequences.

The perils of deepfakes extend beyond mere market concerns. High-profile collaborations are calling for stricter regulations to combat the technologically facilitated dissemination of misleading content. An open letter led by notable AI figures like Yoshua Bengio called for criminalising the creation of harmful deepfakes, highlighting the potential for significant social harm. Indeed, research suggests that AI-driven misinformation could heavily influence public opinion and even political processes—a reality that consumers and creators must reckon with as technology continues to evolve.

Despite the complexities and challenges associated with AI technology, industry leaders like Hulme remain optimistic. He articulates a vision in which companies take proactive measures to ensure that AI serves public good rather than perpetuating biases and creating social bubbles. He stressed, “We have a duty of care to make sure that we’re using these technologies in the right and responsible way,” underlining the need for corporate vigilance in the AI age.

As discussions around AI and its implications for society continue, the focus shifts towards responsible use and regulation. While AI's ability to create content swiftly holds great promise, the technology's dark side must not be ignored. For as long as misuse persists, experts advocate for careful implementation, legislative action, and community education to harness AI's benefits while safeguarding against its threats.

## Reference Map:

* Paragraph 1 – [[1]](https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en), [[2]](https://time.com/6589263/taylor-swift-deepfakes-legal-protections/), [[4]](https://apnews.com/article/c76c46b48e872cf79ded5430e098e65b)
* Paragraph 2 – [[2]](https://time.com/6589263/taylor-swift-deepfakes-legal-protections/), [[6]](https://apnews.com/article/bffbc274dd178ab054426ee7d691df7e)
* Paragraph 3 – [[1]](https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en), [[2]](https://time.com/6589263/taylor-swift-deepfakes-legal-protections/), [[3]](https://www.reuters.com/technology/cybersecurity/ai-godfather-others-urge-more-deepfake-regulation-open-letter-2024-02-21/)
* Paragraph 4 – [[1]](https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en), [[5]](https://www.ft.com/content/8d5bc867-c69d-44df-839f-d43c92785435)
* Paragraph 5 – [[3]](https://www.reuters.com/technology/cybersecurity/ai-godfather-others-urge-more-deepfake-regulation-open-letter-2024-02-21/), [[4]](https://apnews.com/article/c76c46b48e872cf79ded5430e098e65b), [[5]](https://www.ft.com/content/8d5bc867-c69d-44df-839f-d43c92785435)
* Paragraph 6 – [[1]](https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en), [[7]](https://time.com/6590711/deepfake-protection-federal-bill/)
* Paragraph 7 – [[1]](https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en), [[3]](https://www.reuters.com/technology/cybersecurity/ai-godfather-others-urge-more-deepfake-regulation-open-letter-2024-02-21/), [[6]](https://apnews.com/article/bffbc274dd178ab054426ee7d691df7e)

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://news.google.com/rss/articles/CBMifkFVX3lxTE43bVE1OC13UlB3bnRDWllyMG5Vazg2ZlhNLWdQVUtubFN6YVpISjcwNE9xU3otZjUzUW0yLURMZFdNaFh1a0MxWFE2VFU0MzNZX01aVTNERkowb3EzdXI1d0JpZVBFQkhPTDBKTUVrd2xoaTV3cHhUQVFmcnNtQQ?oc=5&hl=en-US&gl=US&ceid=US:en> - Please view link - unable to able to access data
2. <https://time.com/6589263/taylor-swift-deepfakes-legal-protections/> - The sharing of deepfake pornographic images of Taylor Swift on the social media platform X has underscored the need for enhanced legal protections against deepfakes. Despite efforts by X to remove these images, such content often spreads rapidly, making it difficult to eradicate completely. Carrie Goldberg, a lawyer specializing in tech abuse, highlights the increasing prevalence of deepfakes affecting both celebrities and non-celebrities, including children. While some states have laws against deepfakes, there is no federal legislation to address the issue. Bills like the Preventing Deepfakes of Intimate Images Act and the No AI Fraud Act have been proposed but have made little progress. Experts, including Adam Dodge of Endtab, emphasize that deepfakes are part of a broader problem of online violence, particularly against women. They call for stronger legal frameworks and accountability measures for companies that enable the creation and dissemination of deepfakes.
3. <https://www.reuters.com/technology/cybersecurity/ai-godfather-others-urge-more-deepfake-regulation-open-letter-2024-02-21/> - AI experts and industry leaders, including Yoshua Bengio, have signed an open letter calling for stricter regulations on deepfakes. Highlighting the dangers of deepfakes, such as sexual imagery, fraud, and political disinformation, the letter emphasizes the urgent need for safeguards due to rapid AI advancements. Recommendations include criminalizing deepfake child pornography, penalizing the creation or spread of harmful deepfakes, and obligating AI firms to prevent their products from generating harmful deepfakes. The initiative, led by UC Berkeley researcher Andrew Critch, has gained over 400 signatures from various sectors, including academia, entertainment, and politics. Prominent figures like Steven Pinker, Joy Buolamwini, and researchers from Google DeepMind and OpenAI have endorsed the letter. The push for regulation reflects ongoing concerns about AI's societal impact, a topic also addressed in previous warnings by figures like Elon Musk.
4. <https://apnews.com/article/c76c46b48e872cf79ded5430e098e65b> - The Biden administration is urging the tech industry and financial institutions to tackle the proliferation of AI-generated sexually explicit images, known as deepfakes, that are being used abusively. With generative AI tools making it easy to create realistic fake images, the White House is seeking voluntary cooperation from companies to prevent the creation, dissemination, and monetization of such content. The effort is particularly focused on protecting women, girls, and minors from nonconsensual imagery. The administration calls for action from AI developers, cloud service providers, search engines, and payment processors to restrict access and financial opportunities for sites promoting explicit images of minors. Despite these efforts, there is still a need for legislative action to provide comprehensive oversight and enforcement. The issue has gained prominence with cases involving both celebrities and minors, underscoring the urgent need for industry-wide safeguards.
5. <https://www.ft.com/content/8d5bc867-c69d-44df-839f-d43c92785435> - Un estudio de DeepMind, la división de inteligencia artificial (IA) de Google, ha identificado que los 'deepfakes' políticos encabezan la lista de usos maliciosos de la IA. Estas falsificaciones realistas de imágenes, vídeos y audios de políticos y celebridades son mucho más prevalentes que los ciberataques asistidos por IA. El estudio revela que el objetivo principal de quienes utilizan la IA generativa para el mal es influir en la opinión pública, representando el 27% de los casos. Esto ha generado preocupación por su potencial para influir en elecciones a nivel mundial, incluyendo las recientes apariciones de deepfakes del Primer Ministro del Reino Unido, Rishi Sunak, en plataformas como TikTok, X e Instagram. El estudio, realizado junto con la unidad de I+D Jigsaw, analizó aproximadamente 200 incidentes de uso indebido de enero de 2023 a marzo de 2024, y encontró que la mayoría de las herramientas utilizadas son fácilmente accesibles y requieren poca habilidad técnica. DeepMind espera que su investigación ayude a mejorar las evaluaciones de seguridad de los modelos de IA y también influencie a otros actores del sector para abordar los riesgos manifestados.
6. <https://apnews.com/article/bffbc274dd178ab054426ee7d691df7e> - In response to the widespread issue of nonconsensual deepfake images, lawmakers across the United States have been working on measures to combat the creation and distribution of these fake images, particularly following the recent victimization of Taylor Swift. Deepfakes exploit celebrities and non-famous individuals, often creating pornographic content without their consent. Several states have enacted laws targeting nonconsensual deepfake porn, and others are considering similar legislation. Techniques to detect and prevent deepfakes are being developed, including algorithms and digital watermarks. Some proposed federal legislation aims to grant individuals rights over their likenesses used in deepfakes, allowing them to sue offenders. However, there are concerns about balancing regulation with free speech protections and innovation. Law experts and advocacy groups emphasize the need for comprehensive approaches involving technology companies, social media platforms, and government enforcement to address the deepfake problem effectively.
7. <https://time.com/6590711/deepfake-protection-federal-bill/> - Un proyecto de ley introducido por legisladores bipartidistas podría permitir que las víctimas demanden a las personas que crean y distribuyen deepfakes sexualmente explícitos sin consentimiento. Nombrado DEFIANCE Act (Disrupt Explicit Forged Images and Non-Consensual Edits), este proyecto permitiría a las víctimas emprender acciones legales si los creadores de deepfakes sabían o ignoraron imprudentemente que la víctima no consintió. Introducida tras un incidente de imágenes pornográficas deepfake de Taylor Swift en redes sociales, la ley se convertiría en la primera federal para proteger a las víctimas de deepfakes. Actualmente, solo diez estados tienen leyes criminales contra este tipo de medios manipulados. El apoyo público a la legislación contra deepfakes es considerable, con un 84% de estadounidenses a favor de hacer ilegal la pornografía deepfake no consensuada. Expertos destacan la necesidad de que las plataformas sociales también inviertan en evitar la propagación de deepfakes.