Christopher Buenrostro
CST 300 Writing Assessment for Computing & Design
October 17, 2025

Deepfakes and Artificial Intelligence Regulation

Introduction/Background:

In modern society, technology has largely grown and artificial intelligence has become a tool in everyday life. AI can be utilized for things such as problem solving, creativity, and learning, often available right at our fingertips through phones and computers. While this innovative progress is beneficial, it also raises new challenges, particularly in what is called Deepfake technology. Deepfakes are AI-generated content in the form of videos, images, or audio with typically malicious intentions to mimic real people. Artificial intelligence is increasingly being viewed as a concern by governments, the international community, and the public because of its possible impact on both cybersecurity and national security (Buffett Brief, 2023, p.1). This concern brings a two-sided issue where some call for AI regulation at the federal level, and others say it should be treated as free speech, meaning limited to no regulation.

At its roots, deepfake issues derive from how they are made and what is needed to generate content. A user can make a few simple Google searches and find AI generative tools that require little to no power to run. Most deepfake models are created by training LoRA adapters, using as little as 20 images, 25 GB of VRAM, and 15 minutes of time, meaning a low barrier to accessibility (Hawkins, Russell, & Mittelstadt, 2025, p.2). Tools being easily accessible means those with malicious intent no longer need expensive systems or technical skills to take advantage of these technologies. This accessibility sets the stage for misuse, where deepfakes can quickly move from experimental into powerful tools that create issues of erosion of trust, threats to national security, fraud, political misuse, and free speech debates.

The emergence of deepfake technology became a public issue in 2017, when Reddit users were posting videos generated by AI superimposing celebrity faces with startling realism (Asadi, 2025, p.2). Post-2017, deepfakes grew popular in different contexts. In 2019, the Department of Homeland Security reported a deepfake video increase of 84% from the beginning to end of that year (Asadi, 2025, p.2). This growth was prevalent even affecting other countries. In 2022, Russia was a victim where a deepfake video of Ukrainian President Volodymyr Zelenskyy calling for surrender during the war circulated, resulting in social divisions and discredit of leaders (Buffett Brief, 2023). From 2023 to 2025, deepfakes scaled to a global level, increasing by 3,000%, which is the result of the continuation of AI advancements (Asadi, 2025, p.2). In response, deepfakes have moved from an online trend to a regulated risk, influencing Denmark, the United States, China, France, and the United Kingdom to lawfully criminalize harmful deepfakes and mandate takedowns (Patishman, 2025). These technical, historical, and legal developments create the debate: Should governments regulate AI-generated deepfakes, or should free speech and innovation concerns take precedence?

Stakeholder Analysis:

Stakeholder 1 pro-regulation includes lawmakers, regulators, national security and law enforcement officials, newsrooms and fact-checkers, public figures, and victims-rights organizations. This group takes the position of viewing deepfakes as a safety threat globally, affecting governments, economies, and the trust of society advocating lawful intervention as necessary. A majority of individuals who are pro-regulation feel worrisome for digital consumers who are exposed to indefinite amounts of content on a daily basis (Hsu, Thompson, & Myers, 2025, para.5).

Stakeholder 2 anti-regulation or limited-regulation includes technology companies, civil rights organizations, and open-source developers. This group takes the position that placing regulation on AI content breaks First Amendment rights, constrains American competitiveness, and involves policies that are too complex to implement (Hsu, 2025, para.10).

In the conversation of AI deepfake regulation, we have two opposing sides that are directly affected. Stakeholder 1 represents those pro-regulation of deepfakes, while Stakeholder 2 represents those against or favoring limited regulation.

Stakeholder 1: Values

The values guiding Stakeholder 1 revolve around privacy, honesty, dignity, and public trust. These principles represent the foundation of an informed society that holds a level of expectations when it comes to information. They believe that people have the right to digital privacy and protection from exploitation, especially when technology can falsify convincing content without consent. Truth and validity in information are crucial to keeping trust between the public and media sources. What is important to this group is preventing digital harm, valuing a secure environment where individuals can rely on some source of verified information online. To them, maintaining this is essential to bettering society, politics, markets, and many protected individuals.

Stakeholder 1: Position

Stakeholder 1’s position on the issue is that deepfakes must be regulated because the growing misuse of AI-generated content directly threatens these core values. The ease of creating and spreading deepfakes has resulted in widespread misinformation, non-consensual exploitation, and identity theft, all of which cause real harm (Hawkins, Russell, & Mittelstadt, 2025). For this group, unregulated AI isn’t simply a technology problem, it’s a social and ethical crisis. This holds importance because they demonstrate that people's privacy and trust are at risk. Without regulation, those harms will continue to grow. The increasing presence of deepfakes in political campaigns, news media, and online scams proves that societal self-regulation is not enough (Asadi, 2025). This group views government intervention and legal accountability as necessary actions to protect citizens, restore media credibility, and uphold moral responsibility in technology use.

Stakeholders 1: Claims

The claims used by Stakeholder 1 are both factual and ethical with goals of regulating AI effectively. Their claims of fact and value link the rise of misuse of technologies directly to unregulated access. Stakeholder 1 references examples such as Denmark’s copyright protection of one’s likeness and the EU’s AI watermarking laws that require labeling of generated media (Patishman, 2025). In the United States, the Take It Down Act and similar legislation address these harms by enforcing platform responsibility and criminalizing exploitation (Hsu, 2025). The Buffett Brief (2023) also emphasizes that the spread of AI misinformation affects democracy and international security, reinforcing the need for global standards. Altogether, these claims represent an ethical argument for legal implementations that protect trust and prevent societal harm.

Stakeholder 2: Values

Stakeholder 2, the group opposing regulation, values freedom of expression, creativity, and innovation. These individuals consist of civil rights advocates, open-source developers, and tech entrepreneurs who believe regulation threatens the right to free speech and artistic creativity. They value innovation as a main contributor of technological progress and see types of deepfake creation as part of digital expression. This group believes that limiting AI-generated content risks restricting artistic work, satire, and innovation in creative industries.

Stakeholder 2: Positition:

Their position on the issue is that AI regulation contradicts free expression and slows innovation. Stakeholder 2 views deepfakes not only as tools for manipulation but also as opportunities for creativity and accessibility. For example, AI can recreate voices for people who have lost the ability to speak, de-age actors for movies, or simulate historical figures for education (Asadi, 2025). For them, these examples show that AI's benefits outweigh its risks. They also argue that enforcing deepfake laws across global digital spaces is impractical and may lead to overreach. The rapid advancement of AI video generators like OpenAI's *Sora*, which produces hyper realistic videos indistinguishable from real footage, demonstrates that regulation struggles to keep pace with innovation (Hsu, Thompson, & Myers, Chen, 2025).

Stakeholder 2: Claims

Stakeholder 2 uses several types of claims to support their position. Claims of definition describe deepfakes as a form of creative or expressive content, not inherently harmful. Claims of cause argue that strict regulation would harm industries relying on generative AI for design, film, education, and accessibility. Value claims emphasize freedom and innovation over restriction. This group argues that education, labeling tools, and self-governance by platforms are more effective solutions than government control. They see personal responsibility, ethical AI development, and user awareness as the best way to balance innovation with morality. These views are supported by ongoing debates in AI ethics research that highlight open-source development as crucial to innovation and education (Hawkins, Russell, & Mittelstadt, 2025).

Argument Question

Should governments regulate AI-generated deepfakes, or should free speech and innovation take precedence?

Stakeholder 1: Pro-Regulation

Framework

From an ethical perspective, Stakeholder 1’s argument aligns with the Utilitarianism framework founded by Jeremy Bentham and John Stuart Mill. Utilitarianism emphasizes actions that maximize overall happiness and minimize harm across society. The framework judges morality by outcomes, not intentions, focusing on how decisions affect the well-being of the majority. Its main goal is to reduce suffering and promote fairness by ensuring the greatest good is achieved for the greatest number of people.

Tenets

Stakeholder 1 believes regulation of AI-generated deepfakes is ethically justified because it limits harm and improves long-term stability for the public. Deepfakes contribute to misinformation, exploitation, and financial crime, all of which damage social trust (Hawkins, Russell, & Mittelstadt, 2025). Governments applying utilitarian ethics would act to protect citizens through accountability and transparency laws. Requiring watermarking, labeling, and user verification helps prevent abuse, increases safety, and ensures AI tools are used ethically. These steps maximize collective happiness by protecting more people from harm than the few who might lose unrestricted creative freedom.

Action

According to this perspective, governments should implement clear AI-focused laws that punish malicious creation or distribution of deepfakes while still allowing innovation under guidelines. Examples include Denmark’s likeness-protection policy and the European Union’s watermarking requirements (Patishman, 2025). In the United States, the Take It Down Act and related legislation enforce accountability for harmful AI content (Hsu, 2025). Such actions reinforce public confidence in technology and reduce large-scale damage caused by deception and identity theft. By emphasizing harm reduction, regulation fulfills utilitarian ethics by benefiting the majority.

Stakes

What is at stake for Stakeholder 1 is the safety and trust of global information systems. If regulation is enacted, society gains stability, privacy protection, and renewed confidence in media postings. Citizens and policymakers benefit from reduced fraud and misinformation. However, some innovators may experience limited creative freedom or periods of slow development. Despite these small drawbacks, the overall outcome is greater creating social flow, and long-term ethical progres..

Stakeholder 2: Anti-Regulation

Framework

Stakeholder 2's argument aligns with Kantian Ethics, developed by philosopher Immanuel Kant. This framework focuses on duty, moral law, and rational autonomy rather than consequences. Kant's categorical imperative states that one should act only in ways that could become universal laws. The framework values individual freedom, honesty, and moral intention, arguing that ethical behavior arises from respecting others' autonomy and treating them as ends in themselves rather than as means to an end.

Tenets

Stakeholder 2 believes that regulation of AI-generated content violates personal autonomy and limits freedom of expression. From a Kantian perspective, individuals are moral agents capable of making rational choices therefore, they should not be restricted by outside forces like government control. This group argues that people should be trusted to use AI responsibly and that moral development depends on self-governance rather than imposed regulation. Freedom of creativity and innovation are considered universal rights that encourage ethical growth through rational decision-making.

Action

According to this ethical view, the correct course of action is to rely on education, transparency, and ethical standards set by developers instead of government mandates. By allowing individuals and companies to create AI under moral guidelines, innovation can thrive without fear of holdbacks and punishment. Stakeholder 2 maintains that open access to AI tools, similar to open-source platforms, encourages creativity and problem-solving that benefit society (Asadi, 2025). Regulation, on the other hand, is seen as the opposite that could slow progress and discourage legitimate experimentation.

Stakes

For Stakeholder 2, what is at stake is the future of innovation, competition, and creative freedom. If governments impose strict regulations, developers and entrepreneurs could face excessive liability and reduced opportunity to explore new technologies. This could harm economic growth and slow advancements that might otherwise benefit society. However, without oversight, the risk of misuse increases, possibly leading to reputational damage and loss of public trust. Stakeholder 2 must balance these risks while maintaining their commitment to freedom and moral autonomy.

Student Position

Opinion

My position aligns most closely with Stakeholder 1, the pro-regulation side. Deepfakes have shown serious potential to harm individuals and entire societies when left unregulated. Privacy, truth, and safety outweigh unrestricted creative freedom. Using a utilitarian approach, the ethical decision is to protect the greater good through responsible AI regulation that minimizes harm for current and future generations.

Alignment

Countries like Denmark and members of the European Union are already demonstrating balanced approaches through watermarking systems and likeness-protection laws. The United States should adopt similar strategies by enforcing national legislation that holds both creators and platforms accountable (Patishman, 2025; Hsu, 2025). Regulation doesn’t end innovation, it ensures that innovation respects human rights and reduces misuse. With video technologies like Sora and other AI video generators advancing rapidly (Chen, 2025), this balance between creativity and protection is more crucial than ever.

Recommendation

The ethical solution is to guide technology responsibly, making sure a fair balance between creativity and protection so society can continue to enjoy the benefits of AI without sacrificing truth or trust. Responsible AI policy promotes innovation while protecting individuals from harm, aligning technological innovation with ethical duty. Future research and transparent AI design (Hawkins, Russell, & Mittelstadt, 2025) will be key to maintaining this balance and ensuring technology serves while being ethical. Alongside this, public campaigns to keep society informed will be crucial during this period of ethical and legal discovery for AI deepfake technologies.

References

Asadi, O. (2025). Exploring Current and Potential Solutions: The rise of deepfakes in legislative, legal, and technological arenas. *Berkeley Undergraduate Journal,* 39(1). https://doi.org/10.5070/b3.50655

Buffett Brief. (2023). The Rise of Artificial Intelligence and Deepfakes. https://buffett.northwestern.edu/documents/buffett-brief_the-rise-of-ai-and-deepfake-technology.pdf

Chen, B. X. (2025, October 9). what the arrival of A.I. video generators like sora means for us. *The New York Times.* https://www.nytimes.com/2025/10/09/technology/personaltech/sora-ai-video-impact.html?smid=url-share

Hawkins, W., Russell, C., & Mittelstadt, B. (2025). Deepfakes on Demand: the rise of accessible non-consensual deepfake image generators. ArXiv.org. https://arxiv.org/abs/2505.03859

Hsu, T. (2025, May 22). Deepfake laws bring prosecution and penalties, but also pushback. *The New York Times*. Retrieved from https://www.nytimes.com/2025/05/22/business/media/deepfakes-laws-free-speech.html

Hsu, T., Thompson, S. A., & Myers, S. L. (2025, October 3). OpenAI’s Sora makes disinformation extremely easy and extremely real. *The New York Times.* https://www.nytimes.com/2025/10/03/technology/sora-openai-video-disinformation.html?smid=url-share

Patishman, H. (2025, August 12). global legal actions against AI deepfakes: five laws of 2025 [*Review of Global Legal Actions Against AI Deepfakes: Five Laws of 2025*]. https://regulaforensics.com/blog/deepfake-regulations/