

# Ethical and Societal Implications of AI – A Case Study on Deepfakes and AI-generated Media

Alicia Gonzalez Cruz

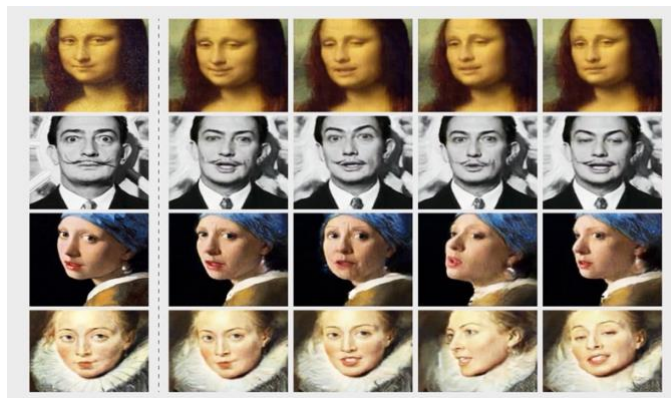
Prof. Anna Devarakonda

ITAI 2372

AI is becoming a big part of our everyday lives as the world moves forward, whether we notice it or not. It is used for online security, helps decide what we see online, plays a significant role in decision-making, and can make things faster and more efficient. However, it also raises some serious ethical and social concerns. One major issue is bias. AI systems learn from data, and if that data has a bias, the AI can make unfair decisions, like favoring one group of people over another. This is where things turn into a negative light- it affects how people are hired or a simple as who gets a loan. From a societal point of view, AI is changing how we live both negatively and positively. However, the main concern is privacy, as the world implements technologies like facial recognition, fingerprint security, and data tracking, it can create social inequalities if it's not used correctly. AI has the potential to help create a more advanced society, but it also comes with risks. What's important is that it's being developed to be used in a fair, transparent, and respectful way.

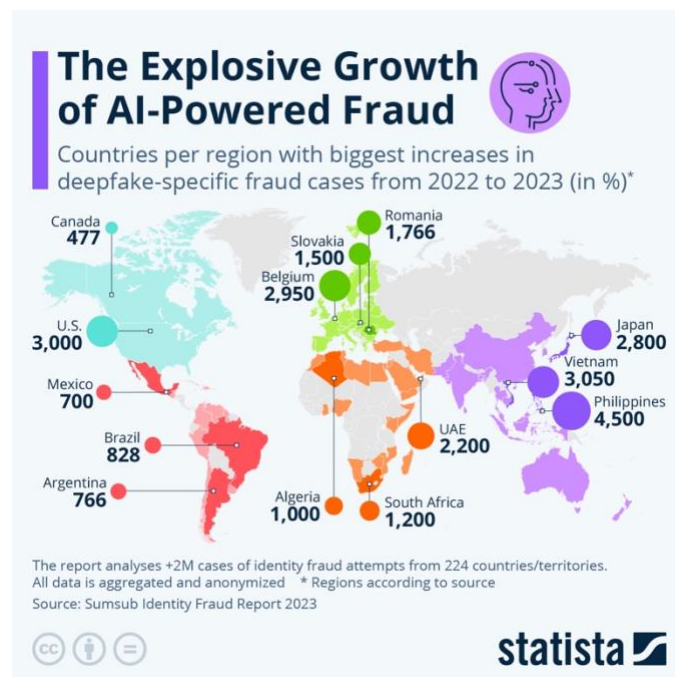
My case study will be based on the ethical and societal implications of deepfakes and the use of AI-generated media. To start things off, what creates deepfakes is Generative AI, which is a type of tech that can make stuff like images, videos, music, or even fake voices. It learns by looking at tons of examples—like photos or audio clips—and then it tries to create new things that look or sound real. So, it can make a fake picture of a person who doesn't even exist or a video that seems real but isn't. The problem here is that it can be used in both positive and negative aspects. While it can bring things like historic images to life, it could be used to mess up someone's image with a fake video or image. The main problem now is that generative AI is made so well that people can't tell the difference between what's real and what's fake anymore.

Deepfakes are becoming more common across several areas of daily life. They are being used as a playful and creative outlet, but are also being used dangerously. The main places deepfakes are showing up are in social media, politics, entertainment, and scams. A positive example would be how platforms such as TikTok and Instagram are filled with deepfake content, which most of the time is for entertainment and fan-made content. While in a negative light, deepfakes are used in scams and fraud. Deepfakes are often used by cloning someone's voice or face to trick people, and as technology gets more realistic, it is harder to verify.



Blurring the line between reality and fiction, AI-generated media like deepfakes pose serious ethical challenges that directly impact people's identities, reputations, and trust in what they see online. One of the biggest ethical problems with deepfakes is that they are often created without a person's knowledge or permission. For example, celebrities are the ones frequently used for deepfakes and the harmful ways they are used include revenge porn, impersonation, and defamation. This leads to Deepfakes making it easy to distort the truth and manipulate public perception. A deepfake with the intent to deceive, intimidate, misattribute, and inflict reputational harm to perpetuate disinformation is unambiguously unethical. This further leads to huge ethical issues of figuring out who is responsible when deepfakes cause harm. Most of the time, the tools used to create deepfakes are freely available for anyone to use, and while platforms are creating a system to target deepfakes, there is still a long way to go.

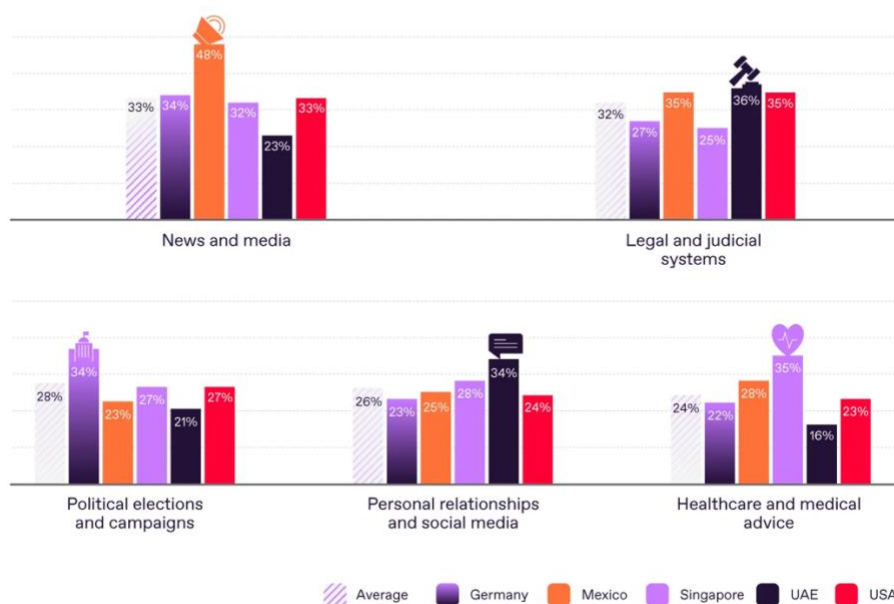
The rise of deepfakes isn't just a tech issue—it affects how people experience truth, safety, and fairness in everyday life. As this technology spreads, it brings a wave of societal challenges that are hard to ignore. For example, when there is a realistic fake video roaming around the internet, it becomes harder to trust what we see online, which creates skepticism about what is real. Deepfakes can also have negative emotional impacts, and most of the time, vulnerable or marginalized groups face greater risk. And while people are trying to figure out a way to create a rule against the creation or spread of deepfakes, there is still a legal gap that makes it easier for bad people to use without consequences.



To better understand the societal risks, looking at real examples where deepfakes have been used to mislead, manipulate, or cause harm is helpful. A real-life example would be the creation of deepfake porn targeting women. For instance, a publication done by Homeland Security stated that “In October 2020, researchers reported over 100,000 computer-generated fake nude images

of women created without their consent or knowledge...approximately 90-95% of deepfake videos since 2018 were primarily based on non-consensual pornography.” The impact of the creation of deepfake porn is that it damages the reputations of women all over the world, and it’s essential for people to understand that this type of deepfake targets women globally and it leads to depression, PTSD, harassment, and defamation. These deepfakes are created without consent, weaponizing someone’s identity and image to harass, silence, or punish them. And in many regions, there are still no clear laws to stop or punish the creators. From an ethical standpoint, this raises serious concerns about digital consent, autonomy, and respect for human dignity. Using AI to fabricate explicit content without someone’s permission is not just harmful—it’s a violation of their basic rights.

### Key areas in society most impacted by deepfakes



While much of the public concern regarding deepfakes focuses on the negative side, it’s also important to recognize that Generative AI also has creative applications. The use of deepfakes has brought some light to accessibility by helping people with speech impairments communicate naturally. Another positive example would be how historical figures are “brought back to life” using deepfake-like tech. There’s an educational project called “Holocaust Education Project” which uses deepfake videos to create interactive interviews with holocaust survivors, which preserves the stories for better understanding and future generations. These benefits highlight that the problem isn’t the technology itself, but how it’s used. Therefore, ethical design and clear boundaries are key to ensuring that deepfakes serve society, rather than harm it.

While we can't stop deepfake technology from evolving, we can design innovative ways to use it more ethically and safely. The innovative proposal that I believe would tackle the ethical and societal harms caused by deepfakes would be to create a system in which AI authenticity tagging is implemented, as well as an awareness campaign. This proposal would be very beneficial because there would be mandatory AI content labeling, which would help people point out what is AI-generated content. It promoted transparency, which combats misinformation and protects people from non-consensual deepfakes. An awareness campaign will also further teach people to recognize deepfakes and understand the technology being used.

All in all, deepfakes aren't just some cool or creepy tech—they're already part of our everyday lives, and they come with some serious risks. From fake videos to non-consensual content, they've shown how powerful and harmful this tech can be when misused. But at the same time, it's not all bad. Deepfakes can also be creative and helpful when used the right way. That's why it's not just about stopping the technology—it's about making sure it's used in ways that are ethical and safe. Proposals like adding labels to AI content and teaching people how to spot deepfakes won't solve everything, but it's a solid start. If we want to keep up with how fast AI is moving, we've got to be just as quick when it comes to setting boundaries and protecting people.

## Citations

- Barron-Lopez, L. (2023a, June 27). *Women face new sexual harassment with deepfake pornography*. PBS NewsHour. <https://www.pbs.org/newshour/show/women-face-new-sexual-harassment-with-deepfake-pornography>
- Caroline Quirk. (2023, June 20). *The High Stakes of Deepfakes: The Growing Necessity of Federal Legislation to Regulate This Rapidly Evolving Technology*. Princeton Legal Journal. <https://legaljournal.princeton.edu/the-high-stakes-of-deepfakes-the-growing-necessity-of-federal-legislation-to-regulate-this-rapidly-evolving-technology/>
- Cohen, A. (2024, May 4). *The Evolution of Deepfake Technologies: Exploring Tools, Applications, and Implications*. Medium; Medium. <https://iamitcohen.medium.com/the-evolution-of-deepfake-technologies-exploring-tools-applications-and-implications-86fc50ae22b0>
- Department of Homeland Security. (2023a). Increasing Threat of Deepfake Identities. In *Department of Homeland Security*. Department of Homeland Security. [https://www.dhs.gov/sites/default/files/publications/increasing\\_threats\\_of\\_deepfake\\_identities\\_0.pdf](https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf)
- Jaiman, A. (2020a, August 27). *Debating the ethics of deepfakes*. ORF. <https://www.orfonline.org/expert-speak/debating-the-ethics-of-deepfakes>
- Sharma, P. B., & Ramteke, Prof. Dr. P. L. (2023a). Recommendation for Selecting Smart Village in India through Opinion Mining Using Big Data Analytics. *INTERANTIONAL JOURNAL of SCIENTIFIC RESEARCH in ENGINEERING and MANAGEMENT*, 07(01). <https://doi.org/10.55041/ijsrem17505>
- The Princeton Review. (2024a). *Ethical and Social Implications of AI Use | The Princeton Review*. [Www.princetonreview.com. https://www.princetonreview.com/ai-education/ethical-and-social-implications-of-ai-use](https://www.princetonreview.com/ai-education/ethical-and-social-implications-of-ai-use)

University, D. (2024, April 9). *Deepfake porn: why we need to make it a crime to create it, not just share it - Durham University*. [Www.durham.ac.uk](http://www.durham.ac.uk).

<https://www.durham.ac.uk/research/current/thought-leadership/2024/04/deepfake-porn-why-we-need-to-make-it-a-crime-to-create-it-not-just-share-it/>

West, D., & Allen, J. (2018, April 24). *How Artificial Intelligence Is Transforming the World*.

Brookings; The Brookings Institution. <https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/>