Newhouse Impact Journal

Volume 1 | Issue 1 Article 4

2024

The Good, the Bad, the Ugly...and the Gray

Gregory Archer
PAR Government Systems

Recommended Citation
Archer, Gregory (2024) "The Good, the Bad, the Ugly...and the Gray," *Newhouse Impact Journal*: Vol. 1: Iss. 1, Article 4.
DOI: http://doi.org/10.14305/jn.29960819.2024.1.1.09
Available at: https://surface.syr.edu/newhouseimpactjournal/vol1/iss1/4

The Good, the Bad, the Ugly... and the Gray

Uses of Generative AI in the Media & Entertainment Industry, Disinformation, and Offensive and Obscene Content

Gregory Archer
Video Producer/Editor & Disinformation Team Lead
PAR Government Systems, USA

Generative AI technology and tools have opened a world of possibilities for users, inside and out of the media industry. These possibilities extend to numerous areas of media and communications, including audio, images, text, and video. However, as with many new technologies, they can be used for both legitimate, benign purposes, or for malicious, harmful, and in some cases illegal purposes. These applications can be broken down into, the good, the bad, and the ugly uses and many of the same generative AI technology spans across all three of these areas. There is also a gray area which consists of all the unanswered questions related to the ethics of generative AI and how, if at all, this technology could or should be governed. There are many examples of each, good, bad, and ugly in the current media landscape, here are a few.

The Good

Generative AI applications have opened a world of possibilities within media and the arts. This can be attributed to how widely and cheaply available they've become, as well as requiring little to no technical expertise to use, allowing users to create, fix, or alter media. On the other hand, for users that do have media production or computer science experience, these tools and technology can unlock a new suite of enhanced capabilities.

Tools like ChatGPT and Google Bard have become increasingly commonplace now and are being incorporated in system level applications like Microsoft Office and internet search engines.[1] These AI text generators can offer several basic benefits by simply helping improve content, structure, and grammar, or improving the efficiency of writers. They can also be leveraged to help facilitate the creative process in brainstorming sessions or work through creative blocks, while still being cost effective and scalable. One example of this is Google's development of an AI tool for news publications that can generate articles and headlines for both news and entertainment stories. Google has stated that they are seeking to partner with news companies to use this tool in their newsrooms.[2]

Images are another type of media where generative AI technology has become very prevalent in the world, initially on social media platforms such as Instagram, and are now starting to show up in the media industry. These applications span every aspect of the art of image production, from traditional concept art, such as drawings or painting, up to the equivalent of photorealistic photography.[3] Some uses in the media industry could be creating ancillary images needed for production, creating still images used in b-roll (secondary/cutaway footage), and creating scenes

for Matte Painting or similar processes. Matte Painting for example takes a considerable amount of skill and time. Now, images to be used for this process can easily be created using AI to be used a backdrop and then combined with a video, eliminating the need for expensive set creation or a matte artist.[4]

Deepfakes, or the ability to swap someone's face and voice for another's in video can be used to replace the skill and cost associated with common post-production VFX and CGI tasks. For example, aging and de-aging actors, using deceased or otherwise unavailable actors' likeness, special effects such as facial wounds, scars, or tattoos, and face swapping applications such as replacing a stunt double's face with an actor's face.

Voice cloning and replacement is another generative AI technology that can replace time consuming and costly post-production. This technology is used in the video and film industry to do things similar to deepfakes such as aging and de-aging actors voices, using deceased or unavailable actors' likeness, special effects such as changing emotion, and replacing an actor's voice with another voice. Advances in the space have continued to evolve to include both replacing words or dialog and language dubbing applications. A great example of this can be seen in the movie *Fall*.[5] In order to get better distribution of the film, the director turned to AI and used the company Flawless AI to replace 30 swear words and other lines of dialog, allowing the rating to be changed from R to PG-13.[6]

The Bad

As with the good, the ease of use and availability of generative AI methods plays a large role in the bad they can be used for. AI image, text, audio, and video generators can be used to create media of almost anything or anyone. This makes the technology a powerful tool capable of accelerating the production of various forms of disinformation across a range of media types.

AI text generation tools such as ChatGPT and Google Bard are being combined with software that can auto generate websites and content, effectively automating the production of spam websites, with some even posing as legitimate news sites.[7] These "news" websites are not regulated or fact-checked and are created, more often than not, to host revenue generating advertising slots. Through tuning, AI text can be trained to produce outputs as desired. New versions of tools created specifically for malicious purposes, such as WormGPT, are being used for hacking and cybercrime and have capabilities for writing code including mal-ware and creating phishing emails and business email compromise (BEC) attacks.

Images are a key part of disinformation, especially in recontextualization. However, in the past to create a convincing manipulated or fake image, it required specific skills and knowledge of manipulation software. AI image generation has solved that and made generating images, from just simple text prompts, much easier. An excellent example of this were images that surfaced in March 2023 of former U.S. President Donald Trump running from and being arrested by police.[8] Another example includes the May 2023 photo which was posted online claiming to show an explosion at The Pentagon. Though these examples were both quickly disproved, they were shared widely on social media and even in investment circles.[9]

Deepfake, though being used for entertainment and things like satire, are continuing to be used in disinformation. Deepfakes can make it look like people are saying or doing things they never did and put them in places they've never been to in real life. It should come as no surprise

that these deepfakes appear in disinformation campaigns and most times, involve politicians. For example, deepfakes of both Volodymyr Zelenskyy[10] and Vladimir Putin[11] have surfaced throughout the ongoing war between Ukraine and Russia, and many experts believe this is just the tip of the iceberg when it comes to deepfake use in disinformation.[12] U.S. politicians often find themselves as targets of deepfakes as well.

The Ugly

These are the topics that most people avoid when talking about generative AI and how it can be used, mainly because the topics are uncomfortable. It would be easy to skip the ugly section and just talk about the good and bad, however I am a firm believer, that at the very least, users or consumers of generative AI media need to be aware of what this technology can, or more accurately, IS, being used for. As discussed, generative AI can be used to create media of almost anything. Images can be made of things like dead bodies, scenes of violence, or the aftermath of violence, scenes depicting racial stereotypes or offensive content, images depicting questionable and/or nonconsensual sexual content, and images of illegal content related to minors and children.

Non-consensual adult content and revenge porn has always been a prominent use of deepfakes. A report by the company Sensity who monitors deepfake use stated that in 2019, 96 percent of deepfakes were used in this way.[13] They first started mainly with celebrities' faces being swapped into pornography,[14] but as the technology became more widely available and easier to use, it progressed to everyday people becoming targets. Deepfake porn primary targets women, the same Sensity report citing that 99 percent of deepfake porn featured women.[15]

Generative AI is also being used to create images and audio for the purposes of graphically retelling the stories of murder victims. Most of these stories currently appearing on TikTok appear to be about children and told from a first-person point of view. AI generated voices, described as "childish," are used along with AI generated images, to create a video, recounting explicit details of the crimes.[16] In this case, the telling of the story itself is not necessarily what pushes this media into the ugly, it is more how it is done and the fact that it is not done in a factual way. In one example, a childish, baby's voice is heard stating "Grandma locked me in an oven at 230 degrees when I was just 21 months old."[17] In another case the same audio about the crime was used for two different images of children of different races and with different names, which makes the claim of "True-Crime" false.[18]

Finally, and the ugliest use to be discussed, is the use of generative AI to create images of child pornography and sexual abuse. Recently the proliferation of content containing children has safety groups, law enforcement and governments concerned. Artificial intelligence could be used to generate "unprecedented quantities" of realistic child sexual abuse material, and The Internet Watch Foundation (IWF) has said it was already finding "astoundingly realistic" AI-made images that many people would find "indistinguishable" from real ones.[19] In just five weeks, from May 24, 2023, through June 30, 2023, the IWF investigated 29 reports of URLs containing suspected AI-generated child sexual abuse imagery.[20] Susie Hargreaves OBE, Chief Executive of the IWF stated that "Our worry is that, if AI imagery of child sexual abuse becomes indistinguishable from real imagery, there is a danger that IWF analysts could waste precious time attempting to identify and help law enforcement protect children that do not exist."[21] Not only does this appear to be an issue of the content being generated, but possibly more

importantly, it has impacts on actual victims as resources are being used to validate if images are fake or real, effectively taking those resources away from helping actual victims.

The Gray

As with any technology, when generative AI becomes more accessible, used, socially accepted, and adopted, questions start to arise regarding how it's being used and how it will be used in the future. We have seen these conversations start within the last few years and as recently as within the last few weeks. These conversations not only revolve around the ability and use of generative AI but also around the standards and ethics of use. To initiate and sustain these discussions, we will need to pose both fundamental and potentially more complex questions such as:

- Who decides how AI is used not only in the creation of media but in the consumption of it? Should it be governments or law makers? Do media or technology companies get to be the ones to decide? Should special interest groups or groups like Unions decide how it is used, at least in their industries?
- Who has the right to use Generative AI? Is the use of generative AI capabilities a Constitutional question of freedom of expression? When is it ok to use some one's likeness in generative AI? What about using someone's likeness in satire or commentary? Should AI generated media be considered art? What if people don't want to use it? Can or should they be forced to use it?
- Should governments be responsible for monitoring AI technology, or should AI companies be responsible for self-monitoring? Should governments restrict the development, release, or use of AI? Should there be only designated AI companies that can do things like create or release AI technology?

The regulation of generative AI presents a multifaceted challenge that cannot be understated. As this technology becomes more accessible and integrated into our daily lives, it raises complex questions about ethics, accountability, and the potential consequences of its misuse. Striking the right balance between innovation and responsible governance is essential. The path to effectively regulating generative AI is a dynamic and evolving one, requiring continual adaptation and collaboration across various stakeholders.

References

1 Kelly, Samantha Murphy. 2023. Microsoft is bringing ChatGPT technology to Word, Excel and Outlook. *CNN*. March 16. Accessed July 22, 2023. https://www.cnn.com/2023/03/16/tech/openai-gpt-microsoft-365/index.html.

2 Fung, Brian. 2023. Google is building an AI tool for journalists. *CNN*. July 20. Accessed july 22, 2023. https://www.cnn.com/2023/07/20/tech/google-ai-tool journalism/index.html.

3 Sajid, Haziqa. 2023. AI in Media: How Is Generative AI Revamping the Media & Entertainment Industry. *United.AI*. June 8. Accessed October 11, 2023. https://www.unite.ai/ai-in-media-how-is-generative-ai-revamping-the-media-entertainment-industry/.

4 McIntyre, Gary. 2023. Matte Painting for a New Generation of Filmmakers. Fstoppers. July 18. Accessed July 24, 2023. https://fstoppers.com/artificial-intelligence/matte-painting-new-generation-filmmakers-636454

5 Fall. 2022. Directed by Scott Mann. Performed by Lionsgate Films, Signature Entertainment.

6 FlawlessAI. 2023. AI-powered tools for filmmakers, studios and distributors. *Flawless AI*. Accessed July 25, 2023. https://www.flawlessai.com/.

7 Crider, Michael. 2023. AI is filling up the internet with garbage spam sites. *PCWorld*. June 28. Accessed July 21, 2023. https://www.pcworld.com/article/1971683/ai-is-filling-up-the-internet-with-garbage-spam-sites.html.

8 Vynck, Gerrit De. 2023. *AI images are getting harder to spot. Google thinks it has a solution.* August 29. Accessed October 11, 2023. https://www.washingtonpost.com/technology/2023/08/29/google-wants-watermark-ai-generated-images-stop-deepfakes/

9 Marcelo, Philip. 2023. FACT FOCUS: Fake image of Pentagon explosion briefly sends jitters through stock market. *Associated Press*. May 23. Accessed July 25, 2023. https://apnews.com/article/pentagon-explosion-misinformation-stock-market-ai-96f534c790872fde67012ee81b5ed6a4.

10 Telegraph. 2022. Deepfake of Volodymyr Zelensky surrendering surfaces on social media. *The Telegraph*. March 17. Accessed July 21, 2023. https://youtu.be/X17yrEV5sl4?si=IRXdX3uYYaN-Y7R4

11 Crux. 2023. Putin 'Deep fake' Speech Aired On Russian State TV After Hack. *CRUX*. June 6. Accessed July 21, 2023. https://youtu.be/t14YaS30h-o?si=LMkcgcRzVyFDmiUg

12 Allyn, Bobby. 2022. Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn. *NPR*. March 16. Accessed July 23, 2023. https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

13 Mahdawi, Arwa. 2023. Nonconsensual deepfake porn is an emergency that is ruining lives. *The Guardian*. April 1. Accessed July 22, 2023. https://www.theguardian.com/commentisfree/2023/apr/01/ai-deepfake-porn-fake-images.

14 Dilworth, Miles. 2023. EXCLUSIVE: Celebrity deep fake porn doubles in a year thanks to sophisticated AI as stars including Taylor Swift, Natalie Portman and Emma Watson are targeted. *Daily Mail*. October 1. Accessed October 11, 2023. https://www.dailymail.co.uk/news/article-12545433/Celebrity-deep-fake-porn-

doubles-year-thanks-sophisticated-AI-stars-including-Taylor-Swift-Natalie-Portman-Emma-Watson-targeted.html.

15 Mahdawi, Arwa. 2023. Nonconsensual deepfake porn is an emergency that is ruining lives. The Guardian. April 1. Accessed July 22, 2023. https://www theguardian.com/commentisfree/2023/apr/01/ai-deepfake-porn-fake-images.; Sturla, Anna. 2021. Pennsylvania woman allegedly created deepfake images of cheerleading gym members to cyberbully them, according to criminal complaint. CNN. March 15. Accessed July 23, 2023. https://www.cnn.com/2021/03/15/us pennsylvania-deepfake-cheerleading/index.html.

16 Dickson, EJ. 2023. AI Deepfakes of True-Crime Victims Are a Waking Nightmare. *Rolling Stone Magazine*. May 30. Accessed July 23, 2023. https://www.rollingstone com/culture/culture-features/true-crime-tiktok-ai-deepfake-victims-children-1234743895/.

17 Ibid.

18 Ibid.

19 Acres, Tom. 2023. "Astoundingly realistic' child abuse images being generated using AI." *Sky News*. July 18. Accessed July 23, 2023. https://news.sky.com/story/ astoundingly-realistic-child-abuse-images-being-generated-using-ai-12920978.

20 IWF. 2023. Prime Minister must act on threat of AI as IWF 'sounds alarm' on first confirmed AI-generated images of child sexual abuse. *Internet Watch Foundation*. July 18. Accessed July 23, 2023. https://www.iwf.org.uk/news-media/news/prime-minister-must-act-on-threat-of-ai-as-iwf-sounds-alarm-on-first-confirmed-ai-generated-images-of-child-sexual-abuse/.

21 Ibid.