Digital Policy Hub — Working Paper

AI and Deepfake Voice Cloning: Innovation, Copyright and Artists' Rights

Harnoorvir Singh Josan

Fall 2023 cohort

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

> ### Key Points
>
> - With the surge in generative artificial intelligence (AI) models, voice cloning bots are rapidly surfacing. These bots allow users to replicate the voice of any artist or celebrity of their choice.
>
> - Many AI-generated songs are going viral on social media, creating confusion among the masses. There have also been instances where these bots have been used to spread misinformation related to the artists. The victims of this technology are not able to prevail in cases of copyright infringement because courts have held that voice is not a copyrightable object.
>
> - If copyright is not possible, one way to ensure fair use of these bots is to restrict them to use only approved audio recordings as training data. A common marketplace of stock recordings provided by artists themselves could be used as a pool by all the AI systems. The law should require these bots to tag the content as AI-generated to avoid any confusion.
>
> - Lastly, the ownership of the copyright should be granted to all human and non-human actors to ensure everyone is compensated fairly and that the development of AI technology continues in a manner that is just and equitable. Current copyright frameworks only allow human authors; altering these to also grant authorship to AI would help ensure the development of the technology in the creative field.

Introduction

AI is a discipline of computer science that allows machines to simulate some aspect of human cognition (Calo 2018). The mind-boggling developments made by AI in recent years have been the reason for the buzz around it. Most of the contemporary excitement is due to a particular sub-field of AI called machine learning. Machine learning refers to the ability of machines to recognize patterns in the training data and, thus, improve their performance in tasks over time (ibid.). The availability of massive amounts of training data and computational resources such as extremely powerful computers have led to breakthroughs in machine learning over the past decade (Sturm et al. 2019). The breakthrough events are not just concentrated in the information technology sector but have bled into other fields such as creative arts, health care, finance and so forth. Recently, the music industry has seen a rapid increase in AI-generated vocals where a song is produced without any actual input from the singer. In these cases, a technique called deepfake voice cloning is used. Deepfake voice cloning refers to when AI bots mimic speech patterns and cadence via exposure to the recordings of human speech (Houser 2019). One side effect of such clones is that they are easy to manipulate and can be used to spread misinformation online. In August 2023, an anonymous account posted alleged "leaked recordings" of the ex-president of Sudan on TikTok, which were later proved to be AI-generated, leading to confusion in a country already torn by strife due to civil war (Goodman and Hashim 2023). Due to this technology being novel, no concrete legal frameworks in the United States and Canada have yet addressed this issue. As more celebrities are falling victim to voice cloning, it is high time to consider the legal avenues of recourse (Wells-Edwards 2022). This paper will decipher the use of AI in creative fields, especially voice cloning techniques, look at some of the key cases that

pertain to this topic and provide some recommendations on policy developments that will prevent the illicit use of deepfake voice clones and assist celebrities in safeguarding themselves from any form of misinformation and disinformation campaigns.

Background

AI has long been used in the music industry but was limited to music and beat production as well as music distribution via streaming services (Sturm et al. 2019). AI voice cloning bots have recently infiltrated the recording industry. A simple Google search using the keywords "voice cloning" generates numerous results for voice cloning bots such as voice.ai, which boasts of generating realistic replicas of the original speaker's voice in real-time for live streaming, group chats or gaming experiences.[1] Some users have used this technology to clone the voices of celebrities and insert them into songs they created in order to garner more streams and views. A recent example of this is an AI-generated song "Heart on My Sleeve" — produced by an anonymous TikToker dubbed Ghostwriter977 — that was uploaded to major streaming platforms and quickly went viral on social media (Blackburn 2023). The song featured AI-generated vocals of two famous music superstars, Drake and The Weeknd, and was even submitted to the Grammy Awards' Best Rap Song and Song of the Year categories (Das 2023). This is not the only example where the voice of a celebrity was used without their consent. Deepfake voice cloning AI models have also previously been accused of depicting a band cancelling its concert that had not actually been cancelled, showing artists endorsing things that they had never recommended and musicians bad-mouthing their fans, resulting in customer deception and undermining the public image of artists (Goldman 2023). Although voices are unique to individuals and are a key aspect of their personality, they are not copyrightable entities because copyright protects only original works of authorship fixed in any tangible form; the sound of a voice is not adequately "fixed" and, hence, cannot be owned by an individual.[2] The lack of a copyright framework to prevent using someone else's voice allows the users of voice cloning bots to use the replicas in any form without facing repercussions.

How to Ensure Voice Cloning Bots Do Not Violate Copyright

The song "Heart on My Sleeve" is not an isolated incident; techies have recreated Ice Spice's "Munch" being sung by Drake, Beyonce's "Cuff It" by Rihanna and Kanye West rapping Travis Scott's "90210," with fans showing they like such music (McGee 2023). If users can clone celebrities' voices easily, it is not difficult for anyone to use the same technology to spread misinformation to portray these celebrities in a bad light. To prevent this from happening, there needs to be a framework that requires a substantial amount of vigilance and ensures that AI-generated music is produced in a way that provides fair compensation to the artists.

---

1   See https://voice.ai/voice-cloning.

2   *Midler v Ford Motor Co*, 849 F (2d) 460, 462 (9th Cir 1988).

The first step in the voice cloning process is training the model to produce the expected results (Koempel 2020). Training machine learning models to produce voice replicas requires a large amount of data, such as pre-existing music, audio recordings, interviews and so forth. Unless an exception applies, music used to train a model is generally protected by copyright (Sturm et al. 2019). Even if the user is prohibited from using the music recordings as training data, they can use a large pool of non-copyrighted data such as voice recordings, audio snippets from interviews and speeches to mimic the voice. Once the voice replica is ready, it may seem that the easiest legal avenue for a victim would be a lawsuit alleging copyright infringement of voice as it is a distinctive feature of one's identity. Unfortunately, a victim of voice cloning cannot prevail in a lawsuit based on copyright infringement as the courts have already held in Bette Midler's case that voices are not fixed in any tangible form and, thus, inherently uncopyrightable.[3] Such a legal avenue is possible only if copyright frameworks are altered to allow for the copyright of voice. If a copyright infringement lawsuit is not an available option, some work can be done to ensure that training data is coming from a source that the voice supplier has approved.

Not all artists are reluctant to share their voices with AI bots. Some artists, such as Grimes, a Canadian synth-pop artist, have allowed the public to use their voices without worrying about copyright infringement or legal enforcement (Shanklin 2023). For such artists, there should be a common marketplace of training data that is not copyrighted, is legally owned by the artist and is free to be used for cloning voices. Only the lawful owner of music should be allowed to supply recordings to the archive. It can be the artists themselves or the recording labels who own any work. Such a bank could act as a government-controlled repository like Getty Images, but for audio recordings where all the stock recordings are available. This archive could provide a catalogue of various artists who have allowed the use of their voice and access to their music and audio recordings. There should also be a regulation for the voice cloning bots to accept training data only from that archive. This would also allow artists or asset management companies to charge for accessing their voice recordings if they want to. Having an archive of data provided by the artists themselves would ensure that users can create voice replicas of only those artists who do not have any objections, without having to worry about copyright infringement or legal enforcement. There should also be a requirement for the bots to automatically tag the output audio as AI-generated so it is easily identifiable and does not cause any confusion. YouTube is currently testing a tool called "Dream Track," which is a voice cloning bot and will be used to generate songs using cloned voices of artists who have collaborated with YouTube and provided permission to use their voice (Savage 2023). YouTube also announced that it will start tagging any content generated by AI so that viewers can easily identify AI-generated content (ibid.). Although the initiative by YouTube is one step toward ensuring that voices are not used without the provider's consent, it can be very disorganized and time-consuming for celebrities to grant permission to individual tools or bots one by one. If there was a bank or marketplace for all the bots, artists would only have to provide their consent once. If this repository comes into play, copyright frameworks could be adjusted later to allow an individual to succeed with a copyright infringement if a deepfake voice is created using their audio recordings outside of the bank of stock recordings.

---

3   *Ibid.*

Who Should Own the Final Product?

Protecting artists from having their music or audio recordings used to deepfake their voice without their consent is not the only problem in the process. Even if the clone is generated with the artist's permission, the question of ownership of the end result still remains. In the United States and Canada, there are three fundamental criteria for granting copyright protection: "originality," "work of authorship" and "fixed in a tangible form."[4] For a work to be original, it must be created independently by the author without the contribution of others and should be the creativity of the author without copying it from someone else's work (ibid.). The original work does not necessarily have to be novel or innovative, but must be more than just mechanical or routine, which requires no creativity. In the case of AI-generated music being discussed here, just the vocals are mechanically generated, and the user has to perform all the other necessary steps such as lyric writing, audio editing, mixing and mastering. As the originality requirement is so low, most works including AI-generated music easily meet this requirement (McJohn 2015). The third requirement is that the generated work is "sufficiently stable to be perceived, reproduced, or communicated for a period of more than transitory duration."[5] As long as it is fixed in a tangible form, music generated with AI vocals meets this requirement for copyright purposes. The only requirement that poses a threat to the ability of AI music to receive copyright protection is the authorship. Copyright protection only applies to work with human authors. Although this requirement is not formally enshrined in the Copyright Act of Canada or the Copyright Law of the United States, the courts have refused to grant protection to works generated by non-human actors.[6] Therefore, AI cannot be considered an author for copyright purposes despite being the most important player in the voice cloning process.

There are three candidates left for the authorship of the AI-generated works. One is the voice cloning software's programmer, another is the user who trains the model and produces the final output, and the last is the artist whose voice is cloned using the audio recordings to train the model. Some authors have supported the claim of granting authorship to the programmers of AI systems, arguing that granting them the rights will encourage the growth of AI technology and systems (Hristov 2017). Some have used the same argument to support the authorship being granted to the end users of the software (Brown 2018). Given that in this case artists play an important role in allowing the use of their identity, they deserve to claim the authorship of the output. This would allow them to decide how and where their voice is used and receive fair compensation if used for commercial uses. Both the United States and Canada grant copyright protection to joint authors, given that each author is human and has contributed substantially to the work.[7] The most equitable output is one that allocates joint authorship to both human and non-human actors (Grubow 2018). This will ensure that all the human

---

4   *Feist Publications, Inc v Rural Tel Serv Co*, 499 US 340 (1991); see https://ised-isde.canada.ca/site/canadian-intellectual-property-office/en/copyright-learn-basics/copyright-learn-basics-protect-your-original-works-learn-why-copyright-matters.

5   *Copyright Law of the United States*, 17 USC 101 (2018) [*Copyright Law*].

6   *Naruto v Slater*, 888 F (3d) 418 (9th Cir 2018).

7   *Copyright Act*, RSC 1985, c C-42; *Copyright Law*, *supra* note 5.

and non-human actors involved such as programmers, end users, artists and AI are compensated fairly and encourage the development of fair AI technology. Once the copyright framework is altered to remove the barriers to AI joint authorship, lawmakers should work to ensure that collaboration with AI is regulated (ibid.).

Conclusion

In the end, AI systems are not something that the public needs to worry about if proper safeguards are put in place. Deepfake voice cloning may sound like something that can be used to portray celebrities and artists in a bad light or steal credit by producing songs using their identity. However, if the users of AI and artists work together synchronously, and some changes are made to the copyright framework, a fair use model of AI can be ensured. The most important step is to ensure that the input data used is approved by the original author, performer or the artist. All the voice cloning systems should be required to input the training data only from a bank of stock recordings that are provided by the artists themselves. Once the model is trained and the final output is ready with the voice replica, it should be tagged as AI-generated so the general public can recognize it as such. To ensure all the involved actors get recognition and a fair share, ownership should be distributed jointly among programmers, users, artists and the AI. The current copyright framework in the United States and Canada does not grant authorship to non-human actors. Altering the copyright frameworks to allow for joint authorship of human actors and AI will allow the artists to be fairly compensated without hindering the development of AI technology. The laws to regulate the collaboration of humans with AI need to be implemented because AI systems, if used fairly, can foster creative intelligence, expand our knowledge and improve access to content.

Acknowledgements

About the Author

Harnoorvir (Harnoor) Singh Josan is a computer science undergraduate student at the University of Alberta with a passion for innovation. Harnoor's research with the Digital Policy Hub will centre on AI, music, and copyright policies in Canada and the United States, furthering his commitment to leave a meaningful impact in the technology and policy creation domains. Previously he worked as a product owner for an enterprise mobility management solution that merged his technical acumen with business insights.

Works Cited

Blackburn, H. Drew. 2023. "The Music Industry Has an AI Problem." *The Washington Post*, May 2. www.washingtonpost.com/business/2023/05/02/ai-drake-heart-on-my-sleeve-and-cultural-appropriation/c236dbfc-e8f1-11ed-869e-986dd5713bc8_story.html.

Brown, Nina I. 2018. "Artificial Authors: A Case for Copyright in Computer-Generated Works." *Science and Technology Law Review* 20 (1): 1–41. https://doi.org/10.7916/stlr.v20i1.4766.

Calo, Ryan. 2018. "Artificial Intelligence Policy: A Primer and Roadmap." *University of Bologna Law Review* 3 (2): 180–218. https://doi.org/10.6092/issn.2531-6133/8670.

Das, Mehul Reuben. 2023. "Music's Digital Nightmare: AI-generated song feat. Drake, The Weeknd in race to win a Grammy." *Firstpost*, September 7. www.firstpost.com/tech/news-analysis/musics-digital-nightmare-ai-generated-song-feat-drake-the-weeknd-in-race-to-win-a-grammy-13091442.html.

Goldman, Sharon. 2023. "'Our life's work': Chorus of creative workers demands AI regulation at FTC roundtable." *VentureBeat*, October 5. https://venturebeat.com/ai/our-lifes-work-chorus-of-creative-workers-demands-ai-regulation-at-ftc-roundtable/.

Goodman, Jack and Mohanad Hashim. 2023. "AI: Voice cloning tech emerges in Sudan civil war." BBC News, October 5. www-bbc-com.cdn.ampproject.org/c/s/www.bbc.com/news/world-africa-66987869.amp.

Grubow, Jared Vasconcellos. 2018. "O.K. Computer: The Devolution of Human Creativity and Granting Musical Copyrights to Artifically Intelligent Joint Authors." *Cardozo Law Review* 40 (1): 387–424.

Houser, Kristin. 2019. "Alarming AI Clones Both a Person's Voice and Their Speech Patterns." Futurism, November 6. https://futurism.com/the-byte/ai-clones-voice-speech-patterns.

Hristov, Kalin. 2017. "Artificial Intelligence and the Copyright Dilemma." *IDEA: The Intellectual Property Law Review* 57 (3): 431–54.

Koempel, Florian. 2020. "From the gut? Questions on Artificial Intelligence and music." *Queen Mary Journal of Intellectual Property* 10 (4): 503–13. https://doi.org/10.4337/qmjip.2020.04.05.

McGee, Noah A. 2023. "AI-Generated Music Is Creating Black Art Without Black Input." The Root, April 19. www.theroot.com/ai-generated-music-is-creating-black-art-without-black-1850353845.

McJohn, Stephen M. 2015. *Examples & Explanations: Intellectual Property*. 5th ed. Philadelphia, PA: Wolters Kluwer.

Savage, Mark. 2023. "YouTube tests AI tool that clones pop stars' voices." BBC, November 16. www.bbc.com/news/articles/c4n9erzrg93o.

Shanklin, Will. 2023. "Grimes invites AI artists to use her voice, promising 50 percent royalty split." *Engadget* (blog), April 24. www.engadget.com/grimes-invites-ai-artists-to-use-her-voice-promising-50-percent-royalty-split-165659578.html.

Sturm, Bob L. T., Maria Iglesias, Oded Ben-Tal, Marius Miron and Emilia Gómez. 2019. "Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis." *MDPI Arts* 8 (3): 115. https://doi.org/10.3390/arts8030115.

Wells-Edwards, Bryn. 2022. "What's in a Voice? The Legal Implications of Voice Cloning." *Arizona Law Review* 64 (4): 1213–39. https://arizonalawreview.org/whats-in-a-voice-the-legal-implications-of-voice-cloning/.

.