Leopold-Franzens Universität Innsbruck
Fakultät für Mathematik, Informatik und Physik

Institut für Informatik

Department of Computer Science

Seminar Paper

Legal challenges to generative AI

von
Luka Mirčeta
(Matr.-Nr.: 12115602)

Submission Date: January 7, 2024

Abstract

The unregulated use of generative artificial intelligence (AI) has raised critical legal questions, particularly concerning the data used for training and the implications of created outputs. This paper examines the two primary legal challenges in the realm of generative AI.

Firstly, it scrutinizes the training data used for the development of generative models, exploring the potential use of publicly available copyrighted data as training data and the associated legal challenges to it.

Secondly, the paper examines the implications of the outputs created by generative AI models, considering issues such as biometric identification and the potential misuse of outputs created. The inclusion of real-world case studies adds relevance to the discussion. The recently introduced "EU AI Act" is also analyzed for its potential to address the issues related to generative AI.

As generative AI continues to advance a clear legal framework is crucial in addressing the challenges related to the field. Continuous cooperation between legal authorities, developers and the civil society is required to ensure that artificial intelligence continues adding a positive benefit to our society, mitigating a potential misuse of this powerful intelligence source.

Contents

1 Introduction

Chat interfaces such as "ChatGPT"[1] or "DALL-E"[2], powered by generative artificial intelligence (AI) models have revolutionized the process of text and image creation, allowing the general public to use such models for a wide range of purposes. Even though the outputs of generative models can be very helpful and innovative, the use of such powerful intelligence raises pertinent legal questions regarding both the data on which the models are being trained and the outputs that are created.

This paper will focus on the two main issues concerning the legal sphere of generative AI: the legal challenges emanating from the data on which those models are being trained on and the other one being the legal challenges in context of the output received by the model.

The core of this paper will be an exploration of two case studies. The first one being the "Doe v. GitHub" case [7], where "OpenAI" and "GitHub" have been alleged to have trained their model using copyrighted data, this can be found in section 3.1. The second case study elaborates the "Smart City" project in the city of Belgrade[2], which includes the implementation of thousands of AI powered surveillance cameras on different locations, this can be found in section 3.2.

In the following sections the paper will examine whether the training datasets adhere to copyright laws. Simultaneously we will also take a look at what legal boundaries the outputs of the generative models might propose, considering issues such as biometric identification and social scoring software using artificial intelligence. Lastly the paper will elaborate the recently published "EU AI Act", the world's first comprehensive AI law [6], aiming to provide a comprehensive understanding of the legal challenges surrounding the field of generative artificial intelligence.

2 Legal challenges

2.1 Legal challenges regarding the training data for generative artificial intelligence models

"OpenAI" has revolutionized the field of artificial intelligence as we have known it so far. As they have stated on their website, their company's goal is to develop algorithms and techniques that endow computers with an understanding of our world[5]. They have also proposed that generative models are one of the most promising approaches towards this goal. Hence, they have made generative AI available to the general public in the fall of 2022, by launching a chat interface known as "ChatGPT". Here is a quote from their website, where they state how

[1] `https://openai.com/chatgpt`

[2] `https://openai.com/dall-e-2`

generative models are being trained[3]:

> *"To train a generative model we first collect a large amount of data in some domain (e.g., think millions of images, sentences, or sounds, etc.) and then train a model to generate data like it.", cf. [5]*

While this process is pivotal for the development of generative AI models, it also poses significant legal challenges. Considering that a very large amount of data is needed for the training of such models, questions arise whether publicly available copyrighted data from the internet has potentially been used as training data. In section 3 real world case studies will be examined.

2.2 Legal challenges regarding the use of outputs created by generative artificial intelligence models

The outputs generated by artificial intelligence models can be very powerful, especially when prompted in the correct way. Dangers arise, when artificial intelligence is misused for malicious purposes. Most of the time a very thin line is drawn between using a software for harmful purposes and using it for the necessary protection of public safety. One example can be AI powered surveillance cameras[1]. When used ethically they could be a very helpful tool for the authorities to track down criminals who have been recorded by such a surveillance tool. However, on the other hand side, imagine an authoritarian regime using the same technology to track down political opponents or ethnic minorities in a country. This can be especially problematic if it is not publicly known what exactly is happening to the data recorded by such a surveillance tool and how exactly the data is being transformed and used. Therefore, a clear set of legal regulations is needed.

The European Union has recently released the "EU AI Act", the world's first comprehensive AI law, which has been designed in such a way, that it categorizes the outputs and usages of artificial intelligence in different risk levels and sets up different sets of rules for each of the levels. [6].

3 Case studies

3.1 The "Doe v. GitHub" case

The first case study to be examined will be the "Doe v. GitHub" case, where four programmers, so far identified as John Does, have sued GitHub, Microsoft and OpenAI, alleging that they have violated laws by using publicly available source

[3]For an in depth explanation on how "OpenAI" is training their models please consider reading the whole article stated in [5]

code as training data for their AI programs. GitHub's "Copilot"[4] uses OpenAI's "Codex"[5] AI model, which has been trained on a very large amount of publicly available source code.

The complaint is saying that "Copilot" and "Codex" are engaged in software piracy at an unprecedented scale. Furthermore, they state that the outputs that are produced by the previously named software programs are substantially similar to open-source code. The plaintiffs also alleged "Copilot" and "Codex" to have wrongfully removed the copyright notices of the source code. Therefore they request a federal court to issue an injunction against those AI systems and to grant the plaintiffs $9 billion in statutory damages.[7]

GitHub and OpenAI have dismissed the claims stated by the plaintiffs. They pointed out that the plaintiffs have not identified any code to which they personally claim rights, nor did they specify any harm which has been caused as a consequence of the actions that have allegedly taken place.

3.1.1 Copyright infringement

Copyright law provides authors of the original content with the exclusive right to control the creation of derivative works[8]. The term derivative work is defined in the U.S copyright statute as[6]:

> *"[...] work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted.".*[7]

Outputs of generative AI Systems fit this definition, since their outputs are definitely based upon previous works. In order to infringe the derivative work right, the created work has to be substantially similar to the original work, meaning that a similarity in idea, methodology or other unprotectable aspects is insufficient for infringement of this right.[8]

Whilst it is possible that works created by generative AI could possibly infringe works created by others, it is the responsibility of the company that trains these models to make sure that non infringing outputs are guaranteed. One way the companies could guard themselves from infringing outputs created by their models is to ensure that duplicate data is being removed and to add some kind of output filters to their models.

[4] `https://docs.GitHub.com/en/copilot/overview-of-GitHub-copilot/about-GitHub-copilot-individual`

[5] `https://openai.com/blog/openai-codex`

[6] The same definition can also be found in the European Union under [3]

[7] `https://www.law.cornell.edu/uscode/text/17/101`

So legally speaking in the above stated case study a copyright infringement could be the case if it is somehow proven that the outputs created by the model really are of substantial similarity, but since GitHub and OpenAI have dismissed the case for now, it will take some time until the verdict is handed down.

3.2 Chinese AI surveillance technology in Serbia

The second case study which this paper is going to examine will be the implementation of the "Safe City" project in the city of Belgrade, Serbia carried out by the Chinese company "Huawei", who is also the leading provider of AI powered surveillance technology[4]. The project involves the implementation of thousands of smart surveillance cameras with object and facial recognition features. This has raised serious concerns, because the scope of the project is not known to the general public.

Numerous questions remain unanswered, pivotal topics such as the storage location of data, the party responsible for data processing, mechanisms in place to prevent misuse, the placement and quantity of cameras, and their designated functions. Serbia's current government is facing constant civil unrest marked by extensive protests, due to critical voices not receiving the desired attention in state-owned media[8]. Civil society representatives fear that due to the poor legal regulations regarding artificial intelligence, the "Safe City" project including its facial recognition technology could be used to strengthen Serbia's current regime in the sense of controlling the general population and intimidating political opponents [2].

3.2.1 The "EU AI Act"

For the purpose of clarity, Serbia is not a member of the European Union, however, in this section we are going to talk about what the "EU AI Act"[6] proposes and if it could resolve some concerns of the civil society activists in Serbia. The act establishes different rules for different risk levels[9].

Looking at the highest risk level, "Unacceptable risk", it says that systems falling under this category are considered to be a threat to people and will therefore be banned.

Let's look at what criteria a system would have to fulfill in order for it to be considered an "Unacceptable risk":

- *"Social scoring: classifying people based on behaviour, socio-economic status or personal characteristics"*
- *"Real-time and remote biometric identification systems, such as facial recognition"*

[8] `https://www.bbc.com/news/business-67817072`

[9] This paper will not go into detail about all the different risk levels. For a detailed explanation please consider reading [6]

- *"Biometric identification and categorisation of people"*
- ... (etc.), cf. [6]

Legally speaking, a project like the "Smart City" project in Belgrade would be banned in the European Union, because it definitely falls into the category "Unacceptable risk". The above quoted risks from the Act prohibit the use of real-time facial recognition systems. The Act would also be able to resolve the concerns of the civil society activists, since it also states that social scoring software, classifying people based on behavior, socio-economic status or personal characteristics is prohibited. Furthermore also biometric identification and categorization of people is prohibited.

Nevertheless, the act also states that exceptions may be allowed for law enforcement purposes saying that:

> *"Real-time remote biometric identification systems will be allowed in a limited number of serious cases, while "post" remote biometric identification systems, where identification occurs after a significant delay, will be allowed to prosecute serious crimes and only after court approval.", cf. [6]*

This could still lead to potential concerns for civil activists living in the European Union, because the spectrum of crimes where such a software could legally be used can be very wide. Hence the risk arises that the limits of this law could potentially be misused for malicious purposes.

4 Conclusion

In conclusion, the paper has explored the legal challenges to generative artificial intelligence, focusing on both the training data used for such models and the outputs created by the same. The examined case studies in section 3 have added real world relevance to the discussion, illustrating the complexities and real-world pitfalls to the current legal landscape.

The "EU AI Act" proposes a legal framework to generative AI, categorizing the use of it into different risk levels and proposing restrictions for each of them. In future works, it would be interesting to investigate the copyrightability of outputs generated by AI models.

As generative AI continues to evolve and awareness about it is being spread, the delicate balance between fostering innovation and preventing misuse becomes increasingly crucial. A clear legal framework is vital to address concerns to intellectual property, privacy and potential social impacts. Continuous cooperation between legal authorities, developers and the civil society is required to ensure that artificial intelligence continues adding a positive benefit to our society.

References

[1] Ahmed Abdelmoamen Ahmed and Mathias Echi. "Hawk-eye: An ai-powered threat detector for intelligent surveillance cameras". In: *IEEE Access* 9 (2021), pp. 63283–63293 (cit. on p. 5).

[2] Maja Bjeloš. "The Sum of All Fears–Chinese AI Surveillance in Serbia". In: *Western Balkans at the Crossroads: Ways Forward in Analyzing External Actors' Influence* (2021), p. 141 (cit. on pp. 4, 7).

[3] Comparative Law EPRS. "Copyright Law in the EU: Salient features of copyright law across the EU Member States". In: (2018) (cit. on p. 6).

[4] Steven Feldstein. *The global expansion of AI surveillance*. Vol. 17. Carnegie Endowment for International Peace Washington, DC, 2019 (cit. on p. 7).

[5] Andrej Karpathy et al. *Generative models*. Accessed on January 1, 2024. 2016. URL: https://openai.com/research/generative-models#going-forward (cit. on pp. 4, 5).

[6] Tambiama Madiega. "Artificial intelligence act". In: *European Parliament: European Parliamentary Research Service* (2021) (cit. on pp. 4, 5, 7, 8).

[7] Pamela Samuelson. "Legal Challenges to Generative AI, Part I". In: *Commun. ACM* 66.7 (June 2023), pp. 20–23. ISSN: 0001-0782. DOI: 10.1145/3597151. URL: https://doi.org/10.1145/3597151 (cit. on pp. 4, 6).

[8] Pamela Samuelson. "Legal Challenges to Generative AI, Part II". In: *Commun. ACM* 66.11 (Oct. 2023), pp. 16–19. ISSN: 0001-0782. DOI: 10.1145/3625251. URL: https://doi.org/10.1145/3625251 (cit. on p. 6).