

# Copyright Protection for Generative Model

Contact Person: Zhuan Shi (Email: [zhuan.shi@epfl.ch](mailto:zhuan.shi@epfl.ch))

Advisor: Prof. Boi Faltings

## 1 Project Overview

### 1.1 Motivation and Background

Recently, generative AI models have been extensively developed to produce a wide range of synthesized content, including text, images, audio, and code, among others. For example, advanced image generative models, such as Diffusion Models (DMs) [1], can produce highly realistic and detailed photographs and paintings. Similarly, large language models (LLMs) like ChatGPT [2] can be leveraged to compose coherent and creative text articles with arbitrary genres and storylines.

However, the widespread deployment of these generative models has brought about significant challenges, particularly in the realm of copyright. These models rely on extensive training data obtained from private sources, some of which may contain original works of the data owners. Therefore, these generative models have the capability to generate synthetic data that closely resembles the copyrighted data, severely undermining copyright protection [3]. In recent years, legal disputes have arisen surrounding the utilization of copyrighted data in AI training. In 2023, The New York Times filed a lawsuit against OpenAI and Microsoft, alleging that their AI products, including ChatGPT [2] and Copilot [4], utilized millions of articles from The New York Times for training generative models. Similar lawsuits have been filed against Stable Diffusion [5] and Midjourney [6], asserting that they generate images using training data from millions of artists, potentially infringing upon their rights. These cases demonstrate the importance of research into copyright protection in generative models.

### 1.2 Research Direction

In this project, we have two research directions, and applicants can choose any one of them:

**Research Direction 1:** Based on copyright protection laws and related regulations, propose a metric of copyright infringement. Based on this metric, propose a new learning algorithm for the process of generative model, which can reduce the probability of generating exactly the same copyright data while ensuring the quality of the generated data.

**Research Direction 2:** Give an economic solution to copyright challenges of Generative AI which propose a framework that compensates copyright owners based on their contributions to the creation of AI-generated content and their copyright loss during the training process of generative model. The metric for contributions is quantitatively determined by leveraging the probabilistic nature of modern generative AI models and using techniques from cooperative game theory in economics. The metric for copyright loss is quantitatively determined by the Copyright protection laws and regulations.

## 2 Prerequisites

- Familiar with Diffusion Models, large language model and have experience programming on these generative models.
- Familiarity with differential privacy and game theory would be a plus;
- Fluent English in writing and communication;
- Self-motivated to learning new things and collaboration in a teamwork. We are targeting at a publishable work, so we hope you are also excited to that ;)
- Hope to have read the following papers [7, 8, 9]

## 3 Postscript

- This project is available for both master thesis and semester project;
- Previous publications would be a plus but not a requirement;
- The host lab will provide the computational resources for the project.

## References

- [1] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.
- [2] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, “Gpt-4 technical report,” *arXiv preprint arXiv:2303.08774*, 2023.
- [3] N. Elkin-Koren, U. Hacoheh, R. Livni, and S. Moran, “Can copyright be reduced to privacy?” 2023.
- [4] I. Perez, F. Dedden, and A. Goodloe, “Copilot 3,” Tech. Rep., 2020.

- [5] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 684–10 695.
- [6] S. Mansour, “Intelligent graphic design: The effectiveness of midjourney as a participant in a creative brainstorming session,” *International Design Journal*, vol. 13, no. 5, pp. 501–512, 2023.
- [7] J. Ren, H. Xu, P. He, Y. Cui, S. Zeng, J. Zhang, H. Wen, J. Ding, H. Liu, Y. Chang *et al.*, “Copyright protection in generative ai: A technical perspective,” *arXiv preprint arXiv:2402.02333*, 2024.
- [8] Z. Deng, H. Chiba-Okabe, B. Barak, W. J. Su *et al.*, “An economic solution to copyright challenges of generative ai,” *arXiv preprint arXiv:2404.13964*, 2024.
- [9] N. Vyas, S. M. Kakade, and B. Barak, “On provable copyright protection for generative models,” in *International Conference on Machine Learning*. PMLR, 2023, pp. 35 277–35 299.