Submission Template for ACM Papers

Creative ownership and control for generative AI in art and design

Adapting generative AI systems for the creative processes of artists and designers

[REMOVED FOR PEER REVIEW]

[Removed for peer review]

[REMOVED FOR PEER REVIEW]

[Removed for peer review]

Generative AI enables exciting new ways of augmenting and enhancing human creative processes. Models like Stable Diffusion were rapidly adopted by the public and creatives alike. The models have been widely described as enabling and even democratizing human creativity. But beyond general astonishment about the models' capabilities and new possibilities they offer, we also observe opposite effects. This especially pertains to the adaptation of these technologies by professional creatives, such as designers and artists. We hypothesize that many of these observations can be connected to matters of creative control and ownership within the creative process. These experiences contrast with the common narrative that these models give creatives "superpowers". We would like to create awareness of the concept of creative ownership, an important element of the usability of generative models for artists and designers.

CCS CONCEPTS • Insert your first CCS term here • Insert your second CCS term here • Insert your third CCS term here

Additional Keywords and Phrases: Generative AI, Human-AI co-creation

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

Generative AI enables exciting new ways of augmenting and enhancing human creative processes. Models like Stable Diffusion [1] have shown a popular existing demand for generative algorithms, as they were rapidly adopted by the public and creatives alike. While issues of copyright and ownership are already subject to the current debate, relatively little has been said about the impact of these tools on the creative process itself. With respect to the new possibility they offer to non-professional creatives and laypeople, the models have been widely described as enabling and even democratizing human creativity [2]. But beyond general astonishment about the models' capabilities and new possibilities they offer, we also observe opposite effects. This especially pertains to the adaptation of these technologies by professional creatives, such as designers and artists. After initial use, as well as after further experimentation, various users expressed their experience using generative models like Stable Diffusion in public. Examples of these negative feelings include the feelings of being overwhelmed, frustration, fear of missing out and even guilt and alienation.

We consider it possible that these experiences emerge as symptoms of insufficient creative control and subsequently difficulties developing a satisfactory level of creative ownership when co-creating with generative AI. This contradicts overly optimistic narratives claiming that generative AI is providing creatives with “superpowers” [3].

In this short paper we investigate the possibility of a connection between the negative reception of generative models for creative use and a possible lack of creative control and ownership enabled by them. We also investigate strategies for designing generative systems with respect to creative control and creative ownership.

2 EXPERIENCES OF CREATIVES WITH GENERATIVE AI

We observed that artists and designers experience challenges in integrating generative AI technology into their creative processes. The described user experiences often include rather strong emotional reactions. The topic was raised at [4] to sensitize the AI research community and is further explored in the creative research project [REMOVED FOR PEER REVIEW], a website which collects and maps statements by creatives about their use of generative AI.

A common experience that users have with generative AI is the feeling of being overwhelmed by the amount and variety of the generated output. One user reported:

> “Feeling dizzy - from the excitement, from impatience, from the number of outcomes, from the size of the latent space. It's like going to a loud shopping mall with thousands of blinking lights, voices, kids crying, awful music mixing with the sound of the air-condition” [5]

A user of text-to-image technology reported:

> “It feels like a kind of creative block, that the generations are quicker than me receiving, evaluating and rethinking what I want to generate. It gets amplified when I change the input slightly and the output changes completely / more than I want” [5]

An unclear relationship between user input and model output introduces randomness that can either create positive anticipation and serendipitous experiences, but also feelings of lack of control and understanding and an impossibility of mastering these tools. As a consequence, the mismatch of generated results and the user's creative intention is often a source of frustration. One user reported about their experience with a text-to-image tool:

> “It's like sitting in the driver's seat of a vehicle, desperately trying to find out what all the levers and knobs do.” [5]

> “I found that the more I had a specific vision in mind of what the output should look like, the more I was disappointed and frustrated. You have to embrace the serendipity.” [6]

These examples of user experiences are suggesting issues with creative control and creative ownership. We define creative control as the need to be in control over the input and output of the generative system, to steer it according to one's creative intentions, to have a balance between predictability and serendipity, to be able to assess the output and to control the pace of co-creation. The subject of creative ownership asserts a personal, maybe even intimate relationship with the technology, the need for personal artistic expression and the expectation that the technology conforms to one's own aesthetic or ethical values. We see creative control as a prerequisite to achieving sufficient levels of creative ownership.

3 CHARACTERISTICS OF GENERATIVE AI

Generative AI has several characteristics that make it fundamentally different from other generative and AI tools that are widely adapted within the creative processes of artists and designers. We identify the following characteristics which pose several challenges for the co-creative adaption of generative models.

Extending the work by Weisz et al. [7] and particularly the concept of "generative variability" which summarizes key characteristics of generative AI such as the probabilistic nature, a large number of outputs, fluctuating quality levels, and limited explainability, we would like to highlight additional characteristics of generative AI that are relevant in our discussion of creative control and ownership:

Ambiguity - This characteristic is closely related to the probabilistic nature and pertains especially to text-to-image models which have become one of the most widely used generative AI systems. The general ambiguity between language-based inputs and the multiplicity of images that can be derived according to a prompt is a challenge: Human language is much more "compressed" than an image, hence the generated images contain much more details that were not explicitly requested by the user as part of the prompt, but emerge as an extrapolation of the model.

Foreign and opaque datasets - Generative AI models require vast amounts of data to be trained on. While it is certainly possible for artists and designers to train models on custom datasets consisting of individually selected assets or even create these assets themselves, the dominating practice nowadays seems to work with existing canonical datasets. The rise of the success of foundation models [8] has further perpetuated this development. This poses several challenges: Often the datasets are not open source such as in the case of DALL-E [9], or they are public such as in the case of LAION [10], the training data set of Stable Diffusion, but offer no accessible interfaces to be examined by the non-technical user. This creates an opaqueness that poses a challenge to creative ownership and control.

Furthermore, the datasets contain inherent biases, both unintended, like racial or sexist biases or intended biases, such as aesthetic filtering (https://github.com/LAION-AI/laion-datasets/blob/main/laion-aesthetic.md). These biases might not be aligned with or even in strong contrast to the creative intentions of a user.

4 THE CREATIVE PROCESS AND HUMAN-AI CO-CREATION WITH GENERATIVE MODELS

Exploration, search and accidents constitute important parts of the creative process. It is common in creative processes to pursue explorations without a specific goal. However, at a certain point, control becomes important again. The co-creative process might also be a constitutive part of the goal-identifying process.

Computational systems that enable both exploration and control would hence be desirable. Generative AI systems are very well suited for supporting exploration since their core features are generative variability and ample output. However, control and steerability are equally important. This is where the shortcomings are located and design solutions are much needed.

5 DESIGNING FOR CREATIVE OWNERSHIP

Extending related work on general design principles by Weisz et al. [7] we see the following approaches as particularly suited for enabling more creative control and creative ownership for generative AI:

Designing for exploration

With exploration being a key component of the creative process, investing in strategies that ease the exploration of generative models seems particularly useful. Among many other solutions, visualizations of the latent space are suited to enhance this exploration and we hope to see more techniques to visualize and interact with this high-dimensional space in the future.

Designing for multiple outputs

We observed that a very common experience of users of generative AI appears to be the feeling of being overwhelmed by the number of outputs and possibilities. This underscores the need for interfaces that can deal with a possible abundance of outputs. Weisz et al. [7] have suggested accommodating interfaces to enable versioning, ease curation and annotation and visualise differences between outputs.

Designing for co-creation

The capability of generative AI tools to be easily integrated in individual workflows is an important part of the co-creation process. “Human in the loop interfaces” such as DALLE-Flow (https://github.com/jina-ai/dalle-flow

) adapt to the iterative nature of human creative processes.

Designing your own controls

Recently exciting progress has been made on extending the architectures of generative image models towards more control. ControlNet [11] represents a new approach for conditioning text-to-image diffusion models via an additional network trained on few examples of control in the form of image pairs. The paradigm “designing your own controls” seems like a promising way to enable more creative control and ownership in the future.

6 DISCUSSION

In this paper, we mostly focused on generative AI in the domain of image generation. However, we think that the topics of creative control and creative ownership are relevant and applicable to other creative domains such as text or music.

Generative AI technology is now being used by a wide range of people. Individual user experiences are deliberately shared in public such as on social media in a much broader way compared to other technologies. This represents an opportunity to further understand the needs of creative users.

Ongoing debates about generative AI regarding IP/Copyright, AI replacement and training data made it evident that generative AI technology is a controversial technology. Attitudes towards this technology range from enthusiasm to fear and refusal. There hasn’t been much research however on how these individual or societal attitudes toward technology affect individual co-creation processes.

7 CONCLUSION

We see a strong need to study the user experience of generative AI with artists and designers as stakeholders. We have observed that the current generative AI tools are lacking important mechanisms for creative control and creative ownership, and thus can’t be justly described as giving “superpowers for creatives”. If generative AI technology should be further developed for designers and artists, more research on mechanisms for creative control and creative ownership is needed.

REFERENCES

[1] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). High-Resolution Image Synthesis with Latent Diffusion Models. arXiv. https://doi.org/10.48550/arXiv.2112.10752

[2] Wong, M. (2022, December 14). The Dawn of Artificial Imagination. The Atlantic. https://www.theatlantic.com/technology/archive/2022/12/generative-ai-technology-human-creativity-imagination/672460/

[3] Heaven, W.D. (2022, December 16). Generative AI is changing everything. But what's left when the hype is gone? MIT Technology Review. https://www.technologyreview.com/2022/12/16/1065005/generative-ai-revolution-art/

[4] XXXX, XXXX. (2022, December 9). XXXX [Conference presentation]. NeurIPS 2022 Workshop on Machine Learning for Creativity and Design, virtual.

[5] Emotional Glossary for Creative AI. Retrieved February 23, 2023, from https://publish.obsidian.md/emotional-glossary-creativeai/

[6] Moulton, R. (2021, August 23). Doorways. Ryan Moulton's Articles. https://moultano.wordpress.com/2021/08/23/doorways/

[7] Weisz, J. D., Muller, M., He, J., & Houde, S. (2023). Toward General Design Principles for Generative AI Applications. arXiv. https://doi.org/10.48550/arXiv.2301.05578

[8] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., . . . Liang, P. (2021). On the Opportunities and Risks of Foundation Models. arXiv. https://doi.org/10.48550/arXiv.2108.07258

[9] Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). Zero-Shot Text-to-Image Generation. arXiv. https://doi.org/10.48550/arXiv.2102.12092

[10] Schuhmann, C., Beaumont, R., Vencu, R., Gordon, C., Wightman, R., Cherti, M., Coombes, T., Katta, A., Mullis, C., Wortsman, M., Schramowski, P., Kundurthy, S., Crowson, K., Schmidt, L., Kaczmarczyk, R., & Jitsev, J. (2022). LAION-5B: An open large-scale dataset for training next generation image-text models. arXiv. https://doi.org/10.48550/arXiv.2210.08402

[11] Zhang, L., & Agrawala, M. (2023). Adding Conditional Control to Text-to-Image Diffusion Models. arXiv. https://doi.org/10.48550/arXiv.2302.05543