Towards Personalizing Generative AI with Small Data for Co-Creation in the Visual Arts

Ahmed M. Abuzuraiq[1,*], Philippe Pasquier[1]

[1]*School of Interactive Arts and Technology, Simon Fraser University , Surrey, Canada*

Abstract

Foundational models such as DALL-E and Stable Diffusion are trained on massive and diverse datasets to fit as many needs and contexts as possible. On the other hand, artists often use creative AI systems for niche and bespoke needs that are personal to them, and prompting alone might not always lead to that level of personalization. Co-creative systems are those in which humans and AI take the initiative to work on a common creative task. Generative AI systems support artistic exploration but their probabilistic and impersonal nature challenge artists' sense of control and ownership. In this work, we argue that artists' sense of agency and ownership can be strengthened in co-creative settings through personal generative AI models that are trained (or fine-tuned) on small datasets. We survey the current state of model personalization on creativity-support tools for the visual arts, whether based on Generative Adversarial Networks or Text-to-Image Diffusion models, and argue that co-creative interfaces should adopt similar patterns of integrating personalization. Consequentially, we also propose including personalization as an integral part of a recent interactions framework for human-AI co-creation. Furthermore, we discuss the challenges for integration, including the computational demand of training models, and suggest that some solutions can be found by adopting "small data" and "slow technology" mindsets. Finally, we explore some concrete opportunities for human-AI interactions that personalizing with small data brings.

Keywords

Generative AI, Visual Arts, Co-Creation, Personalization, Small Data, Slow Technology

1. Introduction

As we pass the anniversary of ChatGPT[1], it is clear that generative AI models are here to stay. Image generation models, like DALL-E[2] and Stable Diffusion[3], have surprised everyone with their potential but left - the already disadvantaged - artists at an unsettling crossroads: their copyrights infringed [1], their jobs at risk, and their sense of ownership and creative control challenged [2, 3]. In this paper, we advocate for integrating the personalization of generative AI models into co-creation interfaces, survey current practices and systems in the art community around that, and discuss how challenges to integrated personalization (esp. via model training) can be addressed by adopting a "small data" and/or "slow technology" mindsets.

[1]https://chat.openai.com/
[2]https://openai.com/dall-e-3
[3]https://stability.ai/stable-image

2. Co-Creation and Creativity-Support Interfaces

Co-creation takes place when at least two agents, human or machine, contribute proactively in a shared creative task. This sets it apart from fully autonomous creative AI systems (computational creativity) and creativity-support systems that simply respond to human requests [4]. Recent generative AI systems produce images that can arguably be considered on par with what human artists can produce, at least in terms of fidelity. As these systems do not actively engage in the creative process, they are categorized as creativity-support systems. Nevertheless, their ability to play a substantial role in synthesizing new content has reignited the interest in co-creation. The same promising potential of generative AI also raises questions on how these systems would fit into or impact creative practices. Visual artists use generative AI systems to automate parts of their process, explore ideas and expand on them, or communicate with others [5]. However, regardless of why artists use generative systems, it's instructive to ask how they relate to those systems. In particular, how does relying on a generative AI system impact the artist's ability to perceive the artistic work as their own (i.e. authorship), and how does the probabilistic and black-boxed nature of those systems [6] impact their sense of control over the results (i.e., agency).

3. Generative AI Models & Personalization

3.1. Generative AI: Between Large and Personal

Large-scale Text-to-image Generation Models (LTGMs) such as DALLE-E and Stable Diffusion, are deep learning models trained on millions of images often utilizing many high-end GPUs. These models enable anyone to explore visuals of high quality in various aesthetic styles [5] by supplying carefully written prompts in a process referred to as prompt engineering.

Prior to the spread of Large-scale Text-to-image models, visual artists relying on AI techniques in their work used generative models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). Some variants of GAN models, such as StyleGAN2 [7], can be trained by artists on their personal machines partially due to their capability to produce results of good quality with relatively small datasets. Being able to personally train a generative model allowed artists to explore different aesthetics by changing how the model is trained, which includes curating the collection of images to train (or fine-tune) the model on, and modifying the model's architecture or hyperparameters. In addition to training, artists could explore different aesthetics by navigating the model's generative space through sampling, creating smooth interpolations, or by following semantic directions [8].

More recently, this line of (personally-trainable) generative models was overshadowed by large generative models, which despite their impressive quality, require massive resources to be trained, making training them inaccessible to most artists. Instead of training, artists achieve different aesthetics by navigating the generative space of large models (e.g. with prompting) and when navigating the model's generative space proves cumbersome or if it consistently fails to produce personal results, artists can fine-tune large models, which helps in narrowing and focusing the scope of navigation.

3.2. Why Personalize?

Whether artists personally train models from scratch, fine-tune large or personally-trainable models, or carefully navigate a generative space (e.g., by prompting), they are personalizing, i.e. making the generated products their own by shaping or influencing the generation. However, these approaches afford artists different means of control over the results. For example, prompting on text-to-image models cannot introduce new concepts, or learn from examples, which is why many artists turn to personalizing their models as we discuss in Section 4.1. We focus here on model personalization with data that can be performed by artists directly, as we find it more empowering and personal for artists to do so. Therefore, we do not consider other cases of model personalization such as online learning and reinforcement learning (where feedback from users indirectly updates a model) or where the system is programmed to personalize itself to users (e.g., based on patterns in their behaviour).

When artists personalize generative models they regain more control, and an increased level of control can contribute to artists' sense of ownership over the results [9, 2] (authorship) and to an increased ability to realize their visions (agency). Furthermore, the ability to personalize the generative model that supports their work (creativity-support), or which they directly or indirectly collaborate with (co-creation), allows artists to push the invisible boundary of what the generative AI model can create [10] in directions of their own choosing, hence boosting their agency. Human collaborators communicate and adapt throughout collaborative creative processes, similarly, supporting personalization (whether a human personalizes an AI model, or the AI partner personalizes itself) can lead to an evolving collaboration between the human(s) and their AI partner(s) in co-creative settings.

Finally, the arguments we make here are not for GAN models over text-to-image models (or vice verse), since GAN models can be large and prompt-based [11] and text-to-image models can be small and personally-trainable [12], and both types of models have their respective communities. Instead, we argue that artist-led personalization offers opportunities for creative control and ownership, and the scale and type of the underlying generative model simply impacts the types of personalization afforded to artists. Having argued for model personalization, we investigate next how personalization is currently integrated into creativity-support and co-creative systems.

4. Current Landscape of Artist-led Personalization

4.1. Personalization in Creativity-Support Systems

Artists with skills and knowledge in coding and deep learning can train or fine-tune generative models at will, potentially with the help of cloud computing. Creativity-support and co-creative systems are often targeted to novices and experts alike [13], and so we focus here on no-coding tools that allow artists to work with and personalize image generation models on the same interface. Table 1 shows different systems used in practice that support training or fine-tuning generative models. There is a larger collection of systems, not included here, that allow users to sample from a generative system but without offering options for personalization.

Table 1
Creativity-support systems that provide no-code or low-code features for training or fine-tuning generative models. T2M denotes text-to-image models, which are commonly based on diffusion techniques. The local T2M platforms are also commonly hosted on on cloud platforms.

System	Algorithm	Training		Navigation Features	Price/Platform
		From Scratch	Fine-tune		
Playform	GAN	Yes	-	Snapshot	Freemium/Cloud
RunwayML (GAN)	GAN	Yes	Yes	Snapshot	Premium/Cloud
Autolume	GAN	Yes	Yes	Semantic Controls, MIDI	Free/Local
RunwayML (T2M)	T2M	-	Yes	Prompting	Premium/Cloud
OpenArt Photobooth	T2M	-	Yes	Prompting	Freemium/Cloud
Automatic1111 WebUI	T2M	-	Yes	Prompting	Free/Local
Invoke AI (community)	T2M	-	Yes	Prompting	Free/Local
Invoke AI (industry)	T2M	-	Yes	Prompting	Premium/Cloud
ComfyAI	T2M	-	Yes	Prompting	Free/Local

4.1.1. GAN-based Systems

Systems such as RunwayML[4], Playform [14], and Autolume [15] are GAN-based systems aimed at artists with no coding experience. These systems offer options for training or fine-tuning GAN models in addition to some process-support features that are often found in digital arts workflows, such as image pre-processing in preparation for training (e.g., cropping, resizing, normalization), visualizing the training progress with charts and snapshots of the results, conditioning generation (e.g., with in-painting or out-painting), or post-processing the results (e.g. up-scaling). We observe that most of the no-code GAN-based systems require payment except for Autolume. Systems such as TorchGAN [16], StudioGAN [17] and the GAN Toolkit [18] enable users to train their own GAN models from scratch by modifying high-level configuration files (e.g. JSON files). These coding-based systems encourage exploring variations to familiar generative models or provide a unified framework to support performance benchmarking. However, they do not offer adequate process-support features as outline above[5].

4.1.2. Text-to-Image Systems

The Text-to-image generative AI platforms listed in Table 1 such as Automatic1111 WebUI [19], ComfyAI [20], InvokeAI (community-version) [21], InvokeAI (industry-version) [22], RunwayML [23], and Photobooth [24] offer similar interfaces where users specify prompts, adjust the diffusion process parameters and generate. The exception here is ComfyUI which provides a graph/nodes interface. Most also support artistic workflows and community-based extensions, but they do not offer the option to train from scratch as it is costly to do so. Instead, artists choose between multiple pre-set models (such as from Stable Diffusion), or models fine-tuned

---

[4]This is an old version of RunwayML where model training was GAN-based: https://app.runwayml.com/train
[5]The same critique applies to generic cloud-based model training platforms for that matter such as AWS or Azure.

by others found on public model repositories like HuggingFace[6], CivitAI[7], or PromptHero[8].

Artists have the option to personalize text-to-image models, locally or on the cloud, on their own datasets through techniques such as Textual Inversion [25], DreamBooth [26], Low-Rank Adaptation (LoRA) [27] and ControlNet [28]. For example, DreamBooth fine-tunes text-to-image diffusion models on few images, which results in associating the subject of those images with a new keyword that can be added to prompts. Finally, in terms of accessibility, some of those text-to-image systems require using a command line to install, and running the personalization techniques effectively requires some experimentation and relying on the experiences of others in the AI art community.
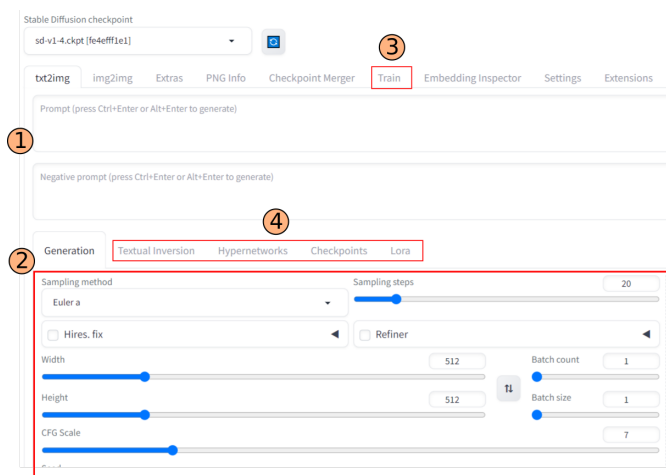
Figure 1: A breakdown of the interface of Automatic1111 WebUI a currently famous interface for art creation with generative AI. (1) users write prompts to describe the desired result. (2) multiple diffusion-specific parameters can be tuned. (3) switching to the Train panel opens options for personalizing a base model. (4) the user's personal models are listed and ready for use in generation.

4.1.3. Integrated Personalization

Personalization through fine-tuning or few-shot adaptation is integrated into all the listed systems in Table 1. As an example, Figure 1 shows how personalization is integrated into the interface of Automatic1111 WebUI. This is can also be seen on GAN-based interfaces such as Autolume [15] which allows artists to train models then swap them as needed. By integrating personalization into these tools, artists can alternate between model selection/personalization and image generation with prompting, and both tasks can shape or validate the other.

4.2. Personalization in the Co-Creation Literature

The systems above come from industry and the AI art community and some of them, e.g. Runway and Automatic1111 WebUI, are quite famous in the community. However, we categorize them as

---

[6]https://huggingface.co/models
[7]https://civitai.com/
[8]https://prompthero.com/

creativity-support tools in that the system is not proactive, i.e. it does not initiate in the creative process. We focus here on how personalization is discussed in the co-creation literature.

Muller et al. [29] describe a framework for interactions between humans and AI in co-creative settings consisting of multiple actions that can be performed by either. Grabe et al. [30] suggest that a smaller and modified list of actions better suit GAN-based co-creative systems, the modified list include: *initialize* (preparing data and choosing models), *learn* (human learns or AI trains), *constrain* (specifying desired characteristics), *create* (generate new artifacts), *select* (choose or exclude), *adapt* (edit one artifact), and *combine* (combine multiple artifacts). Grabe et al. identify four interaction patterns that are common in the literature, such that each pattern traces through a different combination of the actions above. These interaction patterns include *curating*, *exploring*, *evolving* and *conditioning*. For example, curating involves the human initializing the model, AI training the model (learn) and generating new artifacts (create), and the human curating some of them (select).
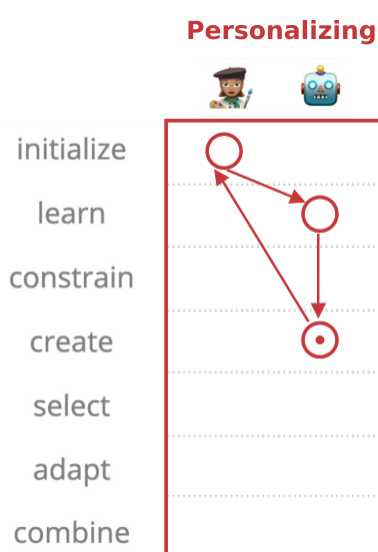
Figure 2: An extension to the interaction patterns between humans and GANs initially proposed by Grabe et al. [30]. Human sets up the training (or fine-tuning) of the model, and the AI is trained based on those settings followed by generating new outputs. If humans are not satisfied with the results they can go back to setting up the model and re-training.

In those interaction patterns identified by Grabe et al. [30], initializing and learning are not revisited during the course of co-creation. However, personalization of the generative model seems to be an essential part of creative AI work in practice as we described in the last section. So in Figure 2, we propose extending Grabe et al.'s taxonomy with a new interaction pattern, namely *personalizing*. We also propose re-defining the *initialize* action to also include choosing among pre-trained generative models, in which case there is no need to pass by the training step. When combining the *personalizing* pattern with others we can describe how Diffusion artists work: they curate (generated artifacts are saved locally), explore (e.g., adapting by upscaling), condition (e.g., tuning model parameters, out-paining and in-painting) and personalize (e.g., fine-tune via textual inversion).

By observing the adoption of personalization on creativity-support systems for image generation (Section 4.1.3), we join other researchers [31, 32] in calling for integrating personalization into co-creative systems. Broad et al. [31] suggest integrating active divergence techniques (e.g., network blending, rewriting and bending among other ways for adapting and personalizing models) into co-creation and creativity-support interfaces. Furthermore, Shimizu and Fiebrink [32] introduce Genny, a live-coding environment for interacting with generative models and suggest integrating the training small-data models into the system in future work. So far we have argued for integrating personalization into co-creation by appealing to practices in the AI art community (Section 4.1.3) and arguing conceptually for the benefits of personalization (Section 3.2). Next, we discuss the challenges and opportunities relating to integrating specific types of personalization into co-creative interfaces, i.e. model training.

5. Personalizing with Model Training: Challenges and Mindsets

Large and personally-trainable models provide different affordances for art creation. Large models may produce higher fidelity results but they are practically impossible to train by individual artists. On the other hand, personally-trainable models offer finer means of control and personalization through model training, but producing good results with them requires time and expertise. For example, training models have a high computational cost, and aside from very simple models like VAEs, training requires hours to days depending on the computational resources available and the size of the model/data. Furthermore, some models such as GANs are notoriously hard to train stably and they suffer from issues like mode collapse, where the model lacks diversity in its outputs. To address those challenges we recommend adopting two mindsets: a "small data" mindset and a "slow technology" mindset.

5.1. A "Small Data" Mindset

As the race for larger generative models trained on massive datasets continues, researchers like Vigliensoni et al. [33] argue that adopting a "small data" mindset might be more befitting to creative domains at times. A "small data" mindset calls for recognizing the value of training (or fine-tuning) models on carefully curated small datasets, particularly within creative domains. Similar calls for "small data" were made by researchers for creative domains where large datasets are not readily available or where artists may prefer to use their own creations, such as VJing [34], choreography [12], and rapping [35].

5.1.1. Small Data is All You Need

Manipulating small training sets alone can be sufficient for co-creation. The case study by Friedman and Pollak [36] on teaching art students about creative deep learning technologies demonstrates how curating training datasets while keeping models fixed, can be sufficient for achieving a variety of creative intents. Due to their size, small datasets can be browsed and manipulated easily which allows artists to appreciate the influence that the training dataset has on the model's behaviour.

5.1.2. Simpler and Data-Efficient Models

A "small data" mindset also advocates for simpler models (or data-efficient models), which can be trained (and retrained) faster. Elgammal and Mazzone [37] discuss how most generative models require training on large datasets (tens of thousands) to produce reasonable results when in practice, artists using the AI Art model training platform (Playform) frequently uploaded datasets of under a 100 images for training. A later work [38] built on this observation and explored a GAN model that trains stably while requiring a few training samples and demanding a low computational cost. Several similar works followed and recent surveys [39, 40, 41, 42] summarize the advances on data-efficient, few-shot and Limited-Data models.

5.1.3. Searching for the Ideal Model

Finding the ideal model for co-creation in the visual arts, i.e., one that features fast training on small datasets, produces good quality images and quick inference for near-real-time interaction, might not be possible as trade-offs do exist between a model's complexity (and hence its training speed), and its quality [43]. Furthermore, for artists, the quality and meaning of generated results are subjective and artists may have different needs and goals from co-creation, in addition to having varying levels of access to the computational resources needed for training models (whether on local machines or on the cloud), and hence different conceptions of what is considered "small". Therefore, we recommend the designers of co-creative interfaces not focus on finding the best model that can be easily integrated and personalized, as much as on giving artists the ability to explore and personalize different generative models at will.

Nevertheless, exploring the spectrum[9] of generative models in terms of their quality, complexity, training stability, and inference speed presents a rich area of future research. This exploration would involve determining how different models balance these aspects, and identifying which types of models are best fitted for various co-creative settings and tasks and the models' characteristics that lead to this fit.

5.1.4. When Simple Models are Better

Purely from an artistic perspective, neither State-of-the-Art models nor simple VAE models are inherently better. In fact, Vigliensoni et al. [33] argue that the unique goals of artists turn issues like overfitting and bias, often associated with training on small datasets or simple machine learning models, into tools for artistic creation. Taking this idea to the extreme, Akten [44] shows that by relying on simple models such VAEs, artists can freely edit a model's hyperparameters and re-train it in real-time, effectively turning those models into *visual instruments*. In another example, Shimizu et al. [12] present an approach for creating custom text-to-media mappings for any generative model using a few mapping examples and relying on a simple multilayer perceptron. Finally, simpler models are particularly suitable for artists who are new to machine learning, as they can be better understood and dissected.

---

[9]Or more technically, the Pareto Frontier, which refers here to the set of models that represents the best trade-off between all measured objectives such as quality and complexity.

5.2. A "Slow Technology" Mindset

Another approach for handling the challenge of slow model training is adopting a different mindset when designing co-creative interfaces, namely to design with slowness in mind. Artists do not necessarily finish their works in one setting, sometimes due to time constraints and other times to create a distance from their work to come back to it with fresh eyes. Furthermore, people often collaborate asynchronously, and this has become more common post-COVID. So why do we expect real-time interaction in co-creative settings? Inspired by "slow technology", which is a mindset of designing for reflection rather than speed, with interactions spanning over days, months or years [45]. An example of slow tech is Photobox [46], which is a wooden box with a printer that looks like a household and prints out pictures from personal albums occasionally over the span of months. As for creative systems designed with slowness in mind, examples include SAGA [47] and Puck [48]. SAGA is an asynchronous collaborative storytelling system where users take turns over time adding to a story, while Puck is an automated game designer that intentionally produces games at "human-like scales of creativity". Similarly, we can envision co-creative interfaces where models are trained over time as both human(s) and/or the AI partner add to the training set, while results are shown on a display placed at home or in public areas. An emphasis can be placed on supporting reflection [49] by keeping a visual record of how the models' outputs change through time as contributions from both sides accumulate.

6. Design Opportunities

By adopting small data and slow technology mindsets, we can explore co-creative interfaces that integrate personalization through model training knowing that an algorithmic or design-based solution will or does exist. Next, we explore some interactions for co-creation with generative AI models that becomes practically possible by adopting a small data mindset.

6.1. Implications for Human Interaction

- **Small Data is All You Need:** Co-creation can happen through changing the training dataset alone. This can be facilitated by dividing the training data into pre-defined sets (e.g., each set denoting a different visual style) and interaction can happen on the level of sets by experimenting with different combinations of sets/style to include in training.
- **Visual Instruments:** Hyperparameter tuning is often seen as a necessary evil, but Akten [44] shows, that iterative and continuous editing of hyperparameters can be a creative instrument in its own right.
- **Unlearning:** Machine Unlearning techniques can be used to remove the effect of a training sample from a trained model. A proposed method for effective unlearning is to split the training dataset between multiple simple models that join together as an ensemble [50], such that the cost of removing the effect of a training sample is lower, i.e. retrain one simple model rather than a large one. Such ensembles of generative models have been shown to model data distributions better than individual large models with a minimal added cost [51] and in a data-efficient manner [52]. Unlearning can be used as an interaction mechanism on a co-creative interface where users add or remove individual

images from the training set at will. It's also common for AI artists to want to remove images that are skewing the results in undesirable directions.

- **Active Divergence:** Artists often seek novelty and surprise from generative models. However, generative deep learning models are, by definition, trained to model a target distribution. Active divergence methods [31], as the name indicates, are a collection of methods that diverge generative models from a target distribution to produce novel and out-of-distribution samples. Simple models can be subjected more easily to to different divergence methods, such as network bending, blending and rewriting. These methods require little or no data to perform, and they can be integrated into co-creative interfaces as Broad et al. suggest [31].

6.2. Implications on Human-AI Interactions and AI's Proactivity

So far we have focused on the agency of human artists when working with generative AI. When humans create together, they evolve and adapt throughout their collaboration and we can expect the same in human-AI co-creation. Koch et al. [53] present an analysis to the concept of agency in a co-creative setting and suggest that AI systems displaying proactivity and adaptability are *perceived* as possessing agency by users, and hence as co-creators. Examples of proactivity can include any of the interactions listed above, e.g. an AI agent can personalize itself on small datasets according to a plan or in reaction to actions taken by the human partner. Koch et al. [53] also speculate that the ability to co-create a design space with a partner may contribute to the perception of agency. Relatedly, Berns et al. [54] discuss a framework where artists can automate parts of the process of creating deep generative models by handing them over to machines. Finally, a recent survey on co-creative systems by Rezwana and Maher [55] points out that in only a few co-creative systems the AI partner contributed to defining the conceptual space to be explored. Giving an AI partner the ability to contribute to building a generative space can lead to conceptual contributions as well. It remains an open question whether *perceiving* AI partners as possessing creative agency diminishes artists' perception of their own agency, or if we can expect similar dynamics as when humans collaborate.

7. Conclusion

In this paper, we argued for model personalization with small data with an emphasis on the role that it could play in strengthening the sense of control and ownership that artists may experience when using co-creative generative AI systems. We surveyed model personalization in multiple creativity-support systems that come from academia, industry and the AI art community, based on our findings we suggested incorporating personalizing into co-creative workflows and frameworks. The computational demands of personalization through model training can hinder integrating them into co-creation, so we discuss some strategies to address that by adopting "small data" and "slow technology" mindsets. Finally, we suggest multiple interactions, initiated by artists or machines, that can become practical to implement after adopting these mindsets. Future work will explore co-creative systems centered around the ideas in this paper.

References

[1] H. H. Jiang, L. Brown, J. Cheng, M. Khan, A. Gupta, D. Workman, A. Hanna, J. Flowers, T. Gebru, AI art and its impact on artists, in: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, 2023, pp. 363–374.

[2] A. Steinbrück, A. Stankowski, Creative ownership and control for generative AI in art and design, in: In Generative AI in HCI Workshop, CHI '23, 2023.

[3] J. Oppenlaender, A. Visuri, V. Paananen, R. Linder, J. Silvennoinen, Text-to-Image Generation: Perceptions and Realities, in: In Workshop on Generative AI and HCI at CHI '23, 2023. `arXiv:2303.13530`.

[4] N. Davis, Human-computer co-creativity: Blending human and computational creativity, in: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, volume 9, 2013, pp. 9–12.

[5] H.-K. Ko, G. Park, H. Jeon, J. Jo, J. Kim, J. Seo, Large-scale text-to-image generation models for visual artists' creative works, in: Proceedings of the 28th International Conference on Intelligent User Interfaces, IUI '23, Association for Computing Machinery, New York, NY, USA, 2023, pp. 919–933. doi:`10.1145/3581641.3584078`.

[6] J. D. Weisz, M. Muller, J. He, S. Houde, Toward General Design Principles for Generative AI Applications, 2023. doi:`10.48550/arXiv.2301.05578`. `arXiv:2301.05578`.

[7] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4401–4410.

[8] E. Härkönen, A. Hertzmann, J. Lehtinen, S. Paris, Ganspace: Discovering Interpretable Gan Controls, Advances in neural information processing systems 33 (2020) 9841–9850.

[9] R. Louie, A. Coenen, C. Z. Huang, M. Terry, C. J. Cai, Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, Association for Computing Machinery, New York, NY, USA, 2020, pp. 1–13. doi:`10.1145/3313831.3376739`.

[10] D. Buschek, L. Mecke, F. Lehmann, H. Dang, Nine potential pitfalls when designing human-AI co-creative systems, in: HAI-GEN 2021: 4th Workshop on Human-AI Co-Creation, College Station, USA, 2021.

[11] M. Kang, J.-Y. Zhu, R. Zhang, J. Park, E. Shechtman, S. Paris, T. Park, Scaling up GANs for text-to-image synthesis, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[12] J. Shimizu, I. Olowe, T. Broad, G. Vigliensoni, P. Thattai Ravikumar, R. Fiebrink, Interactive machine learning for generative models, in: Machine Learning for Creativity and Design Workshop, 2023.

[13] B. Shneiderman, Creativity Support Tools: Accelerating Discovery and Innovation, Commun. ACM 50 (2007) 20–32. doi:`10.1145/1323688.1323689`.

[14] Playform: No-Code AI for Creative People, https://www.playform.io/train, 2024.

[15] Metacreation Lab, Autolume: A Neural-network based Visual Synthesizer, https://www.metacreation.net/autolume, 2024.

[16] A. Pal, A. Das, TorchGAN: A flexible framework for GAN training and evaluation, J. Open Source Softw. 6 (2019) 2606.

[17] M. Kang, J. Shin, J. Park, StudioGAN: A taxonomy and benchmark of GANs for image synthesis, IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (2022) 15725–15742.

[18] A. Sankaran, R. Sinha, N. Panwar, IBM GAN Toolkit, https://https://github. com/IBM/gan-toolki, 2018.

[19] Automatic1111 WebUI, https://github.com/AUTOMATIC1111/stable-diffusion-webui, 2024.

[20] ComfyAI, https://github.com/comfyanonymous/ComfyUI, 2024.

[21] InvokeAI (community-version), https://github.com/invoke-ai/InvokeAI, 2024.

[22] InvokeAI (industry-version), https://invoke.ai/, 2024.

[23] RunwayML, https://runwayml.com/ai-magic-tools/ai-training/, 2024.

[24] Photobooth, https://openart.ai/photobooth, 2024.

[25] R. Gal, Y. Alaluf, Y. Atzmon, O. Patashnik, A. H. Bermano, G. Chechik, D. Cohen-Or, An image is worth one word: Personalizing text-to-image generation using textual inversion, arXiv preprint arXiv:2208.01618 (2022). `arXiv:2208.01618`.

[26] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, K. Aberman, Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 22500–22510.

[27] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, Lora: Low-rank adaptation of large language models, arXiv preprint arXiv:2106.09685 (2021). `arXiv:2106.09685`.

[28] L. Zhang, A. Rao, M. Agrawala, Adding conditional control to text-to-image diffusion models, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 3836–3847.

[29] M. Muller, J. D. Weisz, W. Geyer, Mixed initiative generative AI interfaces: An analytic framework for generative AI applications, in: Proceedings of the Workshop the Future of Co-Creative Systems-a Workshop on Human-Computer Co-Creativity of the 11th International Conference on Computational Creativity (ICCC 2020), 2020.

[30] I. Grabe, M. G. Duque, S. Risi, J. Zhu, Towards a framework for human-AI interaction patterns in co-creative GAN applications, in: Joint Proceedings of Intelligent User Interfaces (ACM IUI) Workshops, 2022, pp. 92–102.

[31] T. Broad, S. Berns, S. Colton, M. Grierson, Active divergence with generative deep Learning–A survey and taxonomy, in: Proceedings of the 12th International Conference on Computational Creativity (ICCC '21), 2021. doi:`arXivpreprintarXiv:2107.05599`.

[32] J. Shimizu, R. Fiebrink, Genny: Designing and exploring a live coding interface for generative models, in: Proceedings of the 7th International Conference on Live Coding (ICLC 2023), 2023.

[33] G. Vigliensoni, P. Perry, R. Fiebrink, A small-data mindset for generative AI creative work, in: In Generative AI in HCI Workshop, CHI '22, New York, NY, USA, 2022, p. 5.

[34] J. Kraasch, P. Pasquier, Autolume-Live: Turning GANs into a Live VJing tool, in: Proceedings of the 10th Conference on Computation, Communication, Aesthetics & X, Coimbra, Portugal, 2022, pp. 152–169.

[35] I. Olatunji, Why try to build try to build a co-creative poetry system that makes people feel

that they have "creative superpowers"?, in: HAI-GEN 2023: 4th Workshop on Human-AI Co-Creation, Joint Proceedings of the ACM IUI Workshops 2023, 2023, pp. 67–80.

[36] D. Friedman, D. Pollak, Image co-creation by non-programmers and generative adversarial networks., in: Joint Proceedings of Intelligent User Interfaces (ACM IUI) Workshops, 2021.

[37] A. Elgammal, M. Mazzone, et al., Artists, artificial intelligence and machine-based creativity in Playform, Artnodes (2020) 1–8.

[38] B. Liu, Y. Zhu, K. Song, A. Elgammal, Towards faster and stabilized gan training for high-fidelity few-shot image synthesis, in: International Conference on Learning Representations, 2020.

[39] M. Abdollahzadeh, T. Malekzadeh, C. T. H. Teo, K. Chandrasegaran, G. Liu, N.-M. Cheung, A survey on generative modeling with limited data, few shots, and zero shot, ArXiv abs/2307.14397 (2023).

[40] Z. Li, X. Wu, B. Xia, J. Zhang, C. Wang, B. Li, A comprehensive survey on data-efficient GANs in image generation, ArXiv abs/2204.08329 (2022).

[41] M. Yang, Z. Wang, Image synthesis under limited data: A survey and taxonomy, ArXiv abs/2307.16879 (2023).

[42] T. Moon, M. Choi, G. Lee, J.-W. Ha, J. Lee, Fine-tuning diffusion models with limited data, in: NeurIPS Workshop on Score-Based Methods, 2022.

[43] G. Menghani, Efficient deep learning: A survey on making deep learning models smaller, faster, and better, ACM Computing Surveys 55 (2023) 1–37.

[44] M. S. Akten, Deep Visual Instruments: Realtime Continuous, Meaningful Human Control over Deep Neural Networks for Creative Expression, Ph.D. thesis, Goldsmiths, University of London, 2021.

[45] L. Hallnäs, J. Redström, Slow technology – designing for reflection, Personal Ubiquitous Comput. 5 (2001) 201–212. doi:10.1007/PL00000019.

[46] W. Odom, M. Selby, A. Sellen, D. Kirk, R. Banks, T. Regan, Photobox: On the design of a slow technology, in: Proceedings of the Designing Interactive Systems Conference, 2012, pp. 665–668.

[47] H. Shakeri, C. Neustaedter, S. DiPaola, Saga: Collaborative storytelling with gpt-3, in: Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing, 2021, pp. 163–166.

[48] M. Cook, Puck: A slow and personal automated game designer, Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment 18 (2022) 232–239. doi:10.1609/aiide.v18i1.21968.

[49] M. Kreminski, M. Mateas, Reflective Creators, in: Proceedings of the 11th International Conference on Computational Creativity (ICCC'21), 2021, pp. 309–318.

[50] N. Aldaghri, H. Mahdavifar, A. Beirami, Coded machine unlearning, IEEE access : practical innovations, open solutions 9 (2021) 88137–88150.

[51] Y. Wang, L. Zhang, J. van de Weijer, Ensembles of Generative Adversarial Networks, in: NIPS Workshop on Adversarial Training, arXiv, 2016. doi:10.48550/arXiv.1612.00991. arXiv:1612.00991.

[52] Y. Du, L. Kaelbling, Compositional generative modeling: A single model is not all you need, arXiv preprint arXiv:2402.01103 (2024). arXiv:2402.01103.

[53] J. Koch, P. T. Ravikumar, F. Calegario, Agency in co-creativity: Towards a structured

analysis of a concept, in: ICCC 2021-12th International Conference on Computational Creativity, volume 1, Association for Computational Creativity (ACC), 2021, pp. 449–452.

[54] S. Berns, T. Broad, C. Guckelsberger, S. Colton, Automating generative deep learning for artistic purposes: Challenges and opportunities, in: Proceedings of the 12th International Conference on Computational Creativity (ICCC '21), 2021.

[55] J. Rezwana, M. L. Maher, Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems, ACM Transactions on Computer-Human Interaction (2022).