Exploring Aesthetic Qualities of Deep Generative Models through Technological (Art) Mediation

Christian Sivertsen
csiv@itu.dk
IT University of Copenhagen
Copenhagen, Denmark

Anders Sundnes Løvlie
asun@itu.dk
IT University of Copenhagen
Copenhagen, Denmark

ABSTRACT

Deep Generative Models (DGM) have had a great impact both on visual art and broader visual culture. In this research-through-design project we investigate the use of a DGM for helping museum visitors explore the aesthetics of Edvard Munch's art. We designed and built an interactive drawing table that allows a user to explore a StyleGAN model trained on sketches by Edvard Munch. The paper makes two novel contributions: 1. It presents a system that allows users to interact with a DGM by drawing on paper (rather than the typical text prompts used by most current systems). 2. We demonstrate how this mode and quality of interaction establish a unique perspective on Munch's drawings as a practice. Through qualitative evaluation, we discuss how this setup led users towards a specific hermeneutic drawing strategy that enables building competency with the model and by proxy the data it is trained on. We suggest that the resulting interaction may contribute to an "education of attention" helping museum visitors to become attentive to certain visual qualities in Munch's drawing practice. Finally, we discuss how the concepts of technological mediation and relationality are useful for designing how the output of a DGM is understood by its users.

CCS CONCEPTS

• **Human-centered computing** → *Interface design prototyping*; **Interaction design**; **Interaction devices**; • **Applied computing** → **Fine arts**; • **Computing methodologies** → *Machine learning*.

KEYWORDS

drawing, interaction design, stylegan, deep generative model, machine learning, aesthetics, postphenomenology, fine art

ACM Reference Format:
Christian Sivertsen and Anders Sundnes Løvlie. 2024. Exploring Aesthetic Qualities of Deep Generative Models through Technological (Art) Mediation. In *Designing Interactive Systems Conference (DIS '24), July 01–05, 2024, IT University of Copenhagen, Denmark*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3643834.3661498

DIS '24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0583-0/24/07
https://doi.org/10.1145/3643834.3661498

1 INTRODUCTION

The recent wave of deep generative models (DGM) capable of synthesizing both convincing text and images has had an extensive impact on visual culture as evidenced by the field of AI Art [8, 75], and has facilitated new practices for the production of images [16, 53]. Generative AI has been described as a new artistic medium that may fundamentally alter artistic processes, as well as raise questions about ownership, copying, and manipulation [25]. While the use of AI technologies for analyzing art collections and creating new forms of art experiences in museums have in recent years been a hot topic for debate [18, 26, 27, 65, 66], the use of generative AI in art museums remains less explored [50], although it has been applied outside the museum context in so-called "immersive" art exhibitions [49, 51, 64].

Some artistic uses of AI have taken a critical stance, exploring biases and other problematic side effects of the technology [20, 69, 71]. Similar concerns have been explored in HCI research relating to Explainable and Interpretable AI, typically aiming to use technical analysis to describe the processes through which machine learning (ML) models derive their output [9, 24, 28, 34, 73]. Benjamin and colleagues argue that ML systems often work behind our perceptual horizon and thus textures how we experience the world [11]. Considering the emerging field of DGMs we find it relevant to bring these systems into our immediate perception so that we can see and experience the work that they do. Not through technical analysis but through aesthetic experience. Aesthetic production is after all the promise of many DGMs.

Benjamin and colleagues [11] also introduce the concept of *pattern leakage*, which offers a way to understand the aesthetic quality and potential of DGMs: Through its output a DGM "leaks" its embedded patterns, which allows us to create designs affording an experiential evaluation from a continuum of perspectives. This calls for a deliberate practice of designing interfaces that support a *hermeneutic relation* with those embedded patterns by exploring and building an understanding of its inherent values and biases.

To explore and evaluate how one might design for this *hermeneutic relation* we will present a Research-through-Design project [74] where we developed an interactive drawing table named *New Snow* allowing a user to explore a StyleGAN model trained on drawings by the Norwegian painter Edvard Munch. The project was carried out in collaboration with the museum MUNCH, using a selection of 5800 line drawings from the museum's digital collection to train a DGM. This model is used in combination with an interactive drawing table, where the user draws with a pen on paper to explore the latent space of the DGM. By tracking the paper in real time and projecting the synthesized image back on the paper, the drawing table supports a fluid iterative interaction between the user and the dataset.

The paper makes two novel contributions:

(1) It presents a system that allows users to interact with a DGM by drawing on paper (rather than the typical text prompts used by most current systems).
(2) We demonstrate how this facilitates a hermeneutic relation that establishes a unique perspective on Munch's drawings as a practice.

Taking a postphenomenological perspective we will analyze how the design mediates particular relations between the user, the design, and the underlying dataset. Through interviews with 20 participants, we present a detailed view of how this mediation plays out in practice and in relation to the specific dataset.

While the approach of presenting synthesized drawings instead of individual artworks might seem controversial in an art museum context, it hinges on the understanding of art mediation as aiming to evoke a directed interest and *educate our attention* [36, 52] towards qualities that might otherwise have been overlooked. In this view, the primary purpose of art mediation is not to transfer knowledge into the head of the user but rather to reveal and magnify qualities and relations that the user may experience that in turn shape their subsequent encounters with art. As the project begins with an interest in a large corpus of works of which many are sketches and unfinished drawings it is relevant to investigate them as a practice rather than as individual works. While DGMs have become popular for their ability to create convincing singular images, this project focuses on the ability of a DGM to produce a smooth latent space through which we can navigate through multi-dimensional representations of visual trends in the dataset. We will end by discussing how the overlap between technological mediation and art mediation shaped the orientation of the research and how the idea of designing for a *hermeneutic relation* to a generative AI system may have relevance in other domains where DGMs are employed.

2 RELATED WORK

Recently DGMs such as StyleGAN [37, 38], Dall-E [55], Midjourney and Stable Diffusion [14], which are capable of generating a wide range of images from verbal prompts, have become popular. Such models have been applied by artists to create new forms of imagery and challenge the relation between artists and technology [2, 3, 8, 15, 16, 32, 33, 75]. Furthermore, these technologies have also entered popular discourse and have been used by amateurs in a wide variety of image-making practices [54].

However, the introduction of these technologies into the mainstream has also sparked discussions about the practices regarding the collection of data that the models are trained on. One critical issue regards the possible infringement of artists' intellectual rights to their own works and style, as they have been included in the training data [5, 17]. Another significant question regards the risk of propagating gender, racial and other stereotypes [7, 13, 20, 23, 43, 67]. In both cases, the opaqueness of the data collection practice as well as the black-boxed nature of the generative models makes them difficult to scrutinize.

Research around *AI as design material* tends to focus on how AI can be put to work solving problems, however, it is also acknowledged to be particularly difficult to work with. Yang et al. [70] argue that two central attributes that make it difficult to design with AI are *capability uncertainty* and *output complexity*. *Capability uncertainty* relates to the difficulty of knowing about the capabilities of an AI system to perform a given task before it is built and the data specific to that task have been sourced. AI systems that develop as they are being used and systems that act in response to contextual factors make this problem even harder. *Output complexity* points to the situation where there are many potential outputs from a system. With DGMs we often talk about many-dimensional variations of the output. Furthermore, Leahu argues that we should be aware of *ontological surprises* when working with machine learning, as particular relations and categories may arise that we did not foresee as a consequence of the specific configurations of technology, humans, and context[45].

2.1 Databases and machine learning in museums

The use of technology in museums has long been an active topic of research in HCI, to the extent that museums are seen as "a great testbed" for trialing novel interactive technologies [35, p. xv]. Technology is used for a range of different purposes in museums, including as a means of digitalizing and archiving collections, as well as a means of communicating, educating, and facilitating experiences for visitors.

In addition to collecting and conserving cultural heritage, an important mission of museums is their ability to exhibit, communicate, and involve the public in our shared cultural heritage. This requires experimenting with how technology can be used to design experiences that are simultaneously engaging, educational, and inspirational. One often cited challenge for museums is the fact that most museums have vast archives of artworks and artifacts that greatly exceed their capacity to exhibit; it is common to estimate that for European museums around 90% of their artifacts are permanently in storage and never exhibited to the public. The Danish National Gallery only exhibits 0.7% of its collection at a given time [63]. Many museum professionals are thus eager to find ways to use digital technology to make the digital versions of these vast collections available to the public.

Many museums offer the public the ability to search their database through a conventional web-based interface with a text query and different options for filtering and sorting on pre-defined parameters. Earlier research has argued for the need for embodied visualization paradigms for large heterogeneous cultural datasets and has proposed immersive, interactive presentations to support navigating collections of thousands of cultural data objects [39, 40, 47, 60]. Some projects create complex spatial visualizations of the data objects to highlight aspects of their individual relations [4, 19, 31]. This approach is in stark contrast to the black-boxed generative models that collapse the data objects into a smooth latent space.

If digital collections should help museums achieve their goals of offering relevant experiences for their audiences, mediation tools are needed to support audiences in developing relevant perspectives. The recent technical developments in image synthesis with DGMs call for an investigation into how such models can participate in mediating relevant perspectives on the collections.

AI technologies have been deployed in museums in a variety of contexts, and the implications of these technologies for museums have caused much debate [10, 27, 65, 66]. While these technologies inspire hope that they can contribute to making collections more searchable and accessible [61], there are also concerns that AI algorithms may perpetuate cultural biases and deal with sensitive issues in a problematic manner, as well as other legal and ethical concerns [18, 26].

However, museums with large digital collections are in a good position to train their own bespoke models, avoiding many of the issues that muddle the ethical implications of the more generalized datasets, because they have ownership or rights of use for large amounts of data that they can correctly attribute. Furthermore, museums often employ domain experts, with deep knowledge about the subjects and historical context of the data objects. With a purposeful data curation practice, the museum is able to control what data goes into the model and potentially put it into play in new and exciting ways.

2.2 Drawing interfaces and machine learning

The most common way of making DGMs synthesize images is through verbal prompts. It requires the user to formulate in words what they want to see, to which the model will then respond with related visual concepts. This translation from words to images is constrained by the *imageability* [48] of the verbal concept. However, certain encoders allow us to stay within the visual domain, where images are used as prompts for other images. This sidesteps the issue of having to bridge the gap between verbal and visual expressions and allows for greater freedom in designing interfaces with attention to the relation between humans, the ML system, and the underlying data. Interfaces that rely on drawing as input to the system have been explored within HCI research [6, 21, 22, 44, 72]. The *Reframer* project by Lawton and colleagues share many similarities with the present project, as it also utilizes a DGM for an interactive real-time drawing experience [44]. However, the purpose of *Reframer* is creativity support and does not attempt to establish connections to a historical practice of art. The projects *Draw to Art* [30] and *Draw to Art: Shape Edition* [29], allow users to make drawings in order to search a large database of artworks. In the first version, the drawing is interpreted as a word, and the system returns artworks related to that word. In the second version, the user draws with simple geometric shapes and the search returns artworks with a similar composition.

2.3 Technology and art mediation

In this paper, we work with DGMs in the context of the art museum, more specifically within the topic of *art mediation*, which can be summarized as supporting art museum visitors' perceptual access to the artworks on display.

Ingold presents *education of attention* in a criticism of a prevailing idea of learning as the transmission of information [36]. Noë talks about a similar relation to art in his book "Learning to Look" where he furthers the point that we need examples in the form of pictures, text, theories, physical instruction, etc. that help us understand where to turn our attention and what to see [52]. Noë presents an example of repair manuals for cars. One car came with a manual with photographs of the car's internals while another manual for another car used line drawings. He argues that the photographs did not manage to pick out what was important, while the drawings were more articulate, bringing your attention to what matters, for the particular purpose of a repair manual [52, p.65]. While this example is based on visual attention, the concept for both Ingold and Noë goes beyond the visual domain and involves all our perceptual capabilities. In this enactivist perspective, learning means becoming attentive to particular features of the environment that are important for solving a given task, such as making sense of an artwork.

Sivertsen and colleagues argue that when using technology for mediation purposes in the art museum, this may constitute an "art critique by other means" [62]. Through this lens, *art mediation* and the postphenomenological concept of *technological mediation* become two sides of the same coin. The purpose of the art critique in this view is to draw the audience into correspondence with the art, rather than transmitting information about it. The art critique does not depend on the original artworks being present, even though engagement with the original that the critique concerns add to the ongoing correspondence. So by training a DGM on carefully curated images, and establishing a *functional perspective* from which the user can engage with it, we might be able to educate the attention of the user to aspects of the images that enter into corresponding with Munch's art.

2.4 Postphenomenology and ML

The term *functional perspective* comes from postphenomenological theory and describes the perspective on the world, that is facilitated by the technology, the physical and social context, and the user's personal context. In this view, technology mediates the world and makes it *legible* in different ways [41]. When working with DGMs we must, as with other technologies, think about how we would like it to make the world legible to us and the people we design for.

Kiran argues that technological mediation can be understood as *revealing* and *concealing* aspects of the world along different dimensions [41, 57]. He presents the *ontological*, the *epistemological*, the *practical*, and the *ethical* dimensions as four that are relevant to consider[41]. The *epistemological* dimension is particularly relevant for the topic of interaction with DGMs because it allows us to consider how the technology employed *magnifies* and *reduces* qualities of the underlying dataset. Through the technological mediation, the generated images *manifest* themselves both in relation to the material properties of the concrete technology as well as in relation to the *task at hand* for the user of the system. This means that the user, the system itself, and the artworks are mutually constrained in facilitating a particular *perspective* on the synthesized output. This perspective is enacted through perceptual actions that, as Scurto et al. point out, are significant for how users of a machine learning system are able to project themselves into it using their body and perception [58].

Postphenomenological literature presents us with several possible relations between technology, humans, and the world that Benjamin and colleagues apply specifically to the relations between humans and ML [11]. In this project, we are interested in establishing a hermeneutic relation to the digitalized drawings mediated by the drawing table. The *hermeneutic relation* is understood as analogous to reading a text. Different experience and skill in reading gives access to the meaning of the text in different ways. The skillful reader almost sees through the letters and perceives the meaning in one quick glance, while a novice reader might start by constructing words from letters, sentences from words, and so on. This hermeneutic relation is also seen with technologies such as maps and thermometers as they make the world legible by presenting the layout of the world through grids or heat phenomena through a scale [57]. Similarly, the DGM has a way of providing a particular perspective on the training data through its output, and we would like the users to gain experience and skill as they interact with it, becoming increasingly better "readers" of Munch's practice.

Another relation that becomes relevant is the *alterity relation*. It describes situations where the interactions between humans and technology are somewhat similar to that of two humans interacting, not necessarily meaning the user is fooled to believe they are interacting with another human, but simply that the technology is seen as acting with a human-like intention, making it a *quasi-other* [57]. In this project, we are interested in the machine acting in the image of Munch, rather than following the whims of the users.

Benjamin and colleagues also present the concept of *pattern leakage*: That is, "the propensity of probabilistic patterns to shape the world they are deployed to represent" [11, p.11]. For instance, a surveillance system classifying events might affect how humans see the world. Generative algorithms on the other hand are designed to produce images and texts as outputs, and as such are designed to contribute to shaping the world through their outputs. We propose that the concept of pattern leakage might be understood as a core quality of generative models. Through their output, they make explicit their inherent patterns and enable us to learn about them, which we can make use of, as we shall discuss further later on.

2.5 Designing for the hermeneutic relation

Deep generative models are capable of producing aesthetically rich images, sounds, and text. Applications often focus on how they can be used in a tool-like capacity to produce media of higher quality or more efficiently. However given the opaque data collection and training practices most often employed the user will have very limited means for understanding its embedded patterns, biases or *intentionality*.

Benjamin and colleagues argue that a typical ML-system establish the following relationality [11, p. 4]:

```
Human - Technology / (Model -> [World]) - World
```

The human is in immediate relation to the technology (Shown with -). Hidden in the background (/) is an embedded model's interpretation (->) of the world that has been *thematized* as data in the model (*[]*).

However, in the present project, we argue that the model's interpretation of the world, which in this case is digitalized drawings, could instead be brought into the immediate perception in the following way:

```
(Human - Technology) -> (Model -> [World])
```

Through designing for a *hermeneutic relation* to the ML model, we can interpret (->) how the model interprets (->) the world through as it has been *thematized* (*[]*) in the data curation for the model. To gain experiential access to the model's behavior we establish an *embodiment relation* (*()*) between the human and the interface, making it a partly transparent extension of the user.

Similar ideas of exposing various ML systems' interpretation of the world are seen in artistic engagements with ML and AI. With *ImageNet Roulette* Crawford & Paglen highlighted the problematic "person" category in the ImageNet dataset by allowing everyone to upload their own image to be classified with the categories from the "person" synset. The system's application of strange, discriminating, and outright offensive labels to the images made the model's interpretation of humans dramatically apparent through its concrete use. In Memo Akten's *Learning to See: Interactive* [1], the audience in the exhibition was able to interrogate five GAN models by showing everyday objects to a video camera. The system would interpret the video feed through one of five models trained on images of water, fire, earth, air, and the cosmos. By manipulating and creating compositions of everyday objects, the audience was able to investigate the patterns and aesthetic qualities of each model. The service LAIKA [42] aims to support creative writing by letting writers interrogate and explore the qualities of different language models. The models can be trained on the work of a famous artist, or a corpus of the user's own text. The proposal is not that the system will generate finished text but rather that it will spark inspiration and reflection on writing style and patterns stimulating the user to write better or more creatively.

In these three examples, the interface provided to the user serves the purpose of giving them access to explore and interrogate the models in question. In this way, the interface supports a use that is hermeneutic with regards to the model, in that it allows for exploring and developing an understanding of its intentionality towards the world. The modalities of the three interfaces - image upload, a video feed, and text prompts - are very different, and facilitate particular perspectives on the models and their mediation of the underlying data. In all three cases, it is the aesthetics of each model that constitute the work it does in the world.

Wolf proposed the term "Explorable AI" [68] arguing for designing AI systems to "to support and empower actors to scrutinize, uncover, and make sense of a variety of dimensions along the broader AI lifecycle" [68, p.15]. In comparison, our focus is on the trained generative model as it concretely mediates a specific situated use.

With the project *Entoptic Field Camera* Benjamin and colleagues presented the *entoptic metaphor* as a way to describe how machine learning systems can give rise to visual phenomena in a way similar to how the human visual system can produce phenomena such as floaters or hallucinations within its system [12]. The introduction of the entoptic metaphor is intended to support designerly inquiry into the materiality of machine-learning systems and the concepts and implications that emerge as a part of a situated investigation. As Benjamin and colleagues write, this does not necessarily mean bringing something hidden to light, but rather producing a multiplicity of perspectives that links particular concerns with technical aspects of ML. They further argue that it moves the orientation from what AI technologies *are* to what they *do*. Benjamin and colleagues relate their approach to *reflective design* [59] which is an approach that aims at enabling designers and users to reflect on values and metaphors embedded in designs. This concerns design processes in a broad sense, however, the interest of this project is more narrowly focused on the designers' and users' interpretation of the intentionality of the particular Munch-oriented DGM.

In this project, we are not interested in making people criticize or unpack DGMs in a general sense. Rather we are interested in letting people interpret how the DGM interprets Munch's drawings. Like a map of a city, the DGM makes an expansive phenomenon, Munch's corpus of drawings, legible through a perspective that emerges from statistics. Through this interpretation, certain aspects of the data will be *revealed* and *magnified*.

A drawing interface provides a way for the user to explore this. The drawing interface should facilitate an *embodiment relation* letting the user act *through* the drawing interface with the DGM. The drawing interface should thus offer a level of transparency that lets the user focus on the output of the DGM rather than the pen and paper itself. The relationship we intend to facilitate can be written up in the following way:

```
(Human - Drawing table) -> (Model -> [Artworks])
```

In the next section, we will present the design of the drawing table and DGM and how it was developed.

3 METHOD & DESIGN PROCESS

This project has been developed following Zimmerman's description of Research through Design [74]. This implies that we present a design process that leads to an invention. We show how we find this invention relevant in addressing a particular situation in the art museum and evaluate our invention in its ability to bring us to a preferred new situation. Finally, we show how the learning from this project can be applied to other design research projects that involve DGMs as part of the design material.

The project has been developed in a number of concurrent trajectories. These will be elaborated below.

(1) Concept development
(2) Data collection
(3) Model training
(4) Table design
(5) Evaluation

Throughout the design process, the design underwent informal evaluations to assess different aspects of the concept and the technical design. These evaluations were used to drive the design forward toward the intended qualities. Finally, the project underwent a summative evaluation with 20 participants who were interviewed about their experience with the system. This will be described in detail later.

3.1 Concept development

The concept was developed in the context of a research collaboration between MUNCH and the IT University of Copenhagen. The goal of this collaboration was to investigate ways of using technology for novel ways of doing mediation in the art museum context. The idea emerged from an interest in activating the digital collection of MUNCH, more specifically the paper-based works. MUNCH is in possession of a large number of drawings, notebooks, and diaries from the hand of Norwegian artist Edvard Munch. Due to the fragility of paper, these are particularly difficult to exhibit, as they are very sensitive to light. The amount of drawings in the paper collection is counted in the thousands, and might not be great works of art in themselves, but nevertheless an interesting entry point into the artistic practice of Munch.

Through an interview with the paper curators at the museum, the design team identified some important qualities of the paper collection. Firstly, Munch was very active in drawing and sketching the world around him. His many drawings range from early sketches of paintings to architectural sketches of his studio, satirical drawings of neighbors and their pets, drawings of everyday scenes in Norway, Norwegian nature, as well as many portraits and studies of models. This apparent interest and involvement with the world around him run counter to a myth that he was a hermit mainly producing somber paintings with dark emotional content. Secondly, the curators spoke of an intimate physical relation to the drawings. Due to the fragility of the drawings, only very few people are allowed to handle them and get the chance to develop this relation. Finally, the approach to artistic process in the museum's department of learning emphasizes process over outcome, reflecting Munch's high productivity (not least in sketches and drawings) and tendency to revisit the same motives throughout his life.

From these points, the first author along with collaborators from the museum developed the idea to let museum visitors explore the vast drawing collection through their own physical engagement with the drawings. Rather than physically handling the drawings, we would let the visitors explore them interactively through drawing. By leveraging the capability of computer vision and DGMs we would let this play out on actual paper, with the synthetic drawings changing in response to the user's drawing actions. The goal of the interaction should be to explore Munch's practice rather than producing new images, and this is where the need to establish a hermeneutic relation comes from.

In addition to the direct interaction with the drawing table, we also considered which atmosphere should surround this drawing activity. To support the idea of the drawings being part of Munch's everyday drawing practice, and standing in contrast to the more emotional tone of some of his other works, we wanted an atmosphere of calmness and serenity. As part of that development, we created a simple soundscape to accompany the drawing experience. The *functional perspective* that the system provides on the drawing also emerges in relation to various contextual factors, such as atmosphere and introduction given to the system.

This soundscape was intended to invite visitors to relax and take their time as they engaged in drawing dialogue through the materially relevant interface. Through their drawing actions, the user's attention would be educated to the images as drawn, as a practice or process, and with visual qualities as magnified by the DGM.

3.2 Data collection

To train a DGM we needed to collect and curate a dataset suitable for training. First, we queried the MUNCH digital archive for all images where the medium was listed as crayon, pencil, ink, or coal. This resulted in approximately 7600 images. The images in the MUNCH

Figure 1: The upper half of the image shows samples from the original sketch data. The lower half shows random synthetic samples from the StyleGAN model

digital archive consist of photographs of the original notebooks and loose paper sheets where the drawings appear. This means that the images also contain table surfaces, paper edges, cataloging labels, torn paper, dirt, and text. To avoid training a model that would synthesize full notebooks, we decided to add bounding boxes to each drawing in an image and extract only the marked areas for training. Furthermore, we distinguished between 4 different kinds of drawings:

- Line drawings with no shading
- Line drawings with some shading
- Heavily shaded drawings
- Colored drawings

The annotation was completed by members of the design team, including the first author and two student assistants. To maintain some stylistic consistency in the model we decided to use the drawings from the first two categories when training the model. As the user of the system uses a black pen we wanted drawings with a matching visual quality. This excluded the colored drawings and those that were drawn with a shading technique rather than through clear lines. These two chosen categories with line drawings cover approximately 5800 drawings.

3.3 Model training

The architecture chosen was a StyleGAN 2 ADA [37] model that was trained on the 5800 selected drawings. When training, the model develops a mathematical space called the *latent space*. This space follows the distribution of images in the dataset, with respect to their visual qualities. This latent space is smooth meaning that it is possible to interpolate between points in the latent space. For each step, the model will create an image coherent in its own right, while morphing gradually between the start and end points.

This architecture requires that all input images must have a 1:1 aspect ratio. This was achieved by stretching the images into shape. While this creates heavy distortion in some images, the effect after training is not very pronounced.

To be able to synthesize images quickly during runtime we have trained the model at 256x256 pixels. After training, the StyleGAN 2 ADA model is capable of synthesizing images that adhere to the visual trends of Munch's drawings in some aspects.

To allow for drawings to be used as input for the model, we trained a pixel2style2pixel (pSp) [56] encoder capable of taking an input drawing and returning a synthetic drawing from the latent space of the StyleGAN model.

The pSp model was trained by using 10000 random synthetic drawings from the StyleGAN model. These 10000 drawings were then processed by a sketchification model, that simplifies the drawings into binary line drawings with a minimal amount of detailing. Then the pSp encoder was trained on the synthetic images and the simplified images in order to learn the mapping between binary input images and images from the latent space of the model.

3.4 Table design

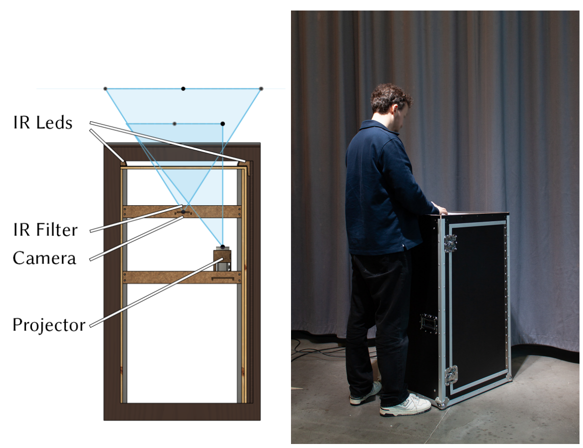

Figure 2: The prototype houses a pico projector and a camera that are aligned with the drawing surface. An infrared filter removes the visible light from the camera input. Right below the top are two strips of infrared LED that illuminate the drawing surface. The lower part of the table can house a PC for processing the images.

The table design was developed to support the use of pen and paper as the input medium. Through testing, it was found that tracing paper and pigment markers provided the best tracking conditions. This selection naturally constitutes a trade-off between getting close to the tools Munch used for his drawings and something that we could track consistently. The table top surface is semi-transparent to allow for recording the paper from below. Two rows of infra-red LEDs light up the tracing paper from beneath, and a camera sensitive to the 850nm wavelength captures the drawn lines from behind a filter blocking visible light (see fig. 2).

The video feed is then processed using TouchDesigner and OpenCV to create a binary input image for the image inference. The input image is then sent to the pSp model, which returns a synthetic image less than a second later. The generated image is then composited with a subtle overlay of the input image and projected back on the tracing paper (fig. 3). Each time a new synthetic image is created the projected image fluidly fades from the previous to the new image. This happens continuously several times per second. The image processing is handled by a PC with a GTX 1070 GPU located inside the table.

While the user draws, the soundscape is played back by a speaker at low volume. The sound is generated live by a Max 8 patch, using three atmospheric digital synths. They each play a slow succession of notes of varying lengths. For each repetition, the relative timing of the notes between the synths shifts leading to a slow atmospheric melody that does not repeat itself.

3.5 Interaction

To use the system, the user picks up a piece of tracing paper and a pen and places it on the drawing surface. In the beginning, a faint pattern is visible on the tracing paper. As soon as the user starts drawing the system starts adapting the generated image to the drawing (fig. 4). The user can choose to draw on top of the lines presented by the system or place their own lines. In any case, the model continuously responds to the lines that are currently on the paper. In addition to drawing more lines, the user also has the option to move the paper around on the drawing surface. As the user moves or rotates the drawing the system interprets it differently and returns new results. This allows the user to investigate a certain "visual space" by slowly moving the paper to gradually see how the input drawing is interpreted differently in different areas of the drawing surface.

Despite being trained on a corpus of drawings that contains a wide variety of subjects, the resulting model has a strong affinity for faces, making them by far the easiest type of subject to evoke. Portraits of various kinds do make up a significant portion of the dataset, and share a visual structure. Through the development of the system, it became evident to the authors that finding the salient lines for various types of faces is much easier than for other types of subjects such as standing figures, animals, or interiors.

4 EVALUATION

The final design was tested on 20 participants recruited at the IT University of Copenhagen during May and June 2023. Each session began by introducing the participant to what was going to happen and asking the participant demographic questions and questions on their familiarity with drawing, machine learning, and art in general. Each participant would then interact with the prototype for approximately 8-10 minutes. As discussed above, the *task at hand* is significant for shaping the *perspective* of the user, therefore all participants were given the same instructions before starting to draw:

> "This system is trained on the drawings of Edvard Munch. You should use the pen as your tool to explore what is hiding in the system. When you draw, the system will attempt to interpret your line as the beginning of an Edvard Munch drawing. The system only knows Edvard Munch's motives and way of drawing, so it will try to lead you into drawing like Edvard Munch. You can draw on the paper, you can move the paper around, and you can have as much paper as you want. To get off to a good start, I suggest that you start by drawing the beginning of a head or a face."

This instruction attempts to shape user interaction in a number of ways. It casts the pen as a tool for exploration rather than self-expression. It emphasizes a narrow focus on Edvard Munch, to tame expectations that the system would have the same capabilities as Stable Diffusion [14] or Dall-E [55]. It indicates that the system has an agency to lead the user. Finally, it suggests that the participants start by drawing a face or head. Due to the tendency of the model to infer faces, this was said to make sure that the participants would quickly get into a dialogue with the model. The participants were allowed to ask clarifying questions while drawing and get as much paper as they wanted. The participants drew between 2 to 6 drawings each. While drawing a video recording was made of the participant's hands.

After drawing, an interview was conducted about their experience that lasted approximately 12-15 minutes. The interview followed the procedure of evocation and explicitation described by Ann Light [46]. First, the interview contract was established letting the participant know that the interviewer was interested in an *account of* their experience, rather than reflections and motivations for why they acted or thought in a certain way. Next, the interviewer encouraged the participant to think about the moment when they took the cap of the marker to start drawing and start telling about their experience from that point onwards. Along the way, the interviewer provided prompts for the interviewer to elaborate on specific experiences, or to gently shift to situations that had seemed significant while the participant was drawing. The interviewer continously attempted to make the participants remain in a state of evocation, and questions were kept open to not lead the participants to judge or fabricate experiences.

Of the 20 participants 12 identified as male, 7 as female, and 1 as non-binary. The average age was 32.2 with a maximum age of 60 and a minimum of 11. The participants were mainly students and faculty, except two participants who were attending elementary school. These two young participants experienced the drawing table together and were interviewed together with a parent present throughout the whole session. 4 participants were graduate students in a computer science program, 3 were graduate students in a design program and the remaining 11 were researchers in the design department.

When asked about their interest in art, 12 stated that they were interested in art. 6 were "partly" interested and 2 did not have an interest in art. Within the last 12 months, the participants had visited art museums and galleries 4.8 times on average (min. 0, max 12).

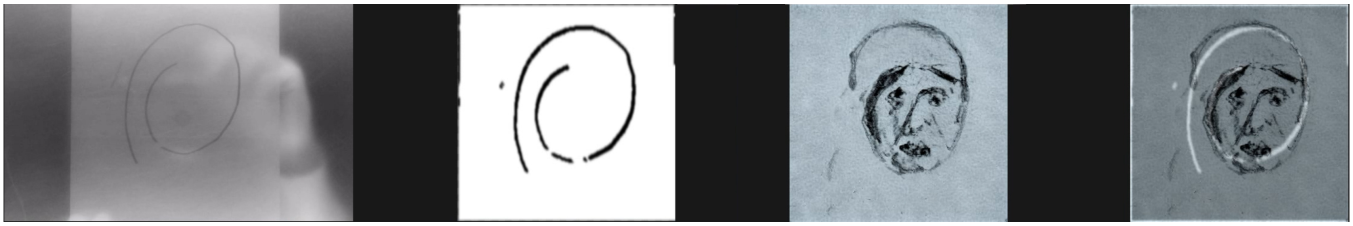

Figure 3: Inferring a Munch-like sketch from a hand-drawn line. From left to right: 1) infrared webcam image from beneath the drawing surface. 2) The cleaned input image. 3) The synthesized sketch. 4) A composite image of the input image over the synthesized sketch that is output as a projection on the tracing paper.

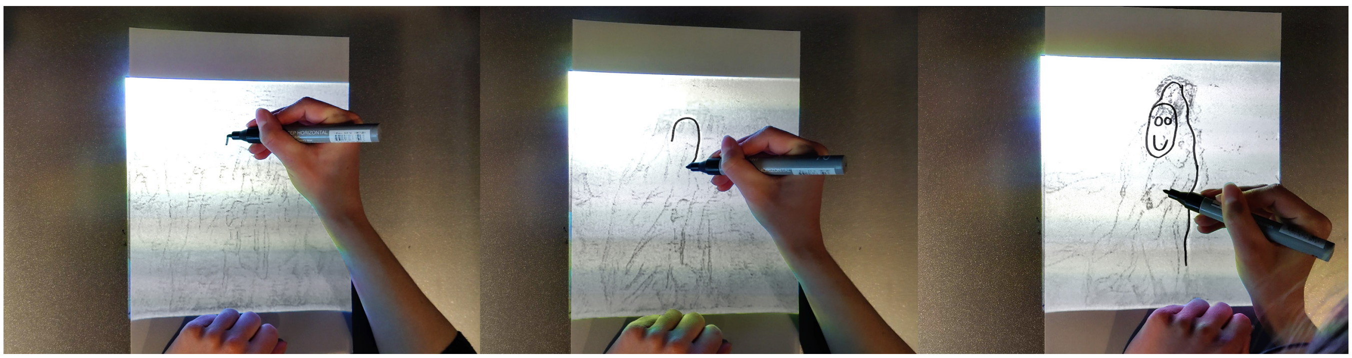

Figure 4: When starting out, the a faint pattern is visible on the tracing paper, but it quickly starts adapting to the participants' drawings. Due to the nature of the projector, the paper appears unevenly lit and with color bands in photographs. Through photo editing we have attempted to limit the effect as it is not visible to the naked eye.

On a scale of 1-5, the participants rated how experienced they felt in drawing at 2.6 on average and machine learning at 3.2. A few participants had machine learning as part of their research area, while none of the participants saw themselves as experts in drawing.

The analysis of the interviews was conducted by first making affinity clusters of statements to allow for emergent themes. Among the themes that emerged were particular drawing strategies and changes between them, descriptions of visual qualities of the model output, and aesthetic qualities of the experience in general. Next, the descriptions of drawing strategies were analyzed further to be able to identify shared patterns across participants. Similarly, statements in the two other overarching themes were analyzed to identify experiences that were shared across participants and those that were exceptional. The results will be presented in the following section.

5 RESULTS

In the following we will present some insights from the observations and interviews with participants, identifying three themes: The aesthetic experience of interaction with the drawing table, the drawing strategies employed by the participants, and the ways in which the system led participants to perceive something about Edvard Munch's drawing style. These three themes are brought into correspondence with the model of relationality presented earlier: the *embodiment relation* between participant and table, the *hermeneutic relation* to the model, and the model's interpretation of the dataset.

5.1 Aesthetic experience

As the drawing session started, the soft ambient soundtrack would start playing. The participants described the music as something that helped create a cozy atmosphere and loosened up the feeling of having to perform while being observed drawing. Many also said that it was "calming", "relaxing" or "meditative" and helped them focus on the drawing experience and get into a flow. Several reported being absorbed by the drawing process. A couple of the participants said that it was not unlike music they could imagine hearing in a museum exhibition.

The physical tools also contributed to the aesthetic of the drawing experience. The participants reported that the marker and paper felt good. The marker produced a solid black line, and the paper felt of high quality. However, several participants mentioned that there was a large discrepancy between the types of lines they were able to produce with the marker, and the quality of the lines produced by the model which were more fuzzy, shaded, and thinner. Several participants expressed a wish to try other drawing tools that would allow them to get closer to the expression of the fuzzy lines and shading in the projections.

These descriptions indicate that the interaction with the drawing table to a large degree was *transparent* in a phenomenological sense as attention was not drawn to the paper, pen, table, or music itself.

Rather these technologies seemed to support the participants in their engagement with the model through an *embodiment relation*, albeit with the limitation that the marker was not a perfect match for the quality of lines produced in the projected drawings.

5.2 Drawing strategies

Through the interviews, it became apparent that the drawing interface invited users to explore a range of approaches. We find that these approaches can be placed along two dimensions. The first dimension is whether the participant was proactive or reactive in relation to the system. Some participants would draw intuitively and expect the system to adapt to their input, while others would consider the current output first, before tracing or drawing in close relation to it. The second dimension is whether the participant expects meaningful drawings to emerge on the paper or in the projection. Some participants would consider the lines on paper the "final" result, while others saw the projection as the result. Over the course of their interactions, most participants changed their strategy multiple times, often starting out being reactive and gradually becoming more proactive as they became more familiar with the system. For some, this also meant becoming more interested in the system output rather than the physical drawing. Others insisted on the physical drawing being the key takeaway from the experience. Below we will highlight some notable strategies used by the participants.

As suggested by the instructions, participants would start by drawing a face or a head (see fig. 4). This typically developed in two different directions. First, for some participants the first lines they drew made the system respond with a shape resembling a face. On seeing this, many participants switched to tracing the lines of the projected face, drawing in eyes, hair, mouth, or other features. Often this became a collaboration between the participant and the system, negotiating which features to add. As the participant traced, the projected face might change in unexpected ways, rendering some of the earlier lines incoherent with the new image. To handle this, a few participants utilized a collage strategy where they traced only the lines that supported the creation of a coherent image on the paper. One participant even attempted to draw faster than the system could update the projected image in order to capture as many lines as possible before they changed (fig. 5).

Second, for some participants, if their first attempt at drawing a face was not sufficiently well-aligned with the model it resulted in a vague or ambiguous response from the model. This caused some participants to request a new paper and start over, while others would continue drawing and partially ignore the projections of the model.

After this first attempt, however, most participants developed some understanding of the model and changed the strategy for their next drawing. A few participants stayed with the idea of the system supporting their self-expression, leading them to try to derive what they could from the projected drawings to support them in creating good-looking drawings on the tracing paper. However, most participants seemed to put less emphasis on the physical drawings and focus more on exploring what they could make the projected drawings become. This varied between a *collaborating* strategy, where the participant switched back and forth between tracing over lines in the projection and adding new lines of their own; to a very deliberate *prompting* strategy where the participant only drew basic shapes in different sizes to explore what the system would make of it.

Since the task imagined by the design team was exactly this exploration of the model, it was positive to see the strategies converging toward this understanding as people became more familiar with the system. This change of strategy was, however, prompted by the tensions between the participants' intentions and the responses from the system.

The system is only capable of generating images that are *within* Munch's practice, as defined by the model. This means that when a participant draws something the system will only change its output to the extent that it reflects lines that are salient in the model and their location in the drawing area. This meant that those participants who tried to draw in their own visual style often only received vague and ambiguous feedback from the model. However, sometimes the system reacted with something coherent in an unexpected way as P14 experienced (fig. 6):

> "And so I was going after [drawing] an owl. I didn't get much guidance from the background [i.e. the projection] so I was just trying to create an owl by myself, and then all of a sudden this face appeared in the ear of the owl, which made me want to create something else."

Like P14, the participants would often change direction with their drawing when something exciting appeared in the projection. This also meant that they would sometimes be drawing on top of existing lines on their paper in an attempt to follow the whims of the model.

This type of exploration was further strengthened by the opportunity to move the paper around, as shifting the paper around on the drawing surface would cause the model to reinterpret the drawing, morphing between different types of faces or shifting into ambiguous shapes and lines. Many participants said that they had fun and found it pleasurable to move the paper around, exploring the different drawings the model could produce (fig. 7):

> "I think it's a fun and playful interaction, me turning and moving the paper and then something new is being drawn. It was just great to see how it morphs from one painting to the next by me moving the paper. I definitely think you could have fun with this for some time." (P5)

However, those invested in creating coherent drawings on the paper found the continuous change a bit chaotic.

This ephemeral nature of the projected drawings and the constant reinterpretations pushed some participants to completely drop the idea of the physical drawings being important in themselves. P11 explains it like this:

> "I started realizing that the sketch is less of a representation of the thing you're trying to produce and more of a kind of fiducial or visual key to something you're looking for, so I started moving the sketches around rotationally or positionally to see if I could explore, given a single starting image, to see what might be out there."

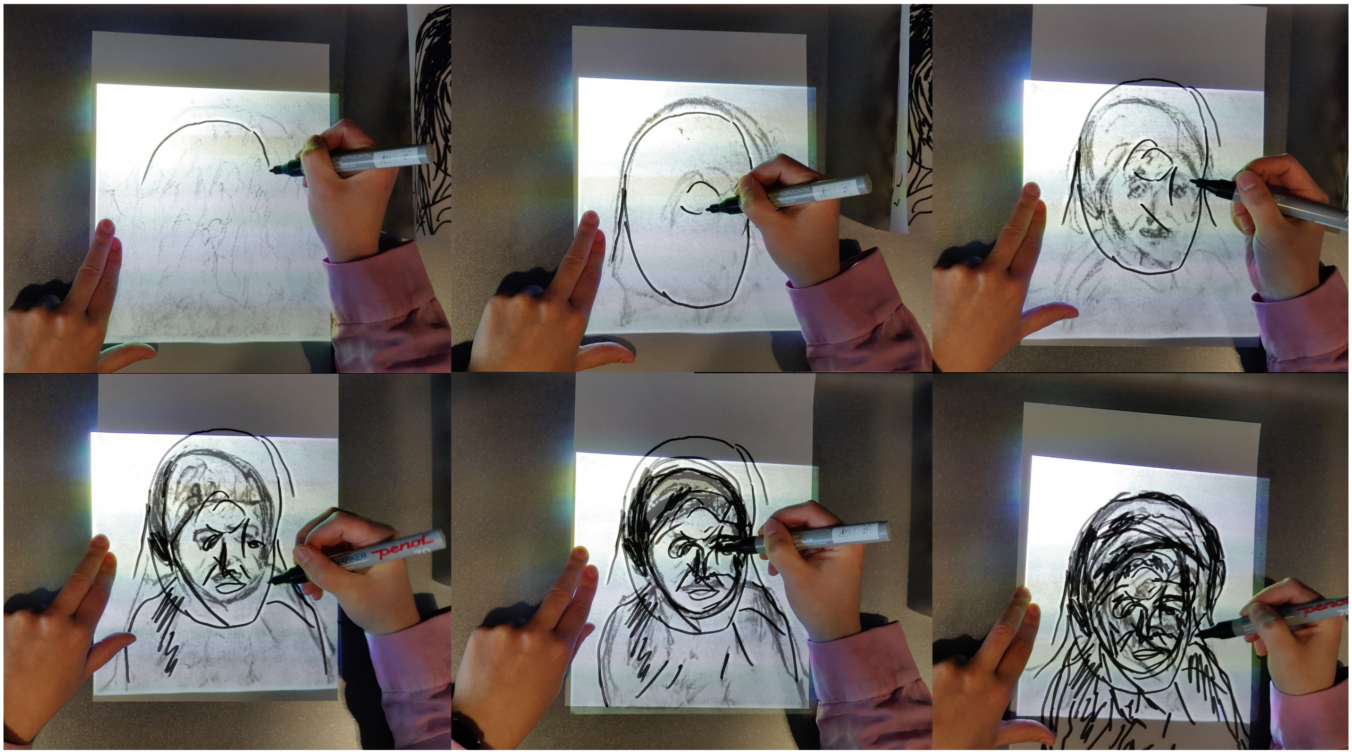

Figure 5: A few participants were mostly oriented towards creating coherent images on the paper. This participant sampled only "useful" lines from the projected drawings in order to end up with a drawing on the paper that was coherent in its own right.

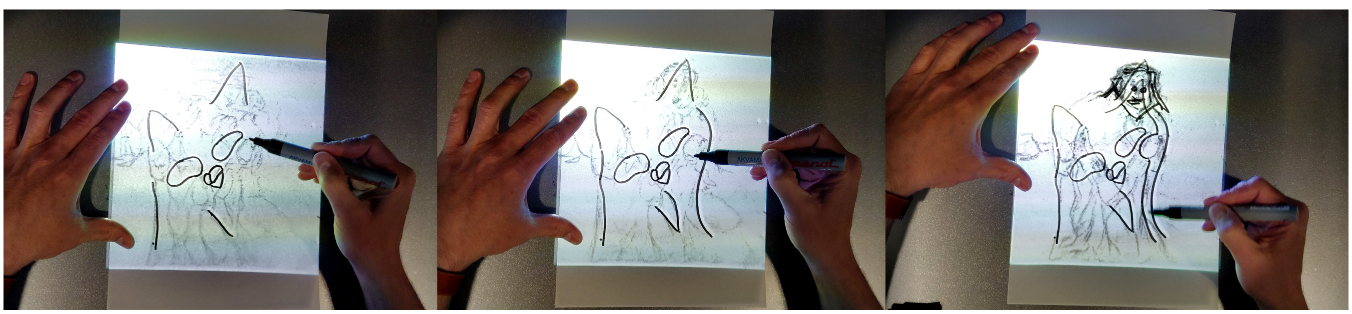

Figure 6: This participant starts by drawing an owl. However the system generates something that the participant interprets as a face, and changes plans with the drawing and adds a small face inside the ear of the owl

In these accounts of the participants' drawing experience, we see their drawing strategies as navigating between different relations to the system. Most participants eventually turn towards a hermeneutic exploration of the qualities of the model, as exemplified by the last quote. However, a few participants remain in a more tool-like relation, where the model is used to support their own drawing practice.

We also see aspects of alterity at play, when participants refer to the model's support of their drawings actions or lack thereof. The model embodies an intentionality derived from its interpretation of Munch's drawings.

5.3 Munch's style as recreated by StyleGAN

Through the drawing engagement with the projected drawings, the participants became attuned to this intentionality and the resulting visual output. The participants described the subjects of the drawings, the compositions, the material quality of the lines, the facial expressions of the faces, and other emotional qualities of the drawings. Some participants described how they understood the aesthetic qualities as a totality: "It's interesting because you enter a universe of these drawings" (P12). Another described it:

> "It was a very interesting way to experience the art, instead of the very static image in the museum, where

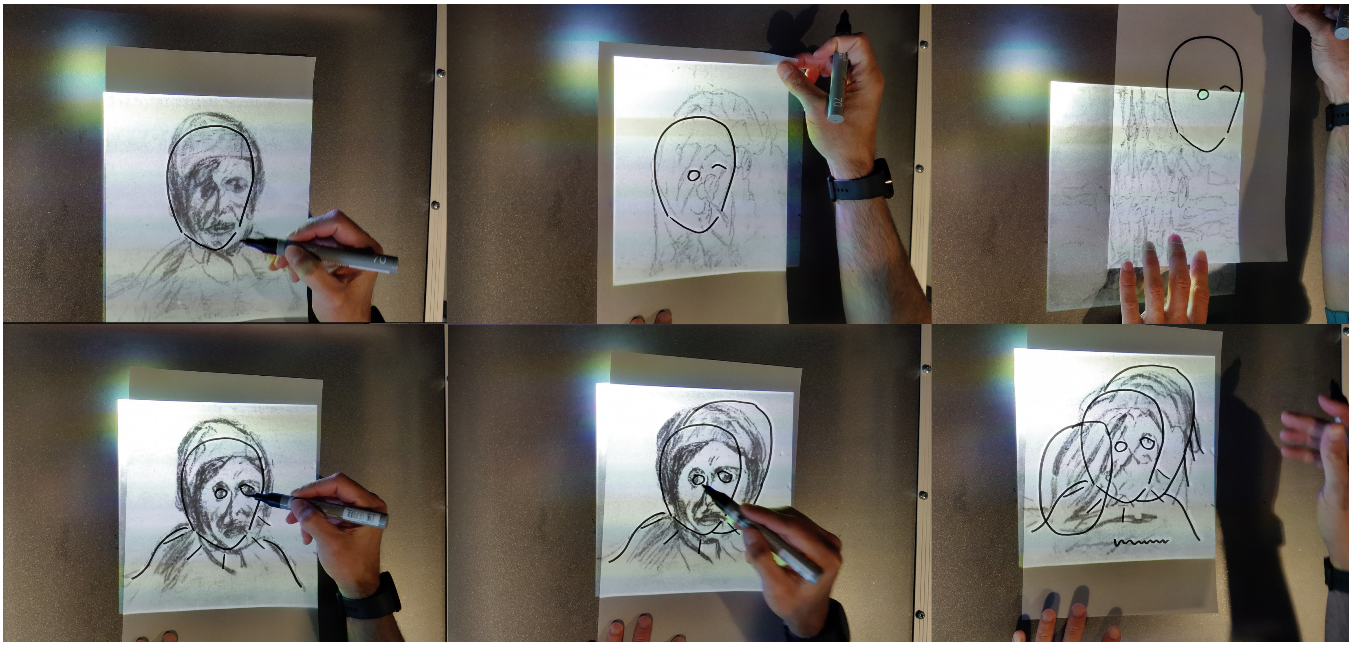

Figure 7: This participant quickly makes the model produce a face. Then the participants experiments with moving the drawing around but decides on the original position. Then the participant adds a line over the hair which makes the face adapts its shape. Next the participant adds a similar line on the other side of the face, which results in the face disappearing

> you can also look at the lines and collect from the different images. It gave a different feeling of the artist when you got to interact with it yourself. There were a lot of things changing but you could still clearly see that it was from the same artists in the same style" (P6)

Participants had different perceptions of the motifs they could recognize in the projections from the system. Some felt that they could identify several motifs and styles: "To me it seems that there are maybe two, three, or four genres. In my head, there are now the turbaned faces, maybe people standing full-body with flowing robes or dresses or cloaks, hats and ponytails, and then perhaps landscapes. I don't know whether Munch enjoyed drawing craggy cliff sides" (P11). No other participants talked about turbans and ponytails, but faces, standing figures and nature such as mountains, cliffs, and trees were common across participants.

Due to the instructions and the model's affinity for faces (fig. 3, 5, 7, 6), almost all participants talked about these, and also in greater detail: "He has this kind of like head where there's usually a hat or something. There's a line connected from the eyes with the nose. One line. It [also] seems like there's a lot of lines, but it's never clean lines." (P13). P19 said about the style of the faces that they were drawn with "maybe not a simple line, but a characteristic line in all the drawings, that had this dark melancholic facial expression. People looking maybe a bit anxious or sad and who had many details, while not having a lot of details". P9 noticed that "the faces had a pained look. It was very dark [...] it was a very strained line in a way. [...] a dark space visually." In these comments we see that the participants noticed particular aesthetic qualities in the lines, and some even experienced a distinct emotional quality in the drawings.

Despite the ephemeral nature of the drawings and their constant instability and tendency to become abstract and ambiguous, we see examples of participants experiencing aesthetic qualities in the drawing technique, the typical subjects, and the drawings' emotional quality. In some cases, however, the ambiguity also led the participants to see things that might not be part of the dataset. While reported by one of the participants, ponytails are not commonly occurring in the experience of the designers. The craggy cliff sides, that several participants refer to, seem to be related to a specific texture that might be an artifact of the training process more than specific drawings (see fig. 4). Like expected, through observing the participants, the first author also noted that many types of drawings did not emerge during the participants' interaction even though we know that they are part of the dataset.

6 DISCUSSION

Through the interviews we see that *New Snow* foregrounds patterns of the DGMs, bringing it to the immediate attention of the participants, rather than being hidden behind our perceptual horizon like Benjamin and colleagues find that they often are [11].

Bringing the model qualities forward opens up a different way of utilizing the qualities of DGMs. It sidesteps some of the problems of black boxing. While it is by no means obvious how the system works, it is at least immediately apparent what it *does*. The narrow selection of images, which is explained in the task, directs the experience towards a small part of Munch's artistic practice. Since all images are sourced from the museum's digital collection, the project avoids uncertainty about the origin of the data. The drawing interface specifically supports the exploration of drawings as a practice of putting pen to paper, rather than through verbal concepts.

6.1 Embodiment and transparency

Three design choices were important in establishing an *embodiment relation* with the drawing interface itself, and establishing the generated drawing as, in fact, *drawn*.

The most important aspect was choosing the right means of prompting the model to generate an output. The pSp encoder enables the mapping from black and white input images to the latent space of the StyleGAN model. This enables an interrogation of the model through the flow of lines, composition, drawing density, location, and scale. The pSp encoder also supports other kinds of image-to-image translation, such as inpainting and generation from segmentation maps. While still in the image-to-image domain, these translations enable very different relations to the images. This is again very different from a system prompted by a text interface, as this would make it much more difficult to express compositions and the quality of the lines, while making it easier to prompt verbal concepts, such as tree, face, and mountain. In short, how this particular use of the encoder enabled navigation in the latent space, was defining for the dimensions through which the model could be interrogated.

The marker and paper interface enabled continuous and slight adjustments to the drawings. The loose paper allowed the participants to slide the drawing around to gradually explore the latent space along two dimensions, while the marker let the participants gradually add to the drawing in response to the system. This gradual change made it possible to uncover the internal relations in the system between the dimensions that the interface afforded. The biggest limitation in the *New Snow* interface in this regard, was that the nature of the pigment marker constrains these adjustments to be additive, as the user must start over on a new paper if they want to remove a line from their drawing.

For *New Snow*, we intentionally picked the StyleGAN architecture for its relative speed in synthesizing images. We also ran the model at 256x256 pixels, for the same reason. The ability to explore was supported by it updating quickly, as it allowed the participants to explore continuously without pausing to wait for the generation to happen. Different architectures have vastly different response times but for *New Snow*, during the development process we found that the fast response was important for creating transparency in the embodiment relation to lead attention past the interface to the behavior of the DGM.

Together the prompting modality, the ability to gradually adjust the prompting, and the fast update speed facilitated a transparency through which the participants could experience the model. We saw, however, situations where this transparency faded. Participants noted a discrepancy between the quality of the lines they could make and what the system generated. Also the many lines drawn on the tracing paper would sometimes get in the way of seeing the output of the model. This very literally made the paper less transparent as well as the experiential access to the model.

6.2 Hermeneutic exploration of a DGM

The *New Snow* drawing table embodies a statistical representation of Edvard Munch's drawings and sketches. While the dataset is made up of distinct originals, the smooth nature of the latent space of the StyleGAN model blends the individual motifs and allows for seamless interpolation between them. The model does not recreate individual drawings from the dataset precisely, even though it might sometimes get close. Each generated frame is an *uncertain entity*. Each synthetic image is fleeting and ephemeral and never solidifies due to the slight noise in the camera feed.

The system relies on a statistical approximation of patterns in Munch's drawings and it is important that this quality is communicated to the users in order to avoid the system being seen as an authoritative representation of Munch's art - as one might expect to meet in a museum. Benjamin and colleagues encourage designers "to not see ML uncertainty as 'to be explained away,' but rather as generative of particular relations that can be designed for."[11, p. 2]. Through the design of *New Snow*, we have attempted to do exactly that by making a system that keeps its users close to the uncertainty by never committing to or concluding what a drawing is.

Fortunately, what happens is that the drawings are ontologically *revealed* as ephemeral, fluid, and malleable, and through the instructions, as being from Edvard Munch's practice. On the other hand, the drawings as discrete physical objects are *concealed*. Through the practical act of drawing, the *drawn* quality of the synthetic images is *magnified*. So are the dynamics and movement of the lines, as interpreted by the StyleGAN model. The paper quality, relative scale, and fine details in the drawings are *reduced*. This particular perspective on the drawings is not arbitrary, but a specific *functional perspective*, which is the only way we are able to know anything at all. In this sense, this manifestation alludes to the intangibility of art as a practice, rather than the drawings as discrete objects. This is the mediation we see reflected in the interviews. The participants speak about the drawings in multiples and do not single in on them as discrete or authoritative objects.

Another important aspect in facilitating the hermeneutic relation was setting up the *noetic context* [57] in which the exploration unfolds. Meaning the mental framing through which the participants experienced the system. This came from the instructions that were given before the drawing activity started. In a museum setting, it is common to provide similar text-based prompts that help direct visitor's attention to particular aspects of the exhibited artworks to support different readings of it.

This is a quite different view of art history than a typical museum installation would offer: Normally museums present art history by exhibiting individual works that are deemed particularly interesting in terms of their unique qualities. Works may also be exhibited as representative of a broader tendency in the work of an artist, a particular style, or period, and sometimes a broad selection of works may be presented together to explore such tendencies - but even then the number of works that can be presented at one time to a visitor is far from the 5800 drawings in the dataset of the DGM model in *New Snow*. As such, we anticipate that this approach toward presenting a large body of artworks might be greeted with some controversy among art historians and curators.

We might not have achieved hermeneutic interpretation in all cases, as some participants appropriated the system for a more tool-like relation, supporting their own drawing practice.

6.3 Creating a Munch-like intentionality

To create a model that warrants the kind of exploration described above requires that special attention is paid to its constitution and ontology. In this art museum domain, this means that the training data must be selected with an eye for what we would like to draw attention to. In this case it is an artists' practice. The data is the matter from which the model is built and from which it derives its form, and therefore also important for how the system, in turn, comes to mediate the world. This presents a design task outside the typical scope of most designers' jobs.

Making a bespoke model is a labor-intensive process. Gathering data and pre-processing it are time-consuming tasks that are hard to evaluate the success of before the first model is trained: Only when the designer can interact with the trained model can they get a sense of whether the model can facilitate a user experience similar to the design vision. This calls for an iterative process in which the data collection and model training are reiterated several times until the desired result is achieved - however, the amount of time and labor it takes to adjust a dataset with thousands - or potentially millions - of entries and retraining through numerous iterations makes the cost of iterating very high. Also in this project, it has limited the amount of iterations we have been able to do on the model.

Training a bespoke model also requires that a certain amount of data is available. Even with a very productive artist like Munch, collecting a sufficient amount of images in the dataset required that we had to accept a certain level of stylistic divergence. We made a selection to include only specific mediums and techniques. Still, we accepted a certain variety in the drawing style due to the fact that Munch's style changed throughout his life and across different types of subjects.

In the end, these are the processes that establish the intentionality of the model, and whether we can take it to be representative of Munch's artistic practice in any aspect. This is important, because the dialogical interaction with the model supports an *alterity relation*, that we see some participants speak about. Through the framing of the task, we have set the participants up to experience drawing *with* some Munch-like entity. This supports the aim of the project, but could also quickly lead to misinterpretation when unexpected artefacts of the training process appear, such as the craggy cliff-faces reported by some participants.

6.4 Limitations and further work

The analysis and discussion of the *New Snow* system in this paper revolves around the participants' experiences with the system and how those relate to the technological mediation happening. An evaluation of the technical performance of the system has thus been outside the scope of this paper. However, the work the system *does* is naturally related to aspects that can be considered from a more conventional perspective of technical performance. We have already discussed the importance of the speed at which the system generates images, and we will briefly present two more aspects that should be investigated further before this system is deployed in an art museum setting.

The first is how well the system is at reproducing the dynamics and quality of Munch's line work. We mentioned earlier, that the images must have a 1:1 aspect ratio before serving as training data for the system. Very few of the drawings have exactly that, which means that the images must be pre-processed before training. There are various strategies to achieve this aspect ratio such as cropping, stretching, or using machine learning assisted in-painting. Each of these will have a different impact on how the system comes to express the relations of relative scale between the X and Y dimensions. Alternatively, other DGM architectures might not have this constraint.

The second issue is the visual subjects that seem to be suppressed in the output of the model. We do not have a concrete measure of which subjects, or other features, like shading or line textures, are magnified or reduced by the model's interpretation of the data. Additionally, we do not know if the perceived bias stems from the StyleGAN model itself or from the way the pSp encoder mediates access to the latent space of the StyleGAN model.

These aspects of the system's performance could be tested quantitatively in various ways. However, considering the purpose of this system for art mediation, a more relevant evaluation would be to invite the paper curators from MUNCH to qualitatively assess the output of the system. Their assessment would be based on how the model draws attention to particular aspects of Munch's drawing practice deemed relevant by the art professionals. The system could then be improved in response to their feedback. This type of evaluation would be more apt for staying focused on what the system *does* in the world, rather than measurements that do not in themselves support doing better art mediation.

7 CONCLUSION

In this paper, we have presented a design that affords a *hermeneutic relation* to a DGM. Through a drawing interface, the system offers the user a *functional perspective* on Munch's sketching practice that is unique to the capabilities of the model we have created for the purpose. By designing for a specific relationality, we have shown that we can utilize DGMs for exploring a corpus of artistic work, that is otherwise challenging to exhibit. We have designed and evaluated how the model reveals Munch's drawing practice to our participants in the study. The *New Snow* system presented magnifies the visual trends and aesthetic qualities of his drawings as a practice while reducing their perceptual presence as objects. The technological mediation becomes a form of art mediation that can *educate the attention* of the user to specific areas of interest in Munch's art.

Our approach can be applied to the practice of other productive visual artists, but also beyond the art museum. Through design for a *hermeneutic relation* models trained on other visual datasets can be explored not as a collection of discrete data objects, but as trends and patterns. We have discussed how we have manifested the synthetic drawings as *uncertain entities* that avoid establishing themselves as authoritative representations of the training data, but as ephemeral expressions of relations across the model. Finally, we have presented three design choices, *prompting modality*, *gradual prompt adjustment*, and *fast updates* that we found important for affording transparency in the drawing relation to allow for a hermeneutic exploration of the qualities of the model.

ACKNOWLEDGMENTS

The authors would like to thank MUNCH, Oslo for help in developing the concept and funding the prototype. We are thankful to the research division at Random International for the development of the first iteration of the tracking system and AIRLab at IT University of Copenhagen for supporting the physical realization of the prototype. Finally a special thanks to René Haas for help with developing the machine learning pipeline.

REFERENCES

[1] Memo Akten. 2017. Learning to See: Interactive. https://www.memo.tv/works/learning-to-see/
[2] Memo Akten. 2021. *Deep Visual Instruments: Realtime Continuous, Meaningful Human Control over Deep Neural Networks for Creative Expression.* doctoral. Goldsmiths, University of London. https://research.gold.ac.uk/id/eprint/30191/
[3] Memo Akten, Rebecca Fiebrink, and Mick Grierson. 2019. Learning to see: you are what you see. In *ACM SIGGRAPH 2019 Art Gallery*. ACM, Los Angeles California, 1–6. https://doi.org/10.1145/3306211.3320143
[4] Jane Alexander. 2014. Gallery One at the Cleveland Museum of Art. *Curator: The Museum Journal* 57, 3 (2014), 347–362. https://doi.org/10.1111/cura.12073
[5] Gil Appel, Juliana Neelbauer, and David A. Schweidel. 2023. Generative AI Has an Intellectual Property Problem. *Harvard Business Review* (April 2023). https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem Section: Intellectual property.
[6] Avital Dell' Ariccia, Alexandra Bremers, Johan Michalove, and Wendy Ju. 2022. How to Make People Think You're Thinking if You're a Drawing Robot: Expressing Emotions Through the Motions of Writing. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 1190–1191. https://doi.org/10.1109/HRI53351.2022.9889638
[7] James Atwood, Yoni Halpern, Pallavi Baljekar, Eric Breck, D. Sculley, Pavel Ostyakov, Sergey I. Nikolenko, Igor Ivanov, Roman Solovyev, Weimin Wang, and Miha Skalic. 2020. The Inclusive Images Competition. In *The NeurIPS '18 Competition (The Springer Series on Challenges in Machine Learning)*, Sergio Escalera and Ralf Herbrich (Eds.). Springer International Publishing, Cham, 155–186. https://doi.org/10.1007/978-3-030-29135-8_6
[8] Sofian Audry. 2021. *Art in the age of machine learning.* The MIT Press, Cambridge, Massachusetts.
[9] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58 (June 2020), 82–115. https://doi.org/10.1016/j.inffus.2019.12.012
[10] Steve Benford, Anders Sundnes Løvlie, Karin Ryding, Paulina Rajkowska, Edgar Bodiaj, Dimitrios Paris Darzentas, Harriet Cameron, Jocelyn Spence, Joy Egede, and Bogdan Spanjevic. 2022. Sensitive Pictures: Emotional Interpretation in the Museum. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–16. https://doi.org/10.1145/3491102.3502080
[11] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3411764.3445481
[12] Jesse Josua Benjamin, Heidi Biggs, Arne Berger, Julija Rukanskaitė, Michael B. Heidt, Nick Merrill, James Pierce, and Joseph Lindley. 2023. The Entoptic Field Camera as Metaphor-Driven Research-through-Design with AI Technologies. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–19. https://doi.org/10.1145/3544548.3581175
[13] Abeba Birhane, Vinay Uday Prabhu, and Emmanuel Kahembwe. 2021. Multimodal datasets: misogyny, pornography, and malignant stereotypes. (2021). https://doi.org/10.48550/ARXIV.2110.01963 Publisher: arXiv Version Number: 1.
[14] Andreas Blattmann, Robin Rombach, Kaan Oktay, Jonas Müller, and Björn Ommer. 2022. Semi-Parametric Neural Image Synthesis. (2022). https://doi.org/10.48550/ARXIV.2204.11824 Publisher: arXiv Version Number: 3.
[15] Margaret A. Boden and Ernest A. Edmonds. 2019. *2 A Taxonomy of Computer Art.* 23–59.
[16] Eva Cetinic and James She. 2022. Understanding and Creating Art with AI: Review and Outlook. *ACM Trans. Multimedia Comput. Commun. Appl.* 18, 2, Article 66 (feb 2022), 22 pages. https://doi.org/10.1145/3475799
[17] Min Chen. 2023. Artists and Illustrators Are Suing Three A.I. Art Generators for Scraping and 'Collaging' Their Work Without Consent. *Artnet News* (Jan. 2023). https://news.artnet.com/art-world/class-action-lawsuit-ai-generators-deviantart-midjourney-stable-diffusion-2246770 Section: Law.
[18] Brendan Ciecko. 2020. AI sees what? The good, the bad, and the ugly of machine vision for museum collections. *The Museum Review* 5, 1 (Jan. 2020). https://themuseumreviewjournal.wordpress.com/2020/04/23/tmr_vol5no1_ciecko/
[19] Richard J. Cole, Frithjof Dau, Jon Ducrou, Peter W. Eklund, and Tim Wray. 2019. Navigating Context, Pathways and Relationships in Museum Collections using Formal Concept Analysis. *International Journal for Digital Art History* 4 (Dec. 2019), 5.13–5.27. https://doi.org/10.11588/dah.2019.4.72070
[20] Kate Crawford and Trevor Paglen. 2021. Excavating AI: the politics of images in machine learning training sets. *AI & SOCIETY* 36, 4 (Dec. 2021), 1105–1116. https://doi.org/10.1007/s00146-021-01162-8
[21] Nicholas Davis, Chih-PIn Hsiao, Kunwar Yashraj Singh, Lisa Li, Sanat Moningi, and Brian Magerko. 2015. Drawing Apprentice: An Enactive Co-Creative Agent for Artistic Collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (Glasgow, United Kingdom) *(C&C '15)*. Association for Computing Machinery, New York, NY, USA, 185–186. https://doi.org/10.1145/2757226.2764555
[22] Edirlei Soares de Lima, Bruno Feijó, Simone D.J. Barbosa, Antonio L. Furtado, Angelo E.M. Ciarlini, and Cesar T. Pozzer. 2014. Draw your own story: Paper and pencil interactive storytelling. *Entertainment Computing* 5, 1 (Jan. 2014), 33–41. https://doi.org/10.1016/j.entcom.2013.06.004
[23] Emily Denton, Alex Hanna, Razvan Amironesei, Andrew Smart, and Hilary Nicole. 2021. On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society* 8, 2 (July 2021), 205395172110359. https://doi.org/10.1177/20539517211035955
[24] Finale Doshi-Velez and Been Kim. 2017. Towards A Rigorous Science of Interpretable Machine Learning. http://arxiv.org/abs/1702.08608 arXiv:1702.08608 [cs, stat].
[25] Ziv Epstein, Aaron Hertzmann, the Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. 2023. Art and the science of generative AI. *Science* 380, 6650 (2023), 1110–1111. https://doi.org/10.1126/science.adh4451 arXiv:https://www.science.org/doi/pdf/10.1126/science.adh4451
[26] Anna Foka, Lina Eklund, Anders Sundnes Løvlie, and Gabriele Griffin. 2023. *Chapter 71: Critically assessing AI/ML for cultural heritage: potentials and challenges.* Edward Elgar Publishing, Cheltenham, UK, 815 – 825. https://doi.org/10.4337/9781803928562.00082
[27] F. Fontanella, F. Colace, M. Molinara, A. Scotto Di Freca, and F. Stanco. 2020. Pattern recognition and artificial intelligence techniques for cultural heritage. *Pattern Recognition Letters* 138 (2020), 23–29. https://doi.org/10.1016/j.patrec.2020.06.018
[28] Adarsh Ghosh and Devasenathipathy Kandasamy. 2020. Interpretable Artificial Intelligence: Why and When. *American Journal of Roentgenology* 214, 5 (May 2020), 1137–1138. https://doi.org/10.2214/AJR.19.22145 Publisher: American Roentgen Ray Society.
[29] Google Creative Lab, Bastien Girschig, and Romain Cazier. 2020. Draw to Art: Shape Edition. https://experiments.withgoogle.com/draw-to-art-shape
[30] Google Creative Lab, Google Art & Culture Lab, and IYOIYO. 2018. Draw to Art. https://experiments.withgoogle.com/draw-to-art
[31] Sergiu Gordea and Michela Vignoli. 2015. CultureCam: Visual exploration of cultural heritage content by professional designers. In *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. 1–6. https://doi.org/10.1109/ICMEW.2015.7169797
[32] Dejan Grba. 2022. Deep Else: A Critical Framework for AI Art. *Digital* 2, 1 (2022), 1–32. https://doi.org/10.3390/digital2010001
[33] Varvara Guljajeva and Mar Canet Sola. 2022. POSTcard Landscapes from Lanzarote. In *Creativity and Cognition*. ACM, Venice Italy, 634–636. https://doi.org/10.1145/3527927.3531191
[34] Joana Hois, Dimitra Theofanou-Fuelbier, and Alischa Janine Junk. 2019. How to Achieve Explainability and Transparency in Human AI Interaction. In *HCI International 2019 - Posters (Communications in Computer and Information Science)*, Constantine Stephanidis (Ed.). Springer International Publishing, Cham, 177–183. https://doi.org/10.1007/978-3-030-23528-4_25
[35] Eva Hornecker and Luigina Ciolfi. 2019. Human-Computer Interactions in Museums. *Synthesis Lectures on Human-Centered Informatics* 12, 2 (April 2019), i–153. https://doi.org/10.2200/S00901ED1V01Y201902HCI042
[36] Tim Ingold. 2001. From the transmission of representation to the education of attention. In *The debated mind: Evolutionary psychology versus ethnography*. Berg, New York, NY, US, 113–153.
[37] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. 2020. Training Generative Adversarial Networks with Limited Data.

In *Proc. NeurIPS*.
[38] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2021. Alias-Free Generative Adversarial Networks. In *Proc. NeurIPS*.
[39] Sarah Kenderdine and Heidi McKenzie. 2013. A war torn memory palace: Animating narratives of remembrance. In *2013 Digital Heritage International Congress (DigitalHeritage)*. IEEE, Marseille, France, 315–322. https://doi.org/10.1109/DigitalHeritage.2013.6743755
[40] Sarah Kenderdine, Jeffrey Shaw, and Tobias Gremmler. 2013. Cultural Data Sculpting: Omnidirectional Visualization for Cultural Datasets. In *Knowledge Visualization Currents*, Francis T. Marchese and Ebad Banissi (Eds.). Springer London, London, 199–220. https://doi.org/10.1007/978-1-4471-4303-1_11
[41] Asle H. Kiran. 2015. Four Dimensions of Technological Mediation. In *Postphenomenological Investigations: Essays on Human–Technology Relations*, Robert Rosenberger and Peter-Paul Verbeek (Eds.). Lexington Books.
[42] LAIKA. 2023. Write with LAIKA - Personalised Artificial Intelligence for Writers. https://www.writewithlaika.com/
[43] Agostina J. Larrazabal, Nicolás Nieto, Victoria Peterson, Diego H. Milone, and Enzo Ferrante. 2020. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences* 117, 23 (June 2020), 12592–12594. https://doi.org/10.1073/pnas.1919012117 Publisher: Proceedings of the National Academy of Sciences.
[44] Tomas Lawton, Kazjon Grace, and Francisco J Ibarrola. 2023. When is a Tool a Tool? User Perceptions of System Agency in Human–AI Co-Creative Drawing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1978–1996. https://doi.org/10.1145/3563657.3595977
[45] Lucian Leahu. 2016. Ontological Surprises: A Relational Perspective on Machine Learning. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, Brisbane QLD Australia, 182–186. https://doi.org/10.1145/2901790.2901840
[46] Ann Light. 2006. Adding method to meaning: A technique for exploring peoples' experience with technology. *Behaviour & Information Technology* 25, 2 (March 2006), 175–187. https://doi.org/10.1080/01449290500331172
[47] Bernd Lintermann. 2012. Beyond Cinema. Hongik University, Korea. https://www.bernd-lintermann.de/papers/Beyond_Cinema_Lintermann.pdf
[48] Nicolas Malevé. 2021. On the data set's ruins. *AI & SOCIETY* 36, 4 (Dec. 2021), 1117–1131. https://doi.org/10.1007/s00146-020-01093-w
[49] Nikita Mathias. 2022. Meta-artistic immersion in digital exhibitions. History – mobilization – spectatorship. *Journal of Aesthetics \& Culture* 14, 1 (Dec. 2022), 2129160. https://doi.org/10.1080/20004214.2022.2129160
[50] Mihaela Mihailova. 2021. To Dally with Dalí: Deepfake (Inter)faces in the Art Museum. *Convergence* 27, 4 (2021), 882–898. https://doi.org/10.1177/13548565211029401 arXiv:https://doi.org/10.1177/13548565211029401
[51] Kate Mondloch. 2022. The Influencers: Van Gogh Immersive Experiences and the Attention-Experience Economy. *Arts* 11, 5 (Sept. 2022), 90. https://doi.org/10.3390/arts11050090
[52] Alva Noë. 2021. *Learning to Look: Dispatches from the Art World*. Oxford University Press.
[53] Jonas Oppenlaender. 2022. A Taxonomy of Prompt Modifiers for Text-To-Image Generation. http://arxiv.org/abs/2204.13988 arXiv:2204.13988 [cs].
[54] Jonas Oppenlaender, Rhema Linder, and Johanna Silvennoinen. 2023. Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering. arXiv:2303.13534 [cs.HC]
[55] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical Text-Conditional Image Generation with CLIP Latents. (2022). https://doi.org/10.48550/ARXIV.2204.06125 Publisher: arXiv Version Number: 1.
[56] Elad Richardson, Yuval Alaluf, Or Patashnik, Yotam Nitzan, Yaniv Azar, Stav Shapiro, and Daniel Cohen-Or. 2021. Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
[57] Robert Rosenberger and Peter-Paul Verbeek. 2015. A Field Guide to Post-Phenomenology. In *Postphenomenological Investigations: Essays on Human-Technology Relations*, Robert Rosenberger and Peter-Paul Verbeek (Eds.). Lexington Books, 9–41.
[58] Hugo Scurto, Baptiste Caramiaux, and Frederic Bevilacqua. 2021. Prototyping Machine Learning Through Diffractive Art Practice. In *Designing Interactive Systems Conference 2021 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 2013–2025. https://doi.org/10.1145/3461778.3462163
[59] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility (CC '05)*. Association for Computing Machinery, New York, NY, USA, 49–58. https://doi.org/10.1145/1094562.1094569
[60] Jeffrey Shaw, Neil Brown, Dennis Del Favero, Matt McGinity, and Peter Weibel. 2006. T_Visionarium II. https://www.jeffreyshawcompendium.com/portfolio/t_visionarium-ii/
[61] Eugenia Sinatti, Simon Weckert, and Ewelina Dobrzalski. 2022. Image Garden: Curating Collections and Designing Smart Exhibitions with AI-Based Tools. *i-com* 21, 1 (2022), 83–94. https://doi.org/doi:10.1515/icom-2022-0010
[62] Christian Sivertsen, Maarten Smith, and Sander van der Zwan. 2023. Art Critique by Other Means. In *Designing Interactive Systems Conference (DIS '23)*. ACM, Pittsburg, PA, 10. https://doi.org/10.1145/3563657.3596069
[63] SMK. 2018. SMK Open: Setting Art Free. https://www.smk.dk/en/article/smk-open/
[64] Jamie Tolentino. 2022. Dalí Cybernetics: The Immersive Experience. *London Unattached* (Dec. 2022). https://www.london-unattached.com/dali-cybernetics-the-immersive-experience/
[65] Elena Villaespesa and Seth Crider. 2021. Computer Vision Tagging the Metropolitan Museum of Art's Collection: A Comparison of Three Systems. *J. Comput. Cult. Herit.* 14, 3, Article 28 (jul 2021), 17 pages. https://doi.org/10.1145/3446621
[66] Elena Villaespesa and Oonagh Murphy. 2021. This is not an apple! Benefits and challenges of applying computer vision to museum collections. *Museum Management and Curatorship* 36, 4 (2021), 362–383. https://doi.org/10.1080/09647775.2021.1873827 Publisher: Routledge _eprint: https://doi.org/10.1080/09647775.2021.1873827.
[67] Angelina Wang, Alexander Liu, Ryan Zhang, Anat Kleiman, Leslie Kim, Dora Zhao, Iroha Shirai, Arvind Narayanan, and Olga Russakovsky. 2022. REVISE: A Tool for Measuring and Mitigating Bias in Visual Datasets. *International Journal of Computer Vision* 130, 7 (July 2022), 1790–1810. https://doi.org/10.1007/s11263-022-01625-5
[68] Christine T. Wolf. 2021. Towards "Explorable" AI: Learning from ML Developers' Sensemaking Practices. (2021). https://doi.org/10.18420/ECSCW2021_N28 Publisher: European Society for Socially Embedded Technologies (EUSSET).
[69] Niels Wouters, Ryan Kelly, Eduardo Velloso, Katrin Wolf, Hasan Shahid Ferdous, Joshua Newn, Zaher Joukhadar, and Frank Vetere. 2019. Biometric Mirror: Exploring Ethical Opinions towards Facial Analysis and Automated Decision-Making. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 447–461. https://doi.org/10.1145/3322276.3322304
[70] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376301
[71] Nushin Isabelle Yaszdani. 2018. Machine Bias. https://nushinyazdani.com/Machine-Learning-Bias
[72] Paulina Yurman and Anuradha Venugopal Reddy. 2022. Drawing Conversations Mediated by AI. In *Proceedings of the 14th Conference on Creativity and Cognition* (Venice, Italy) *(C&C '22)*. Association for Computing Machinery, New York, NY, USA, 56–70. https://doi.org/10.1145/3527927.3531448
[73] Jichen Zhu, Antonios Liapis, Sebastian Risi, Rafael Bidarra, and G. Michael Youngblood. 2018. Explainable AI for Designers: A Human-Centered Perspective on Mixed-Initiative Co-Creation. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, Maastricht, 1–8. https://doi.org/10.1109/CIG.2018.8490433
[74] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 493–502. https://doi.org/10.1145/1240624.1240704
[75] Joanna Zylinska. 2020. *AI Art: Machine Visions and Warped Dreams.* Open Humanites Press. http://www.openhumanitiespress.org/books/titles/ai-art/