Machine Learning in Art: Tools, Techniques, and Implications for Conservation
By Deena Engel for *The Electronic Media Review*

ABSTRACT
The software applications that artists use for creating works of art that integrate or are based on machine learning fall into several categories reflecting different taxonomies.

One taxonomy describes artworks by identifying the technical skills and resources needed to create the artworks. Some artists use text-to-image applications, which allow them to create images through textual description without programming. Artists with intermediate programming skills may prefer writing original code such as Python scripts to generate new images. Building an original machine learning application, however, requires great resources and advanced programming skills.

A second taxonomy is based on presentation format, as artworks created using machine learning may be exhibited and sold as digital video, web pages, physical prints, and/or other media.

A third taxonomy is based on the artist's data source for the artwork: whether the artist is using a proprietary machine learning model; working with an open-source machine learning model; or building a bespoke model.

Each of these artistic approaches raises specific conservation concerns. Conservation strategies for these fragile and complex artworks are needed as artists continue to explore the use of machine learning as an artistic medium.

INTRODUCTION
Machine learning is a new technology that in turn has given birth to a range of new art mediums. As artists learn how to take advantage of machine learning, gallerists, institutions, and conservators grapple with how to collect and care for the resulting artworks.

WHAT IS MACHINE LEARNING?
Machine learning refers to the methods by which computers "learn" from data in order to make predictions or decisions without being explicitly programmed as they would be in classical or traditional programming. In classical or traditional computer programs we begin with rules and data, and we write code, a process also called "coding," to arrive at our answers. For example, we start with an input—such as all of the financial transactions for a given bank account over a given period of time—and we apply the rule that credits are positive amounts and debits are negative amounts to reach the solution, the current balance on the account, through addition and subtraction.

However, in machine learning, we start with the data, such as the Netflix library, and an answer, such as everything you have watched on Netflix over time, and the software generates "rules" in order to make predictions and suggest new movies for you, based on its interpretation of your tastes.

Note that while the terms "artificial intelligence" and "machine learning" are often used interchangeably, they refer to different concepts. While some artificial intelligence applications do not use machine learning, all machine learning methods are considered part of the family of artificial intelligence methods.

To extend this analogy to the visual arts, consider Sol LeWitt's (1928-2007) *Wall Drawing 273* (September 1975. Graphite and crayon on seven walls, dimensions variable) at SFMOMA (LeWitt 1975). This is "rules-based art:" the instructions for this specific wall drawing are available for study (SFMOMA 2020) as are the instructions for many of Sol LeWitt's other wall drawings.

Another example of rules-based or algorithmic art is *Color Panel 1.0* (*Edition 3/12*. 1999. Software, modified Apple Macintosh Powerbook 280c, and acrylic (plastic), overall dimensions: 13 1/2 × 10 1/2 × 3 in. (34.3 × 26.7 × 7.6 cm) by John F. Simon, Jr. (b. 1963). If we examine the source code of this artwork, which was written in the language C++, we can predict everything that will happen in this artwork, even though the piece itself will not precisely repeat over a time frame that approaches eternity (Simon Jr. 1999).

Mosaic Virus (2019), by Anna Ridler (b. 1985), on the other hand, looks like a video, and the exhibition format that one sees is indeed a video. However, the artwork was made using machine learning. The artist describes *Mosaic Virus* as "…an image not of a real tulip, but what [the software] thinks a tulip should be, based on all of the tulips that are contained in the datasets …[by] forming a concept of a flower and recording it as such." (Ridler 2019.)

CATEGORIZING WORKS OF ART BASED ON MACHINE LEARNING

The terms "taxonomies" and "controlled vocabularies," sometimes used interchangeably, are controlled lists of terms used to organize information (University of Pittsburgh Library System, 2025). Through the following examples, I develop three taxonomies to categorize the diverse artworks created using machine learning.

In an earlier study, published in the *Leonardo* journal, "A Concise Taxonomy for Describing Data as an Art Material," Freeman et al. (2018) define taxonomies for data as used for art. This study forms a basis for the following standards, as it predates consideration of data used with machine learning in creating works of art.

Building taxonomies for categorizing works of art based on machine learning can facilitate defining approaches to their conservation by highlighting commonalities among these vastly different artworks to capture standards and to build guidance towards their care.

A TAXONOMY BASED ON TECHNICAL SKILLS AND RESOURCES REQUIRED TO RENDER THE ARTWORK

One taxonomy to categorize artworks based on machine learning reflects the artist's skill level, collaboration with engineers or computer scientists, and/or the resources required to produce the respective works of art. This taxonomy is important to a technical art history approach as well as to a curatorial and art historical understanding of how and where the artworks were made.

For the purpose of discussing conservation needs, these artworks fall into three categories: (1) artworks created using software that does not require programming skills; (2) artworks created using Python or another programming language to create images requiring skills ranging from beginner to intermediate; and (3) artworks created through the use of large custom models such as GANs (Generative Adversarial Networks) which require extensive hardware and software resources and advanced engineering skills.

Artworks created using software that does not require programming skills
There are many software applications for generating images that do not require the artist to write code, such as the popular "text-to-image" packages, including OpenAI's DALL•E 2, DALL•E 3, and tools available through Adobe's Creative Cloud. The following examples are based on the author's digital photographs, scanned physical sketches, and software experiments.

Adobe's Firefly (Adobe 2025) offers the artist software that "learns" from the structure or the style of an existing image. In this case, the photograph on the left (see Fig. 1) set as a structure model, along with the prompt of "large pink flower with green leaves," results in the four images on the right (see Fig. 2):

		
Figure 1: Digital photograph by the author.		Figure 2: Using the photograph in Fig. 1 as a structure reference and the text prompt "large pink flower with green leaves" resulted in these four images.

However, using the same photograph as a reference for structure (see Fig. 3), along with my original watercolor for style (see Fig. 4) and the same text prompt of "large pink flower with green leaves," yields this second set of four images (see Fig. 5):

 	
Figure 3: Digital photograph by the author. Figure 4: Watercolor and ink sketch by the author.	Figure 5: Using the photograph in Fig. 3 as a structure reference, the watercolor in Fig. 4 as a style reference, and the text prompt “large pink flower with green leaves” resulted in these four images.

Note that the software transformed many tight layers of petals in the rosette in the sketch into a flower with one or two loose layers of petals, and only the first software-generated design displays the correct leaves for a rose. Although a rosarian or a gardener would be likely to notice this right away, as a society, we tend to take images “at face value,” and it is important to bear in mind the propensity to errors in all aspects of working with machine learning.

In the next example, I used one image (see Fig. 6) for both the structure and style models, along with a text prompt of “California mountains in the rain,” with the following result (see Fig. 7):

	
Figure 6: Watercolor sketch by the author of California mountains in the rain.	Figure 7: Using the watercolor sketch in Fig. 6 as a model for both style and structure, and the prompt "California mountains in the rain done by watercolor," resulted in the four images in Fig. 7.

For comparison, OpenArt (OpenArt 2025) offers a different interpretation. Using the prompt "large pink flower with green leaves" on OpenArt, without an image model, results in the following image (see Fig. 8):

Figure 8: The prompt "large pink flower with green leaves" using OpenArt resulted in this image.

However, using the same flower watercolor as above for style (see Fig. 9) yields a very different result (see Fig. 10) than the result using Firefly.

	
Figure 9: Watercolor and ink sketch by the author.	Figure 10: Using the sketch in Fig. 9 as a style reference yields this result in OpenArt.

In another series of examples, the monthly rate of approximately USD 12.00 on OpenArt allows users to train up to two models per month using 4-64 images for this purpose. I uploaded 61 images of myself to train a model for character; other options include style, face, and object models. Uploading the images and training the model took well over an hour.

Here are some of the examples that I generated from this model and the associated text prompts in figures 11, 12, and 13. The website offers links to training materials for beginners (Diab et al. 2022; Nozem et al. 2023).

	<mymodel>**Royal Princess**, flowing silk gown, medieval fantasy, intricate crown, majestic posture, soft and ethereal lighting, high quality, fantasy, regal, detailed gown, royal setting, elegant, graceful, luxurious fabrics, majestic, soft pastel tones, atmospheric lighting
Figure 11: OpenArt portrait and the prompt, Example 1.	

	<mymodel>**Playing a silver flute** on a stage in a concert hall wearing a blue dress and high heel shoes.
Figure 12: OpenArt portrait and the prompt, Example 2.	

	<mymodel>**Wearing hiking boots, jeans, winter jacket, hat, and gloves**, hiking on a snowy trail in a Connecticut pine forest in the winter afternoon light.
Figure 13: OpenArt portrait and the prompt, Example 3.	

Artworks created using Python or another programming language to create images requiring beginner to intermediate programming skills

Using elementary Python, one can create text prompts to download selected images from the web, also referred to as “scraping” the web. Artists can then use these images as a basis for further work. This approach may introduce copyright concerns.

The six images that you see here in Figures 14-1 through 14-6 are accompanied by the associated metadata that can be downloaded simultaneously through the original code written by the author for this purpose. (See Note #1 for the metadata.) In this tiny sample, only the first and fifth images contain an attribution within the image itself. This download of six images took approximately 3 seconds under MacOS running Monterey 12.7.5 with 8GB of RAM and a 1.8 GHz Dual-Core Intel Core i5 processor. Using the same computer and original code as above required 21 seconds to download 100 images at 21 MB; and 3 minutes and 8 seconds to download 1,000 images for a total of 142 MB.

Once the files have been downloaded, one can write elementary Python code or use another programming language to organize or select images based on file type by identifying the suffix (e.g., .gif, .jpg, .png, etc.), to resize the images, to rename the images, and to handle other file management tasks programmatically to efficiently manage a potentially large collection of image files.

An artist who requests 5,000 or more images could therefore download these images for further work on their local personal computer. Note that the artist would still be well advised to visually examine the results. This approach also raises many questions of attribution, authenticity, and other curatorial and art historical questions. From a conservation perspective, however, storing and archiving collections of digital images is a well-documented task.

 [14-1]	 [14-2]	 [14-3]
 [14-4]	 [14-5]	 [14-6]

Figures 14-1 through 14-6: The results of running a custom Python script to download six images from the web. In this case, the text prompt is "watercolor of pink flowers with green leaves." (See Note #1.)

Google's Colab and similar environments allow programmers to write and execute Python in a browser, without any configuration changes to their own computer. This allows the artist access to the computing power needed for complex computational graphics tasks, such as style transfer routines using Python, and requires an elementary to intermediate knowledge of programming.

In following example, I used my original digital photo for the content and my original watercolor for style, running code through a Google Colab Python demo to generate a "digital watercolor" based on the photo as though I were the artist (see Fig. 15). I was mindful to use as much of the provided Python code as possible, modifying it where necessary, to see how it might feel to an artist with limited coding experience experimenting with this code to generate images. This screenshot displays the original photograph that I used for content, along with a digital scan of the original watercolor painting of mountains in the rain for style. The resulting stylized image mimics the ethereal quality of the wet-on-wet watercolor adapted to the subject of the flowers from the photograph, without any text prompts.

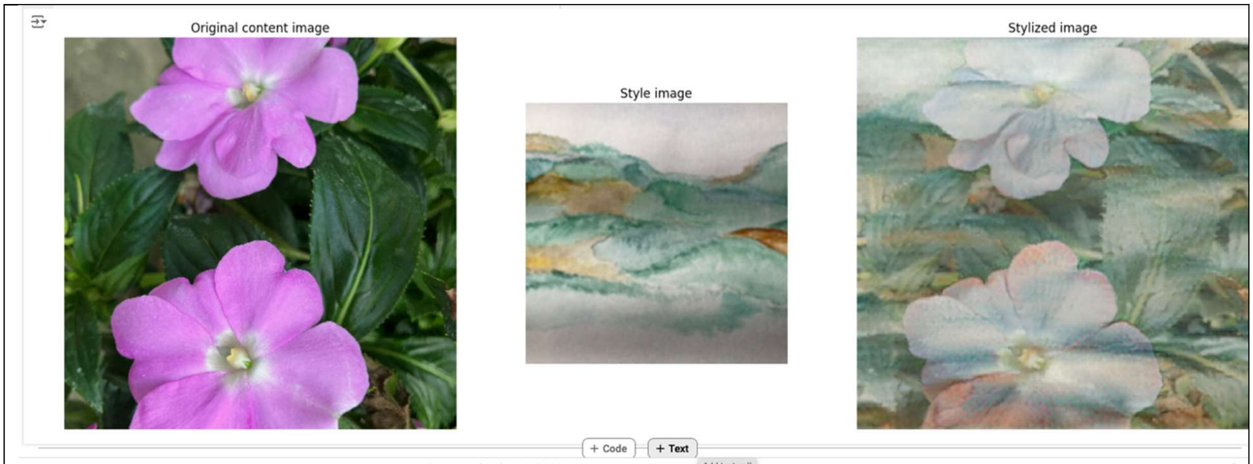

Figure 15: Screenshot excerpt of the result of running Python code in the Google's Colab environment. The author's photograph on the left was used for content and the author's watercolor landscape in the middle was used for style to produce the image on the right.

For comparison, using the same photograph of flowers for content, along with an original watercolor of similar flowers for style, yields the result seen in a second screenshot (see Fig. 16).

Figure 16: Screenshot excerpt of the result of running Python code in the Google Colab environment. The author's photograph on the left was used for content and the author's watercolor pencil and ink sketch in the middle was used for style to produce the image on the right.

Inpainting is a technique to update an image, such as by filling in gaps, while maintaining consistency with the existing image. This technique uses text-to-image to support image generation. To begin this experiment, I started with an original image to provide a basis for style transfer and then added a masked area that I wished to computationally fill.

Figure 17: This is a cropped copy of a watercolor drawing by the author, accompanied by a mask to the right to be filled computationally, using the digital technique of inpainting.

In order to use inpainting, I needed further resources, specifically a machine learning model to implement this technique. Hugging Face (Hugging Face 2025) is a platform where people interested in machine learning can collaborate on models, datasets, and applications. The model web pages on the Hugging Face website provide documentation, instructions, and sample code for artists and programmers to explore inpainting and many other techniques guided by machine learning. The following example based on intermediate-level Python code uses the Hugging Face Stable Diffusion 2 model; specifically, the model for text-guided image inpainting. The software is described as "[letting] you edit specific parts of an image by providing a mask and a text prompt using Stable Diffusion." (Hugging Face Inpainting 2025).

Documentation on Hugging Face and other models is an important component to consider when studying and conserving works of art based on these models. For example, Hugging Face offers guidance to users on Stable Diffusion 2 with respect to appropriate usage, limitations, bias, and other contextual information (Hugging Face Diffusion 2 2025) with advice such as "The model was trained mainly with English captions and will not work as well in other languages;" "Using the model to generate content that is cruel to individuals is a misuse of this model;" and "While the capabilities of image generation models are impressive, they can also reinforce or exacerbate social biases," among others.

The following three examples were created by running code in Python based on the image and mask seen in Fig. 17 and the text prompt listed with each example. The results are subtly different, based on the season (winter and spring in Figs. 18 and 19 respectively, and no season listed in Fig. 20), and in all three cases, the wet-on-wet watercolor style and color palette are consistent with the excerpt in Fig. 17.

Figure 18: Using an inpainting technique in Python based on the model in Fig. 17 and the text prompt “watercolor painting of nature landscape with mist, mountains, lake in winter.”

Figure 19: Using an inpainting technique in Python based on the model in Fig. 17 and the text prompt “watercolor painting of nature landscape with mist, mountains, lake in spring.”

Figure 20: Using an inpainting technique in Python based on the model in Fig. 17 and the text prompt “watercolor painting of nature landscape with mist and mountains.”

Artworks created through the use of large custom models such as GANs (Generative Adversarial Networks) which require extensive hardware and software resources and advanced engineering skills

Figure 21: Refik Anadol. *Unsupervised — Machine Hallucinations — MoMA*, 2021–2022. Custom hardware and software (color, sound, generative algorithm with artificial intelligence). Dimensions variable. The Museum of Modern Art, New York. Gift of 1of1 led by Ryan Zurrer and the RFC Collection led by Pablo and Desiree Rodríguez-Fraile, 2022. Installation view, *Refik Anadol: Unsupervised*, The Museum of Modern Art, New York (November 19, 2022–October 29, 2023). Digital image © 2026 The Museum of Modern Art, New York. Photo: Robert Gerhardt.

Creating a large original model can be very expensive, as it requires a great deal of hardware and other resources. The artist Refik Anadol (b. 1985) reports working with billions of images over many months with a large staff in his studio to create his works such as *Unsupervised—Machine Hallucinations*, which was exhibited at the Museum of Modern Art in New York City (Anadol 2022; see Fig. 21). Anadol describes his work *Large Nature Model: Coral* (2024. Immersive Installation at United Nations Headquarters) as "[u]tilizing a machine learning algorithm trained on a staggering 100 million coral reef images" and "trained on an extensive and ethically sourced dataset of the natural world, [this model] serves as a bridge between scientific research, environmental advocacy, and artistic expression." (United Nations General Assembly 2024; Anadol 2024).

Artists with coding expertise or who are working with computer scientists or engineers in a lab or studio with the necessary computing power to handle machine learning have the option to build their own datasets for training, and to run custom software to process the images and create their artworks. Due to the resources required to create artworks on this scale, I was not able to provide an example for this study. Many artists also do not have access to bespoke software and the kinds of studio requirements and technology resources needed to create artworks such as Anadol's *Unsupervised*, nor do they have access to the resources to attain a digital literacy to support or create works based on machine learning outside of open-source models and cost-effective development environments such as Google's Colab.

Figure 22: American Artist. *Security Theater* (2023). CCTV cameras, AI, acrylic, cable, 4K monitors, television mounts, computer, tandem sling seating, phone security pouches, messenger bags, aluminum, steel, hardware, desk, and signage. Source: https://www.guggenheim.org/audio/track/a-guide-to-engaging-with-security-theater-2023

Some artists use professionally written commercial software. American Artist's (b. 1989) *Security Theatre* (2023. CCTV cameras, AI, acrylic, cable, 4K monitors, television mounts, computer, tandem sling seating, phone security pouches, messenger bags, aluminum, steel, hardware, desk, and signage), seen at the Solomon R. Guggenheim Museum in 2024 in New York City, uses machine learning techniques to support image recognition (American Artist 2023; see Fig. 22). This artwork identifies figures in video footage taken in the museum by distinguishing visitors from architectural elements, other artworks, and objects in the surrounding areas.

TAXONOMY BASED ON PRESENTATION FORMAT

A second taxonomy categorizes artworks based on machine learning by identifying presentation format in order to inform conservation needs. A given artwork can be sold in more than one format, and some works are sold with NFT registration. Conservation planning is well documented elsewhere for works that are collected or that enter an institution as a digital video; digital image or series of digital images; a physical image such as a silkscreen print; a web page;

or other presentation formats common to time-based media collections. Machine learning can also be used to create one or more elements in a larger time-based media art installation.

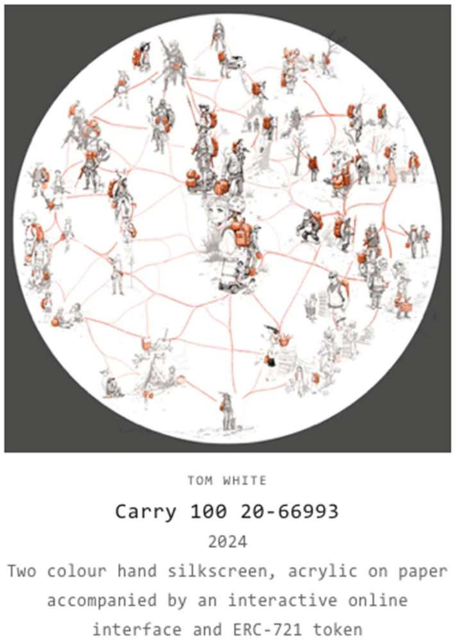

Figure 23: Tom White. *Maximum* Activations. Two colour hand silkscreen, acrylic on paper accompanied by an interactive online interface and ERC-721 token. Source: https://gazell.io/collections/tom-white-maximum-activations

The Aotearoan (New Zealander) artist Tom White (b. unknown) sold his artworks in the show "Maximum Activations" at the Gazelli Art House in London in early 2025 in several formats described as "Two colour hand silkscreen, acrylic on paper accompanied by an interactive online interface and ERC-721 token" (White 2025; see Fig. 23). The prints are listed in both metric and inches (e.g., "Print: 45 18/25 x 45 18/25 cm | 18 x 18 in") and the digital images are given in pixels (e.g., "Digital Image: 5400×5400 px"). They are further described as "Visual collages, created by rendering relevant computer memories—datasets, activations, and connections in the neural network—into artwork." (White 2024).

In another example of wide-ranging presentation formats, bitforms gallery in San Francisco presented the show "Artificial Imagination," the first DALL•E inspired art exhibition, from October 26–December 29, 2022 (bitforms gallery 2022). All of these works were created using machine learning, but the presentations varied widely. Artworks in this show were offered for sale in different media, including a digital image with NFT registration; a video (with color and sound) in a 30-second loop and NFT registration; a digital collage on archive Bamboo with NFT registration; a Giclée print on archival rag with NFT registration; and custom software and internet-connected computer with NFT registration, among other presentations (bitforms gallery [Catalog], 22).

Figure 24: Holly Herndon and Mat Dryhurst. *Embedding Study 1 & 2 (xhairymutantx series)* (2024). Thermal dye diffusion transfer prints each: 47 3/4 × 71 5/8 in. (119 × 180 cm.). Executed in 2024, this work is number one from an edition of three plus two artist proofs. Source: https://onlineonly.christies.com/s/augmented-intelligence/holly-herndon-b-1980-mat-dryhurst-b-1984-3/249745 .

On March 5, 2025, Christie's held a first-ever AI-based art auction called "Augmented Intelligence." The work *Embedding Study 1 & 2 (xhairymutantx series)* (2024. Thermal dye diffusion transfer prints each: 47 3/4 × 71 5/8 in. (119 × 180 cm.)) by Holly Herndon (b. 1980) and Mat Dryhurst (b. 1984) was put up for sale as "Non fungible token (NFT) & physical object" (Herndon and Dryhurst 2024; see Fig. 24).

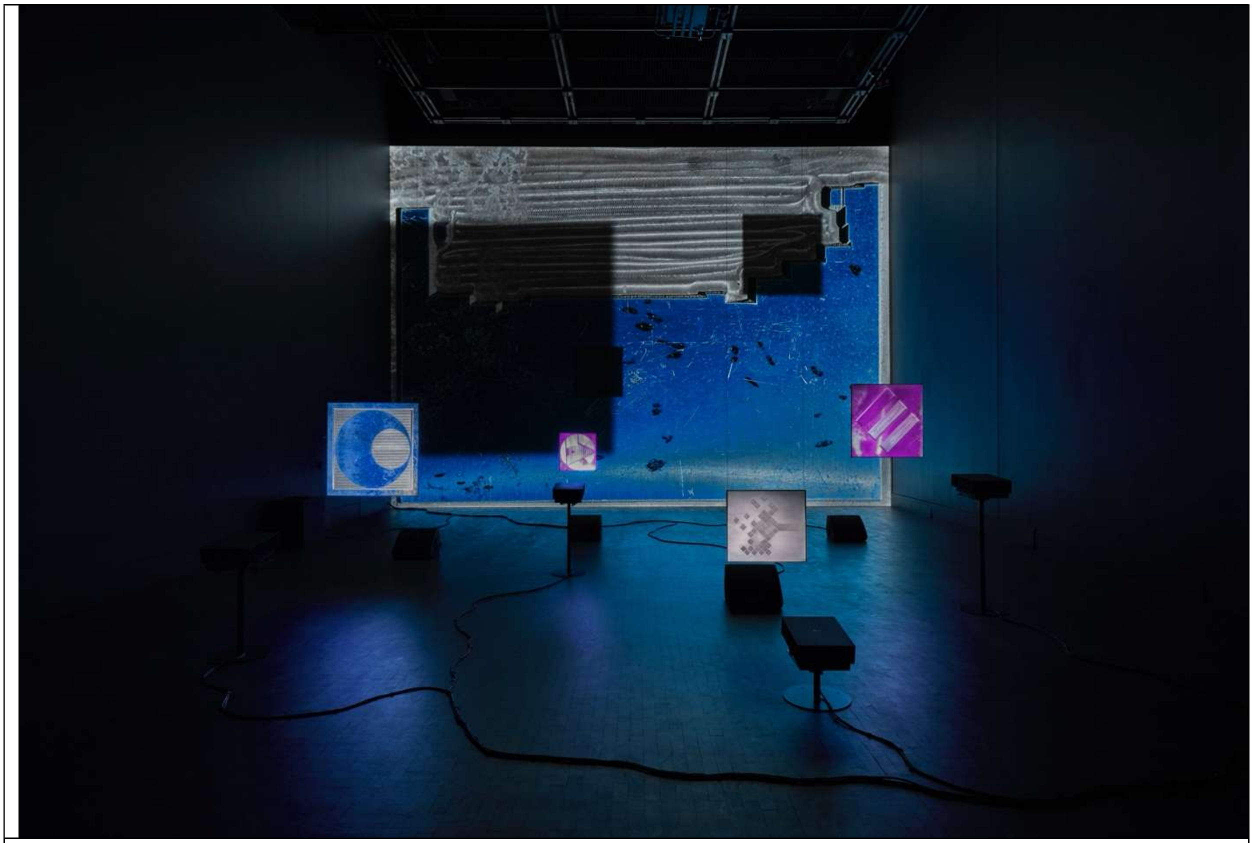

Figure 25: Alexandre Estrela. *Flat Bells*, 2023. Five-channel computer-generated video, eight-channel audio, and four engraved aluminum plates. Dimensions Variable. The Museum of Modern Art, New York. Fund for the Twenty-First Century, 2024. © 2026 Alexandre Estrela. Installation view, *Alexandre Estrela: Flat Bells*, The Museum of Modern Art, New York (November 4, 2023–January 7, 2024). Digital image © 2026 The Museum of Modern Art, New York. Photo: Jonathan Dorado.

Readers are encouraged to navigate to the MoMA webpage at https://www.moma.org/calendar/exhibitions/5562 to watch a YouTube video of *Flat Bells* to better understand the artwork and the role of sound.

Flat Bells (2023. Five-channel computer-generated video, eight-channel audio, and four engraved aluminum plates, dimensions variable) by Alexandre Estrela (b. 1971) at the Museum of Modern Art in New York City (Estrela 2023) is an example of a multi-media installation that integrates machine learning through one of its components. (See Fig. 25.) The artist and his engineer made use of a machine learning library, FluCoMa, designed for media artists who work with sound (FluCoMa 2025). In these cases, conservation of the specific elements of the installation is not necessarily straightforward and requires conservators to undertake further research.

Defining a taxonomy for artworks created using machine learning according to the final presentation poses important questions when collecting these works: what exactly gets sold with respect to digital and/or physical objects with each artwork and how many of each, if more than one format is collected? This taxonomy is also crucial for the conservators who install, de-install, and care for these works. While the conservation for the standard presentation formats, e.g., digital video, silkscreen prints, and others, is well-documented, there may be significant challenges in caring for the multimedia installations that incorporate machine learning elements in whole or in part.

TAXONOMY BASED ON DATA SOURCE

With respect to data sources and machine learning models, following is a description of a taxonomy based on three categories of data sources. Each of these categories is associated with questions of data and privacy; data and ethics; the technology and resources required to manage the datasets; and other data-related issues. These three categories include proprietary machine learning models; open-source machine learning models; and artist-curated datasets for building bespoke machine learning models.

The “Artificial Imagination” DALL•E-inspired art exhibition in San Francisco mentioned above (bitforms gallery 2022) exhibited artworks which were created using DALL•E 2, proprietary software running a machine learning model that was trained on unfiltered data from the Internet. DALL•E 2’s reliance on public datasets therefore influences the results, leading to algorithmic bias. Artists working in this medium are therefore working “blind” from the point of view of not being familiar with the training model sources and relying on proprietary software to generate their images.

Tom White’s “Maximum Activations,” shown at the Gazelli Art House (White 2025), is based on an open-source documented machine learning model. (See Fig. 23.) The artist includes websites as a presentation format for his artworks (White Online 2025) and posts a link on each artwork website that leads to documentation where the artist specifies the machine learning model that he used; in this case, the artist used the gemma-2-2b-it model on Hugging Face in “Maximum Activations” (Hugging Face Gemma-2-2b-it 2025). This model in turn is an open-source model and users can dive deeper into its contents and read the documentation on the limitations, bias, and other contextual aspects of this model. Artists using open-source machine learning models such as those available on the Hugging Face website have access to a great deal of information about the model they are using, in comparison with the artists using DALL•E or other proprietary models who cannot study their models in detail.

The third category in this taxonomy includes artist-curated datasets for building bespoke machine learning models. Héctor González (b. 1978) explores working with data through his artworks and text project “Computational Aesthetics: The Creation of Philosophical Forms with Artificial Intelligence” by drawing upon a vast database of thousands of images of Ancient Greek artworks curated by the artist (González 2020).

Figure 26: Anna Ridler, *Mosaic Virus* (2019, 3-screen GAN video installation.) Source: https://annaridler.com/mosaic-virus .

The English artist Anna Ridler wrote, "I am aware that the control that comes to me in this process really comes from what I do with the dataset. As someone who is interested in the hidden and the forgotten, it makes me uncomfortable to use someone else's dataset without properly exploring it. By creating my own dataset, it forces me to examine each image and inverts the usual process for creating this type of dataset." To this end, this artist curates her own datasets such as in her work *Mosaic Virus* (2019, 3-screen GAN video installation; see Fig. 26). Ridler uses Python libraries to generate images from her bespoke machine learning models (Ridler 2019).

Figure 27: Anna Ridler, *Myriad (Tulips)*, (2018). C-type digital prints with handwritten annotations, magnetic paint, magnets. Source: https://annaridler.com/myriad-tulips .

Ridler further created an artwork specifically based on the dataset she used for *Mosaic Virus* which she calls *Myriad (Tulips)* (2018. C-type digital prints with handwritten annotations, magnetic paint, magnets; see Fig. 27). The artist writes that "Myriad (Tulips) is an installation of thousands of hand-labeled photographs of tulips; these photographs were later used as the dataset for *Mosaic Virus 2018* and *Mosaic Virus 2019*. By choosing to make the dataset an artwork it draws attention to the skill, labour and time that goes into constructing it, whilst also helping to expose the human element in machine learning, usually hidden by algorithmic processes." (Ridler 2018).

Understanding an artist's approach to working with data and the sources for data and the machine learning models are important to a curatorial understanding of the artworks and a relevant technical art history. Conservators can capture the limited documentation, if any, on proprietary machine learning models and relevant documentation on open-source machine learning models. Bespoke datasets and models present challenges to collectors and institutions in conserving the artwork and in making decisions as to how and whether the dataset itself and the model should also be conserved as part of that conservation effort. Understanding the data and model source also helps collectors and conservators to raise and address issues of ethics, bias, privacy, limitations, and more.

CONCLUSIONS

These taxonomies will, I hope, inform and assist conservators to better understand and be able to care for these new and important works of art based on machine learning.

Of technical complexity and resources required: no programming skills required and minimal hardware and software requirements; intermediate-level programming/engineering skills with access to academic or online computational resources; advanced or professional-level programming/engineering skills and extensive hardware and software resources for running a complex resource-intensive computational environment.

Of presentation format: Digital video; Digital image or series of still images in digital format; Physical image(s), e.g., silkscreen or print; web page(s); multimedia installation; machine learning as one computational element in a larger time-based media art installation.

Of data source: Proprietary hidden machine learning model, without access to the data; open-source machine learning model with documentation; artist-curated dataset for building a bespoke machine learning model.

D. Graham Burnett, a science historian, wrote the following in a recent essay: "What it is like to be us, in our full humanity—this isn't out there in the interwebs. It isn't stored in any archive, and the neural networks cannot be inward with what it feels like to be you...That can only be lived.…This remains to us. The machines can only ever approach it secondhand.…The work of being here—of living, sensing, choosing—still awaits us. And there is plenty of it." (Burnett 2025). As artists continue to explore machine learning as a medium and push the boundaries of expression, art historians, curators, conservators and all those charged with interacting with and caring for these fragile artworks will be challenged to confront the boundaries of art and technology as well.

ACKNOWLEDGMENTS
Many thanks to Professors Jennifer Mass and Ivan Gaskell at the Bard Graduate Center; and to Professors Ernie Davis and Marsha Berger in the Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, for their guidance on my research and dissertation.

Many thanks to Thiago Hersan, Tandon School of Engineering, New York University, who co-taught the New York University Tandon School of Engineering class DM-GY 9103 Conservation of AI-Based Artworks in the Spring 2025 semester with me, as we encouraged students to experiment with some of these tools.

Many thanks also to my colleagues Agathe Jarczyk, Solomon R. Guggenheim Museum, New York; Patrica Falcão, Tate, London; Kate Lewis, Amy Brost and Peter Oleksik, Museum of Modern Art, New York; Joanna Phillips, Director, Conservation Center, Düsseldorf, Germany; Samrudha Dixit, Bard Graduate Center; and Sharon Nettles for their feedback and encouragement.

NOTES

Note 1. Metadata for the six images in Figure 14 retrieved along with the images.
[14-1] File Name=watercolor+of+pink+flowers+with+green+leaves_100.jpeg, Downloaded from https://c8.alamy.com/comp/2PYA80D/set-of-illustrations-of-watercolor-flower-bouquet-pale-pink-green-pink-flower-green-leaf-leaves-branches-of-bouquets-collection-wedding-station-2PYA80D.jpg .

[14-2] File Name=watercolor+of+pink+flowers+with+green+leaves_3c1.jpeg, Downloaded from https://img.freepik.com/premium-photo/watercolor-pink-flowers-green-leaves-bouquet-illustration-isolated_202769-4762.jpg .

[14-3] File Name=watercolor+of+pink+flowers+with+green+leaves_11.jpeg, Downloaded from https://img.freepik.com/premium-photo/watercolor-light-pink-flowers-green-leaves-bouquet-illustration-isolated-white-wedding_202769-6730.jpg .

[14-4] File Name=watercolor+of+pink+flowers+with+green+leaves_20a.jpeg, Downloaded from https://img.artpal.com/404302/145-22-4-4-7-54-30m.jpg .

[14-5] File Name=watercolor+of+pink+flowers+with+green+leaves_25c.jpeg, Downloaded from https://as1.ftcdn.net/v2/jpg/03/33/81/18/1000_F_333811889_IaHAdQ3i1bzZF9XyXLzYuKKuykngUtJc.jpg .

[14-6] File Name=watercolor+of+pink+flowers+with+green+leaves_84.jpeg, Downloaded from https://img.freepik.com/premium-photo/bouquet-made-pink-watercolor-flowers-green-leaves-wedding-greeting-illustration_202769-9147.jpg .

REFERENCES

Adobe. "Firefly [Software]." Accessed September 5, 2025. https://www.adobe.com/products/firefly/landpa.html .

American Artist. *Security Theater [Artwork]*. 2023. CCTV cameras, AI, acrylic, cable, 4K monitors, television mounts, computer, tandem sling seating, phone security pouches, messenger bags, aluminum, steel, hardware, desk, and signage, dimensions variable. https://www.guggenheim.org/audio/track/a-guide-to-engaging-with-security-theater-2023

Anadol, Refik. *Large Nature Model: Coral*. 2024. Immersive Installation at United Nations Headquarters. https://refikanadol.com/works/un-2024/ .

Anadol, Refik. *Unsupervised—Machine Hallucinations (Exhibition: Nov 19, 2022–Oct 29, 2023)*. 2022. Custom hardware and software (color, sound, generative algorithm with artificial intelligence), dimensions variable. https://www.moma.org/calendar/exhibitions/5535 .

Anadol, Refik. "Works." Refik Anadol. Accessed September 11, 2025. https://refikanadol.com/works/ .

bitforms gallery. "DALL•E: Artificial Imagination, October 26–Dec 29, 2022." Bitforms Gallery, 2022. https://bitforms.art/exhibition/dall·e-artificial-imagination/ .

bitforms gallery. "DALL•E: Artificial Imagination, October 26–Dec 29, 2022 [Catalog]." Bitforms Gallery, 2022. https://www.bitforms.art/wp-content/uploads/2024/09/bitforms-sf-group-artificial-imagination.pdf .

Burnett, D. Graham. "Will the Humanities Survive Artificial Intelligence?" *The New Yorker*, April 26, 2025. https://www.newyorker.com/culture/the-weekend-essay/will-the-humanities-survive-artificial-intelligence .

Diab, Mohamad, Julian Herrera, Musical Sleep, Bob Chernow, and Coco Mao. *Stable Diffusion Prompt Book*. OpenArt, 2022. https://openart.ai/promptbook .

Estrela, Alexandre. *Flat Bells*. 2023. Five-channel computer-generated video, eight-channel audio, and four engraved aluminum plates, dimensions variable. https://www.moma.org/calendar/exhibitions/5562 .

Freeman, Julie, Geraint Wiggins, Gavin Starks, and Mark Sandler. "A Concise Taxonomy for Describing Data as an Art Material." *Leonardo* 51, no. 1 (2018): 75–79. https://doi.org/10.1162/LEON_a_01414 .

González, Héctor. "Computational Aesthetics: The Creation of Philosophical Forms with Artificial Intelligence." 2020. https://drive.google.com/file/d/1R5EkmPqYrt3_j483wr0reyxfHvmbgU3v/view .

Google Colab. "Google Colab." Accessed February 1, 2025. https://colab.research.google.com .

Herndon, Holly, and Mat Dryhurst. *Embedding Study 1 & 2 (Xhairymutantx Series); Christie's Augmented Intelligence Sale*. 2024. Thermal dye diffusion transfer prints each: 47 3/4 × 71 5/8 in. (119 × 180 cm.) (2) Executed in 2024, this work is number one from an edition of three plus two artist proofs. https://onlineonly.christies.com/s/augmented-intelligence/holly-herndon-b-1980-mat-dryhurst-b-1984-3/249745 .

Hugging Face. "Gemma-2-2b-It Model." Accessed October 20, 2025. https://huggingface.co/google/gemma-2-2b-it .

Hugging Face. "Inpainting." Accessed September 9, 2025. https://huggingface.co/docs/diffusers/en/api/pipelines/stable_diffusion/inpaint .

Hugging Face. *Stable Diffusion 2*. n.d. Accessed September 10, 2025. https://huggingface.co/stabilityai/stable-diffusion-2 .

Hugging Face. "The AI Community Building the Future." Accessed February 19, 2025. https://huggingface.co/ .

LeWitt, Sol. *Wall Drawing 273*. September 1975. Graphite and crayon on seven walls, dimensions variable. https://www.sfmoma.org/artwork/FC.512/ .

Nozem, Ivan Filho, PublicPrompts, and Coco Mao. *Model Training Book for Beginners*. OpenArt, 2023. https://openart.ai/modeltrainingbook .

OpenArt. "OpenArt." 2025. https://openart.ai .

Ridler, Anna. *Mosaic Virus*. 2019. 3-screen GAN video installation. https://annaridler.com/mosaic-virus .

Ridler, Anna. *Myriad (Tulips)*. 2018. C-type digital prints with handwritten annotations, magnetic paint, magnets. https://annaridler.com/myriad-tulips .

San Francisco Museum of Modern Art (SFMOMA). "Drawing with Instructions: Inspired by Sol LeWitt's Wall Drawing 273." San Francisco Museum of Modern Art (SFMOMA), May 2020. https://www.sfmoma.org/read/drawing-with-instructions/ .

Simon Jr., John F. *Color Panel 1.0, Edition 3/12*. 1999. Software, modified Apple Macintosh Powerbook 280c, and acrylic (plastic), Overall: 13 1/2 × 10 1/2 × 3 in. (34.3 × 26.7 × 7.6 cm). https://whitney.org/collection/works/12273 .

The FluCoMa Project. *FluCoMa*. Released September 8, 2025. https://www.flucoma.org/ .

United Nations General Assembly (UNGA). *Unprecedented AI-Powered Art Brings Digital Environmentalism to the United Nations [Press Release]*. Press Release. United Nations, 2024. https://www.un.org/sustainabledevelopment/blog/2024/09/press-release-large-nature-model-coral-unga-exhibit/ .

University of Pittsburgh Library System. "Metadata & Discovery @ Pitt [Course & Subject Guides]." University of Pittsburgh Library System. Accessed September 15, 2025. https://pitt.libguides.com/metadatadiscovery/controlledvocabularies .

White, Tom. *Maximum Activations*. 2024. Two colour hand silkscreen, acrylic on paper accompanied by an interactive online interface and ERC-721 token. https://gazell.io/collections/tom-white-maximum-activations .

White, Tom. "Maximum Activations: Online." Gazelli Art House, 2025. https://got.drib.net/maxacts/online/100.html .

CONTACT

Prof. Deena Engel
Clinical Professor Emerita
Department of Computer Science
Courant Institute of Mathematical Sciences
New York University
New York, New York, USA
Mobile: 646-263-0196
Email: deena.engel@nyu.edu