

Hertziana Studies
in Art History



BIBLIOTHECA HERTZIANA
MAX PLANCK INSTITUTE
FOR ART HISTORY

From Hype to Reality

Artificial Intelligence in the Study of Art and Culture

Edited by Eva Cetinić & Darío Negueruela del Castillo



BIBLIOTHECA HERTZIANA
MAX PLANCK INSTITUTE
FOR ART HISTORY



Hertziana Studies in Art History

Editors

Tanja Michalsky & Tristan Weddigen

Managing Editor

Elisa Bastianello

All articles of the present volume have undergone an anonymous scholarly peer-review by international experts to whom we are grateful for their generous support.

Graphic Project

Sergio Polano (2019)

Michele Elia (2023)

Page Layout

Michela Maguolo

ISBN | 978-3-949381-02-7

DOI | <https://doi.org/10.48431/hsah.03>

Published online on 31 March 2024

© 2024 Bibliotheca Hertziana

Cover image: Eva Cetinić's prompt for DALL·E: "The persistence of memory in artificial intelligence, abstract, double exposure, award winning photography, colorful, macro-shot."



From Hype to Reality

Artificial Intelligence in the Study of Art and Culture

edited by

Eva Cetinić & Darío Negueruela del Castillo

Contents

7

Eva Cetinić

Introduction

13

Amanda Wasielewski

The Reification of Style in AI Image Generation

23

Nuria Rodríguez-Ortega

Exploring the Concept of Transduction to Interpret AI-Generated Images

31

Bokar Lamine N'Diaye

Stochastic Parrots in the Ancient City?

39

Dejan Grba

Incidental Reverberations: Poetic Similarities in AI Art

47

Arif Kornweitz

Inside the High-Dimensional White Cube

57

Dominik Vrabič Dežman

Interrogating the Deep Blue Sublime: Images of Artificial Intelligence in Public Media

69

Pepe Ballesteros Zapata

Evolving Methodologies: Computation in Art History

75

Valentine Bernasconi

Automated Body Pose Estimation on Painting: A Crisis of Uncertainty

83

Piera Riccio & Nuria Oliver

A Techno-Feminist Perspective on the Algorithmic Censorship of Artistic Nudity

91

Anne-Laure Oberson

Making Sense of AI Generated Images in the Process of Imagination

Abstract

Amanda Wasielewski

The Reification of Style in AI Image Generation

Keywords | AI technologies, Digitized artworks,
Text-to-image generation tools



Over the past decade, an active subfield in computer vision and artificial intelligence research has sought to identify and assign style in image data. Collections of digitized artworks have been central to much of this research. More recently, text-to-image generation tools, such as DALL·E, Midjourney, and Stable Diffusion, have made the replications of artistic style central to their operation. In art history, style is a highly contested category and the terminology used is understood to be historically situated. In other words, style terms are dependent on context, reflecting the differing purposes of art historians, artists, or critics at different points in history. When style terms become data and enter into automated processes through machine learning, these nuances are often lost. This paper addresses how contemporary artificial intelligence methods reify the Western concept of style and what that means for the study of visual culture.

Cite as: Amanda Wasielewski, "The Reification of Style in AI Image Generation" in *From Hype to Reality: Artificial Intelligence in the Study of Art and Culture*, ed. Eva Cetinić, Darío Negueruela del Castillo, 2024, *Hertziana Studies in Art History*, vol. 03, 2024, pp. 13–20. Doi: <https://doi.org/10.48431/hsah.0302>.

1. Introduction

The minimal interface of Open AI's popular text-to-image generation platform DALL·E shows, at the time of writing, a blank off-white background and a text input box (fig. 1). Inside the text box, there is a sample input prompt: "An Impressionist oil painting of sunflowers in a purple vase..." Between this opening suggestion and the name of the system itself (a play on the names of Surrealist artist Salvador Dalí and Pixar's robot character WALL·E), it is clear that fine art is an essential part of Open AI's branding for DALL·E. How is art instrumentalized here and to what ends?

As I have written elsewhere,¹ fine art can have a sanitizing effect when deployed in products, services, or public projects by corporations or governments and is often used to make them more palatable to mainstream audiences or middle-class tastes. This is known as "artwashing" and can, in recent times, be seen in the deployment of art in AI and machine learning projects. Beyond its general public relations utility, however, art (and the study of art history) has a more immediate, concrete use for text-to-image generators: it supplies a range of style terms and visual forms to go with them.

As DALL·E's sample prompt suggests, specifying an image style is a core component of the prompt, i.e., the text that initiates the image generation process. Of course, style terms in this context expand far beyond those used in the canon of art history. However, given that art historians have long debated the concept of visual style,² this disciplinary perspective serves as a useful starting point to begin to understand what style means for text-to-image generators like DALL·E. This article argues that such systems reify artistic styles, stripping them of their historical context and the sense of time and place essential to understanding them. Every style is thus treated as interchangeable and equivalent to one another.³ Style terms become objective qualities assigned to image data in training sets, rather than debatable and competing subjective categories as they are in the history of art. They are thus fused to their digital image-based representation and its visual and material characteristics alone.

The consequences of this reification process can be subtle. For example, it may replicate stereotypes and biases from the historical moments in which they were coined, or it may erase the political or social dimension of a particular style. On a deeper level, it strips style terms of meaning and context, along with the artworks grouped under them.

¹ Wasielewski 2023, pp. 1–2.

² Riegl 1893, Schapiro 1953, Persinger 2007,

Cahn 2002, Gombrich 1960, Wölfflin 1915, Lang 1987.

³ Wasielewski 2023, p. 50, Meyer 2023, p. 107.

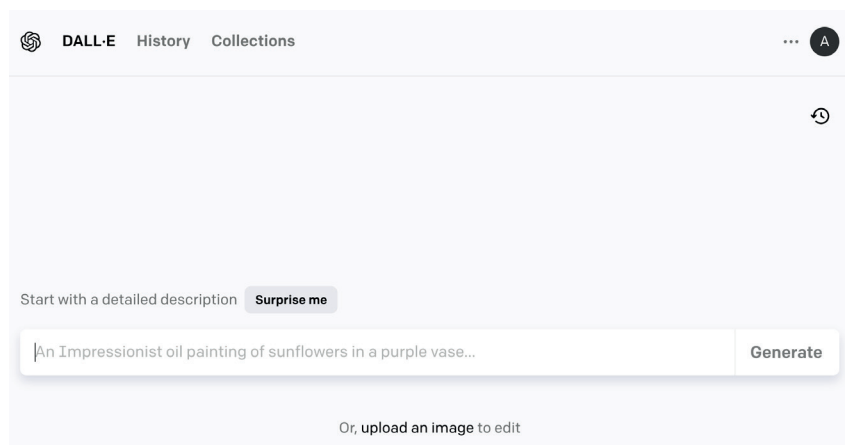


Figure 1. The opening screen of the DALL·E text-to-image generation platform.
Source: Open AI (<https://openai.com/dall-e-2>).

2. Style as Generative

Art historians have, since the nineteenth century, tasked themselves with cataloging and organizing artworks, grouping them under a given textual term to indicate a particular artistic style.⁴ Style terms have also historically originated with art critics or artists themselves.⁵ Thus style is not a unified concept but rather a reflection of disparate motivations, both in the moment the artwork is created and retrospectively. Style terms can indicate time periods (of varying lengths), visual qualities, or contextual factors around the creation of the work. However, there is no consistent or methodical standard for defining what style is exactly. When it is framed as a singular category, comparing these terms or the works grouped under them to one another can sometimes be akin to comparing proverbial apples and oranges.

Style describes not only the form something takes but implies an agent or cause, in other words, a kind of intentionality. In referring to Richard Wollheim, Richard Neer writes: “Individual style is [...] *generative*: as part of Rembrandt’s technical repertoire, it stands causally to his output.”⁶ This contrasts sharply with much of the contemporary rhetoric around AI, which holds that forms created through generative AI do not have intentionality. Emily Bender et al. caused a tech industry firestorm with their now-infamous “stochastic parrot” article, which argues that large language models

are very good at imitating form but have no “communicative intent.”⁷ Looking at the case of image generation and the unification of text and image in multimodal models, this raises the question: does reducing artists’ individual styles to taxonomic categories fundamentally disconnect their former generative potential?

Or, phrased another way: does style still mean something apart from the context of its development? Is a contemporary painting done in an Impressionist style an Impressionist painting? Most art historians would likely answer that it is not, as the painting is not grounded in the late nineteenth-century Parisian milieu where the style originated. However, as Neer argues, “saying something is baroque suggests nothing about its origins [...] a fake Tiepolo is, potentially, just as baroque as a real one” (fig. 2).⁸ The category of Post-Impressionism exists in the first place because of the widespread and far-flung influence of Impressionism itself, which spawned both imitators and opponents (all of whom were labeled Post-Impressionists).⁹ Are contemporary imitators, including those who use generative AI, to be wholly dismissed? In this new landscape of image generation, style terms are not used primarily as a kind of content or meaning in and of themselves but, rather, act as empty vessels to receive the content provided in the text prompt.

The design of the DALL·E interface is reminiscent of Google and other search engines,

4 Wölfflin 1915, Riegl 1893.

5 Greenberg 2003, Greenberg 1964, Fry 1911.

6 Neer 2005, p. 11.

7 Bender et al. 2021. For more on this from the perspective of text, see also the debate articles in the journal *Critical Inquiry*’s blog on Computational Literary Studies: [https://](https://critinq.wordpress.com/2023/06/)

critinq.wordpress.com/2023/06/.

8 Neer 2005, p. 11, n. 28.

9 Rose 2019.



Figure 2. An AI-generated “fake” Tiepolo image created with Midjourney using the prompt “a Baroque oil painting by Giovanni Battista Tiepolo depicting the Apotheosis of the Spanish Monarchy.”

Source: generated by the author, April 2023.

but the process the user initiates in this case is different.¹⁰ Generators like DALL·E do not search existing data points but rather create complex syntheses based on them. In order to do this, the machine-learning model must be ‘trained’ or ‘learn’ to connect text and image data together.¹¹ Like a search engine, DALL·E can only work with the data it has. This means that it cannot generate something that is not represented in the text and image pairings of

the dataset. However, unlike the typical search engine, the output it generates is not identical to any individual data point or pairing. Roland Meyer calls this process “generative search.”¹²

However, perhaps ‘search’ is not the most effective conceptual framing of these technologies. Googling “a painting of sunflowers” invariably brings up Vincent van Gogh’s famous works from the late nineteenth

¹⁰ Meyer 2023, p. 103.

¹¹ Radford et al. 2021.

¹² Meyer 2023, p. 106.



Figure 3. A selection of output images from DALL-E using the prompt “an Impressionist oil painting of sunflowers in a purple vase.” Source: generated by the author using DALL-E 2, April 2023.

century. DALL-E’s opening prompt seems to invite the user to recreate this motif as if it had been depicted by other artists. In practice, the output still suggests Van Gogh (fig. 3), not only in the sunflowers themselves but in the style of vases, the thickness of the use of line, and the layering of complementary colors.

The overwhelming fame of Van Gogh’s sunflowers factors in here. Scraping the internet for images, as Open AI and other companies have done to produce the datasets behind text-to-image generators, would not only come up with the one entry of a painting in a museum’s online catalog but many copies and iterations elsewhere around the web. Prompting for a less iconic motif would likely less directly resemble a particular existing painting or set of paintings. The Impressionist painter Claude Monet actually painted a vase of sunflowers in 1881, about seven years before Van Gogh’s attempts at the motif.¹³ However, the weight of the fame of Van Gogh’s exemplars is, in many ways, reflected in the end depictions from DALL-E, reminding us that AI-generated depictions are

still the product of averages and accumulations rather than singular or distinctive entities.

Yet, this statistical weighing of image content is further complicated by the fact that, apart from the canonical exemplars of sunflower paintings that can be found online, there are countless images of “fake” Impressionist paintings (to use Neer’s characterization) that will end up associated with any given style term in the training of a model. In other words, a web scrape would likely collect more images of amateur paintings than canonical ones. These amateur works may be labeled as Impressionist or other historic style terms. A search of the open source LAION dataset (fig. 4), which was used to train the DALL-E competitor Stable Diffusion, using the terms “oil painting sunflowers purple vase” shows primarily amateur artworks. Many of these are created in styles that could be deemed Impressionist or Post-Impressionist. The way that these models are trained, therefore, means that a “fake” Impressionist painting is, indeed, an Impressionist painting. No distinction is made.

¹³ Bailey 2019, p. 68.

Figure 4. A search of LAION 5B-H-14 for “oil painting sunflowers purple vase.”
Source: <https://rom1504.github.io/clip-retrieval/>

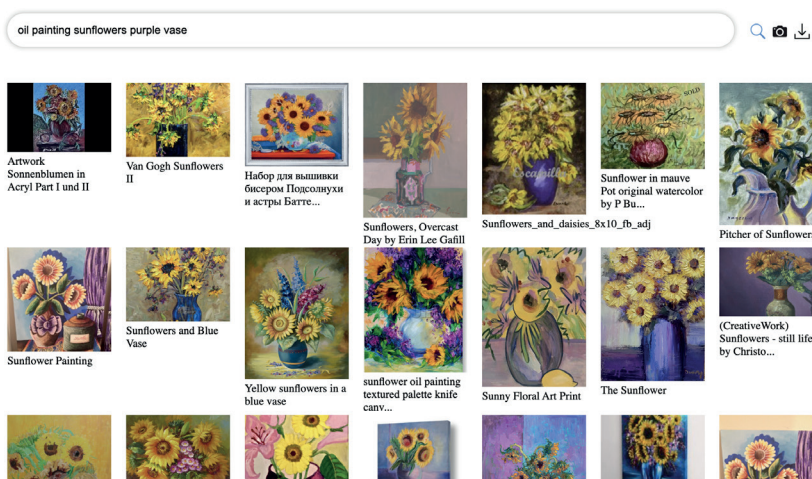


Figure 5. DALL·E output images created using the prompts “A Cubist oil painting of sunflowers in a purple vase,” “An Expressionist oil painting of sunflowers in a purple vase,” and “A Renaissance oil painting of sunflowers in a purple vase” (L-R) respectively.
Source: generated by the author, April 2023.



Figure 6. DALL·E output images created using the prompts “Vaporwave sunflowers in a purple vase” and “Claymation sunflowers in a purple vase” (L-R) respectively.
Source: generated by the author, April 2023.



Analyzing the anatomy of prompts in text-to-image generators, it seems that style is designed to be a filter rather than content as such. The word “Impressionist” in DALL·E’s sample prompt is endlessly interchanged with other style terms, not only Western art history terms, such as Cubism, Expressionism, Renaissance, etc. (fig. 5), but also more recent popular culture and design terms, such as vaporwave and claymation (fig. 6). Something as complex and context-dependent as style thus becomes, when entered in a dataset, static and unchanging. The image data connected to such terms is similarly frozen in relation to them.

The naming of categories or groups of people and objects matters. There is a reason most

style categories are assigned evocative names rather than, for example, numbers.

The conventions for naming styles and artistic periods constantly evolve and change over time as new thinkers and cultural commentators respond to and revise what came before them. For example, the term “Pre-Columbian” has been (and still often is) widely used in art history to group the cultural products and artworks of Central and South America that were made before Christopher Columbus crossed the Atlantic in 1492. In recent decades, however, scholars have questioned what it means to describe a hugely diverse cultural heritage produced over millennia in terms that imply European influence—and one man’s



Figure 7. A selection of DALL-E output images created using the prompt “An Impressionist oil painting.”

Source: generated by the author, April 2023.

influence in particular—as the most important factor to consider when describing it.¹⁴

Adding an additional wrinkle to this discussion of style for text-to-image generators is the issue of default style. When style terms are absent or no style is specified, text-to-image generators tend to default to a form that evokes the quotidian documentary mode of digital photography.¹⁵ This is unsurprising given that the vast majority of image training data available today is digital photographic imagery (i.e., images of artworks are also photographs but do not have a photographic ‘look’). As Roland Barthes wrote, “a photograph is always invisible: it is not it that we see.”¹⁶ In computer science, photographic images are referred to as ‘natural images’, which re-enforces the idea that photography is not a representational, human-inflected process but rather a direct transfer of nature to the image without mediation. Nevertheless, photography remains one of the many representational forms available to us and acts as a sort of default style for text-to-image generators.

The typical prompt for a text-to-image generator requires words that indicate both style and

content. Lev Manovich asserts that not all style and content pairings produce successful results, namely, if they are too “different” from one another.¹⁷ Nevertheless, style and content are the typical components of a prompt. In a pinch, style can be disregarded in favor of simple content, relying on the “photographic” default style. The same is not true if we eschew content in favor of style terms, however. Shortening DALL-E’s sample prompt to “An Impressionist oil painting,” we see that the words “oil painting” have become the content or the object of the prompt (fig. 7). Simplifying further to just “Impressionist” produces a confusing array of responses that are sometimes reminiscent of Impressionist artwork but often not (fig. 8). This exercise further demonstrates the extent to which reified style has little meaning apart from form. That is, a filter struggles to produce its effect when it is divorced from any content. However, style terms *do* have content, context, and meaning beyond their visual effects. In text-to-image generators, the meaning of style and its generative potential is collapsed.

3. Conclusion

The training datasets for popular text-to-image generators like DALL-E are time capsules. The further we move away from the time of the dataset’s collection, the more this will become apparent. The proliferation of AI-generated content online means that future data scraping activity will be tainted by the effects of recursion. According to Shumailov et al., this will lead to “model collapse.”¹⁸ Any errors in the output of AI-generation models will thus be continually amplified, as more synthetic content is used to train them, to the extent that the products of the models may be rendered completely unusable.

In formulating the concept of model collapse, Shumailov et al. were inspired by the concept of “mode collapse” in generative adversarial networks (GANs) and the issue of “catastrophic forgetting” in training neural

¹⁴ Klein 2002.

¹⁵ One notable exception to this was the text-to-image generation tool Midjourney prior to version 5. With the release of version

5, its default image style was photographic in the absence of other style cues. In release v.5.1, this aspect was dialed back again.

¹⁶ Barthes (1980) 2000, p. 6.

¹⁷ Manovich 2023, p. 40.

¹⁸ Shumailov et al. 2023.



Figure 8. A selection of DALL-E output images created using the prompt “Impressionist.”
Source: generated by author, April 2023.

networks. Commentators on social media have begun to call this idea “Habsburg AI,” after the famously inbred European monarchical dynasty that was affected by multiple hereditary diseases. Over generations of inbreeding, the genetic defects of the Habsburgs became amplified. So too, the theory goes, may the defects of AI-generated content.

What this means for style is that even more variability, nuance, context, and history may be lost. One might assume that AI-generated style will be less easily identified as ‘wrong’ or ‘faulty’ compared to a string of text or a naturalistic depiction of an object or person. However, if enough training recursion occurs, style too may begin to lose its specificity or distinction. The challenge for using and replicating style in machine learning applications lies, therefore, in addressing its dynamic, evolving, and interconnected relationships. Currently, AI applications of style too often flatten it to an ahistorical and static object class.

References

- Bailey 2019**
Martin Bailey, *The Sunflowers Are Mine: The Story of Van Gogh's Masterpiece*, White Lion Publishing, London, 2019.
- Barthes 2000**
Roland Barthes, *Camera Lucida. Reflections on photography*, Vintage, London, 2000.
- Bender et al. 2021**
Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery 2021, pp. 610–623. Doi: <https://doi.org/10.1145/3442188.3445922>.
- Cahn 2002**
Walter B. Cahn, "Schapiro and Focillon," *Gesta, International Center of Medieval Art* vol. 41, 2, (2002), pp. 129–136. Doi: <https://doi.org/10.2307/4126578>.
- Fry 1911**
Roger Fry, "Post-Impressionism," *The Fortnightly Review*, (1911), pp. 856–867. URL: <https://fortnightlyreview.co.uk/2017/10/post-impressionism/>.
- Gombrich 1960**
Ernst Hans Gombrich, *Art and Illusion. A Study in the Psychology of Pictorial Representation*, Phaidon, London, 1960. URL: https://archive.org/details/artillusionstudy-0000gomb_09d9. Doi: <https://doi.org/10.1515/9780691252742>.
- Greenberg 1964**
Clement Greenberg "Post Painterly Abstraction," *Art International*, vol. 8, 5-6, (1964), pp. 63–65.
- Greenberg 2003**
Clement Greenberg, "Modernist Painting," in *Art in theory, 1900-2000: An anthology of changing ideas*, ed. Charles Harrison and Paul Wood, Blackwell, Malden, MA, 2003, pp. 773–779.
- Klein 2002**
Cecelia F. Klein, "Not Like Us and All the Same: Pre-Columbian art history and the construction of the nonwest," *RES: Anthropology and Aesthetics*, [President and Fellows of Harvard College, Peabody Museum of Archaeology and Ethnology] 42, (2002), pp. 131–138. URL: <https://www.jstor.org/stable/20167573>.
- Lang 1987**
The Concept of Style. Revised and Expanded Edition, ed. Berel Lang, Cornell University Press, Ithaca, NY, 1987.
- Manovich 2023**
Lev Manovich, "AI Image Media Through the Lens of Art and Media History," *IMAGE: Zeitschrift für interdisziplinäre Bildwissenschaft*, vol. 37, 1, (2023), pp. 34–41. URL: <https://image-journal.de/ai-image-media-through-the-lens-of-art-and-media-history/> Doi: <https://doi.org/10.1453/1614-0885-1-2023-15448>.
- Meyer 2023**
Roland Meyer, "The New Value of the Archive: AI Image Generation and the Visual Economy of 'Style,'" *IMAGE: Zeitschrift für interdisziplinäre Bildwissenschaft*, vol. 37, 1, (2023), pp. 100–111. URL: <https://image-journal.de/the-new-value-of-the-archive/> Doi: <http://doi.org/10.1453/1614-0885-1-2023-15458>.
- Neer 2005**
Richard Neer, "Connoisseurship and the Stakes of Style," *Critical Inquiry*, The University of Chicago Press vol. 32, 1, (2005), pp. 1–26. Doi: <https://doi.org/10.1086/498001>.
- Persinger 2007**
Cynthia L. Persinger, *The politics of style: Meyer Schapiro and the crisis of meaning in art history* Ph.D Thesis, University of Pittsburgh 2007. URL: <https://d-scholarship.pitt.edu/10313/1/main-file-etd-12112007-110642.pdf>.
- Radford et al. 2021**
Alec Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," in *Proceedings of the 38th International Conference on Machine Learning*, PMLR 2021, pp. 8748–8763. URL: <https://proceedings.mlr.press/v139/radford21a.html>.
- Riegl 1893**
Alois Riegl, *Stilfragen: Grundlegungen zu einer Geschichte der Ornamentik*, G. Siemens, Berlin, 1893. Doi: <https://doi.org/10.11588/diglit.8075>.
- Rose 2019**
Sam Rose, *Art and Form: From Roger Fry to Global Modernism*, Penn State University Press, University Park, 2019.
- Schapiro 1953**
Meyer Schapiro, "Style," in *Anthropology Today: An Encyclopedic Inventory*, ed. Alfred Louis Kroeber, University of Chicago Press, Chicago, 1953, pp. 287–311.
- Shumailov et al. 2023**
Iliia Shumailov et al., "The Curse of Recursion: Training on Generated Data Makes Models Forget," arXiv 2023. Doi: <https://doi.org/10.48550/arxiv.2305.17493>.
- Wasielewski 2023**
Amanda Wasielewski, *Computational Formalism: Art History and Machine Learning*, The MIT Press, Cambridge, MA, 2023. URL: <https://direct.mit.edu/books/oa-monograph/5587/Computational-Formalism-Art-History-and-Machine>. Doi: <https://doi.org/10.7551/mitpress/14268.001.0001>.
- Wasielewski 2024**
Amanda Wasielewski, "The Reification of Style in AI Image Generation," in *From Hype to Reality: Artificial Intelligence in the Study of Art and Culture*, ed. Eva Cetinić and Darío Negueruela del Castillo, 2024, Doi: <https://doi.org/10.48431/hsah.0302>.
- Wölfflin 1915**
Heinrich Wölfflin, *Kunstgeschichtliche Grundbegriffe: Das Problem Der Stilentwicklung in der neueren Kunst*, Bruckmann, Munich, 1915. URL: <https://digi.ub.uni-heidelberg.de/diglit/woelfflin1915>. Doi: <https://doi.org/10.11588/diglit.27250>.

