

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

Eva Cetinić

Introduction

Keywords | AI, Digital humanities, AI Arts,
Computer sciences



As technologies related to AI (artificial intelligence) are progressing intensively, the hype surrounding the acronym “AI” have become an evolving but persistent phenomenon. However, the reality of working with AI technologies often remains distant from its idealized representation and involves many constraints and difficulties, which become particularly obvious in interdisciplinary research contexts. In recent years, the critical discourse on AI has also gained momentum, partly due to the increasing interest of scholars in the arts and humanities in exploring AI technologies and the implications of their widespread integration. This collection comprises short articles from scholars who undertook the challenging task of examining various aspects of AI in the context of art and culture. It combines approaches that explore how AI technologies can be applied to study art and culture, as well as how arts and humanities-based theories and methodologies contribute to contextualizing and critically assessing AI technologies. This volume consolidates contributions from scholars from diverse disciplines, including art history, media studies, philosophy, anthropology, digital humanities, and computer science. It offers a unique selection of insightful articles that provide valuable cross-disciplinary perspectives, fostering a deeper understanding of the dynamic intersection between AI, art, and culture.

Cite as: Eva Cetinić, “Introduction,” 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. 7–10. Doi: <https://doi.org/10.48431/hsah.0301>.

From Hype to Reality

The hype associated with the acronym “AI” has been a persistent phenomenon emerging in various contexts over the last several years. As the development of technologies related to AI has been progressively accelerating in the last decade, we have witnessed a growing interest in these technologies and their integration in our scientific, economic, social, and cultural frameworks. Technological breakthroughs, usually followed by intensive media coverage and often overhyped representations of AI technologies, have driven both the genuine fascination and the sensationalism surrounding AI. Meanwhile, the technologies themselves have reached a quite sensational level of innovation, particularly in the last two years with the most recent advances in generative and multimodal AI techniques. What was considered to be just a hype only a few years ago, recently became a technological reality. In parallel with the advancements of technologies, the hyped narrative surrounding AI has also kept evolving. For example, in 2018, with the rise of popularity of image-based generative AI models (e.g. the famous GAN model),¹ machines were granted “the gift of imagination.”² Nowadays, as generative AI models are progressing rapidly in producing semantically and visually appealing synthetic content, the impact of these models on human creativity and imagination have started to become a more seriously discussed topic. Meanwhile, the media attention shifted to artificial general intelligence and to even more radical speculations of attributing advanced human traits to

AI, like consciousness for example.³ And while consciousness has become the new imagination in the hyped narrative of AI media coverage, the reality of working with AI technologies is still far from its idealized portrayal and involves many limitations and challenges. In parallel with the techno-enthusiastic hype, the critical discourse on AI has also progressed and intensified in the last few years. One of the reasons for this trend is the growing interest of scholars from the arts and humanities in exploring the phenomenon of AI and the implications of its widespread adoption.

The tendency to integrate computational methods and digital tools in arts and humanities studies has been ongoing for more than a decade and has resulted in numerous research and educational initiatives united under the umbrella term “digital humanities.” As a result, a dichotomy between the “digital” and “traditional” modes of scholarship was established and often prevailed as the main starting point in discussing the role and impact of technology in the study of art and culture. However, the recent advancements in AI and the enormous public attention that was triggered by technologies such as large language models (e.g. ChatGPT)⁴ or text-to-image models (e.g. DALL·E,⁵ Midjourney,⁶ and Stable Diffusion⁷) contributed to a rise of interest in AI among scholars belonging to both the “digital” and “traditional” factions. The perspective on the use of AI in the context of arts and humanities therefore has started to transition from the initial opposition between (techno-) enthusiasm and (techno-) reluctance, towards

1 Goodfellow et al. 2015.

2 Giles 2018. URL: <https://www.technologyreview.com/2018/02/21/145289/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/> (last accessed: January 9 2024).

3 Huckins 2023. URL: <https://www.technologyreview.com/2023/10/16/1081149/ai-consciousness-conundrum/> (last accessed: Jan. 9 2024).

4 OpenAI 2022. URL: <https://openai.com/blog/chatgpt/> (last accessed: Jan. 9 2024).

5 OpenAI 2021. URL: <https://openai.com/research/dall-e> (last accessed: Jan. 9 2024)

6 Midjourney 2023. URL: <https://www.midjourney.com/> (last accessed: Jan. 9 2024).

7 Rombach et al. 2022.

an increased awareness of the necessity for collaborative critical reflection.

However, establishing and developing a meaningful theoretical contextualization and critique of the various aspects of AI technologies requires transcending the boundaries of one research discipline and engaging in an interdisciplinary exploration. Although “interdisciplinarity” is often used as an appealing buzzword, the actual practice of interdisciplinary research at the intersection of AI and arts and humanities reveals numerous challenges such as cross-domain knowledge gaps, terminological misunderstandings, oversimplifications, a highly dynamic pace of development, lack of transparency, issues of explainability, and the AI “black box” problem. The attempt to overcome and critically engage with such challenges relies on nurturing cross-disciplinary dialogues. It also requires scholars to take the risk of stepping out of the comfort zone of their core disciplines and it might even require a readiness to engage in speculative and fragmentary explorations in order to tackle the often unexplainable and transient manifestations of constantly progressing AI technologies.

This volume presents a collection of short articles from scholars who undertook the challenging task of reflecting on the various aspects of AI technologies in the context of their relation to art and culture. It combines a twofold approach by focusing both on how AI technologies can be used to study art and culture, and how arts and humanities-based theoretical and methodological approaches can contribute to the contextualization and critical reading of AI technologies. This volume features contributions from scholars who are experienced with the challenges of using computational methods in arts and humanities research, as well as from scholars who reflect on contemporary AI developments from the more traditional standpoint of their respective discipline. The volume consolidates contributions from scholars belonging to different disciplines such as art history, media studies,

philosophy, anthropology, digital humanities, and computer science.

The contributions in this volume represent a unique collection of short articles that discuss a variety of contemporary relevant topics and case studies that emerge at the intersection of AI and arts and humanities. The article by Amanda Wasielewski⁸ discusses the implications of AI-based text-to-image models on the study of visual culture by focusing on how such models reify the Western concept of style. Nuria Rodríguez-Ortega⁹ elaborates on the concept of transduction as a framework for explaining, analyzing, and interpreting AI-generated images. The article by Bokar N'Diaye¹⁰ examines myths of the Greco-Roman world as a case study for contextualizing perspectives on human agency and creativity with relevance for debates on contemporary generative AI technologies. Dejan Grba¹¹ discusses the poetic similarities of contemporary AI artworks as a starting point for examining the strengths and weaknesses of AI art and for initiating a more comprehensive critical discourse on art and creativity. The article by Arif Kornweitz¹² explores the intersection of contemporary art practices and AI by focusing on case studies of two artists and drawing analogies between the concept of the latent space and the white cube gallery setting. Dominik Vrabčič-Dežman¹³ addresses the clichéd portrayal of AI in public media and draws on a range of interpretive theories of images to establish a framework for the examination of images of AI as a reciprocal link between aesthetics and politics. The article by Pepe Ballesteros Zapata¹⁴ elaborates on the role of computational methods in the methodological evolution of art history, focusing on several research examples in digital art history. Valentine Bernasconi¹⁵ discusses the challenges of interdisciplinary research in digital art history through the specific case study of machine learning-based body pose estimation in digitized paintings. The article by Piera Riccio and Nuria Oliver¹⁶ addresses the growing challenge of algorithmic censorship of artistic content that exhibits nudity on

8 Wasielewski 2024.

9 Rodríguez-Ortega 2024.

10 N'Diaye 2024.

11 Grba 2024.

12 Kornweitz 2024.

13 Vrabčič Dežman 2024.

14 Ballesteros Zapata 2024.

15 Bernasconi 2024.

16 Riccio / Oliver 2024.

online social platforms from a techno-feminist perspective. Anne-Laure Oberson's¹⁷ contribution reflects on AI generated images in the context of understanding the meaning of an image in the process of imagination.

This diverse collection of articles evolved into its current form from an initial attempt to facilitate interactive scholarly exchange and discussion on the currently relevant topic of AI and its relation to art and culture. More specifically, the articles presented in this volume represent a selection of extended contributions presented at the symposium "From Hype to Reality: AI in the Study of Art and Culture,"¹⁸ which took place at the University of Zurich in April 2023. The authors featured in this volume accepted the invitation to extend their initial symposium contributions into peer-reviewed short articles. The goal of the symposium was to bring together diverse perspectives and encourage cross-disciplinary discussions on the globally relevant topic of AI and its far-reaching implications on the future of art and culture. This volume therefore represents a unique selection of insightful articles that continue the dialogue initiated at the symposium, providing valuable insights and fostering a deeper understanding of the dynamic intersection between AI, art, and culture.

As the organizer of the symposium and co-editor of this volume, I would like to thank the whole editorial team—managing director Prof. Dr. Tristan Weddigen, co-editor Darío Negueruela del Castillo, and digital publications manager Elisa Bastianello—for the effort they invested in making the publication of this volume possible. On behalf of the editorial team, I would especially like to thank all the reviewers who dedicated their time and expertise to evaluate the submissions and provide valuable feedback. This volume and the preceding symposium represent results of the "From Hype to Reality: Artificial Intelligence in the Study of Art and Culture" project, which was funded by the UZH Global Strategy and Partnerships Funding Scheme, as well as supported by the Center for Digital Visual Studies.

¹⁷ Oberson 2024.

¹⁸ From Hype to Reality: AI in the Study of Art and Culture 2023. URL: [https://web.](https://web.archive.org/web/20240118144352/https://fromhypetoreality.com/)

[archive.org/web/20240118144352/https://fromhypetoreality.com/](https://web.archive.org/web/20240118144352/https://fromhypetoreality.com/) (last accessed: January 9 2024).

References

- Ballesteros Zapata 2024**
Pepe Ballesteros Zapata, "Evolving Methodologies: Computation in Art History," 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.0308>.
- Bernasconi 2024**
Valentine Bernasconi, "Automated Body Pose Estimation on Painting: A Crisis of Uncertainty," 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.0309>.
- Giles 2018**
Martin Giles, "The GANfather: The man who's given machines the gift of imagination," 2018. URL: <https://www.technologyreview.com/2018/02/21/145289/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/>.
- Goodfellow et al. 2015**
Ian Goodfellow et al., "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems*, ed. Z. Ghahramani et al., Curran Associates, Inc. 2015. URL: https://papers.nips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html.
- Grba 2024**
Dejan Grba, "Incidental Reverberations: Poetic Similarities in AI Art," 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.0305>.
- Huckins 2023**
Grace Huckins, "Minds of machines: The great AI consciousness conundrum," 2023. URL: <https://www.technologyreview.com/2023/10/16/1081149/ai-consciousness-conundrum/>.
- Kornweitz 2024**
Arif Kornweitz, "Inside the High-Dimensional White Cube The Ideology of Latent Space," 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.0306>.
- Midjourney 2023**
Midjourney, "Midjourney," 2023. URL: <https://www.midjourney.com/>.
- N'Diaye 2024**
Bokar N'Diaye, "Stochastic Parrots in the Ancient City?," 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.0304>.
- Oberson 2024**
Anne-Laure Oberson, "Making Sense of AI Generated Images in the Process of Imagination," 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.0311>.
- OpenAI 2021**
OpenAI, "DALL·E: Creating images from text," 2021. URL: <https://openai.com/research/dall-e>.
- OpenAI 2022**
OpenAI, "Introducing ChatGPT," 2022. URL: <https://openai.com/blog/chatgpt>.
- Riccio / Oliver 2024**
Piera Riccio and Nuria Oliver, "A Techno-Feminist Perspective on the Algorithmic Censorship of Artistic Nudity," 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.0310>.
- Rodríguez-Ortega 2024**
Nuria Rodríguez-Ortega, "Exploring the Concept of Transduction to Interpret AI-Generated Images," 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.0303>.
- Rombach et al. 2022**
Robin Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," in 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE 2022, pp. 10674–10685. URL: <https://ieeexplore.ieee.org/document/9878449/>. Doi: <https://doi.org/10.1109/cvpr52688.2022.01042>.
- Vrabič Dežman 2024**
Dominik Vrabič Dežman, "Interrogating the Deep Blue Sublime: Images of Artificial Intelligence in Public Media," 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.0307>.
- 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>.

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.

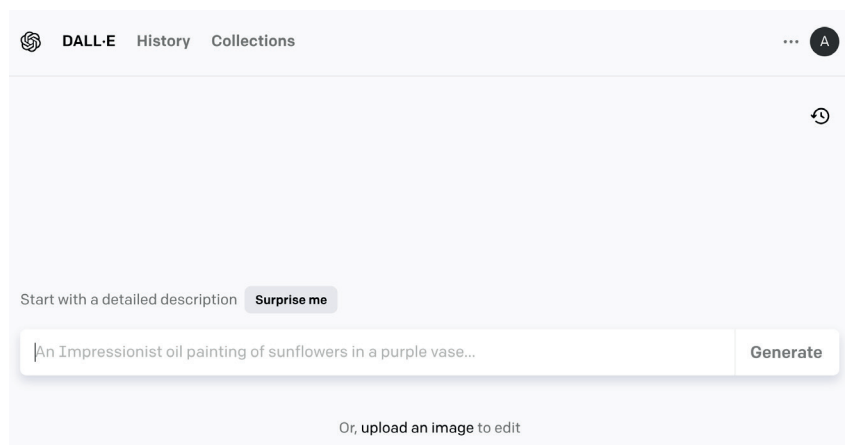


Figure 1. The opening screen of the DALL·E text-to-image generation platform.

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/>

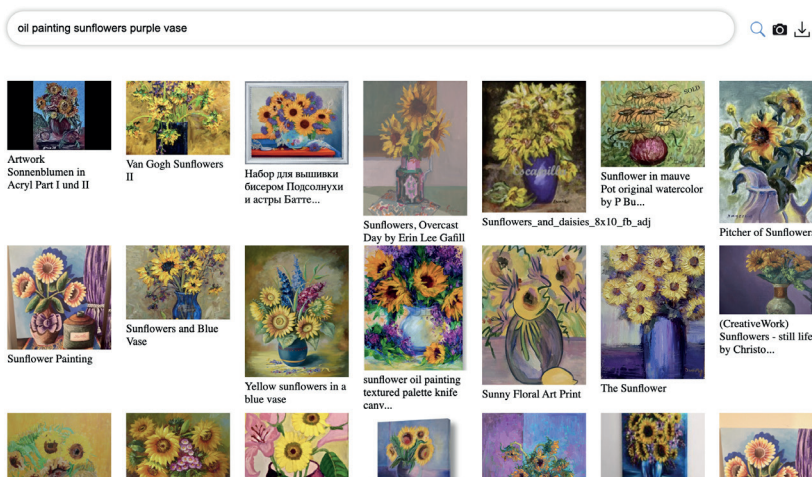


Figure 5. DALL·E output images created using the prompts “A Cubist oil painting of sunflowers in a purple vase,” “An Expressionist oil painting of sunflowers in a purple vase,” and “A Renaissance oil painting of sunflowers in a purple vase” (L-R) respectively.
Source: generated by the author, April 2023.



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>.

Abstract

Nuria Rodríguez-Ortega

Exploring the Concept of Transduction to Interpret AI-Generated Images

Keywords | *Techno-concept, Transduction, AI-generated images, Digital art history, Digital visual studies*



With the aim of contributing to the epistemological turn in the field of digital art history and cultural heritage studies, in a recent paper I introduced the idea of a techno-concept, defined as a co-production between machine rationality and human thought/ imagination. Within this framework of discussion, this paper argues that the computational operations and corresponding information transformation processes that take place in latent spaces, especially multimodal latent spaces, can be explained as a transductive process, so that AI-generated images can also be understood as transductive phenomena. In this way, the concept of transduction, widely used in various scientific and philosophical fields, becomes a potential theoretical category for interpreting AI-generated images. The argument is based primarily on two concepts of transduction: The concept of signal transduction as used in the fields of biomedicine and biochemistry, and the concept of transduction proposed by Gilbert Simondon as part of his theory of individuation and differentiation. The paper concludes with some of the interpretive implications of approaching generative images as transductive phenomena.

Cite as: Nuria Rodríguez-Ortega, "Exploring the Concept of Transduction to Interpret AI-Generated Images" 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. 23–28. DOI: <https://doi.org/10.48431/hsah.0303>.

1. Techno-Concepts for a Renewed Epistemology of Culture

The progressive consolidation of Digital Art History as a specific field of study forces us to elaborate a new (or renewed) epistemic apparatus that provides us with intellectual tools to explain and interpret the new cultural phenomena that emerge as technology continues to evolve. To this end, in a recent paper I have introduced the idea of a techno-concept as a notion that could help shape such epistemological endeavors.¹

Based on the well-established notion of technogenesis, the idea of a techno-concept postulates that the production of concepts cannot be considered independent of the technologies used to produce knowledge in any historical period, since they have co-evolved. Thus, a techno-concept refers to those intellectual productions in the form of ideas and/or concepts whose conceptual elaboration is mediated by the technology used for their operationalization. For example, our current way of understanding the mathematical concept of vector space cannot be separated from the kind of vector and latent spaces produced by deep learning technologies, and from the tools we create to visualize and understand how they are organized and operate. This understanding, in turn, shapes how and for what purposes we use them in the field of cultural analysis. For this reason, I argue that a techno-concept should be understood as a complex co-production between humans (the human/thought imagination) and technological entities.

In this context, this paper aims to make a preliminary approach to the concept of transduction as a techno-concept that could become a fruitful category for the explanation, analysis, and interpretation of AI-generated images. My argument is that the computational processes that take place in vector and latent spaces can be assimilated to transductive processes, or, in other words, that transductivity seems to be one of the main qualities of the production of generative images, that is, of their ontogenesis, but also of their modes of existence and operating logics. If this argument proves to be valid, then the concept of transduction, with all its epistemological and theoretical possibilities, could be incorporated into the explanation, analysis, and interpretation of generative images.

2. Concepts of transduction

The term and the concept of transduction have been used with different meanings in different fields over the course of time, including in the field of cultural analysis and interpretation. In the field of literary studies, the term has been widely used since Dolezel (1986) proposed a theory of literary transduction. In general, it is understood that the transformation of meaning that a given text undergoes when it is interpreted in different historical or social contexts involves a transductive component, since “a transmission of information with modification of meaning” is produced.² The concept of transduction has also found fruitful application in the arena of multimodal semiotics.³ Yet, in these approaches, the distinction with the processes of intersemiotic translation is often unclear, since it is subsumed

¹ Rodríguez-Ortega 2022.

² See also Izquierdo Arroyo 1980, Albaladejo

1998 and Albaladejo 2005.

³ Kress 1997 and Kress 2010.

within the concept of transduction. Likewise, its contiguity with Julia Kristeva's concepts of intertextuality or Gérard Genette's transtextuality has also been argued. Similarly, the concept of transduction has been used in the field of the performing arts to explain the process of materialization, of bringing into action, theatrical, filmic or musical texts.⁴

Although a very common definition of transduction is to say that transduction is the process of transforming one type of energy into another, I would like to focus attention primarily on two conceptions of transduction that I believe have rich conceptual implications for understanding and interpreting the field of AI-generated images.

2.1. Signal transduction

The first concept is associated with the use of the term in the fields of biochemistry and biomedicine, and more specifically with what is called 'signal transduction'. Signal transduction refers to an energetic transmission between cells that involves the transfer of new information into the cell. The integration of this new information with the information already present in the cell provokes intracellular changes that in turn generate a response (an action, a new signal, etc.). Thus, transduction is both the transmission of information and, more importantly, the process of transformation and production of a new signal/response. Generally speaking, in the case of dealing with generative processes of images (for example, images generated by Generative Adversarial Networks—GANs), it could be said that the computational operations trigger the transformation of the numerical information encoded in the vector and latent spaces into a visual sign (the image we see on the screen). This process cannot be considered as a translation or transposition of a previous sign, as happens, for example, when a cultural object is recoded as a matrix of numerical information during the digitization process. In other words, the

AI-generated image process seems to be better described as a transduction process rather than an intersemiotic process.

However, I think this phenomenon becomes even clearer in the context of text-to-image models. Take the stable diffusion model as an example: In the stable diffusion model, text embeddings are integrated into the latent space during the denoising process, which consists of a sequence of convolutions, with each convolution incorporating textual information that transforms the image embeddings (the latent visual information). This transformation then produces a new latent image (the new embeddings), which is processed by the next convolution, and so on. Thus, in this process, the former visual information is transformed by the incorporation of linguistic information, resulting in a new latent sign at the end. In a way, we could say that the final image embeddings produced conditioned by the text embeddings when trying to get an image that responds to a prompt (the stimulus) is the result of a transductive process.

2.2. Simondon's concept of transduction

The second concept of transduction that I find extremely relevant is the one proposed by French philosopher of science Gilbert Simondon in the 1950s.⁵ Transduction is the central component of Simondon's theory of ontogenesis. Indeed, transduction is the operation that makes possible the processes of individuation and differentiation, or in other words, the production of new realities in all dimensions (physical, biological, psychological, social). It is a conceptually very rich notion that conveys different but complementary meanings in Simondon's work.

On the one hand, transduction refers to a "dynamic operation by which a potential field of energy is actualized, moving from one state to the next in a process that individuates new materialities."⁶ If latent space

4 Maestro 1996, Naves 2006, Llanos López / Piñera Tarque 2002, De Assis 2017.

5 Simondon addresses the concept of transduction in three main texts: in the introduction to *L'individuation à la lumière*

des notions de forms et d'information (1958, published in 2005); in a 1960 lecture entitled "Forme, information et potentiels" (Eng. tr. "Form, information and potentials" 2019); and in a 1962 lecture entitled "L'amplification

dans les processus d'information" (Simondon 2015b, Spa. tr. Heredia 2016). My main source of reference has been *L'individuation* in its Spanish version, Simondon 2015a. 6 De Assis 2017; Simondon 2015a.

can be informally defined as an internal and compressed representation of all the relevant features of a domain (encoded in numerical vectors), and thus as a model that encodes a representation of a cultural domain as a whole, then latent space could be understood as a field of (real) virtuality, as a field of potentialities or possible actualizations (the pre-individual reality or metastable system in Simondonian terminology) that will begin to happen when computational agents start to operate. This computational operation will produce concrete actualizations of “forces” in a particular form or shape. In fact, it is precisely the countless deployment of possible actualizations that defines the essence (the nature) of AI-generative works beyond their instantiation in a particular form or shape.

On the other hand (and complementarily), transduction is the dynamic operation that governs the process of becoming, which is the core of Simondon’s theory of ontogenesis. According to Simondon and in contrast to Aristotelianhylomorphism—which still sustains our usual approach to cultural objects as entities that materialize and/or embody a pre-existing idea or content—individuals are not preconstituted, that is, there are no predefined entities (such as the matter and form of Aristotelianhylomorphism) that determine the substance or being of individuals; on the contrary, there are only processes of in-formation through which matter organizes itself by passing from one state (structural and energetic) to another. Thus, Simondon’s focus is on the process of individuation and in-formation through which individuals come into being. In this context, transduction is the very process of passing from one state to another.⁷ Moreover, the individual is constituted through a process of individuations that is never definitive (especially in the case of living forms). The individual is always in a process of becoming. Each individuation, understood as the concrete actualization of a field of possibilities, is a moment in a chain (past and future) of transformations.

In the light of this conception, I believe that the realm of generative images can be clearly explained as a process of becoming form, and thus as a transductive phenomenon. As I have said before, within the horizon of possibilities encoded in latent spaces (the metastable system), computational operations (acting as a structural germ) produce concrete actualizations of possibilities in a particular form or shape. Consequently, the generated images (the forms and shapes that we see and identify) could be considered as individuated forms that emerge from a field of numerical information through a process that we could consider immanent, to the extent that the latent space already has the potentiality of becoming that form. Moreover, the generated image (the individuated form) is never a final result, but only a point in a sequence, to the extent that the actual image can be transformed into another indefinitely. Generative images are in a continuous state of becoming, which has a very clear visual manifestation in the generative videos, where we literally witness the act of continuous transformation, the state of continuous becoming.

Likewise, as Deleuze points out in relation to the concept of transduction, the dichotomy between the general and the particular, which implies a vertical distribution of forms and matter, is replaced by a multiplicity of individual singularities resulting from the process of individuation. Thus, in this approach, transformation occurs by moving from one individual singularity to another individual singularity following a certain direction. That’s why the exploration of latent spaces through specific paths, which involves moving from one particular latent embedding to another particular latent embedding could also be explained (in my view) as a transductive process.

3. Some interpretive implications

Thinking of AI-generated images as transductive phenomena has a number of implications for their interpretation that deserve more detailed exploration. In this paper I will mention only some of them.

⁷ Simondon 2015a.

Since there is no longer a substantive separation between form and matter, but rather a flow of information that individuates (acquires) form, the focus shifts from the forms and structures per se, what they are and what they mean, to the conditions that lead to a particular form and structure. The central question then is not how a content is materialized in a form, but how one meaningful form becomes another meaningful form, how signs and meanings are produced through the transductive process that goes from one singular materiality to another singular materiality. This means locating the problem of interpretation in the field of morphogenesis, topology, and ontogenesis.

Since, as a transductive phenomenon, the generative image we see is only a momentum between the past and a possible future, the center of interest is not (or not only) in analyzing or interpreting what the individual is here and now, but in what it will be in the future. That is, in its transformative capacity to become something new, a new reality. Thus, the meaning of an image does not lie in its origin, but in its destiny, in the future towards which it is projected.

Since transduction is fundamentally characterized by its processual nature, a performative approach to AI-generated images seems crucial. This means interpreting AI-generated images in their condition of events, of actions that occur as an entanglement of relations between different dimensions, systems, and elements. Indeed, in Simondon's thought, transduction also implies the relation between different orders of scale.⁸ Individuation occurs as a process in which a heterogeneity of dimensions, systems, and elements of different scales enter into relationship. Thus, the individual is never an autonomous entity, but a complex system "in" relation, defined "in" such a relation.

This last question also has implications for the task of interpretation. In fact, transduction itself involves an interpretive process. The transducer has to interpret the meaning of the received sign/signal/information

in order to act, to produce a certain response and/or action. But such interpretive action does not take place on a single level or field of possibilities, but on many different levels and scales simultaneously. When dealing with AI-generated signs, a number of transducers (human and non-human) and interpretive processes can be identified: The recoding of cultural objects into a set of vectors of numerical information and the constitution of latent spaces (both of which are in themselves the results of interpretive acts); the computational logic that interprets the information encoded in latent space to produce a latent sign (the image/text embeddings); the decoding of the final image embeddings into a RGB visual sign that we see on the screen; and the human subject that transduces this sign in two ways: The transformation of light energy into bioelectrical signals; and the transformation of the computational sign into a cultural and subjective sign, when he/she appropriates the newly produced form in order to give it a meaningful sense within the framework of a particular cultural and social context (or, in other words, as a cultural and social subject that he/she is), that is, when the generative form becomes a real sign for human beings. Therefore, there is a chain that combines intersemiotic and transductive processes, where interpretation, the production of meaning, is the result of complex assemblages of human and computational forces, and where the computational operations that produce the latent space and operate in the latent space are transformed (in another kind of transductive process) into subjective meaning.

Throughout this text, I have tried to show that transduction, a concept that is part of our intellectual and scientific tradition, can be reformulated as a techno-concept insofar as the computational processes involved in the production of AI-generative images can be assimilated or described as a kind of transductive phenomenon. Thus, the concept of transduction, with the heuristic and epistemological possibilities associated with Simondon's concept of transduction in particular, could be used as a category for

⁸ Simondon 2015a.

analyzing and interpreting generative images in accordance with their own transductive nature. In this way, the problem of interpretation is located in the realm of the processes of individuation (ontogenesis), becoming form (morphogenesis, topology, flux), and transitionality (change and transformation).

References

Albaladejo 1998

Tomás Albaladejo, "Del texto al texto: transformación y transferencia en la interpretación literaria," in *Estudios de lingüística textual: homenaje al profesor Muñoz Cortés : [XVIII Curso de Lingüística Textual]*, ed. Estanislao Ramón Trives, Herminia Provencio Garrigós and Manuel Muñoz Cortés, Universidad de Murcia 1998, pp. 31–46.

Albaladejo 2005

Tomás Albaladejo, "Transducciones en la obra de Vila-Matas," in *La novela contemporánea española*, ed. María Dolores de Asís Garrote and Ana Calvo Revilla, Instituto de Humanidades Ángel Ayala-CEU y Universidad San Pablo-CEU (Fundación Universitaria San Pablo-CEU), Madrid, 2005, pp. 11–31.

De Assis 2017

Paulo De Assis, "Gilbert Simondon's 'Transduction' as Radical Immanence in Performance," *Performance Philosophy*, vol. 3, 3, (2017), pp. 695–716. URL: <https://www.performancephilosophy.org/journal/article/view/140/200>. Doi: <https://doi.org/10.21476/pp.2017.33140>.

Dolezel 1986

Lubomir Dolezel, "Semiotics of Literary Communication," *Strumenti critici*, vol. 1, (1986), pp. 5–48.

Heredia 2016

Juan Manuel Heredia, "La invención de la individuación a la luz de una problemática histórico-epistemológica," *Páginas de Filosofía*, 20, (2016), pp. 59–82. URL: <https://revel.uncoma.edu.ar/index.php/filosofia/article/view/975>.

Izquierdo Arroyo 1980

José Izquierdo Arroyo, "Sobre la transducción (meditaciones semiológicas). I. Transmisión y De-sustanciación," *Boletín Millares Carlo*, vol. 2, (1980), pp. 323–406.

Kress 1997

Gunther R. Kress, *Before writing: Rethinking the paths to literacy*, Routledge, London, 1997. Doi: <https://doi.org/10.4324/9780203992692>.

Kress 2010

Gunther R. Kress, *Multimodality: A Social Semiotic Approach to Contemporary Communication*, Routledge, Taylor & Francis Group, London, 2010.

Llanos López / Piñera Tarque 2002

Rosana Llanos López and Ismael Piñera Tarque, "Transducción dramática y transducción fílmica," in *Del teatro al cine y la televisión en la segunda mitad del siglo XX.*, ed. José Romera Castillo, Visor Libros 2002, pp. 381–397.

Maestro 1996

Jesús G. Maestro, "Lingüística y poética de la transducción teatral," in *Problemata Theatralia. El signo teatral: texto y representación*, ed. Jesús G. Maestro, Servizo de Publicacións da Universidade de Vigo 1996, pp. 175–211.

Naves 2006

María del Carmen Bobes Naves, "El proceso de transducción escénica," *Dialogía. Revista de lingüística, literatura y cultura*, vol. 1, (2006), pp. 35–53. URL: <https://journals.uio.no/dialogia/article/view/4039>.

Rodríguez-Ortega 2022

Nuria Rodríguez-Ortega, "Techno-Concepts for the Cultural Field: N-Dimensional Space and Its Conceptual Constellation," *Multimodal Technologies and Interaction*, vol. 6, (2022), p. 96. URL: <https://www.mdpi.com/2414-4088/6/11/96>. Doi: <https://doi.org/10.3390/mti6110096>.

Rodríguez-Ortega 2024

Nuria Rodríguez-Ortega, "Exploring the Concept of Transduction to Interpret AI-Generated Images," 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.0303>.

Simondon 1958

Gilbert Simondon, *L'individuation à la lumière des notions de forme et d'information* Ph.D Thesis, Faculté des lettres de Paris 1958.

Simondon 1960

Gilbert Simondon, "Forme, information, et potentiels," *Bulletin de la Société française de philosophie*, vol. 54, 5, (1960). URL: <https://www.sofrphil.fr/forme-information-potentiels/>.

Simondon 2005

Gilbert Simondon, *L'individuation à la lumière des notions de forme et d'information*, Krisis, Millon, Grenoble, 2005. URL: <http://catalogue.bnf.fr/ark:/12148/cb400814710>.

Simondon 2015a

Gilbert Simondon, *La individuación a la luz de las nociones de forma e información*, 2nd ed., Editorial Cactus, Buenos Aires, 2015.

Simondon 2015b

Gilbert Simondon, *Communication et information: cours et conférences*, ed. Nathalie Simondon, Presses Universitaires de France 1re édition: 2015, octobre., PUF, Paris, 2015.

Simondon 2015c

Gilbert Simondon, "L'amplification dans les processus d'information (1962)," in *Communication et information: cours et conférences*, ed. Nathalie Simondon, Presses Universitaires de France, Paris cedex 14, 2015, pp. 157–176. URL: <https://www.cairn.info/communication-et-information-9782130631293-p-157.htm>.

Simondon 2019

Gilbert Simondon, "Form, Information, and Potentials: February 27, 1960, Session of the Société française de philosophie," *Philosophy Today*, vol. 63, 3, (2019), pp. 571–583. Doi: <https://doi.org/10.5840/philtoday20191028281>.

Abstract

Bokar Lamine N'Diaye

Stochastic Parrots in the Ancient City?

Keywords | Generative AI, Authorship, Creativity, Critical AI



Recent advances in generative AI have highlighted how crucial perspectives on the temporality of technological change are to their reception. While machine-learning models are known to be the result of decades of steady progress in engineering circles, both enthusiastic and reactionary critics seem to privilege the narrative of an unprecedented ontological rupture, either extolling an imminent revolution or warning against the dissolution of the supposedly fixed human institutions of authorship and creativity. A core concern for the humanities should then be the contextualization of perspectives on human agency and creativity as contingent and culture-specific, as evidenced by historical and anthropological comparative approaches. This grounding of a disruptive technology as the latest for long-existing cultural mechanisms might help inform debates about generative AI, and counter excessive optimism or pessimism surrounding it. The Classical Greco-Roman world provides an interesting backdrop against which to examine these questions, as its conception of reality is both temporally distant and different from ours, but often invoked as the source of much of our philosophical outlook.

Cite as: Bokar N'Diaye, "Stochastic Parrots in the Ancient City?" 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. 31–37. Doi: <https://doi.org/10.48431/hsah.0304>.

1. Introduction

First off: I'm a historian. [...] It makes me think of the collapse of the Roman Empire. If you're not familiar: A once vibrant world power lost their way. This is unstoppable and we may as well accept the US's fate.

This is not assembly line work that's being threatened, this is human creative jobs. This has no historical precedent because art is not an industrialized product that's expected to be the same all the time. AI is not a power loom.¹

A popular definition of hype as “a state of anticipation generated through the circulation of promotion, resulting in a crisis of value,”² fairly describes the tone of this tweet by concept artist and anti-AI activist RJ Palmer. Faced with the perceived risks that Machine Learning (ML) generative models incur to his profession, he emphasizes his perspective as grounded in history, with a central unresolved paradox: Right before presenting AI as a sudden and unprecedented phenomenon, he characterizes it as a symptom of an inevitable historical trend of his home country towards decline—with the Roman Empire as an archetypal example. This centrality of the discontinuity introduced by AI is not unique to Palmer: Some, like the French philosopher Éric Sadin, even openly bemoan an “injunctive turn”: the alleged reversal of the proper relationship between humans and machines, the latter establishing and providing commands to the former.

Musa Victoriosa, an artwork of March 2023 by concept artist and anti-AI advocate Karla Ortiz, goes a step further from Palmer's pessimistic

yet passive witnessing of the AI timeline: Crafted as advertisement for the software Glaze, aiming to exclude digital artworks from machine learning training datasets through invisible blurring filters, this painting aims to extoll the newfound agency this move will supposedly bring traditional artists. One can note the reference to the classical Muse as a conventional symbol of artistic inspiration, the choice of depicting the titular figure as *victoriosa*, the purple, laurel crown and soaring eagles distantly recalling the Roman triumph, reframing the work as a clear martial retaliatory strike. Several pictures of Ortiz painting the canvas on an easel in her workshop (for what she is mostly known) consolidate the idea of material and traditional craft as the normative standard of artistic production.

Palmer and Ortiz's construction of Classical Antiquity as an idealized antithesis to AI art obviously reflects current concerns regarding the impact of generative ML on their livelihood: Radical upheaval of the job market, clogging of the canals of communication under a flood of generated content, and devaluation or obsolescence of human creative faculties.³ In their discourse, AI is portrayed as a jarring break in the very continuity of art. Beyond the awkward framing of art as an objective and preferably immutable cultural standard, the subordination of contemporary or digital artists to their painterly forbearers is starkly reminiscent of the long-established tendency of reactionary commentators to see Neo-Classical artworks as the sole worthy “approved art,” and later artworks as mere degenerate expedients. In the words of Florence Dupont,

¹ RJ Palmer (@arvalis), X (formerly Twitter), September 19, 2022, <https://x.com/arvalis/status/1571821476243984389> and <https://x.com/arvalis/status/1571823987252170755> [deleted by the author].

² Powers 2012, p. 863.

³ Vesa / Tienari (2020) 2022, p. 1134.



Figure 1. Karla Ortiz, *Musa Victoriosa*, March 2023.

Source: <https://x.com/kortizart/status/1636136818914762752>.

[T]oday, more than ever, antiquity plays the role of origin... At the same time, these are origins that have been neglected, increasingly ignored and therefore fantasized about. Escaping the critical scrutiny of history, Antiquity is, for some, the origin of the values of Western civilization, preserving its essence and quintessence. Antiquity becomes the stronghold to defend for those who believe in an essentialized European culture.⁴

2. A Comparative Anthropology of Generative AI?

Some form of entanglement of a somewhat skewed perception of art in Greco-Roman Antiquity (and temporality at large)⁵ and AI doom appears blatant in the aforementioned cases. In the same way as a growing number of scholars in the humanities explore “how to think from within the actual techniques, tools, and technologies of ML and how to leverage that practical knowledge in the development of new critical frameworks and methods, counter-mythologies, and epistemologies that might help enact a different way of living with AI,”⁶ the exploration of tangible and contextualized information from this time period might constitute a welcome way to lessen overenthusiastic or anxious discourse surrounding AI.

Naturally, even with the best interests of comparative research at heart, the risk of temporal bias does not spare the researchers themselves. Explicitly directing one’s focus to narratives that present robotic or android-like features runs the risk of tilting the conversation towards matters of AI sentience,⁷ which fall outside the boundaries of this study, or overriding original categories with our own operative concepts, such as the resolutely non-Ancient neologism of “*biotechné*” constructed by Adrienne Mayor in *Gods and Robots*.⁸ Consequently, the sample myths discussed below have been selected in order to present as few of the usual trappings

of technology as we currently understand it (mechanistic automata, human simulacra, etc.) as possible, instead focusing directly on fables and practices exploring the rhetoric themes that serve as the rationale for AI opponents.⁹

The idiom “stochastic parrot”¹⁰ initially proposed as a concise and evocative way to underline the factual limitations of Large Language Models (LLMs), is increasingly used as a retort to overly enthusiastic reviews of AI capabilities on social media. This expression provides us with a tentative framework of thematic comparison, by establishing which features of AI the modern West tends to imagine as both integral to the pursuit of a functional artist and unreachable for now by a machine: Namely, randomness as opposed to agency (“stochastic”), and mindless regurgitation over inspired imitation (“parrot”). As this article will try to prove, both of these dichotomies are not necessarily as irreconcilable in the ancient mindset as they sometimes are portrayed today.

3. Mimesis and Divination in Classical Discourse

3.1. Zeuxis’ Helen and Butades’ Soldier: Where is the Inspiration?

Would the Ancient Greeks have favored the easel over the digital? Is Ortiz correct in mobilizing the figure of the inspiring Muse as the sole privilege of the manual, self-expressing artist, relegating users of generative AI to mere commissioners or directors by default? Ancient discourses surrounding two myths relative to the visual arts may shed light on the ancient associations between artistic depiction and some of the most criticized features. In one of them, Zeuxis, painter *par excellence* in several myths of 5th-century Greece¹¹ is faced with a conundrum: Replicating the beauty of Helen, the most beautiful being to ever have lived, yet dead long before his time. His process, which Cicero incites his disciples to apply to their own

4 Dupont / Colonna d’Istria / Taussig 2013, p. 14.

5 Zeitlyn 2015.

6 Raley / Rhee 2023.

7 A typical example would be Roberts 2018.

8 Mayor 2018.

9 The very nature of the myth itself, as a non-committal call to a body knowledge assumed as commonly shared and always answering to the specific rhetorical goal of the utterer, is

a further conceit we sadly won’t expand upon here, but features prominently in Veyne 1983.

10 Bender et al. 2021.

11 Mansfield 2007, pp. 19–39.

literary endeavors, consists in the creation of a “corpus” of five women, chosen *in absentia* based on the beauty of their male relatives, the physical features of which he later assembles into a perfect whole—a route to which his very name, a form of the verb *zeugnumi* (yoking, fastening together) is certainly not unrelated.

He selected five, whose names many poets recorded because they were approved by the judgement of him who must have been the supreme judge of beauty. He chose five because he did not think all the qualities which he sought to combine in a portrayal of beauty could be found in one person, because in no single case has Nature made anything perfect and finished in every part...¹²

The peculiar status of the young women, which Cicero insists are remembered in name for their contribution, but not for the specific features they lent to the sculptures, strikes an odd resemblance to the training data of generative models, collected and annotated, yet diffused and unrecognizable in the outputted results.

A further source of information lies in the tale of Butades of Sicyone, a sculptor credited as the first depicter of a human likeness—the lover of his unnamed daughter, a departing soldier whose shadow she traced in coal for her father to fill in with clay. Frontisi-Ducroux demonstrates that Pliny’s account of the tale credibly hinges on wordplay—the Latin word *pupilla*, just like the Greek *kore*, designating both a young woman or the pupil of one’s eye—to establish a dual model of artistic *poiesis*: The daughter stands for the artist’s senses and subjectivity, sensitive to beauty yet unskilled yet also subordinated to the technical craft of the rational mind that presents them.¹³ Frontisi-Ducroux evokes in closing the possibility of the myth aptly describing the relationship between a photographer and a camera:¹⁴ Extending the parallel to the complementary abilities of the AI artist and the machine does not seem like a wild stretch.

Neither of the quoted accounts center inspiration or emotional expression, instead underlying a starkly different set of skills: Zeuxis as the master combiner, assembling details from a corpus he indirectly selected, or Butades, skilled executor of the requests of the separate “apple of his eye.” Looking at this evidence, the dichotomous opposition between “classical” or “traditional” artistic methods and the novelty of AI, as a technological and data-sampling field, does not appear as self-evident.

3.2. Agency in Praxis: The Divinatory Flight of Words and Birds as Ancient Techne

One may object that both Zeuxis and Butades operate deliberate choices that a ML model would enact in their stead. To use Bender’s terminology, beyond the *parroted* nature of the content, the further crux of the issue would lie in the *stochastic* nature of the non-sentient machine that produces it indiscriminately, and the resulting alleged divestment of human agency. While the works of anthropologists of art, most notably Alfred Gell in *Art and Agency*¹⁵ already do a great deal in recasting the artwork as an entanglement of social factors whose directionality isn’t necessarily always so obvious, we can find a further wrinkle in the institution of divination, especially in cases where it involves random processes through which human subjects apparently delegate their agency. Indeed, divination figures in good place in Aeschylus’ *Prometheus Bound*, when the eponymous Titan of forethought and human savviness enumerates the different *tekhnai* (“crafts,” “know-hows”) he lent to primitive mankind to ensure its survival:

First among all beings, I learned to sift dreams for things that will necessarily happen the next day; I familiarized them with the chance utterances (*kledônes*), an art (*tekhnê*) so tricky to interpret (*duskritous*), as well as with their own chance encounters (*sumboloi*) during their daily travels. The flight of the birds (*ptesin oionon*) with hooked talons I delineated with precision, between the one which by nature is auspicious and the one which is not to be

¹² Cicero, *De Inventione*, II, 1, 3, in Cicero 1949, p. 169.

¹³ Frontisi-Ducroux 2007, pp. 39–44.

¹⁴ Frontisi-Ducroux 2007, p. 55.

¹⁵ Gell 1998.

named... In a brief sentence, learn the summary of all this: all the crafts (*tekhnai*) of mortals, it is from Prometheus that they get them.¹⁶

Just as central to human agency as medicine or metallurgy, then, are the *kledônes* and *sumboloi*—chance utterances and random meetings—as well as the witnessed flight of birds. The onset of these events is enough to trigger a shift in a decision-making process or ground a consensus, usually in favor of the individual with the wherewithal to verbally accept the omen as such.¹⁷ In Aristophanes' comical play *The Birds*, from the same period, the humor primarily stems from the transfer of responsibility to a self-important and rebellious avian race. They argue that their essential role in the human society justifies their elevation to almost divine status. A key piece of evidence they present for their supreme nature is mankind's tendency to consult the *ptesin oionon* before undertaking important projects. This practice has led to the colloquial use of the term "bird" (*ornis*) to refer to other kinds of divinatory signs that could similarly influence someone's action:¹⁸

THE BIRDS: We are to you Ammon, Delphi, Dodona, Phoebus Apollo. For you first turn to birds before undertaking any course, whether that be trade, and towards acquisition of livelihood, and towards marriage of a man. And you label as "birds" (*órnin te nomízete*) all the things you employ to discern matters of divination (*diakrínei manteías*): to you, an ominous rumor (*phémē*) is a "bird" (*órnin*); you call a sneeze (*ptarmón*) a "bird," a chance encounter (*súmbolon*) a "bird," a sound (*phōnēn*) a "bird," [encountering] a servant (*therápon*) a "bird," a dream (*ónon*) a "bird." Are we not clearly the prophetic (*manteĩos*) Apollo to you? If then you recognize us (*nomísēte*) as gods, we shall be your divining Muses, through us you will know the winds and the seasons, summer, winter, and the temperate months. We shall not withdraw ourselves to the highest clouds like Zeus, but shall be among you.¹⁹

The humor of this comedic passage lies at least partly in the Athenian audience's perception of the tacit ambiguity at the heart of their practice of divination: while an essential piece of meaning- and decision-making, even closer to human senses than deities due to their tangible existence, divinatory vessels such as birds are ultimately only an extension of mankind's agency, with the idea of their actual supremacy relegated to the satirical realm. One could even surmise that this scenario of the divinatory vessels exploiting this reliance of men to claim agency of their own, here played for laughs by Aristophanes, constitutes an early comparanda to the "rebellion of the machines" trope, common in both fictional depictions and current safety polemics surrounding artificial intelligence.

As we can see from the examples above, the institution of divination alone challenges deeply set dichotomies of the modern epistemological model: proactivity versus passivity, randomness versus deliberate course of action, arbitrary noise versus meaningful signal. Such evidence comforts us in abandoning the idea of a sudden "injunctive turn" tied to modernity (as Sadin claims), or the association of these themes solely with the mechanical or biological sciences (as does Mayor).

4. Conclusion

While insufficient to fully encompass the diversity and mutability of conceptions of agency and the creative process in the Greco-Roman ancient world, this short inventory of "counter-myths" proves the richness of contrast that such comparison can bring. Instead of an idealized apex of aesthetics and expressivity, forever marred by modernity, several indirect clues and testimonies from the Ancient themselves help us piece together a far more nuanced picture, and a number of epistemological constructs (the diffused body of Zeuxis' Helen, or the agentic recasting of random divinatory outcomes) that may inform our own behavior regarding artificial intelligence.

¹⁶ Aeschylus, *Prometheus Bound*, v. 484-499. Translation adapted from Calame 2017.
¹⁷ Lateiner 2005.

¹⁸ For more on this double meaning of *ornis*, and the divinatory role of birds in general, see Mynott 2018.

¹⁹ Aristoph., *Birds* 717-731. Translation adapted from Aristophanes 1938, vv. 717-731.

This article in no way seeks to censor the likes of Ortiz or Palmer, nor to enshrine this particular angle on Antiquity as the enlightened standard in these matters. In a similar way as the Ancient scholars and playwrights wielded the narrative of folktales, the constructions of expressions, either idiomatic or narrative, is a crucial feature for scholars to monitor and engage with the full breadth of the “data arena” surrounding the applications of machine learning.²⁰ As such, the most defining staple of critical AI as a growing discipline might well reside in its embrace of these semantic or polemic imprecisions as spurious entry points rather than unwelcome noise, starting with the deliberate academic use of the polythetic and ever vaguer term of “artificial intelligence”:

Critical AI, while recognizing the reductive, even absurd aspects of the term “AI” and the magical thinking it perpetuates, nonetheless allows for a kind of linguistic pragmatism, treating the term metonymically and engaging AI as an assemblage of technological arrangements and sociotechnical practices, as concept, ideology, and *dispositif*.²¹

²⁰ Ślosarski 2023.

²¹ Raley / Rhee 2023, p. 188.

References

- Aristophanes 1938**
Aristophanes, "Birds," in *The complete Greek drama. All the extant tragedies of Aeschylus, Sophocles and Euripides, and the comedies of, Aristophanes and Menander, in a variety of translations, Vol II*, ed. Whitney J. Oates and Eugene Jr. O'Neill, Random, New York 1938, pp. 733–801. URL: <https://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0019.tlg006.perseus-eng1>.
- 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>.
- Calame 2017**
Claude Calame, "À propos des pratiques divinatoires et de la « raison des signes » en Grèce classique," in *Dossier : Soigner par les lettres*, Éditions de l'École des hautes études en sciences sociales, 2017, pp. 393–410. URL: <http://books.openedition.org/editionsehess/4777>. Doi: <https://doi.org/10.4000/books.editionsehess.4777>.
- Cicero 1949**
Marcus Tullius Cicero, "De Inventione," in *De Inventione. De optimo genere oratorum, Topica with an English translation by H.M. Hubbell*, Harvard University Press, William Heinemann, Cambridge Mass. and London, 1949, pp. 1–345. URL: <http://archive.org/details/in.ernet.dli.2015.189092>.
- Dupont / Colonna d'Istria / Taussig 2013**
Florence Dupont, Pauline Colonna d'Istria and Sylvie Taussig, *L'Antiquité territoire des écarts : Entretiens avec Pauline Colonna d'Istria et Sylvie Taussig*, Albin Michel, Paris, 2013.
- Frontisi-Ducroux 2007**
Françoise Frontisi-Ducroux, "La fille de Dibutade, ou l'inventrice inventée," *Cahiers du Genre*, Association Féminin Masculin Recherches vol. 43, 2, (2007), pp. 133–151. URL: <https://www.cairn.info/revue-cahiers-du-genre-2007-2-page-133.htm>. Doi: <https://doi.org/10.3917/cdge.043.0133>.
- Gell 1998**
Alfred Gell, *Art and Agency: An Anthropological Theory*, Oxford University Press, Oxford, New York, 1998.
- Lateiner 2005**
Donald Lateiner, "Signifying Names and Other Ominous Accidental Utterances in Classical Historiography," *Greek, Roman, and Byzantine Studies*, vol. 45, 1, (2005), pp. 35–57. URL: <https://grbs.library.duke.edu/index.php/grbs/article/view/211>.
- Mansfield 2007**
Elizabeth C. Mansfield, *Too Beautiful to Picture: Zeuxis, Myth, and Mimesis*, University of Minnesota Press, Minneapolis; London, 2007. URL: <https://www.jstor.org/stable/10.5749/j.ctttspvd>.
- Mayor 2018**
Adrienne Mayor, *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology* | Department of Classics, Princeton University Press, Princeton, 2018.
- Mynott 2018**
Jeremy Mynott, *Birds in the ancient world: Winged words*, Oxford University Press, Oxford, 2018. URL: <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1809382>.
- N'Diaye 2024**
Bokar N'Diaye, "Stochastic Parrots in the Ancient City?," 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.0304>.
- Pietrobelli et al. 2017**
Antoine Pietrobelli et al., *Dossier : Soigner par les lettres : La bibliothérapie des Anciens*, Métis, Éditions de l'École des hautes études en sciences sociales, Paris, 2017. URL: <https://books.openedition.org/editionsehess/4619>.
- Powers 2012**
Devon Powers, "Notes on Hype," *International Journal of Communication*, vol. 6, 0, (2012), pp. 857–873. URL: <https://ijoc.org/index.php/ijoc/article/view/1441>.
- Raley / Rhee 2023**
Rita Raley and Jennifer Rhee, "Critical AI: A Field in Formation," *American Literature*, vol. 95, 2, (2023), pp. 185–204. Doi: <https://doi.org/10.1215/00029831-10575021>.
- Roberts 2018**
Stuart Roberts, "From Homer to HAL: 3000 years of AI narratives," 2018. URL: <https://www.cam.ac.uk/stories/ai-narratives>.
- Ślosarski 2023**
Bartosz Ślosarski, "Data arenas: The relational dynamics of data activism," *Big Data & Society*, vol. 10, 1, (2023), Doi: <https://doi.org/10.1177/20539517231177617>.
- Vesa / Tienari 2022**
Mikko Vesa and Janne Tienari, "Artificial intelligence and rationalized unaccountability: Ideology of the elites?," *Organization*, vol. 29, 6, (2022), pp. 1133–1145. Doi: <https://doi.org/10.1177/1350508420963872>.
- Veyne 1983**
Paul Veyne, *Les Grecs ont-ils cru à leurs mythes ?*, Seuil, Paris, 1983.
- Zeitlyn 2015**
David Zeitlyn, "Looking Forward, Looking Back," *History and Anthropology*, vol. 26, 4, (2015), pp. 381–407. Doi: <https://doi.org/10.1080/02757206.2015.1076813>.

Abstract

Dejan Grba

Incidental Reverberations: Poetic Similarities
in AI Art

Keywords | AI art, Creative thinking, Poetic
similarities, Computational art



Contemporary AI art's diverse and widely recognized repertoire features numerous artworks that share conceptual, thematic, narrative, procedural, or presentational properties with other artworks across disciplinary and historical spectrums. AI artists occasionally leverage well-sanctioned poetic referencing as an asset but when obvious or easily discoverable similarities remain unacknowledged, they may become liabilities. Lurking behind the hype waves in the media, art world, and academia, these liabilities shape contemporary AI art's cultural identity and affect its social impact. As part of a broader study of poetic contingencies, in this paper I discuss selected AI art exemplars whose multifaceted expressive parallels are symptomatic of the field and beyond. I argue that expressive similarities in AI art are as detrimental to its cultural value as they are avoidable in its variety of important topics addressable with a wide range of creative affordances. My critique takes the well-informed autonomy of expression and the socially responsible freedom of creative thinking as the tenets of artmaking to indicate some of AI art's related issues and challenges induced by its entanglements with AI science, technology, and industry. In conclusion, I suggest that poetic similarities open a valuable perspective for studying AI art's strengths and deficiencies and for articulating a broader critical discussion of art and creativity.

Cite as: Dejan Grba, "Incidental Reverberations: Poetic Similarities in AI Art" 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. 39–44. Doi: <https://doi.org/10.48431/hsah.0305>.

1. Introduction

Since its outset in the 1970s, AI art remained largely obscure until the expansion of digital infrastructures and the availability of versatile computational tools in the early 2000s facilitated its poetic diversification. The confluence of increasingly accessible machine learning (ML) techniques and the rising social impact of AI since the 2010s, have allowed it to gain productive momentum, cultural presence, the AI industry's patronage, and the art market's recognition.¹ Contemporary AI art features experimental, exploratory, and speculative practices that emerge from and respond to modern AI's phenomenological realms and social implications.² However, these diverse practices often result in artworks that share conceptual, thematic, narrative, procedural, or presentational properties with other artworks across disciplinary and historical spectrums. While some AI artworks leverage well-sanctioned poetic referencing as an asset, in cases where obvious or easily discoverable similarities remain unacknowledged, they may become liabilities.

Lurking behind the hype waves in the media, art world, and academia, these "incidental reverberations" frequently happen because various traits and drives, such as audaciousness and calculated ambitions, compel or lure artists to disregard the extent and open-endedness of informed external interpretation and intervention. They belong to a corpus of intellectual blunders and methodological miscalculations whose unforeseen consequences are usually undesired by the artists but always instructive for their audience. They affect AI art's cultural

identity and social impact by raising concerns about the artists' contextual awareness of current art production, art-historical knowledge, or professional ethics. As part of a broader critical study of poetic contingencies in computational art,³ this paper summarizes the issue of widespread expressive similarities in AI art by discussing selected exemplars whose multifaceted parallels are emblematic of the field.

2. Incidental Reverberations

Conceptual parallels, thematic repetitions, methodological similarities, and presentational likeness in AI art are hardly surprising since artmaking inevitably entails some degree of obvious or implied creative processing of artistic and other cultural artefacts. Various modes of this creative processing have been validated in different ways throughout 20th-century art, from Cubism and Dada, through Pop-Art, Fluxus, and Conceptual Art, to Postmodernism.⁴ Widely accepted and most recognizable as part of the remix culture,⁵ artefactual creativity⁶ permeates all contemporary art disciplines and plays an important role in exploratory applications of computation for transforming existing data, ideas, relations, and phenomena.

However, artefactual creativity involves a deceptively smooth continuum of procedures ranging from interpretation, free copy, reprise, remake, allusion, citation, dedication, derivation and détournement, through mashup, remix, pastiche, reference, reminiscence, homage, and parody, to imitation, plagiarism, and forgery.⁷ Expressive values of this procedural realm unfold in a grey zone of

1 Cetinić / She 2022, Browne 2022.

2 Grba 2022.

3 Grba 2023.

4 Haber n.d.

5 Navas / Gallagher / burrough 2016.

6 Artefactual creativity, also called (re) creativity Lessig 2006, is the application of combinatorial inventiveness to the specific

qualities, meanings, contexts, or implications of existing artefacts to produce interesting new artefacts Grba 2015, Grba 2020, p. 74.

7 Boon 2013, Grba 2015.

cultural inertias, dispersed knowledge, subtle influences, fuzzy ethical notions, and slippery moral categories, which fundamentally relativize the concept of (and to some degree the requirement for) authenticity or originality.⁸ Furthermore, the expressive undercurrents, tendencies, and trends are closely interwoven with the fabrics of artists' professional lives and can be hard to identify and resist. Within such context, the grayest aspect of artefactual creativity concerns the imitation of another creation's poetic arc that synergizes a particular idea or topic with the most suitable, economic, or elegant elements such as formal rendering, production technique, or presentation format. Although the variational repetition of this synergetic relationship is unavoidable and necessary for apprenticeship in many areas of creativity, in artmaking it establishes the artwork's core identity and can be considered exploitative if the new artwork does not enhance the imitated poetic arc with meaningful values. The abstractness of these meaningful values sometimes makes it difficult to define the exploitation objectively.

For all these reasons, the assessment of expressive similarities navigates a fine and often blurry line of distinction meandering around fraudulent, flawed, and proper artistic strategies and always risks turning out as hasty, biased, uninformed, nitpicking, or preachy. Nevertheless, when an apparent similarity of relevant creative factors or a strong poetic parallel between a new artwork and a reasonably knowable referent remains unacknowledged by the new artwork's creator, comparative criticism is legitimate. It is invaluable for the maturation of AI art, whose originality-related mishaps often indicate deeper related issues and challenges caused by its techno-cultural entanglements and other factors.

AI art and its superset computational art have a long history of poetic similarities imposed by cognitive requirements, such as procedural literacy,⁹ or induced by common methodologies. By their coding skills, AI artists range from creative directors who employ programmers or collaborate with computer scientists, through *bricoleurs* and tinkerers with various types and levels of technical expertise, to artist-engineers/scientists. They usually appropriate and sometimes modify the existing software packages or code libraries¹⁰ and train them with publicly available datasets, which leads to aesthetic homogeneity. Attempting to break out of it, some artists pursue authenticity through obsessive experimentation with code,¹¹ thus revealing the lack of appreciation that originality is highly contextual, so its overidentification with technical or formal properties is largely misconceived.¹² Such superficial approaches relate to a more problematic type of expressive normalization that results from coupling AI systems with simplistic conceptual frameworks.

Formal repetitiveness, trivial contextualization, and shallow conceptualization are evident in Inceptionism, GANism, AI-derivative mainstream art, spectacular large-scale installations, and other subdomains.¹³ Crucially, regardless of their causes, many AI artworks' similarities remain undisclosed by their authors. The prevalence of these "phantom" or "orphan" references merits a detailed study that would substantially exceed this paper's available volume, so we must be content here with a few multifaceted exemplars.

For the short film *Sunspring*,¹⁴ Ross Goodwin trained one ML system on 162 science fiction (SF) movie scripts found online to generate the screenplay and screen directions, and another one on a folk songs database to generate the

⁸ Although many poetic similarities are intentional (rip-offs), some result from (a combination of) enthusiasm, ignorance, carelessness, disregard, overconfidence, or negligence, which makes it difficult to categorize them with certainty.

⁹ Procedural literacy is the ability to think, read, and write processes to engage procedural representation and aesthetics. It

is closely associated with programming skills Reas / McWilliams / LUST 2010.

¹⁰ For instance, the widely-discussed case of the Obvious' *Portrait of Edmond Belamy* (2018) laid the ground for questions about the nature of authenticity, the "localization" of expressive agency (in humans or AI systems), the creative autonomy, and the relativity of property liabilities in AI art

McCormack / Giffors / Hutchings 2019, pp. 1, 3, Browne 2022, Audry 2021, pp. 164, 250 and Cetinić / She 2022, pp. 10–11.

¹¹ Miller 2019.

¹² Saltz 2019.

¹³ Grba 2022, p. 19.

¹⁴ Directed by Oscar Sharp, 2016.

song lyrics.¹⁵ Oscar Sharp used this material to produce the film. Brimming with plot inconsistencies and awkward dialogues, *Sunspring* touches upon several issues of its underlying cultures. The artists' satirical application of ML to filmmaking reverses the logic of corporate movie search algorithms, playfully mimics Hollywood's screenwriting paradigm of regurgitating themes from earlier films, and anticipates the current application of ML for screenplay design. It simultaneously exemplifies the power and the perils of using statistics to trace the "cloud" of common ideational threads in a specific cultural domain. *Sunspring*'s incongruity when compared with conventional SF narratives also offers an analogy for the frivolity or nonsensicality of SF imaginaries in real life.

Two years later, Alexander Reben appropriated Hollywood strategies and regurgitated *Sunspring*'s concept and methodology to produce *Five Dollars Can Save the Planet* (2018)—"the world's first TED talk written by an AI." The text of this 3-minute TEDx talk was generated by training an ML model on "all TED talks."¹⁶ Aiming to join the ongoing critique of TED's intellectual sharing model,¹⁷ this work's humorous take on the trend of "robotization" of sales-pitch public lectures echoes Doug Zongker's more radical comic act *Chicken Chicken Chicken* (2007)—a 4-minute lecture in which the whole narration and different types of slides contain one word: Chicken.¹⁸ Moreover, Reben's satirical logic and production methodology duplicate Goodwin and Sharp's while his choice of auto-recursive format (critiquing TED talks in a TED talk) mirrors Benjamin Bratton's 2013 TEDx talk *New Perspectives: What's Wrong with TED Talks?*¹⁹ Although *Sunspring* is conceptually akin to SF parodies such as *Dark Star*²⁰ and its implication that the palatability of popular expressive forms partly relies on clichés nods

toward Jennifer and Kevin McCoy's works with pop-cultural sampling,²¹ its creators authentically activated one of the SF tropes—artificial intelligence—to make these points "mathematically." *Five Dollars...* uses an identical approach for a parodic statement about corporate public talks, but reveals or adds nothing new.²²

Libby Heaney's two-channel video *Elvis* (2019) further illustrates the delicate dependency between an artwork's conceptual, topical, or methodological authenticity and critical cogency. Featuring a portrait of Heaney deepfaked as Elvis Presley and Presley's portrait deepfaked as Heaney, it directly copies (but does not acknowledge) the emblematic Gavin Turk's *POP* (since 1993).²³ In a series of selfie-pop-icon chimeras, *POP* addresses the same topics of individual identity and cultural mechanisms of celebrity mythmaking, involves the same pop icon, and applies the same formal method (face swapping) albeit in different media (sculpture, photographs, and prints) and a more complex chain of artistic/cultural allusions (acknowledged by Turk): For instance, a figure of Sid Vicious with Turk's face posing as Andy Warhol's *Elvis Presley* (1963). The sole diversion in Heaney's *Elvis* is the introduction of an AI technique (deepfaking) into the tactical repertoire of gender construction within digital technologies. Even in this aspect, *Elvis* is not too convincing when compared with the earlier uses of digital technologies to question the visual aspects of gender construction, such as image manipulations by Inez van Lamsweerde in the mid-1990s (V.A 1995).²⁴

3. Conclusion

To whatever degree the persuasive weight of liminal differences in *Sunspring* vs. *Five Dollars...*, *Elvis* vs. *POP*, and numerous other cases²⁵ may be considered an open question or a matter of individual interpretation, poetic

¹⁵ Goodwin 2016.

¹⁶ Reben 2018.

¹⁷ Morozov 2012, Harouni 2014.

¹⁸ Bauman 2007.

¹⁹ Bratton 2013.

²⁰ Directed by John Carpenter, 1974.

²¹ See, for example, McCoys' *Every Shot, Every Episode* (2001) McCoy / McCoy 2023a

and *Every Anvil* (2002) McCoy / McCoy 2023b.

²² The conceptual and methodological cloning of *Sunspring* continued on a Gesamtkunstwerk level with the project *Legend of Wrong Mountain* (2018). Its central part is a generative video of a traditional Chinese *Kunqu* opera produced by a team of computer engineers, artists, and designers

who trained an assortment of ML models on four different datasets about the forms of *Kunqu* opera to make the script (libretto), musical score, gesture choreography, and a woodcut book Huang 2019.

²³ Heaney 2019, Turk n.d.

²⁴ Brusatin / Clair 1995.

²⁵ For further consideration of questionable expressive similarities in AI art, whose listing

similarities profoundly affect AI art. They signal a nonchalance toward both legacy and current creative landscapes, which is perhaps the most frustrating weakness of modern AI art and one of the key aspects for its constructive critique. We can hardly attribute it to the field's youth because AI art is more than 50 years old and shares all major poetic features with experimental arts whose history reaches back to at least the late 19th century.

Contemporary AI art's diverse affordances for addressing a vibrant ecosystem of important topics should make expressive inertias superfluous. Yet, "incidental reverberations" persist not only as accidental byproducts of the spontaneous convergence of ideas,²⁶ cognitive requirements, or technical constraints but also due to less justifiable psychological factors such as carelessness, indolence, egoism, arrogance, narcissism, and vanity mixed with professional deficiencies such as inexperience, contextual ignorance, or sketchy art-historical knowledge. They happen when artists disregard the essential requirement for multi-directional questioning and iterative critique of conceptual cogency, thematic relevance, and formal soundness in their creative processes. The artists' lack of transparency about their "phantom references" degrades AI art's images in the eyes of the general audience, which is chronically suspicious of the flimsiness, flippancy, and widely (mis)perceived lack of authenticity in contemporary art.

Some artworks, such as *Disnovation.org's Predictive Art Bot* (since 2017)²⁷ or Adam Basanta's *All We'd Ever Need Is One Another* (2018),²⁸ facetiously leverage the formal ambiguities of originality and authorship in modern ML to broach the vague terrain of ideational homogeneity in AI art and popular culture, expose the professional priorities of technical merit and monetary gain, and challenge the worldviews and interests that shape the intellectual property regulation in the digital

age.²⁹ Their critique underlines that curious, earnest, and respectful exploration of general and field-specific art history is the basic—and empowering—artistic method for bringing up new works responsibly.³⁰

But such tactical provocations are exceptional. The specific cases of AI art's unacknowledged poetic similarities and ethically charged slip-pages remain seldom exposed and openly discussed in the professional community probably because its members prefer to stay out of the reputational minefield that opens by clearly expressing potentially confrontational opinions. This self-protecting leniency goes in tandem with the equally present, but even less discussed, meritocratic inconsistency imposed by cultural hegemonies, power games, and systemic injustices of the contemporary art world. Indicating darker shades of human nature, it retains the accidents of birth, nationality, language, or geographical location as important vectors of career trajectories and recognition.

Along with other creative contingencies,³¹ poetic similarities place AI art firmly within the modern context of existential crises, cultural contradictions, and societal tensions. They help us infer artists' knowledge and skills in light of their personal qualities (which inform poetics as much as any other expressive factor) and reaffirm that understanding artists equally as creators and human beings should be integral to the art appraisal. They provide an invaluable perspective for studying the strengths and deficiencies of AI art and for articulating the critical discussion of art and creativity in general.

would exceed this paper's available volume, see Grba 2023, p. 213.

²⁶ The somewhat analogous legal category of "independent creation" applies to cases when an existing creative expression gets reproduced without awareness of the

original, so the new expression may not legally constitute a copy (Zeilinger 2021).

²⁷ Maigret / Roszkowska 2017.

²⁸ Basanta 2018.

²⁹ See the discussion of the *Predictive Art*

Bot in Grba 2021, p. 247 and the discussion of *All We'd Ever Need Is One Another* in Zeilinger 2021, pp. 94–108.

³⁰ Grba 2023, pp. 217–218.

³¹ See Grba 2023.

References

- Audry 2021**
Sofian Audry, *Art in the Age of Machine Learning*, The MIT Press, Cambridge Mass., 2021. Doi: <https://doi.org/10.7551/mitpress/12832.001.0001>.
- Basanta 2018**
Adam Basanta, "Adam Basanta - All We'd Ever Need Is One Another," 2018. URL: <https://adambasanta.com/allwedeverneed>.
- Bauman 2007**
Yoram Bauman, "Chicken Chicken Chicken," 2007. URL: https://youtube/yL_-1dgOSdk.
- Boon 2013**
Marcus Boon, *In Praise of Copying*, Harvard University Press, Cambridge, 2013.
- Bratton 2013**
Benjamin Bratton, "New Perspectives: What's Wrong with TED Talks?," 2013. URL: <https://youtu.be/Yo5cKRmJafo>.
- Browne 2022**
Kieran Browne, "Who (or What) Is an AI Artist?," *Leonardo*, vol. 55, 2, (2022), pp. 130–134. Doi: https://doi.org/10.1162/leon_a_02092.
- Brusatin / Clair 1995**
Manlio Brusatin and Jean Clair, *Identity and alterity: Figures of the body 1895/1995 : La Biennale di Venezia*, 46. *Esposizione internazionale d'arte*, 1st ed., Marsilio, Venezia, 1995.
- Cetinić / She 2022**
Eva Cetinić and James She, "Understanding and Creating Art with AI: Review and Outlook," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 18, 2, (2022), pp. 1–22. Doi: <https://doi.org/10.1145/3475799>.
- Goodwin 2016**
Ross Goodwin, "Sunspring," 2016. URL: <https://youtu.be/LY7x2lhqjmc>.
- Grba 2015**
Dejan Grba, "I Cite (Very) Art: (Re) creativity in Contemporary Art," in *Going Digital: Innovations in Contemporary Life Conference Proceedings*, STRAND 2015, pp. 29–36.
- Grba 2017**
Dejan Grba, "Avoid Setup: Insights and Implications of Generative Cinema," *Leonardo*, vol. 50, 4, (2017), pp. 384–393. Doi: https://doi.org/10.1162/leon_a_01456.
- Grba 2020**
Dejan Grba, "Alpha Version, Delta Signature: Cognitive Aspects of Artefactual Creativity," *Journal of Science and Technology of the Arts*, vol. 12, 3, (2020), pp. 63–83. URL: <https://revistas.ucp.pt/index.php/jsta/article/view/9491>. Doi: <https://doi.org/10.34632/jsta.2020.9491>.
- Grba 2021**
Dejan Grba, "Brittle Opacity: Ambiguities of the Creative AI," in *xCoAx 2021 Proceedings of the Ninth Conference on Computation, Communication, Aesthetics & X*, ed. Mario Verdicchio et al., i2ADS Research Institute in Art, Design and Society 2021, pp. 235–260. URL: <https://2021.xcoax.org/>. Doi: <https://doi.org/10.5281/zenodo.5831884>.
- Grba 2022**
Dejan Grba, "Deep Else: A Critical Framework for AI Art," *MDPI Digital*, vol. 1, 2, (2022), pp. 1–32. URL: <https://www.mdpi.com/2673-6470/2/1/1>. Doi: <https://doi.org/10.3390/digital2010001>.
- Grba 2023**
Dejan Grba, "Renegade X: Poetic Contingencies in Computational Art," in *Proceedings of xCoAx 2023, 11th Conference on Computation, Communication, Aesthetics & X*, ed. Miguel Carvalhais et al., i2ADS Research Institute in Art, Design and Society 2023, pp. 201–221. Doi: <https://doi.org/10.34626/xcoax.2023.11th.201>.
- Grba 2024**
Dejan Grba, "Incidental Reverberations: Poetic Similarities in AI Art," 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.0305>.
- Haber n.d.**
John Haber, "The Reusable Past (Rosalind E. Krauss: The Originality of the Avant-Garde)". URL: <https://www.haberarts.com/krauss.htm>.
- Harouni 2014**
Houman Harouni, "The Sound of TED: A Case for Distaste," *The American Reader*, (2014). URL: <https://theamericanreader.com/the-sound-of-ted-a-case-for-distaste>.
- Heaney 2019**
Libby Heaney, "Elvis," 2019. URL: <https://www.art-ai.io/programme/elvis>.
- Huang 2019**
Lingdong Et al. Huang, "Legend of Wrong Mountain: AI Generated Opera," in *ISEA2019 Gwangju Lux Aeterna: 25th International Symposium on Electronic Art Proceedings*, ed. Juyong Park, Jin Wan Park and Juhan Nam, ISEA International 2019, pp. 255–261. URL: https://isea-archives.siggraph.org/wp-content/uploads/2019/08/ISEA2019_Proceedings.pdf.
- Lessig 2006**
Lawrence Lessig, *Copyright and Other Fairy Tales*. Hans Christian Andersen and the Commodification of Creativity, ed. Helle Porsdam, Edward Elgar, Cheltenham, UK; Northampton, MA, USA, 2006. Doi: <https://doi.org/10.4337/9781781951019>.
- Maigret / Roszkowska 2017**
Nicolas Maigret and Maria Roszkowska, "Predictive Art Bot V4 (2017)," 2017. URL: <http://peripheriques.free.fr/blog/index.php?/works/2017-Predictive-Art-Bot-V4/>.
- McCormack / Giffors / Hutchings 2019**
Jon McCormack, Toby Giffors and Patrick Hutchings, "Autonomy, Authenticity, Authorship and Intention in computer generated art," in *Computational Intelligence in Music, Sound, Art and Design 8th International Conference, EvoMUSART 2019, Held as Part of EvoStar 2019, Leipzig, Germany, April 24–26, 2019, Proceedings*, ed. Anikó Ekart, Antonios Liapis and María Luz Castro Pena, Springer Nature Switzerland 2019, pp. 35–50. URL: <https://arxiv.org/abs/1903.02166>. Doi: https://doi.org/10.1007/978-3-030-16667-0_3.
- McCoy / McCoy 2023a**
Jennifer McCoy and Kevin McCoy, "Every Shot, Every Episode," 2023. URL: <https://mccoyspace.com/project/51>.

McCoy / McCoy 2023b

Jennifer McCoy and Kevin McCoy, "Every Anvil," 2023. URL: <https://www.mccoyspace.com/projects/every-anvil>.

Miller 2019

Arthur I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity*, The MIT Press, Cambridge Mass., 2019. Doi: <https://doi.org/10.7551/mitpress/11585.001.0001>.

Morozov 2012

Evgeny Morozov, "The Naked and the TED," *The New Republic*, (2012). URL: <https://newrepublic.com/article/105703/the-naked-and-the-ted-khanna>.

Navas / Gallagher / burrough 2016

The Routledge Companion to Remix Studies, ed. Eduardo Navas, Owen Gallagher and xtine burrough, Routledge, New York, 2016. Doi: <https://doi.org/10.4324/9781315879994>.

Reas / McWilliams / LUST 2010

Casey Reas, Chandler McWilliams and LUST, *Form+Code in Design, Art, and Architecture*, Princeton University Press, Princeton, 2010.

Reben 2018

Alexander Reben, "Five Dollars Can Save the Planet," 2018. URL: <https://areben.com/project/five-dollars-can-save-planet-earth>.

Saltz 2019

Jerry Saltz, "What the Hell Was Modernism?," *New York Magazine*, (2019). URL: <https://www.vulture.com/2019/10/jerry-saltz-new-mo-ma-modernism.html>.

Turk n.d.

Gavin Turk, "Artworks- Pop". URL: <https://www.gavinturk.com/artworks/1993/pop>.

Zeilinger 2021

Martin Zeilinger, *Tactical Entanglements: AI Art, Creative Agency, and the Limits of Intellectual Property*, meson press, Lüneburg, 2021. URL: https://meson.press/wp-content/uploads/2021/03/9783957961846_Zeilinger.pdf. Doi: <https://doi.org/10.14619/1839>.

Abstract

Arif Kornweitz

Inside the High-Dimensional White Cube

Keywords | *Latent space, Generative machine-learning art, White cube*



Drawing on O'Doherty's analysis of the white cube gallery setting, this paper examines how the high-dimensional latent space of machine learning has become a material to be altered by art. By analyzing artists' utilization of this space, particularly Refik Anadol and Nora Al-Badri, who employ generative machine learning for creating art, the paper delves into how these artists manipulate and populate the multi-dimensional space. It contends that the ideology of latent space, its construction through training data, model choices, and the artist's interventions, underscores a cultural moment where taxonomical systems have evolved into predictive and productive tools, creating transient, canonical models for art classification and generation. Moreover, the paper suggests that while latent space operates as a technique for abstraction, it simultaneously challenges artistic practices to invent gestures that reintroduce the concrete, transcending mere classification or generation.

Cite as: Arif Kornweitz, "Inside the High-Dimensional White Cube The Ideology of Latent Space" 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. 47–55. Doi: <https://doi.org/10.48431/hsah.0306>.

Arif Kornweitz Inside the High-Dimensional White Cube

Introduction

There is no such thing as computer art, as the artist Frieder Nake (1971) put it. And still today, “computers do not make art, people do.”¹ But while this observation frees us from having to deal with ‘AI Art’ proper, it certainly does not change the fact that people make art using computers. How, then, are we to understand such practices?

In 1976, Brian O’Doherty published “Inside the White Cube,” a collection of essays that examine how the neutral, white-walled setting “made the gallery space itself the primary material to be altered by art.”² The white cube functions as a container that both influences and restricts the interpretation and reception of artworks. Its sterile setting commodifies art and detaches it from its original context, seemingly eliminating the outside world but letting the “dealer and curator enter from the wings.”³

Crucially, the white cube did not ‘cause’ a change in art-making. Instead, O’Doherty “argues that the communal mind of our culture went through a significant shift that expressed itself in the prominence of the white cube as a central material and expressive mode for art, as well as a fashionable style of displaying it.”⁴ The white cube has to be understood as an expression of particular mode, as O’Doherty proposes:

It often feels as if we can no longer experience anything if we don’t first alienate it. In fact, alienation may now be a necessary preface to experience. Anything too close to us bears the label, “Objectify and Re-ingest.” This

mode of handling experience—especially art experience—is inescapably modern.⁵

Processes of alienation and commodification tend to expand and thus, in what follows, I argue that the high-dimensional latent space of machine learning can be understood as a contemporary white cube, a space that appears to be neutral but has become a material to be altered by art. A transitional device that speaks of another world⁶ and that allows us to objectify, re-ingest, and generate.

Latent space, the higher-dimensional space from which AI Art⁷ is generated, is arguably fundamentally different from a white cube exhibition space. The former is a cultural technique expressed in the technical form of a multi-dimensional space where each dimension represents a specific feature of a dataset, a virtual space of mathematical abstraction. The physical space of the white cube is not only constrained by ordinary dimensionality, it also requires a wholly different form of labor to be populated and allows for different sets of encounters, not least of bodies. Accordingly, I do not propose we view the two spaces as equal: I merely transpose the analysis of one to the other.

Doing so implies attending to the ideology⁸ of latent space in order to understand which cultural mode it is an expression of, and how it shapes and is shaped by artistic practices. To analyze artistic practices that partly play out in latent space as an expression of a cultural mode implies understanding them not in terms of form, as medium-based AI art, but in terms of gesture. As O’Doherty writes, “a gesture wises

1 Hertzmann 2020.

2 McEvilley 1986.

3 O’Doherty 1986, p. 27.

4 McEvilley 1986, p. 11.

5 O’Doherty 1986, p. 52.

6 McEvilley 1986, p. 11.

7 McCormack / Giffors / Hutchings 2019
utilize a useful working definition: “art that is

made using any kind of AI technique.”

8 The full title of O’Doherty’s essay is “Inside the White Cube - The Ideology of the Gallery Space.”

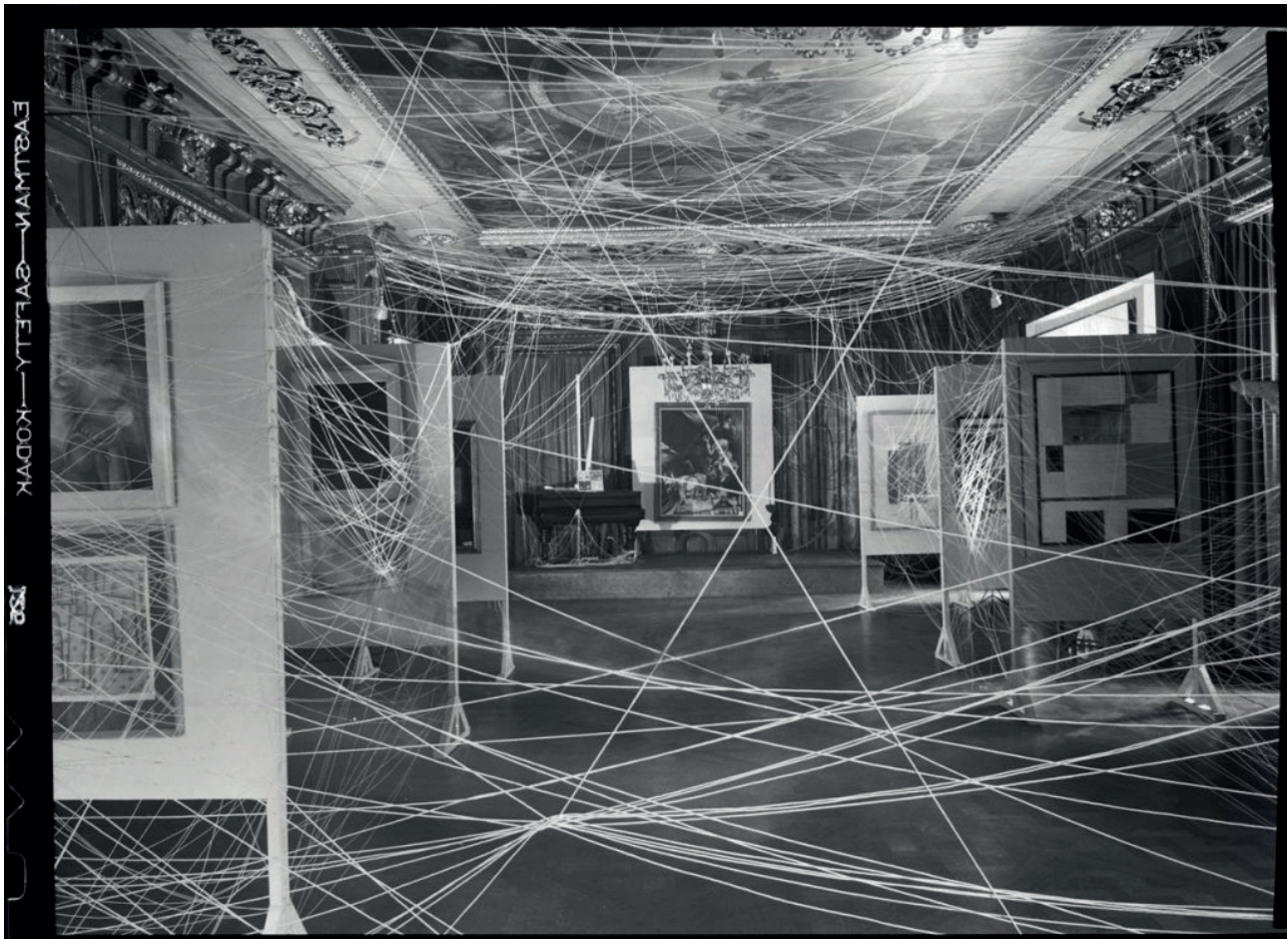


Figure 1. One of the works by Duchamp discussed by O'Doherty is the installation *His Twine* (also referred to as *Mile of String*), which used the gallery space as material. Viewing the criss-cross of string running through an exhibition, one might be reminded of vectors in latent space. Credit: John D. Schiff. Installation view of First Papers of Surrealism exhibition, showing Marcel Duchamp's *His Twine* 1942. Gelatin silver print. Philadelphia Museum of Art.

you up. It depends for its effect on the context of ideas it changes and joins.”⁹

Deconstruction of Categories

To begin to craft such an understanding, I argue that the use of latent space by artists is an expression of the contemporary mode of art that Peter Osborne has described as post-conceptual, under which the ontological status of the work of art is transcategorical. The term post-conceptual does not indicate an art-historical concept at the level of medium, form or style, but rather functions “as the critical register of the historical destruction of the ontological significance of such categories.”¹⁰ Osborne argues that the contemporary work of art routinely involves gestures¹¹ that rely on the deconstruction or destruction of the ontological weight of art historical categories.

O'Doherty's analysis of the white cube offers us a tool to distinguish and understand relations between artistic gestures and latent space. This requires both analyzing the technical constitution of latent space as well as how it is incorporated into practices. Latent space prescribes certain digital input and output modalities, but it can hardly be called an invariant in artistic processes. One key aspect to pay attention to, therefore, is the way that artists populate and manipulate multidimensional space. As Cetinić and She write:

Perhaps the most novel aspect of AI Art is this possibility to venture into that abstract multi-dimensional space of encoded image representations. From the artist's perspective, the latent space is neither a space of reality nor imagination, but a realm of endless

⁹ O'Doherty 1986, p. 70.

¹⁰ Osborne 2013.

¹¹ Or consciously negates such gestures, we might add.

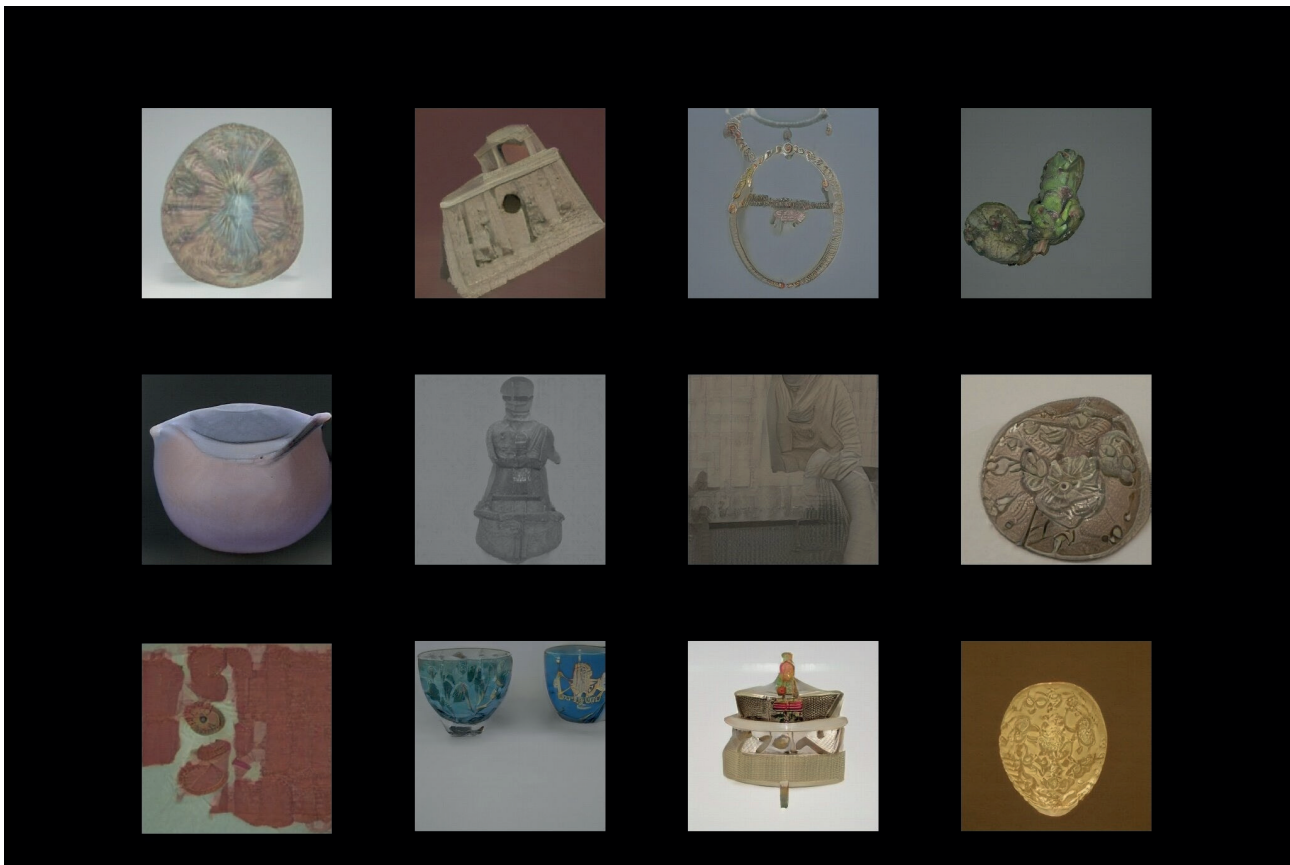


Figure 2. Nora Al-Badri. *Neuronal Ancestral Sculptures Series*. 2020. GAN (General Adversarial Network) art. Copyright Nora Al-Badri. Source <https://www.nora-al-badri.de/works-index#neuronal-ancestral-sculptures-series>

suggestions that emerge from the multi-dimensional interplay of the known and unknown. How one orchestrates the design of this space and what one finds in it, eventually becomes the major task and distinctive “signature” of the artist.¹²

To illustrate this task of orchestrating and traversing latent space, I will introduce the work of Refik Anadol and Nora Al-Badri, two artists who use images of existing works of art as training data in order to create new imagery using generative machine learning. These practices are of interest here because both artists incorporate imagery of existing works of art into their process, allowing us to observe how a work is manipulated in latent space to be incorporated into a new piece.

Refik Anadol and Nora Al-Badri

For *Unsupervised*, Refik Anadol employs machine learning algorithms, such as deep neural networks and generative adversarial networks (GANs), to construct models from training data sets based on the collection of the Museum of Modern Art in New York. Prominently installed in the museum’s foyer in 2022, the resulting video installation visualized a flight through the abstract latent space of the models. The result is a continuously morphing imagery that somehow mimics features of modernist art works, such as the use of color or the arrangement of shapes. Viewing it feels like flicking through an endless array of vaguely familiar pictures, an experience art critics have likened to watching a “lava lamp”¹³ and a “screensaver.”¹⁴ This critique notwithstanding, Anadol’s work is widely popular, has been purchased by the MoMA as well as multiple buyers on the blockchain, and has led him to become the epitome of an AI artist.

¹² Cetinić / She 2022.

¹⁴ Saltz 2023.

¹³ Davis 2023.

For the *Neuronal Ancestral Sculptures Series*, Nora Al-Badri composed training data sets with “10.000k images from 5 different museum collections with the largest collections of Mesopotamian, Neo-Sumerian and Assyrian artefacts.”¹⁵ Notably, her data acquisition included scraping images of artefacts from museum websites, highlighting her critical stance on Western museums’ claims of ownership over dispossessed artefacts and their digital counterparts.¹⁶ By generating alternative versions of these artefacts, Al-Badri addresses and questions these claims from a decolonial perspective. The resulting artefacts represent snapshots of an imaging process that are grounded in a latent space comprising abstractions that meet specific inclusion criteria, effectively combining digital and material archeological methods. As the curator Nadim Samman puts it, each of these “ancestral neuronal sculptures stands for a captured instant of imagination (a process or movement),” a dynamic interplay of human, material, and synthetic memories.¹⁷

Canon, Worldview and Mimesis

How do the artists work with and in latent space? Both Al-Badri and Anadol populate latent space with information derived from training data sets that contain images and metadata of works of art, and employ it in the generation of further artefacts. This exemplifies the post-conceptual tendency of contemporary artists to incorporate other works into their own, e.g. by setting up environments in which works of art function as an index.¹⁸ When combined with the specific and efficient generative qualities of latent space, the index becomes productive of new artefacts. How this productivity is enabled deserves some attention.

Peli Grietzer states that machine learning algorithms learn “how to create facsimiles through a process of trial and error.”¹⁹ The author offers

the conceptual triplet of canon–worldview–mimesis to make sense of this process. According to Grietzer, the training data of a machine learning algorithm can be considered a ‘canon’ on the basis of which facsimiles are created. The feature function, which describes a manifold in latent space, constructs a ‘worldview’. And finally, the projection function, which generates new artefacts, can be considered a method of ‘mimesis’. Accordingly, the use of machine learning can be understood “as the process of determining a worldview and a canon in developing a method of mimesis.”²⁰ Crucially, these three aspects are interconnected in so far as the geometric properties determined by the input function correspond to the feature and projection functions. Taking into account a certain error, the canon is equivalent to the worldview.

Following Grietzer, the worldview a machine learning model has learned can be understood as “an abstractum that cannot be separated from its concreta.”²¹ Machine learning thus allows us to handle an abstraction of a work of art, one that is determined and limited by the technical, taxonomical processing of flattened, digital version of an artefact.²² In latent space, the *abstractum* plays out within the taxonomies inferred from the concrete material provided. Based on this *abstractum*, a new concrete artefact can be generated, for example using style transfer.²³

When trained on images and metadata of works of art, a machine learning algorithm projects abstracted features of a concrete canon into latent space. The algorithm constructs a manifold that functions as a specific worldview, a model of the world. Now, this model can be used to classify other works of art as belonging to the same category or style. It can also be used to generate synthetic images based on the abstractions in latent space, such as the series of images in the works of Anadol and Al-Badri that mimic a specific style.

¹⁵ Al-Badri 2023.

¹⁶ Samman 2023.

¹⁷ Samman 2023.

¹⁸ Lütticken 2023.

¹⁹ Grietzer 2017a.

²⁰ Grietzer 2017a.

²¹ Grietzer 2017b.

²² Krämer 2023.

²³ Gatys / Ecker / Bethge 2015.

Training Data and Model Choices

What comprises a specific worldview depends on what kind of data a machine learning algorithm has been trained on to construct a model, and what kind of patterns can be extracted from that data. In other words, the properties of a manifold in latent space are largely determined by the training data and by the model's algorithm, therefore these need to be taken into account.

In Anadol's and Al-Badri's works, the training data consists of a particular set of images and metadata related to works of art of specific periods. But while Anadol used the MoMA's publicly available dataset of works, an index of the museum's collection,²⁴ Al-Badri did not have access to all of the ancient artefacts she aimed to incorporate, because the objects *and* the digital images representing them are owned by museums and are not always freely available for use. Therefore, the artist resorted to scraping these images from the websites. Here, the gathering of training data becomes part of the work, a decolonial gesture. This illustrates how composing a specific training data set, a canon, has become part of the registers available to artists.

The choice of model influences how artists can interact with latent space as well as the quality of output. Artists may train their own models or use pre-trained models, with different models producing diverging aesthetics. For example, StyleGAN3²⁵ produces imagery that is more realistic than StyleGAN2 by solving the issue of texture sticking and mimicking camera movement better.²⁶ While the results are more realistic, they are also less easily identified as GAN imagery, indicating that we cannot rely on latent space to produce a stable and distinct aesthetic.

Neither Anadol nor Al-Badri communicate clearly what kind of models they have used. In some cases, it may be possible to "attribute a generated image to a particular network."²⁷

Also, it might be helpful to compare output resolution with model cards.²⁸ However, often we are left guessing what kind of approach was used specifically. For example, Anadol states that he has used DCGAN,²⁹ PGAN,³⁰ StyleGAN,³¹ StyleGAN2,³² and StyleGAN2 ADA³³ for the *Machine Hallucinations* series,³⁴ but the artist does not state exactly which models were used for which tasks.

Sometimes the model can be inferred from the date of a work. According to Al-Badri, for the *Neuronal Ancestral Sculptures Series*, "a pre-trained neural network based on [GAN technology]" was used.³⁵ The work was realized in 2019, which may be an indication that the first StyleGAN model was used.³⁶ The use of a pre-trained model indicates that the specific latent space Al-Badri employed to generate artefacts already contained features from one or more data sets, in the case of StyleGAN these would be LSUN³⁷ and CelebA-HQ.³⁸ Using pre-trained models may be a stylistic choice, but it may also simply be too costly to train models from scratch. In any case, latent space is not always initially void and we are facing many different possible latent spaces that artists may use as a starting point.

Partisan Models

To shape the output of a model, artists may want to decide how well a model extracts features. As Casey Reas describes: "In research, the goal is always fidelity. It's always realism. But for example, what Refik was doing with this work is he stopped training early, because he wanted it to be going more toward abstraction, you didn't want anything from the MoMA collection to actually appear in the results."³⁹ Hence, the degree of equivalence of a canon and a worldview is subject to manipulation by the artist.

Finally, there is the issue of bias. Srinivasan & Uchino (2021) investigate biases, such as gender, in "the generative art AI pipeline"

24 MoMA 2023.

25 Karras et al. 2021.

26 See <https://nvlabs.github.io/stylegan3/>.

27 Karras et al. 2020.

28 Mitchell et al. 2019.

29 Radford / Metz / Chintala 2015.

30 Karras et al. 2018.

31 Karras / Laine / Aila 2020.

32 Karras et al. 2020.

33 Karras / Aittala et al. 2020.

34 Anadol 2023.

35 Al-Badri 2023.

36 Karras et al. 2018.

37 Yu et al. 2016.

38 Karras et al. 2018.

39 Anadol et al. 2021.

and note that there is an “increased risk for generative art to be biased.” While this does apply to instances such as a hypothetical use case of an educational software that generates examples of works of a distinct period, it does not apply to artistic practices that consciously make use of latent space. Art needs to be biased because bias is an expression of aesthetic and political choices. Bias is a disposition that marks and testifies to a specific worldview, a stance, not a fault.

If we understand the orchestration of latent space as a way of constructing a specific canon-based manifold, as outlined above, then we need to delegate the mitigation of bias. It is up to the artist to decide which kind of biases needs to be reduced or amplified. Latent space is not autonomous; it is consciously used by artists who make decisions throughout their processes. While the issue of bias certainly matters a great deal when discussing the application of machine learning in general, in contemporary art the issue has to be discussed as one of the aspects of latent space subject to artistic gestures, as part of the “primary material to be altered by art.”⁴⁰

Perhaps the reason that the work of Refik Anadol has been received unfavorably by some critics lies in the fact that the viewer cannot discern such critical gestures in the work. By employing readily available and fairly generic training data, a cross section of modern art, and rendering a flight through latent space, what the artist achieves is an illustration of that space. The viewer primarily witnesses the demonstration of a cultural technique, one of continuous, interchangeable resemblance. We simply see features of a canon mimicked.

Boris Groys (2023) proposes that the degree to which a new work breaks with a canon can be considered the degree of its authenticity. Al-Badri’s *Neuronal Ancestral Sculptures* seem to play on this notion of authenticity by creating non-authentic artefacts. According to the artist, “through GANs, one generates

new materialities, which rise to the surface as the affective qualities of the original in a post-original form.”⁴¹ For the artist, resorting to a “post-original form” is necessary because the original artefacts are unavailable in different ways. As such, the choice of generating new but morphed versions of ancient sculptures can be understood as a pointer to a political issue beyond latent space. Hence, if machine learning can be understood as a device to construct worldviews based on canons,⁴² we need to pay attention to how artists choose, render, situate, elevate or repudiate a canon.

The Ideology of Latent Space

What then, in the spirit of O’Doherty, is the ideology of latent space? What might latent space be an expression of? One answer lies in the paradox of the prominence of the taxonomical technique that is machine learning at a time where works of art defy boundaries of medium and form, where art is post-conceptual and transcategorical. This problem resolves once we understand latent space not as reductive,⁴³ but as a cultural technique for the operationalization of specific and transient canons for the classification and generation of artefacts. The cultural moment that this technique expresses is one where taxonomical systems have become predictive and productive in the service of concrete ends. The triplet of canon-worldview-mimesis put forward by Grietzer⁴⁴ summarizes the ideological mechanism at play.

Latent space is a passageway; a space where two problems, classification and generation, come together through abstraction. Although it is high-dimensional, latent space is a technique for flattening⁴⁵ which sets up abstractions of artefacts to be manipulated further, also beyond computation. Therefore, what matters are the vectors that are drawn beyond latent space, the gestures that reintroduce the concrete. To use latent space as a material for contemporary art thus necessitates the continuous invention of gestures

⁴⁰ McEvilley 1986, p. 10.

⁴¹ Al-Badri 2021.

⁴² Grietzer 2017a.

⁴³ Chiang 2023.

⁴⁴ Grietzer 2017a.

⁴⁵ Krämer 2023.

that situate abstractions. Such gestures are external to latent space: They cannot be neatly classified, subsumed, or generated, unlike an image. Gestures depend on the ideas they join and change.⁴⁶

By simultaneously demonstrating the arbitrariness and efficiency of classification, latent space invites gestures that illustrate “the classically Hegelian function of exceeding a limit in its established form (the aesthetic) in such a way as to render it visible and thereby reinstitute it on new grounds.”⁴⁷ Exemplified by Al-Badri’s work, such gestures largely rely on a critical reparation of established aesthetic categories, e.g. that of sculpture, for a critical purpose.

Latent space augments the white cube as its logical yet fleeting continuation. Logical because it is one of the ways in which contemporary artists may go beyond the white cube, just as they have used the gallery space to go beyond the frame. And fleeting because, as of yet, we can only ever grasp parts of latent space simultaneously. It evades our perception. The eye and the spectator are partly barred from the space, to use the words of O’Doherty. Only when a feature of latent space is rendered in reduced dimensionality, do we get a glimpse of its contents.

As such, the inaccessibility of latent space is deceiving. The challenge it poses to artistic practices is that “we see not the art but the space first,”⁴⁸ an opportunity that Anadol’s *Unsupervised* confidently ignores. Works like that of Anadol might be akin to Yves Klein exhibiting an empty gallery, a gesture that was successful because it made itself superfluous. Now, instead of representing it, we need to understand how latent space joins the white cube “and the picture plane as a unit of discourse.”⁴⁹

Latent space never fails to produce ‘something’. If the white cube implied neutrality, latent space suggests indifference to style

and overabundance of respective expressions. Its automated processing tantalizes with the promise of revealing unexpected connections, pouring them into new mimetic artefacts that could be commodified. Positioning the viewer at a critical point within this automated stream of objectification, ingestion and generation might just be the task in front of us.

⁴⁶ O’Doherty 1986.

⁴⁷ Osborne 2013.

⁴⁸ O’Doherty 1986, p. 14.

⁴⁹ O’Doherty 1986, p. 39.

References

- Al-Badri 2021**
Nora Al-Badri, "The Post-truth Museum – Open Secret," 2021. URL: <https://opensecret.kw-berlin.de/essays/the-post-truth-museum/>.
- Al-Badri 2023**
Nora Al-Badri, "Works," 2023. URL: <https://www.nora-al-badri.de/works-index>.
- Anadol 2023**
Refik Anadol, "Machine Hallucinations : ISS Dreams," 2023. URL: <https://refikanadol.com/works/machinehallucinations-iss/>.
- Anadol et al. 2021**
Refik Anadol et al., "Modern Dream: How Refik Anadol Is Using Machine Learning and NFTs to Interpret MoMA's Collection," *The Museum of Modern Art Magazine*, (2021). URL: <https://www.moma.org/magazine/articles/658>.
- Cetinić / She 2022**
Eva Cetinić and James She, "Understanding and Creating Art with AI: Review and Outlook," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 18, 2, (2022), pp. 1–22. Doi: <https://doi.org/10.1145/3475799>.
- Chiang 2023**
Ted Chiang, "ChatGPT Is a Blurry JPEG of the Web," *The New Yorker*, (2023). URL: <https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web>.
- Davis 2023**
Ben Davis, "An Extremely Intelligent Lava Lamp: Refik Anadol's A.I. Art Extravaganza at MoMA Is Fun, Just Don't Think About It Too Hard," 2023. URL: <https://news.artnet.com/art-world/refik-anadol-unsupervised-moma-2242329>.
- Diel / Lewis 2022**
Alexander Diel and Michael Lewis, "Structural deviations drive an uncanny valley of physical places," *Journal of Environmental Psychology*, vol. 82, (2022), p. 101844. URL: <https://www.sciencedirect.com/science/article/pii/S0272494422000895>. Doi: <https://doi.org/10.1016/j.jenvp.2022.101844>
- Gatys / Ecker / Bethge 2016**
Leon Gatys, Alexander Ecker and Matthias Bethge, "A Neural Algorithm of Artistic Style," *Journal of Vision*, vol. 16, 12, (2016), p. 326. Doi: <https://doi.org/10.1167/16.12.326>.
- Gatys / Ecker / Bethge 2015**
Leon Gatys, Alexander S. Ecker and Matthias Bethge, "Texture Synthesis Using Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, ed. C. Cortes et al., Curran Associates, Inc. 2015. URL: https://proceedings.neurips.cc/paper_files/paper/2015/file/a5e00132373a7031000fd987a3c9f87b-Paper.pdf.
- Grietzner 2017a**
Peli Grietzner "A Theory of Vibe," 2017. URL: <https://www.glass-bead.org/article/a-theory-of-vibe/>.
- Grietzner 2017b**
Peli Grietzner, *Ambient Meaning: Mood, Vibe, System* Ph.D Thesis, Harvard University 2017. URL: <https://dash.harvard.edu/handle/1/39988028>.
- Groys 2023**
Boris Groys, "From Writing to Prompting: AI as Zeitgeist-Machine," 2023. URL: <https://www.e-flux.com/notes/553214/from-writing-to-prompting-ai-as-zeitgeist-machine>.
- Harvey 2022**
Adam Harvey, "Creative Commons and the Face Recognition Problem," 2022. URL: <https://adam.harvey.studio/creative-commons/>.
- Hertzmann 2020**
Aaron Hertzmann, "Computers do not make art, people do," *Communications of the ACM*, vol. 63, 5, (2020), pp. 45–48. Doi: <https://doi.org/10.1145/3347092>.
- Karras et al. 2018**
Tero Karras et al., "Progressive Growing of GANs for Improved Quality, Stability, and Variation," in *6th International Conference on Learning Representations, Vancouver April 30–May 2 2018*, *OpenReview.net* 2018. URL: <https://openreview.net/forum?id=Hk99zCeAb>.
- Karras / Aittala et al. 2020**
Tero Karras, Miika Aittala, et al., "Training generative adversarial networks with limited data," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, ed. H. Larochelle, M. Ranzato and R. Hadsell, Curran Associates Inc. 2020, pp. 12104–12114. Doi: <https://doi.org/10.5555/3495724.3496739>.
- Karras et al. 2021**
Tero Karras et al., "Alias-Free Generative Adversarial Networks," arXiv 2021. Doi: <https://doi.org/10.48550/arxiv.2106.12423>.
- Karras / Laine / Aila 2020**
Tero Karras, Samuli Laine and Timo Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, 12, (2020), pp. 4217–4228. URL: <https://www.computer.org/csdl/journal/tp/2021/12/08977347/1h2AHNHb-gbW>. Doi: <https://doi.org/10.1109/tpami.2020.2970919>.
- Karras / Laine et al. 2020**
Tero Karras, Samuli Laine, et al., "Analyzing and Improving the Image Quality of StyleGAN," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Proceedings*, IEEE Computer Society 2020, pp. 8107–8116. URL: <https://www.computer.org/csdl/proceedings-article/cvpr/2020/716800i107/1m3nK5FM-mWl>. Doi: <https://doi.org/10.1109/cvpr42600.2020.00813>.
- Kornweitz 2024**
Arif Kornweitz, "Inside the High-Dimensional White Cube The Ideology of Latent Space," 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.0306>.
- Krämer 2023**
Sybille Krämer, "Should we really 'hermeneutise' the Digital Humanities? A plea for the epistemic productivity of a 'cultural technique of flattening' in the Humanities," *Journal of Cultural Analytics*, vol. 7, 4, (2023). URL: <https://culturalanalytics.org/article/55592-should-we-really-hermeneutise-the-digital-humanities-a-plea-for-the-epistemic-productivity-of-a-cultural-technique-of-flattening-in-the-humanities>. Doi: <https://doi.org/10.22148/001c.55592>.
- Lütticken 2023**
Sven Lütticken, *Objections, Forms of Abstraction. Volume 1*, Sternberg Press, London, 2023.

McCormack / Giffors / Hutchings 2019

Jon McCormack, Toby Giffors and Patrick Hutchings, "Autonomy, Authenticity, Authorship and Intention in computer generated art," in *Computational Intelligence in Music, Sound, Art and Design 8th International Conference, EvoMUSART 2019, Held as Part of EvoStar 2019, Leipzig, Germany, April 24–26, 2019, Proceedings*, ed. Anikó Ekart, Antonios Liapis and María Luz Castro Pena, Springer Nature Switzerland 2019, pp. 35–50. URL: <https://arxiv.org/abs/1903.02166>. Doi: https://doi.org/10.1007/978-3-030-16667-0_3.

McEvilley 1986

Thomas McEvilley, "Introduction," in *Inside the White Cube: The Ideology of the Gallery Space*, The Lapis Press, San Francisco, 1986, pp. 7–12.

Mitchell et al. 2019

Margaret Mitchell et al., "Model Cards for Model Reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery 2019, pp. 220–229. Doi: <https://doi.org/10.1145/3287560.3287596>.

MoMA 2023

MoMA, "The Museum of Modern Art (MoMA) Collection," MoMA 2023. URL: <https://github.com/MuseumofModernArt/collection>.

Nake 1971

Frieder Nake, "There should be no Computer Art," *PAGE - Bulletin of the Computer Arts Society*, 18, (1971), p. n.p. URL: <https://computer-arts-society.com/casarchive/cas/uploads/page-18.pdf>.

O'Doherty 1986

Brian O'Doherty, *Inside the White Cube: The Ideology of the Gallery Space*, The Lapis Press, San Francisco, 1986.

Osborne 2013

Peter Osborne, *Anywhere or Not at All: Philosophy of Contemporary Art*, Verso Books, London; New York, 2013.

Radford / Metz / Chintala 2015

Alec Radford, Luke Metz and Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," arXiv 2015. Doi: <https://doi.org/10.48550/arxiv.1511.06434>.

Rohrmeier 2022

Martin Rohrmeier, "On Creativity, Music's AI Completeness, and Four Challenges for Artificial Musical Creativity," *Transactions of the International Society for Music Information Retrieval*, vol. 5, 1, (2022), pp. 50–66. URL: <https://transactions.ismir.net/articles/10.5334/tismir.104>. Doi: <https://doi.org/10.5334/tismir.104>.

Saltz 2023

Jerry Saltz, "MoMA's Glorified Lava Lamp Refik Anadol's Unsupervised is a crowd-pleasing, like-generating mediocrity.," 2023. URL: <https://www.vulture.com/article/jerry-saltz-moma-refik-anadol-unsupervised.html>.

Samman 2023

Nadim Samman, *Poetics of Encryption: Art and the Technocene*, Hatje Cantz, Berlin, 2023.

Srinivasan / Uchino 2021

Ramya Srinivasan and Kanji Uchino, "Biases in Generative Art: A Causal Look from the Lens of Art History," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery 2021, pp. 41–51. Doi: <https://doi.org/10.1145/3442188.3445869>.

Weizman 2012

Eyal Weizman, *The Least of All Possible Evils: Humanitarian. Violence from Arendt to Gaza*, Verso, London; New York, 2012.

Yu et al. 2016

Fisher Yu et al., "LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop," arXiv 2016. Doi: <https://doi.org/10.48550/arxiv.1506.03365>.

Abstract

Dominik Vrabič Dežman

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

Keywords | Artificial Intelligence (AI), Visual Culture, Imaginaries, Aesthetics, Politics



Artificial Intelligence (AI) is becoming quickly entangled into the fabric of social life. In its growing scope of implementation, AI has been critiqued for being promoted as an inevitable general-purpose technology, and at other times as “practice-based speculation.” Within the current “AI hype,” deep-blue sci-fi stock images employed to represent AI have flooded the mediasphere, closely accompanying the sensationalized language employed to mediate the technology to the public. In liaison with media outlets, search engines and online stock photography marketplaces form sites of fabrication of the visions that constitute the socio-technical imaginary surrounding AI, torqueing the collective archive of public future visions of AI. In this paper, I build on the normative imperative of studying the “ethics [and] politics of AI images.” I draw on a range of interpretive theories of images to establish a framework for the examination of images of AI as a reciprocal link between aesthetics and politics, and to problematize images of AI as operational, performing active work. In doing so, I engage a range of case studies from art and design to explore what responses have emerged to the dominant clichés in the public images of AI. I extend the claim that image-making is always embedded in a socio-technical context, and that every act of visualizing AI is essentially an artistic gesture. The public image of AI rests on the production and dissemination of media images. The mechanisms through which divergent image-making enters the public eye are becoming increasingly crucial to the study of the mediation of AI to the public.

Cite as: Dominik Vrabič Dežman “Interrogating the Deep Blue Sublime: Images of Artificial Intelligence in Public Media” 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. 57–66. Doi: <https://doi.org/10.48431/hsah.0307>.

1. Introduction

When looking up the term “Artificial Intelligence” (AI) and inspecting the image search results, the mood quickly turns blue (fig. 1). In recent years, sensationalist stock images employed to mediate AI to the general public have pervaded digital media. Search engines serve as windows onto the digital-media image economy, showcasing the ubiquity of futuristic, sci-fi images of AI across leading media outlets and institutional, and governmental websites.¹ The images advance the narrative of AI as a frictionless, weightless, and imminent force, while concealing the fact that the global enterprise of AI is made up of “material resources, human labor, and data.”² Understanding AI as a global resource-intensive material enterprise, it quickly becomes clear that these images are not engaging in a representational relationship: The signifiers (anthropomorphic robots, binary rain, floating sci-fi interfaces, religious iconography, infinite deep-blue space) bear no ‘necessary’ connection to the signified (“AI”). While images of AI have been studied for perpetuating prototypical whiteness,³ racism,⁴ sensationalized religious iconography,⁵ sci-fi tropes and clichés,⁶ fear mongering,⁷ employment of metaphors,⁸ and ethical implications,⁹ theoretically grounded accounts squaring the images’ aesthetic and political dimensions from the perspective of image-making remain largely absent.

In this study I investigate the dominant images employed to represent AI in public digital media, such as those returned when a person looks up the keyword “Artificial Intelligence”

online. If the dominant images do not meaningfully represent AI, it becomes germane to ask: What are these images and what are they doing? And who might be served by these images? *Cui bono*? And furthermore, what could it mean to contest them and ‘do otherwise?’

I advance a framework for considering public images of AI as a link between aesthetics and politics.¹⁰ Following a discussion of the dominant images of AI in public media, I turn to cohere a range of responses from art and design into an account of aesthetic resistance to the hegemonic visual regime. In doing so, I probe the tension between AI’s empirical and normative dimensions and ultimately explores the role of images of AI as both representational and agentic, in service of a broader socio-economic operation. The study seeks to problematize the responsibility of image-makers and all other agents involved in the proliferation and circulation of these images.

2. Images of AI as Matter of Politics

The hegemony of sci-fi images of AI, combined with sensationalist writing and the indiscriminate use of the term in the media,¹¹ ostensibly precludes a more nuanced social relationship with AI innovation: “One can watch thousands of similar images of AI without having to develop any critical reasoning about AI.”¹² I have previously proposed the ‘deep blue sublime’ as a provisional concept to grapple with sensationalist popular media images of emergent technologies such as AI, the cloud, bitcoin, blockchain.¹³ The deep blue sublime picks out both an aesthetic category, defined by

¹ Vrabič Dežman 2023.

² Crawford / Joler 2018, p. 31.

³ Cave / Dihal 2020.

⁴ Mhlambi 2019.

⁵ Singler 2020.

⁶ Steenson 2018.

⁷ Dihal 2020.

⁸ Wallenborn 2022.

⁹ Romele 2022.

¹⁰ Rancière 2013, Lavin 2001, Romele 2022.

¹¹ Devedzic 2022.

¹² Romele 2022, p. 12.

¹³ Vrabič Dežman 2023.

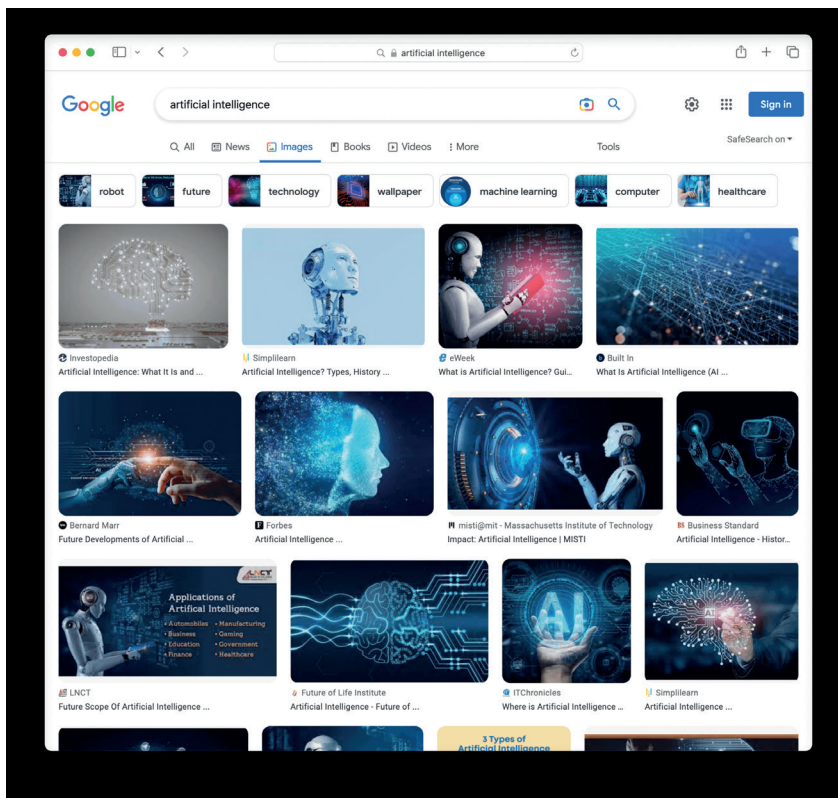


Figure 1. Google Images search results for “Artificial Intelligence”.

The search was conducted on 30 July 2023 from Amsterdam, Netherlands from a newly instantiated Safari browser instance.

oft-blue sci-fi visuals, and an epistemic orientation homologous to the algorithmic sublime, gathered around “‘astonishment, awe, terror, and psychic distance’— feelings once reserved for natural wonders or intense spiritual experiences.”¹⁴ As critical discourse surrounding unchecked industry promises of AI grows,¹⁵ a groundswell of increased scrutiny and familiarity with the technology has largely reduced the sublime hold of AI images to bathos.¹⁶

Employing sci-fi stock images as “[short-cuts] to a common ideal”¹⁷ aids the top-down permeation of a socio-technical and corporate¹⁸ imaginary about AI, imposing ethico-political “collective [future] visions of the good society.”¹⁹ In historical bids to secure funding and further AI development,²⁰ normative visions have represented “conditions of possibility for the existence and development [of AI].”²¹ The dominant images of AI orient our gaze towards an “ever-receding horizon of the future,”²² brought ever closer by AI innovation and development. Rather than mirroring societal hopes

or beliefs surrounding AI, these images illuminate how “few [actors] furnish the collective imaginary on behalf of [the] many.”²³

The status quo of the visibility of AI is fundamentally a matter of politics. Following Rancière, politics revolve “around what is seen and what can be said about it, around who has the ability to see and [...] speak.”²⁴ A central concept for organizing the relationship between ‘aesthetics,’ understood as “configurations of experience that create new modes of sense perception and induce novel forms of political subjectivity,”²⁵ and politics, is the notion of the “distribution of the sensible”: “a set of relations between the perceptible, the thinkable and the doable that defines a common world.”²⁶ The term picks out the dynamic inner organization and interlinkage between our sensory environment, and its link with hegemonic political organization.²⁷ Images of AI in public media, almost exclusively stock images, form the backbone of our ‘public visual world,’²⁸ and gradually torque

¹⁴ Ames 2018, p. 2.

¹⁵ Powles / Nissenbaum 2018, Burrell 2023, Larson 2022, Whittaker / O’Leary 2023, Siegel 2023, Bender 2022.

¹⁶ Craft-Jenkins 2012, p. 4.

¹⁷ Johnson, qtd. in Steenson 2018.

¹⁸ Crawford 2021, p. 47.

¹⁹ Jasanoff / Kim 2015, p. 123, Jasanoff / Kim 2009.

²⁰ Cardon / Cointet / Mazières 2018.

²¹ Romele 2022, p. 2.

²² Burrell 2023.

²³ Vrabčič Dežman 2023.

²⁴ Rancière 2006, p. 13.

²⁵ Rancière 2006, p. 9.

²⁶ Rancière 2013, p. 11.

²⁷ Rancière 2006, p. 40.

²⁸ Aiello 2016.

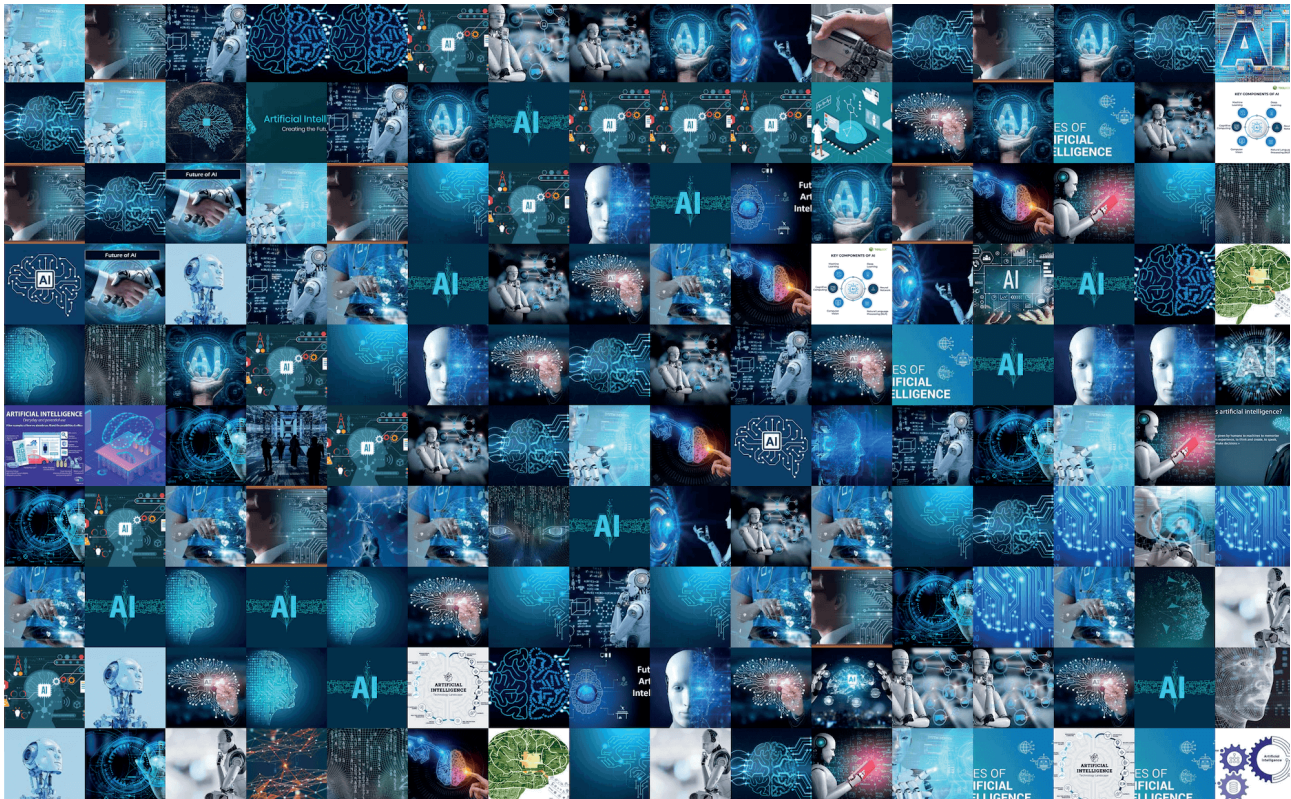


Figure 2. Sample of top-ranking Google Images search results for “Artificial Intelligence”.

July-August 2022, searching from Amsterdam, Netherlands. The figure shows (from top to bottom) the first-ranked search results. The images were cropped to fit into squares. The entire dataset spanning 29 days and 2654 images is archived and may be accessed at <https://perma.cc/8PUN-CHHW>.

the contours of public discourse surrounding emergent tech, by rendering visible certain interpretations of AI while foreclosing other by the same token. The images represent entryways into the examination of the politics of AI.

Jacques Rancière’s political ontology rests on oppositional political organizations of ‘consensus’ and ‘dissensus.’ Consensus is characterized by a general agreement on the distribution of the sensible, limiting the range of possible political action and expression. On the other hand, dissensus challenges the existing distribution of the sensible, creating new possibilities for political action and expression,²⁹ while Rancière mainly elaborates on the radical potential of art to expose “a gap in the sensible.”³⁰ I contend that design *qua* applied art, operating at the frontline of media production, may more directly enact political dissensus. However, in the context of the contemporary production of images used to represent AI,

most image-makers’ political power remains ‘hamstrung’ in the service of corporate incentives.³¹ Hence, the importance of independent artistic production and institutional support and funding becomes central to the potential of design to create spaces of dissensus.

3. Interrogating the Deep Blue Sublime

With few exceptions, public images of AI equate to stock images woven from sci-fi tropes and clichés, employing ‘yesterday’s metaphors to sketch out tomorrow’s technologies’ (fig. 2).³² From a redemptive perspective, the exalted sci-fi visuals might be considered benign, “problematizing human issues through AI as a metaphor.”³³ Considering their circulation and dominance in public media, these images “[promote] forms of consensus about the general hopes and fears about AI,”³⁴ reducing their political power. Through their incessant reproduction and circulation, the currently

²⁹ Rancière 2013, p. 26.

³⁰ Rancière / Corcoran 2010, p. 38.

³¹ Lavin 2001, p. 2, Frosh 2003.

³² Vrabčič Dežman 2023.

³³ Hermann 2020.

³⁴ Romele 2022, p. 13.

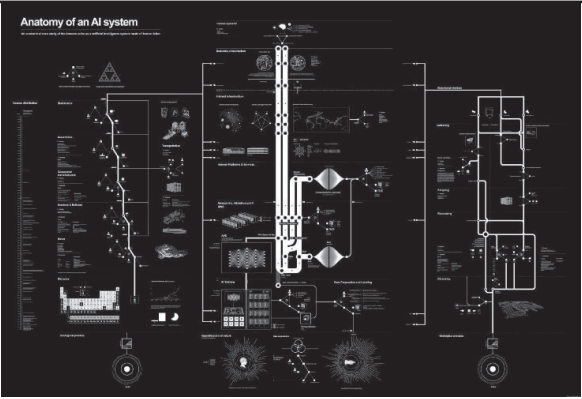
Type of Strategy	Visualization Example
1. Using Symbolic Language Otherwise Science fiction and film (e.g. Interstellar 2014, annihilation 2018) (Kahn 2016, Better Images of AI 2023)	 Adam Ferris, from Kahn 2016.
2. Documentary 'realism' Journalistic outlets (e.g. Yuan 2018) (Better Images of AI 2023)	 Data annotators labeling data. Credits: Nacho Kamenov & Humans in the Loop / Better Images of AI from https://betterimagesofai.org/images?artist=NachoKamenov&title=Dataannotatorslabelingdata. Licenced by CC-BY 4.0.
3. Cognitive Mapping (Pasquinelli / Joler 2020, Crawford / Jowler 2018)	 Crawford, Kate, and Vladan Joler. Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources. 2018. Source: Crawford / Joler 2018, https://anatomyof.ai/img/ai-anatomy-map.pdf.

Table 1. Typology of heterogeneous directions of aesthetic resistance towards dominant tropes in the visualization of AI.

dominant images “anesthetize any form of ‘disagreement’,”³⁵ and foreclose trajectories of looking elsewhere and doing otherwise.

The AI image problem rests on the fact that AI “is difficult to visually represent, but is represented nevertheless.”³⁶ Understood as a scientific discipline or global material enterprise, visualization reveals a part, at best. Showing AI is an artistic gesture, including via photography or critical cartography and their attendant processes of framing and selection. If the dominant images are “unpolitical,” as they “anesthetize any debate and disagreement about AI,”³⁷ and perpetuate the same regime of signification by employing the same metaphors,³⁸ “aesthetic resistance” could be understood as the production of such images of AI that may generate spaces of disagreement or dissensus.³⁹ Such a trajectory might be spearheaded by “pensive” images, which promote political polyvocality by “[bringing] together different regimes of expression without homogenizing them.”⁴⁰ Recent years have seen the production of a range of artistic and design interventions seeking to decrease the stronghold of the deep-blue public visibility of AI (table 1).

First, using symbolic language in innovative ways may dilute the current monolithic visibility while inaugurating avenues to alternative ways of being – and seeing – with AI. In doing so, ‘interesting’ new visions can emerge, encouraging curiosity, discursivity, exchange, and social learning.⁴¹ Here, success might be defined by visualizations that advance visions by means of *poiesis*,⁴² the critical political capacity of artistic production to bring into being something not-yet-seen or done, redispensing social relations in a consequential manner and opening spaces of dissensus, disturbing the distribution of the sensible. Second, drawing attention back to the ‘daily’ materiality of AI draws awareness to its bounded terrestriality, reinforcing the established conception of AI as a global infrastructural enterprise. This

reading foregrounds the relationship between human labor, extractionist industrial processes, and incessant volumes of data required to power the global AI operation.⁴³ Simultaneously, reducing AI to a sum of its planetary infrastructure dissimulates its crucial normative dimension “while the currently prevalent way of showing the infrastructure tends to anonymise the actors involved.”⁴⁴ Third, interventions seeking to map out the machinations of the global AI enterprise orient investigations towards the complex web of material constituents of AI. Such interventions may be prone to a “fantasy of overview,” attempting to produce a “seeing from nowhere,” as if situated outside of historical context and interest.⁴⁵ The value of overview mappings might lie in the ways such interventions underwrite the limitation, if not impossibility, of rendering AI sensible through a discrete image.

Visualizing AI differently and diversely helps to re-sensitize our existing interpretations of AI and opens up onto alternative modalities of social relation, with the potential of diffusing “the monopoly of the forms of description”⁴⁶ of currently hegemonic imagery. Changing one node within the interconnected matrix of the distribution of the sensible also changes the interpretative means by which to perceive others. Although the interventions examined within this study are mostly funded projects—freed from the corporate strictures and market forces binding commercial stock image producers⁴⁷—their power to rework the public visibility of AI remains ‘hamstrung.’ Employing prudence in choosing and using AI imagery⁴⁸ meets resistance as alternative figurations—mostly stemming from small-scale art and design initiatives—remain hidden from view. After being brought into existence, these divergent figurations face a range of hurdles before being seen in the public eye, and possibly torquing the hegemonic visual regime. Media outlets often have exclusive contracting agreements with major stock image platforms, which have

35 Romele 2022, p. 12.

36 Romele 2022, p. 4.

37 Romele 2022, p. 17.

38 Wallenborn 2022.

39 Ranci re 2013, Romele 2022.

40 Romele 2022, p. 15.

41 Ngai 2015.

42 Ziarek 2004.

43 Cf. Crawford / Joler 2018.

44 Demos 2017, p. 11, Romele 2022.

45 Toscano / Kinkle 2015, p. 354, Haraway 1988.

46 Ranci re 2013, p. 27.

47 Lavin 2001, Frosh 2003.

48 Romele 2022, p. 4.



Figure 3. DALL-E image outputs for query “Artificial Intelligence”. July 2023. Prompted by the author.

become the monopolistic provider of digital media imagery.⁴⁹ The stock image platforms’ recommendation and sorting algorithms, and the search engine’s own politics and ranking hierarchies⁵⁰—which tend to ossify the archive of already existing top-ranked images and prioritize already-dominant sources—further hinder the interventions’ public visibility.

The increasing adoption of AI-generated stock imagery represents another affront to the possibility of aesthetic resistance.⁵¹ Generative image synthesis networks are largely trained on online stock imagery, cementing the existing visuals through an operation of recursive aesthetics, in which the inputs become the outputs (fig. 3). A media environment consisting of the automated production and dissemination of AI media imagery earmarks a new visual frontier of algorithmic governmentality.⁵² Furthermore, as future models continue to be trained on existing web stock images, the proliferation of AI-generated stock imagery might lead to generative AI “model collapse”⁵³

threatening further development of the technology. The ontological connection between AI image generation and stock imagery, in which a given image represents an interchangeable semantic unit based on a “visual essence”⁵⁴ merits an independent future study.

4. From Surface to Action

By remaining fixated on the surface, either affectively or analytically through hermeneutic interpretation, one may be diverted from material operations happening behind the scenes of such an image economy. Rather than representing AI per se, the dominant images might be reframed as “rendering visually a set of otherwise invisible forces,”⁵⁵ namely the “economic and algorithmic logic behind stock images [of AI],”⁵⁶ consisting of socio-political operations to fortify public trust and expedite the widespread development and adoption of AI products. Rather than representing “the world out there,”⁵⁷ operational images actively participate in an operation.⁵⁸ The hegemonic

⁴⁹ Frosh 2020.

⁵⁰ Rogers 2018.

⁵¹ Edwards 2022, Attié 2022.

⁵² Celis Bueno 2017, p. 748.

⁵³ Shumailov et al. 2023.

⁵⁴ Crawford / Paglen 2021, Frosh 2020.

⁵⁵ Celis Bueno 2017, p. 744.

⁵⁶ Romele 2022, p. 3.

⁵⁷ Flusser (1983) 2000, p. 15.

⁵⁸ Farocki 2004 cited in Hoel 2018, p. 12, *Photography off the scale* 2021, p. 217.

images of AI seek to decrease technological literacy while carving out space for the unfettered growth and integration of these technologies into the socio-economic fabric of our daily lives.⁵⁹ Another core tenet of operational images is their creation by machines for consumption by other machines.⁶⁰ A reading of the dominant stock images as either human-made or algorithmically-made diminishes the deep entanglement “between algorithms put into practice and the social tactics of users who take them up.”⁶¹ The public (in)visibility of images of AI is deeply intertwined with algorithmic logics, as outlined in the preceding section.

I contend that the sci-fi visuals “respond to a specific need that [other types of images of AI] are unable to address.”⁶² If we treat the contents of the currently dominant images as a window onto the normative dimension of AI, consisting of a historical progression of promises made to a diverse set of stakeholders with material commitments, these images ‘do’ represent AI and, furthermore, index the current status quo of ‘AI hype.’ Under this reading, the deep blue sublime earmarks the main distinction between the imagined, normative dimensions of AI, and the empirical, material operations of machine learning (ML).⁶³ Hence, AI could be reframed as a series of socio-technical and corporate imaginaries,⁶⁴ speculative promises made to a diverse range of stakeholders, while ML might refer to the terrestrial, data-hungry infrastructure of these operations.

5. Conclusion

Ultimately, the question of the production and dissemination of images of AI is a question of political translation, of “what happens to private visions in public forums.”⁶⁵ This emphasizes a reading of politics organized around the control of images: Which private visions become public, who gets to see what, and who has a say in what can be seen.⁶⁶ While it may be tempting

to point towards the dominant images of AI as reflecting societal beliefs, hopes, or fears, it is crucial to read them as the estuary of a longer, value-laden politico-economic flow. They are partially of their makers, mostly stock image producers, and importantly of their consumers: Editorial actors in the media, with attendant commercial ties and imperatives.

I concur that “one cannot develop any comprehensive understanding of AI without taking into account [its imaginaries].”⁶⁷ The question of public images of AI is fundamentally a question of “the production of meaning as a mechanism of social subjection.”⁶⁸ The dominant images serve to persuade rather than comprehend,⁶⁹ and aid the delivery rather than deliberation of AI innovation. Several critical initiatives have responded to the tired blue visuals: Using symbolic language and fiction differently to define new speculative modes of relating to AI; documentary ‘realism,’ an optical reorientation towards the daily, terrestrial, lived reality of AI; and cognitive mapping, attempting to map out the sprawling network of global AI operations. What remains relevant to explore is which kind of images could open up the “possibility of thinking and building a new form of (technological) democracy.”⁷⁰ Every act of visualizing AI, read as normative visions enacted through material infrastructure, is inherently an artistic gesture. Importantly, artistic production does not happen in a socio-historical vacuum. Future research might undertake historical analysis into the origins of the current AI visibility or track ‘AI hype’ over time as this has been expressed in the images used to mediate the technology in public media. At present, strategies of aesthetic resistance are faced with a series of hurdles towards reaching the public eye. Changing the visions means agitating the systems through which these visions become public.

⁵⁹ Vrabčič Dežman 2023.

⁶⁰ Hoel 2018, p. 13, Pantenburg 2016, p. 58.

⁶¹ Gillespie 2014, p. 183.

⁶² Romele 2022, p. 10.

⁶³ As the search query approaches a specific level of ML application, the results lose their sublime qualities. For example, a

search for “Neural Network,” a widespread machine learning architecture, is far less sensationalist than “Artificial Intelligence.”

⁶⁴ Sartori / Bocca 2023.

⁶⁵ Lavin 2001, p. 2.

⁶⁶ Cf. Rancière 2006 Rancière / Corcoran 2010, Rancière 2013.

⁶⁷ Romele 2022, p. 2.

⁶⁸ Celis Bueno 2017, p. 744.

⁶⁹ Romele 2022, p. 8.

⁷⁰ Romele 2022, p. 11.

References

- Aiello 2016**
Giorgia Aiello, "Taking Stock," 2016. URL: <https://ethnographymatters.net/blog/2016/04/28/taking-stock/>.
- Ames 2018**
Morgan G. Ames, "Deconstructing the algorithmic sublime," *Big Data & Society*, vol. 5, 1, (2018), p. 4. Doi: <https://doi.org/10.1177/2053951718779194>.
- Attié 2022**
Ivanna Attié, "AI Generated Images: The Next Big Thing in Stock Media," 2022. URL: <https://www.stockphoto secrets.com/stock-agency-insights/ai-generated-images.html>.
- Bender 2022**
Emily M. Bender, "Look behind the curtain: Don't be dazzled by claims of 'artificial intelligence' | Op-Ed," *The Seattle Times*, (2022). URL: <https://www.seattletimes.com/opinion/look-behind-the-curtain-dont-be-dazzled-by-claims-of-artificial-intelligence/>.
- Better Images of AI n.d.
Better Images of AI, "Better Images of AI". URL: <https://betterimagesofai.org>.
- Better Images of AI 2021**
Better Images of AI, "Better images of AI," 2021. URL: <https://www.bbc.co.uk/rd/blog/2021-12-artificial-intelligence-machine-stock-image-library>.
- Burrell 2023**
Jenna Burrell, "Artificial Intelligence and the Ever-Receding Horizon of the Future," 2023. URL: <https://www.techpolicy.press/artificial-intelligence-and-the-ever-receding-horizon-of-the-future/>.
- Cardon / Cointet / Mazières 2018**
Dominique Cardon, Jean-Philippe Cointet and Antoine Mazières, "Neurons spike back. The invention of inductive machines and the artificial intelligence controversy," *Réseaux*, vol. 211, 5, (2018), pp. 173–220. URL: <https://www.cairn-int.info/journal-reseaux-2018-5-page-173.htm>. Doi: <https://doi.org/10.3917/res.211.0173>.
- Cave / Dihal 2020**
Stephen Cave and Kanta Dihal, "The Whiteness of AI," *Philosophy & Technology*, vol. 33, 4, (2020), pp. 685–703. Doi: <https://doi.org/10.1007/s13347-020-00415-6>.
- Celis Bueno 2017**
Claudio Celis Bueno, "Harun Farocki's Assignifying Images," *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, vol. 15, 2, (2017), pp. 740–754. URL: <https://www.triple-c.at/index.php/tripleC/article/download/874/1039>. Doi: <https://doi.org/10.31269/triplec.v15i2.874>.
- Craft-Jenkins 2012**
Kyle Craft-Jenkins, *Artificial Intelligence and the Technological Sublime: How Virtual Characters influence the Landscape of Modern Sublimity* Ph.D Thesis, University of Kentucky 2012. URL: https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1000&context=english_etds.
- Crawford 2021**
Kate Crawford, "Earth," in *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence*, Yale University Press, New Haven; London, 2021, pp. 23–51.
- Crawford / Joler 2018**
Kate Crawford and Vladan Joler, "Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources," 2018. URL: <https://anatomyof.ai/>.
- Crawford / Joler 2019**
Kate Crawford and Vladan Joler, "Anatomy of an AI System," *Virtual Creativity*, vol. 9, 1, (2019), pp. 117–120. URL: https://intellectdiscover.com/content/journals/10.1386/vcr_00008_7. Doi: https://doi.org/10.1386/vcr_00008_7.
- Crawford / Paglen 2021**
Kate Crawford and Trevor Paglen, "Excavating AI: The politics of images in machine learning training sets," *AI & SOCIETY*, vol. 36, 4, (2021), pp. 1105–1116. Doi: <https://doi.org/10.1007/s00146-021-01162-8>.
- de Mul 2011**
Jos De Mul, "The Technological Sublime," 2011. URL: <http://www.nextnature.net/2011/07/the-technological-sublime/>.
- Demos 2017**
Thomas J. Demos, *Against the anthropocene: Visual culture and environment today*, Sternberg press, Berlin, 2017.
- Devedzic 2022**
Vladan Devedzic, "Identity of AI," *Discover Artificial Intelligence*, vol. 2, 1, (2022), p. 23. Doi: <https://doi.org/10.1007/s44163-022-00038-0>.
- Dihal / Duarte 2023**
Dr Kanta Dihal and Tania Duarte, "Better Images of AI. A Guide for Users and Creators," The Leverhulme Centre for the Future of Intelligence and We and AI. 2023. URL: <https://blog.betterimagesofai.org/wp-content/uploads/2023/02/Better-Images-of-AI-Guide-Feb-23.pdf>.
- Dihal 2020**
Kanta Dihal, "Can artificial superintelligence match its hype?," *Physics Today*, vol. 73, 2, (2020), pp. 49–50. URL: <https://pubs.aip.org/physicstoday/article/73/2/49/914456/Can-artificial-superintelligence-match-its-hype>. Doi: <https://doi.org/10.1063/pt.3.4410>.
- Edwards 2022**
Benj Edwards, "Adobe Stock begins selling AI-generated artwork," *Ars Technica*, (2022). URL: <https://arstechnica.com/information-technology/2022/12/adobe-stock-begins-selling-ai-generated-artwork/>.
- Farocki 2004**
Harun Farocki, "Phantom Images," *Public*, (2004). URL: <https://public.journals.yorku.ca/index.php/public/article/view/30354>.
- Flusser 2000**
Vilém Flusser, *Towards a Philosophy of Photography*, Reaktion Books, London, 2000.
- Frosh 2003**
Paul Frosh, *The image factory: Consumer culture, photography and the visual content industry*, New technologies/new cultures series, Berg, Oxford, 2003.
- Frosh 2020**
Paul Frosh, "Is Commercial Photography a Public Evil? Beyond the Critique of Stock Photography," in *Photography and Its Publics*, ed. Melissa Miles and Edward Welch, Routledge, Taylor & Francis Group, 2020, pp. 187–206. Doi: <https://doi.org/10.5040/9781350054998.ch-010>.
- Garland 2018**
Alex Garland, "Annihilation," Paramount Pictures, Skydance Media, Scott Rudin Productions 2018.

Gillespie 2014

Tarleton Gillespie, "The Relevance of Algorithms," in *Media Technologies: Essays on Communication, Materiality and Society*, ed. Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot, The MIT Press, 2014, pp. 167–194. URL: <https://direct.mit.edu/books/edited-volume/3021/chapter-abstract/82555/The-Relevance-of-Algorithms?redirectedFrom=fulltext>. Doi: <https://doi.org/10.7551/mitpress/9780262525374.003.0009>.

Haraway 1988

Donna Haraway, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective," *Feminist Studies*, vol. 14, 3, (1988), pp. 575–599. Doi: <https://doi.org/10.2307/3178066>.

Hermann 2020

Isabella Hermann, "Beware of fictional AI narratives," *Nature Machine Intelligence*, vol. 2, 11, (2020), pp. 654–654. Doi: <https://doi.org/10.1038/s42256-020-00256-0>.

Hoel 2018

Aud Sissel Hoel, "Operative Images. Inroads to a New Paradigm of Media Theory," in *Image – Action – Space*, ed. Luisa Feiersinger, Kathrin Friedrich and Moritz Queisner, De Gruyter, 2018, pp. 11–28. Doi: <https://doi.org/10.1515/9783110464979-002>.

Jasanoff / Kim 2009

Sheila Jasanoff and Sang-Hyun Kim, "Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea," *Minerva*, vol. 47, 2, (2009), pp. 119–146. Doi: <https://doi.org/10.1007/s11024-009-9124-4>.

Joler / Pasquinelli 2020

Vladan Joler and Matteo Pasquinelli, "The Nooscape Manifested: AI as Instrument of Knowledge Extractivism," 2020. URL: <https://web.archive.org/web/20230626222127/https://nooscape.ai/>.

Kahn 2016

Nora Kahn, "Towards a Poetics of Artificial Superintelligence," 2016. URL: <https://medium.com/after-us/towards-a-poetics-of-artificial-super-intelligence-ebff11d2d249>.

Larson 2021

Erik J. Larson, *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*, Belknap Press of Harvard University Press, Cambridge, Mass., London, 2021.

Larson 2022

Erik J. Larson, *The myth of artificial intelligence: Why computers can't think the way we do*, First Harvard University Press paperback edition., The Belknap Press of Harvard University Press, Cambridge, Massachusetts, 2022.

Lavin 2001

Maud Lavin, *Clean New World: Culture, Politics, and Graphic Design*, The MIT Press, Cambridge, Mass., London, 2001.

Mhlambi 2019

Sabelo Mhlambi, "God in the image of white men: Creation myths, power asymmetries and AI," 2019. URL: <https://cyber.harvard.edu/story/2019-03/god-image-white-men-creation-myths-power-asymmetries-and-ai>.

Ngai 2015

Sianne Ngai, *Our Aesthetic Categories: Zany, cute, interesting*, Harvard University Press, Cambridge, Massachusetts London, England, 2015.

Nolan 2014

Christopher Nolan, "Interstellar," Paramount Pictures, Warner Bros., Legendary Entertainment 2014.

Pantenburg 2016

Volker Pantenburg, "Working images: Harun Farocki and the operational image," in *Image Operations. Visual Media and Political Conflict*, ed. Jens Eder and Charlotte Klöck, Manchester University Press, 2016, pp. 49–62. URL: <https://academic.oup.com/manchester-scholarship-online/book/16019/chapter-abstract/170983252?redirectedFrom=fulltext>.

Parikka 2023

Jussi Parikka, "Operational Images: Between Light and Data," *e-flux Journal*, 133, (2023). URL: <https://www.e-flux.com/journal/133/515812/operational-images-between-light-and-data/>.

Pasquinelli / Joler 2020

Matteo Pasquinelli and Vladan Joler, "The Nooscape manifested: AI as instrument of knowledge extractivism," *AI & SOCIETY*, vol. 36, 4, (2020), pp. 1263–1280. Doi: <https://doi.org/10.1007/s00146-020-01097-6>.

Photography off the scale 2021

Photography off the scale: Technologies and theories of the mass image, ed. Tomáš Dvořák and

Jussi Parikka, Technicities, Edinburgh University Press, Edinburgh, 2021.

Powles / Nissenbaum 2018

Julia Powles and Helen Nissenbaum, "The Seductive Diversion of 'Solving' Bias in Artificial Intelligence," *Medium*, (2018). URL: <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>.

Rancière 2006

Jacques Rancière, *The Politics of Aesthetics. The Distribution of the Sensible*, ed. Gabriel Rockhill, Revelations, Bloomsbury Academic, London, New York, 2006.

Rancière 2010

Jacques Rancière, *Dissensus: On politics and aesthetics*, ed. Steve Corcoran, Continuum, London ; New York, 2010. URL: <http://artsites.ucsc.edu/sdaniel/230/Ranciere%20-%20Dissensus%20-%20On%20Politics%20and%20Aesthetics.pdf>.

Rancière 2012

Jacques Rancière, "In What Time Do We Live?," in *The State of Things*, ed. Maria Kuzma, Pablo Lafuente and Peter Osborne, Office for Contemporary Art (OCA) Norway, Koening Books, London, 2012, pp. 8–37.

Rancière 2013

Jacques Rancière, "In What Time Do We Live?," *Política Común*, vol. 4, (2013), Doi: <https://doi.org/10.3998/pc.1232227.0004.001>.

Rogers 2018

Richard Rogers, "Aestheticizing Google critique: A 20-year retrospective," *Big Data & Society*, vol. 5, 1, (2018). Doi: <https://doi.org/10.1177/2053951718768626>.

Romele 2022

Alberto Romele, "Images of Artificial Intelligence: A Blind Spot in AI Ethics," *Philosophy & Technology*, vol. 35, 4, (2022). Doi: <https://doi.org/10.1007/s13347-022-00498-3>.

Sartori / Bocca 2023

Laura Sartori and Giulia Bocca, "Minding the gap(s): Public perceptions of AI and socio-technical imaginaries," *AI & SOCIETY*, vol. 38, 2, (2023), pp. 443–458. Doi: <https://doi.org/10.1007/s00146-022-01422-1>.

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>.

Siegel 2023

Eric Siegel, "The AI Hype Cycle Is Distracting Companies," *Harvard Business Review*, (2023). URL: <https://hbr.org/2023/06/the-ai-hype-cycle-is-distracting-companies>.

Singler 2020

Beth Singler, "The AI Creation Meme: A Case Study of the New Visibility of Religion in Artificial Intelligence Discourse," *Religions*, vol. 11, 5, (2020), p. 253. Doi: <https://doi.org/10.3390/rel11050253>.

Stallabrass 2007

Julian Stallabrass, "What's in a Face? Blankness and Significance in Contemporary Art Photography," *October*, The MIT Press vol. 122, (2007), pp. 71–90. URL: <https://www.jstor.org/stable/40368490>.

Stenson 2018

Molly Wright Stenson, "A.I. Needs New Clichés," 2018. URL: <https://maximolly.medium.com/ai-needs-new-clich%C3%A9s-ed-0d6adb8cbb>.

Toscano / Kinkle 2015

Alberto Toscano and Jeff Kinkle, *Cartographies of the absolute*, Zero Books, Alresford, Hants, 2015.

Vrabič Dežman 2023

Dominik Vrabič Dežman, "Defining the Deep Blue Sublime," *SETUP* 2023. URL: <https://www.deepbluesublime.tech/>.

Vrabič Dežman 2024

Dominik Vrabič Dežman, "Interrogating the Deep Blue Sublime: Images of Artificial Intelligence in Public Media," 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.0307>.

Vrabič Dežman / Vis 2022

Dominik Vrabič Dežman and Dirk Vis, *Encountering the Deep Blue Sublime: Artificial Intelligence in the Media Imaginary*, self-published, Tržič, 2022. URL: https://books.google.it/books?id=gRU_zwEACAAJ.

Wallenborn 2022

Johnna T. Wallenborn, "AI as a flying blue brain? How metaphors influence our visions of AI," 2022. URL: <https://www.hiig.de/en/ai-metaphors/>.

Whittaker / O'Leary 2023

Meredith Whittaker and Lizzie O'Leary, "What the President of Signal Wishes You Knew About A.I. Panic," *Slate*, (2023). URL: <https://slate.com/technology/2023/05/meredith-whittaker-interview-geoffrey-hinton-ai-threats.html>.

Yuan 2018

Li Yuan, "How Cheap Labor Drives China's A.I. Ambitions," *The New York Times*, (2018). URL: <https://www.nytimes.com/2018/11/25/business/china-artificial-intelligence-labeling.html>.

Zeilinger 2021

Martin Zeilinger, "Generative Adversarial Copy Machines," *Culture Machine*, ed. Peter Jakobsson, Anne Kaun and Fredrik Stiernstedt, vol. 20, (2021). URL: <https://culturemachine.net/vol-20-machine-intelligences/generative-adversarial-copy-machines-martin-zeilinger/>.

Ziarek 2004

Krzysztof Ziarek, *The Force of Art*, Stanford University Press, Redwood City CA, 2004.

Abstract

Pepe Ballesteros Zapata

Evolving Methodologies: Computation in Art History

Keywords | Digital Art history, Visual arts, Computational technologies



This essay delves into methodology as the core element to integrate computation within the practice of art history. A cursory examination over time reveals how art history has witnessed various methodological shifts that guide the study of cultural objects. The integration of computation resemblances are yet another iteration echoing past transitions in methodological approaches. Through illustrative research examples inside digital art history, the essay highlights computation as a shift in the point of departure, used to add layers of interpretation which can iteratively contribute to a deeper understanding and interpretation of art. Moreover, the article examines the methodological dichotomy inherent in combining computational and traditional art historical methods. This centaur-like methodological fusion has the opportunity of enhancing the art historical practice by recognizing quantifiable features in traditional non-quantifiable cultural objects.

Cite as: Pepe Ballesteros Zapata, "Evolving Methodologies: Computation in Art History" 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. 69–73. Doi: <https://doi.org/10.48431/hsah.0308>.

Introduction

Just as art historians may feel uneasy when computer scientists use the term ‘art history’ to refer to paintings, computer scientists and digital humanists also face debates when seeking a suitable term for applying computer-based analysis to art studies. Scholars in the field have attempted to resolve this issue. In the well-known article “Is There a ‘Digital’ Art History?,” Johanna Drucker differentiates between a “digitized” art history and a “digital” art history. The former refers to art historical practices using digitized collections, while the latter refers to “analytic techniques enabled by computational technology.”¹ It seems like the naming relates to the medium, while the definition relates to the methods. Following this line of thought, a tentative name for the field might be ‘Computational art history,’ a term I have strikingly only seen being used by Richard Johnson.² The reason for my surprise lies in the fact that other research fields like linguistics, biology, or sociology had already chosen the term ‘computational’ to differentiate between established and computational methods specifically. Just as many other fields of research seamlessly integrate computation into their practice without the need for a separate identity, so too may digital art history follow suit. As discussed by Schich,³ computation could evolve to become an inseparable part of art history, unifying naturally without the burden of a distinct name. Nuria Rodríguez recognizes Drucker’s terminology distinction but suggests an updated definition. She defines digital art history “as a series of epistemological and methodological transformations.”⁴ She says that “there is not an essential distinction between digital art history and other art histories – there is simply

doing art history.” The goal is the same, what changes is the method. This article explores how the choice of a particular method guides and constrains the layers of interpretation given to the object of study. Arguably, that has been the case already in art history, and computation is simply another iteration. Seeking to reduce art history to a limited—and also incomplete—taxonomy of methods carries the risk of yielding banal assumptions and conclusions, and will most likely annoy the cultivated reader. However, amidst these hazards, there is value in contextualizing art historical methods as a preliminary step towards understanding the role of computational thinking in the field.

Visual Analysis of Art

One of the first relevant methodologies to create art history was bibliographical. Painters and art connoisseurs like Vasari, Van Mander, or Bellori, exploited documentary evidence to report the life of artists. The bibliographical approach also sought to differentiate the arts from the crafts, alluding to their role as a liberal art in contrast to a mechanical work. The method is based on a close examination of individuals and their work, imposing their own ideals of artistic quality and influencing the creation of a canon. Emphasis on individual artists and their works as embodiments of intellectual and aesthetic ideals set the stage for the centrality of visual analysis in art historical practice. As the discipline became more established, art historians like Giovanni Morelli or Bernard Berenson⁵ used their educated eye to diagnose works of art (e.g., assess authenticity, style, or workshop practices), and assess their artistic significance in relation to the broader artistic landscape.

¹ Drucker 2013.

² Doing / Johnson 2015.

³ Schich 2016.

⁴ Rodríguez-Ortega 2019.

⁵ For canonical examples of art historical connoisseurship: Morelli 1900. This

technique was further developed by Berenson in Berenson 1948.

Historians and scientists have a long tradition of working together to empower visual analysis techniques, where technology is used as a microscope to date, authenticate, or discover enigmatic clues hidden inside works of art. In multispectral imaging, ultraviolet light is employed to discover extraneous paint marks, potentially from later restorations aiming to repair damages (e.g., cracks on the wood panel). Infrared imaging unveils previously hidden underdrawings, obscured by pigment degradation and varnishing. X-rays provide insights into a painting's genesis and original composition, revealing its initial outlines.⁶ The combination of multispectral images with computational software makes it possible to study how colors may have changed over time, bringing the original painting back to life—at least digitally.⁷ Not limited by digital restoration, less invasive computational methods like image processing have been used to help historians date books and prints,⁸ or to identify watermarks in artworks for authentication.⁹

Layers of Meaning

Since the late 20th century, the 'new art history' incorporates intellectual thought and ideological movements to further integrate art history within the framework of cultural history. The method I see applied here is the incorporation of theoretical frameworks into their practices, for example, interpreting art history through the lens of historical materialism (like the work of Meyer Schapiro or Arnold Hauser), or feminism (Griselda Pollock or Chris Weedon), among many others.¹⁰ This is strong evidence of the richness of art history, where the entire tapestry of artistic evolution and individual artworks can be perceived and interpreted in a myriad of ways. Like the meaning of words depends on context, artworks embody multiple dimensions of meaning. The act of interpreting an artwork invites a spectrum of possible understandings rather than deterministic conclusions. Meaning

is defined by the space between the artwork and the chosen method of analysis, subject to the particular perspective of an individual or a collective, and their cultural traditions. If meaning is multidimensional, computational methods allow the historian to add another layer to the high-dimensional space of meaning.

Body-pose estimation algorithms can be used to capture gesture in art, a method used to reinterpret Aby Warburg's concept of *Pathosformeln*.¹¹ Measures could identify distinct figures and their corresponding *Pathosformeln*, and highlight notable contrast between the movements of the upper and lower body parts, reflecting the signature of *Pathos* as a break in the body's unity. Madhu and colleagues used computational methods to study compositional structures in art historical images inspired by Max Imdahl's theories on image composition.¹² They used pose and gaze estimation to extract artwork's structures, tracing focal areas or regions of action to understand narratives in art historical scenes. A distant viewing study of landscape painting reveals dominant dissection ratios that serve as unique signatures across the history of western art.¹³ Researches have used computational methods to discuss art historical inquiries in the absence of contextual information. Using image processing and optimization algorithms to calculate the position of the light source in paintings, researchers presented a tentative rebuttal to the Hockney-Falco thesis,¹⁴ claiming that local light sources were insufficiently bright at the time to project virtual images on a canvas.¹⁵ Facial recognition frameworks are also used to enhance art historical discussions when only visual evidence is available. For example, to suggest the possible appearance of Leonardo da Vinci in his old age,¹⁶ or to identify key figures in imperial portraits to trace systems of lineage.¹⁷ Data-driven methods represent yet another possibility to add new layers of meaning. For instance, a distant reading of artwork titles

6 Refer to <https://www.metmuseum.org/articles/vermeer-conservation-for-an-example>.

7 Cotte / Dupraz 2006.

8 Blair Hedges 2006.

9 Johnson 2020.

10 Harris 2002. I see the way intellectual

frameworks guide the study of cultural phenomena as a method. This is evident in recent art historical trends that question the validity of western-based theories as a practice to study cultures outside their geographical borders.

11 Impett / Süssstrunk 2016.

12 Madhu et al. 2020.

13 Lee et al. 2020.

14 The controversial Hockney-Falco thesis argues that old masters used optical aids to project virtual images on the canvas.

15 Kale / Stork 2009.

16 Tyler / Smith / Stork 2012.

17 Zhong 2023.

prompts inquiries into the formation and evolution of artistic canons as reflected in museum collections.¹⁸ Network analysis is used to disambiguate the complex social dynamics of buying, brokering, and selling in the art market.¹⁹ The hidden patterns found across large amounts of data unveil underlying structures that inevitably lead to new questions. For example, what can explain a pattern of synchronization found across different markets at the same time?

Methodological Dichotomy

Jacob Burckhardt reacted to Hegelian metaphysics when he described philosophy of history as a centaur made up of two incompatible parts, one rational and the other irrational.²⁰ The metaphor yields skepticism about the blending of history—chronological evidence as occurred—with philosophy—arrangement of events within a theoretical framework to derive meaning. The dichotomy between disciplines echoes Snow's delineation of the division between the scientific and humanistic cultures.²¹ This division, often perceived as an incompatibility, might more accurately describe fundamental differences in the method and ultimate goal: The understanding of cultural phenomena as opposed to the understanding of natural phenomena. The examined 'what'—the object of study—naturally affects the 'how'—the methodological approach. But what happens when we maintain the 'what' and change the 'how'? This is what digital art history explores inside the field of art history.

The art historical community can agree on the explanation of a phenomenon, or argumentation can reach a general consensus if the argumentation is convincing enough—that is, contextualizing the problem and supporting the commentary with some sort of evidence. In art history, empirical evidence is often merged with the interpretation of cultural contexts, semiotics, aesthetic experience, and other complex features that are inherently unquantifiable. Digital art history preserves the empirical evidence—the artwork, archival materials, etc.—but transforms the way to

approach it—the method. The application of computational methods—inherited from scientific methods—to art historical evidence suggests that traditional non-quantifiable cultural objects may have quantifiable features. This suggestion carries stereotypes associated with scientific methods, like 'objectivity' or 'techno-positivism.' The study of art as something computable has often been perceived as an opportunistic approach linked to neo-liberalization that aims to produce 'useful' outputs.²² However, transforming cultural artifacts into metrics can be seen as having the potential to shift the point of departure, to reveal information not immediately apparent, and to add layers of information that iteratively might lead to meaning. The resulting centaur—part computational, part art historical—is not a creature of oversimplification but a hybrid entity capable of synthesizing two different yet compatible realms of understanding.

Conclusion

Historically, art historians have aligned their approaches with the evolving intellectual and technological landscapes. The research examples showcased in this article are intended to position digital art history not as a fracture but as a continuum in the methodological evolution of art history. The intersection between art historical knowledge and computational methods offers an opportunity to add new layers of meaning by analyzing visual evidence in the absence of contextual information, reinterpreting art historical theories, directly assessing art historical claims, or studying complex cultural dynamics at scale. The methodological dichotomy inherently present in digital art history guides researchers to recognize two situations: The first one recognizes that complex cultural objects may also exhibit attributes amenable to quantification, the second one recognizes that computation serves as an alternative method and cannot solely generate art historical meaning.

¹⁸ Bowman 2023.

¹⁹ Schich et al. 2017.

²⁰ Fernie 1995.

²¹ Snow 1959.

²² Drucker / Bishop 2019.

References

- Ballesteros Zapata 2024**
Pepe Ballesteros Zapata, "Evolving Methodologies: Computation in Art History," 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.0308>.
- Berenson 1948**
Bernard Berenson, *The Italian Painters of the Renaissance*, Oxford University Press, London, 1948.
- Blair Hedges 2006**
Stephen Blair Hedges, "A method for dating early books and prints using image analysis," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 462, 2076, (2006), pp. 3555–3573. Doi: <https://doi.org/10.1098/rspa.2006.1736>.
- Bowman 2023**
Mike Bowman, "Text-mining metadata: What can titles tell us of the history of modern and contemporary art?," *Journal of Cultural Analytics*, vol. 8, 1, (2023). URL: <https://culturalanalytics.org/article/74602-text-mining-metadata-what-can-titles-tell-us-of-the-history-of-modern-and-contemporary-art>. Doi: <https://doi.org/10.22148/001c.74602>.
- Cotte / Dupraz 2006**
Pascal Cotte and Damien Dupraz, "Spectral imaging of Leonardo Da Vinci's Mona Lisa: A true color smile without the influence of aged varnish," in *CGIV 2006 3rd European Conference on Colour in Graphics*, Society of Imaging Science and Technology 2006, pp. 311–317. Doi: <https://doi.org/10.2352/cgiv.2006.3.1.art00064>.
- Criminisi / Kemp / Zisserman 2002**
Antonio Criminisi, Martin Kemp and Andrew Zisserman, "Bringing Pictorial Space to Life: Computer techniques for the analysis of paintings," Microsoft Research 2002. URL: <https://www.microsoft.com/en-us/research/publication/bringing-pictorial-space-to-life-computer-techniques-for-the-analysis-of-paintings/>.
- Doing / Johnson 2015**
Park Doing and C.Richard Jr Johnson, "On Applying Signal Processing to Computational Art History: An Interview," *International Journal for Digital Art History*, 1, (2015), pp. 88–96. URL: <https://journals.ub.uni-heidelberg.de/index.php/dah/article/view/21637>. Doi: <https://doi.org/10.11588/dah.2015.1.21637>.
- Drucker 2013**
Johanna Drucker, "Is There a 'Digital Art History?'," *Visual Resources. An international journal on images and their uses*, vol. 29, 1–2, (2013), pp. 5–13. Doi: <https://doi.org/10.1080/01973762.2013.761106>.
- Drucker / Bishop 2019**
Johanna Drucker and Claire Bishop, "A Conversation on Digital Art History," in *Debates in the Digital Humanities 2019*, ed. Matthew K. Gold and Lauren F. Klein, University of Minnesota Press, Minneapolis, 2019, pp. 321–334. Doi: <https://doi.org/10.5749/j.ctvg251hk>.
- Fernie 1995**
Eric Fernie, *Art history and its methods: A critical anthology*, Phaidon, London, 1995.
- Harris 2002**
Jonathan Harris, *The New Art History: A Critical Introduction*, Routledge, Taylor & Francis Group, London; New York, 2002. Doi: <https://doi.org/10.4324/9780203466780>.
- Impett / Süssstrunk 2016**
Leonardo Impett and Sabine Süssstrunk, "Pose and Pathosformel in Aby Warburg's Bilderatlas," in *Computer Vision – ECCV 2016 Workshops Amsterdam, The Netherlands, October 8–10 and 15–16, 2016. Proceedings, Part I*, ed. Gang Hua and Hervé Jégou, Springer International Publishing 2016, pp. 888–902. Doi: https://doi.org/10.1007/978-3-319-46604-0_61.
- Johnson 2020**
C.Richard Johnson, "Decision Trees for Watermark Identification in Rembrandt's Etchings," *Journal of Historians of Netherlandish Art*, vol. 12, 2, (2020). URL: <https://jhna.org/articles/decision-trees-for-watermark-identification-in-rembrandts-etchings>. Doi: <https://doi.org/10.5092/jhna.12.2.5>.
- Kale / Stork 2009**
David Kale and David G. Stork, "Estimating the position of illuminants in paintings under weak model assumptions: An application to the works of two Baroque masters," in *SPIE Proceedings - Human Vision and Electronic Imaging XIV*, ed. Bernice E. Rogowitz and Thrasyvoulos N. Pappas, 2009, Doi: <https://doi.org/10.1117/12.806267>.
- Lee et al. 2020**
Byunghwee Lee et al., "Dissecting landscape art history with information theory," *Proceedings of the National Academy of Sciences*, Proceedings of the National Academy of Sciences vol. 117, 43, (2020), pp. 26580–26590. URL: <https://www.pnas.org/doi/full/10.1073/pnas.2011927117>. Doi: <https://doi.org/10.1073/pnas.2011927117>.
- Madhu et al. 2020**
Prathmesh Madhu et al., "Understanding Compositional Structures in Art Historical Images Using Pose and Gaze Priors," in *Computer Vision – ECCV 2020 Workshops*, ed. Adrien Bartoli and Andrea Fusiello, Springer International Publishing 2020, pp. 109–125. Doi: https://doi.org/10.1007/978-3-030-66096-3_9.
- Morelli 1900**
Giovanni Morelli, *Italian painters; critical studies of their works*, J. Murray, London, 1900.
- Rodríguez-Ortega 2019**
Nuria Rodríguez-Ortega, "Digital Art History: The Questions that Need to Be Asked," *Visual Resources. An International Journal on images and their uses*, vol. 35, 1–2, (2019), pp. 6–20. Doi: <https://doi.org/10.1080/01973762.2019.1553832>.
- Schich 2016**
Maximilian Schich, "Figuring Out Art History," *International Journal for Digital Art History*, 2, (2016), pp. 42–66. URL: <https://journals.ub.uni-heidelberg.de/index.php/dah/article/view/24761>. Doi: <https://doi.org/10.11588/dah.2016.2.24761>.
- Schich et al. 2017**
Maximilian Schich et al., "Network Dimensions in the Getty Provenance Index," arXiv 2017. Doi: <https://doi.org/10.48550/arxiv.1706.02804>.
- Snow 1959**
Charles Percy Snow, *The two cultures and the scientific revolution*, The Rede Lecture, Cambridge University Press, New York, 1959.
- Tyler / Smith / Stork 2012**
Christopher W. Tyler, William A.P. Smith and David G. Stork, "In search of Leonardo: Computer-based facial image analysis of Renaissance artworks for identifying Leonardo as subject," in *SPIE Proceedings – Human Vision and Electronic Imaging XVII*, ed. Bernice E. Rogowitz, Thrasyvoulos N. Pappas and Huib de

Ridder, SPIE Press 2012, Doi: <https://doi.org/10.1117/12.904749>.

Zhao / Akdağ Salah / Salah 2023
Shu Zhao, Almila Akdağ Salah and Albert Ali Salah, "Automatic Analysis of Human Body Representations in Western Art," in *Computer Vision – ECCV 2022 Workshops Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part I*, ed. Leonid Karlinsky, Tomer Michaeli and Ko Nishino, Springer Nature Switzerland 2023, pp. 282–297. Doi: https://doi.org/10.1007/978-3-031-25056-9_19.

Zhong 2023
Grace Zhong, "A computer vision-aided analysis of facial similarities in Song dynasty imperial portraits," *Electronic Imaging*, Society for Imaging Science and Technology vol. 35, (2023), pp. 1–6. Doi: <https://doi.org/10.2352/ei.2023.35.13.cvaa-212>.

Abstract

Valentine Bernasconi

Automated Body Pose Estimation on Painting:
A Crisis of Uncertainty

Keywords | Digital art history, AI art research,
Computer sciences



Defining a project in digital art history is both exciting and promising, offering the potential to uncover artistic trends and inspirations over time and to train machines to see what has been overlooked by traditional art historical practices. However, the reality of these interdisciplinary projects often falls short of expectations due to the limitations of large digital corpuses and computational methods. Machine learning models, typically trained on natural images from contemporary life, do not achieve the same accuracy with paintings, partly because the ideal corpus of paintings does not yet exist. Many paintings remain undigitized, and not all collections are prepared for computational analysis, lacking the annotations needed for effective machine learning training. Furthermore, assumptions about the digital material and fundamental notions of art history can lead to a narrow, formalist vision of the field. Drawing on a project focused on the computational and historical analysis of hands in Early Modern art, this paper reflects on the gap between initial research questions and the adjustments needed when applying AI and art historical methodologies. It highlights the challenges of navigating between the divergent and demanding fields of art history and computer science, especially when confronting preconceptions and uncertainties. The paper questions the feasibility of finding a middle ground between these disciplines, suggesting that the integration of AI into art historical research requires not only technical adjustments but also a reevaluation of methodological approaches to accommodate the complexities of art.

Cite as: Valentine Bernasconi, "Automated Body Pose Estimation on Painting: A Crisis of Uncertainty" 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. 75–81. DOI: <https://doi.org/10.48431/hsah.0309>.

Automated Body Pose Estimation on Painting: A Crisis of Uncertainty

1. Conducting a digital project in art history

Starting a project in digital humanities often follows a standard procedure which consists of gathering a large collection of data and then cleaning and processing it with computational models for information extraction and statistical analysis. It is the data processing pipeline.¹ From a digital art history perspective, the process also implies adapting to specificities of the field of art history through the definition of research questions, with a core understanding of the corpus of study, as well as the development of methodological solutions to adapt computational methods.

It is within this context that this research on the understanding of hand gestures in early modern paintings was conducted. Painted hands represent a complex object of study from an art historical perspective,² and constitute a challenging field of exploration from a digital perspective³ with the recent development of performative computer vision methods for body-pose estimation.⁴ At the intersection of art history and computer science, the research project primarily aimed to outline the evolution and use of specific hand poses in time and place through the creation of interactive tools for visualization. It explored the possibility to reach a comprehensive understanding of hand poses in the early modern period, with the analysis of their specific codification and potential gestural systems. Quickly reduced to a more humble scope after the first curatorial steps towards the creation of a corpus of hands, the project had to align with the

reality of working with artistic and art historical material, and address important and too often hidden issues of the framework imposed by the digital environment. Above all, it raises questions about the possibility of globally studying hand gestures on a corpus limited by its metadata and poor computer vision results, as well as the way results can be properly addressed when built upon curatorial biases. These two primary factors also outline fundamental issues in the way we conduct projects in digital art history and how some assumptions on both the digital and art historical realm impact research.

2. Implications on processing historical data

The dataset and its curation represent a foundational layer to research conducted in the digital realm. The very first step consists of understanding the material, and to perform adequate cleaning in order to further process it with computational models. In our case, the dataset consists of the photographic collection from the Bibliotheca Hertziana:⁵ Digitized photographs of paintings broadly ranging from the 14th to the 17th Century, and their corresponding metadata. This metadata includes the titles, authors, and dates of creation of the artworks, and were inconsistently produced in German, English, French, and Italian. ‘Inconsistency’ is a keyword in such an interdisciplinary context as it expresses the diverging reality of information representation between art history and computer science: The transcription of the knowledge produced in art history does not naturally meet the digital

¹ Kaplan 2015.

² See the works of Dimova 2020, Barasch 1987, Gombrich / Huxley 1966 Wittkower 1977.

³ See the works of Impett 2020, Madhu et al. 2022, Bell / Schlecht / Ommer 2013, Impett / Süssstrunk 2016.

⁴ Cao et al. 2021, Güler / Neverova /

Kokkinos 2018, Jin et al. 2020.

⁵ See more details of the photographic collection at URL: <https://foto.bibl.hertz.it>.

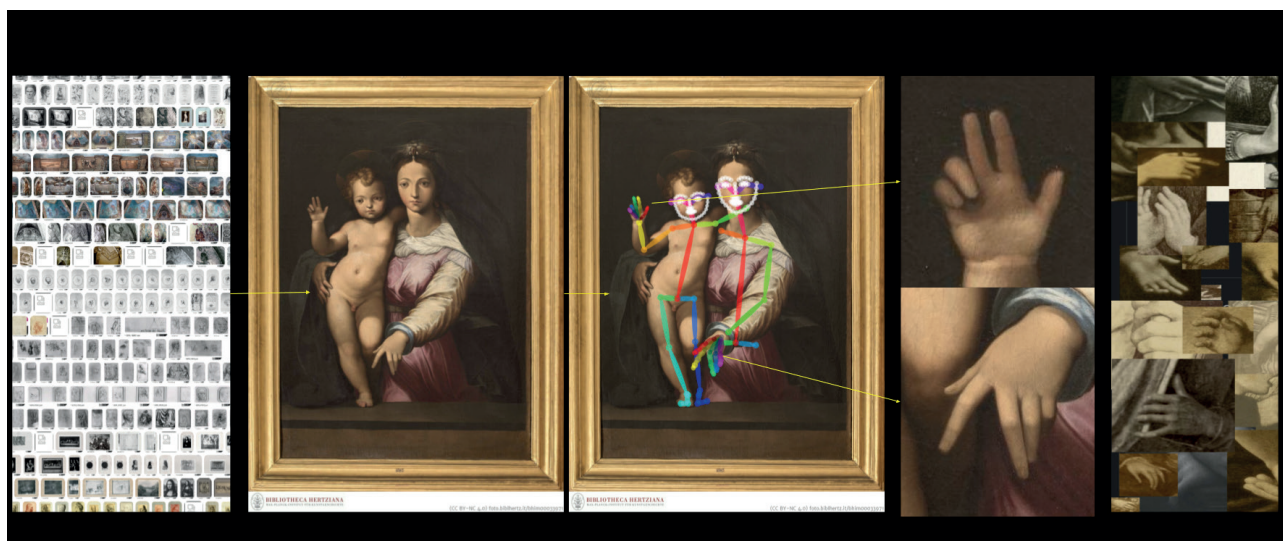


Figure 1. Different steps for the extraction of hands in paintings
Composition by the Author.

requirements. Still, the photographic collection of the Bibliotheca Hertziana is an extreme case: Titles come in different languages and are not standardized according to specific iconographies, the names of artists are not consistent from one annotator to another and may be referred to with a nickname or surname, dates come with many question marks and with non-numerical formats and broad formulations, and the place of creation of an artwork is usually missing.

The inspection of this metadata established the initial curatorial steps, where titles and authors had to be standardized with the help of the Integrated Authority File⁶ and Iconclass norms,⁷ question marks removed, and dates formatted into brackets of times. In the process, information was lost, such as the markers of uncertainty of some data, as well as the typical model of knowledge produced by art historians. The example of the question mark used to address the uncertainty of date attributions is strongly evocative of this loss. Based on the favored type of representation system, the information becomes right or wrong, and the removal of a question mark at

the end of a year means shifting the data from the unsure into a true and established fact.⁸

The next step consisted of extracting hands from the paintings with the help of a machine learning model for human body-pose estimation. The process involved a model called OpenPose,⁹ which uses a system of keypoints to describe the position of the bodies detected on any given image. As we can see in fig. 1, based on these keypoints, the hands were automatically extracted from the original images in order to build a collection of hands. Additionally, a set of features were created from the keypoints to better characterize the pose of the hands, such as the use of unit vectors for the direction of the fingers, and the angles shaped by the joints. Such geometric characterization better helps the comparison of different poses, regardless of their formal aspect.¹⁰

As it has been shown and discussed before,¹¹ paintings are difficult material from a computer vision perspective. OpenPose was pre-trained on natural images¹² and gives a lower accuracy when applied to artworks. Hence, many bodies were not found in images,

⁶ Also called GND (Gemeinsame Normdatei), see URL: https://gnd.network/Webs/gnd/EN/Home/home_node.html for more information.

⁷ Brandhorst / Posthumus 2017.

⁸ Standards have been established for the representation of temporal uncertainty in computer science. Yet, as one of the primary

goals was to create a chronological order for the representation of the hand pose, a choice was made to remove this uncertainty, thus shifting the uncertain information to a form of truth.

⁹ Cao et al. 2021.

¹⁰ For example, the features were used for the creation of an artwork browser based

on the hand pose input of the user. See Bernasconi 2022.

¹¹ Bernasconi / Cetinić / Impett 2023, Madhu et al. 2022.

¹² By natural images, we here refer to photographs representing scenes of our everyday and contemporary life.

and consequently not all the hands are present in the final collection. Furthermore, many hands detected were not correctly estimated. Based on the visual information provided by the keypoints, the collection of hands had to be manually cleaned, reducing the corpus of study from 18,641 detected hands to a total of 5,993 correctly estimated hands.

3. Addressing uncertainties of machine curation

These two simple curatorial steps—a preamble of further research on hands—show important issues that impact potential results of the work and the way research questions should be addressed from a digital art history perspective. First, there is the question of uncertainty induced by the results from the machine learning model. The material produced, a total of 5,993 hands, is not only a fairly small amount of data from a computational perspective but it is also the expression of missing information when put back in the context of their original artworks. Indeed, it is difficult to speak about gestural patterns within a composition when barely two hands out of five bodies were detected on a single painting. Consequently, the primary question about outlining gestural systems has to be rephrased to interrogate the way a narrative can be properly built from the lack of accuracy of the computer vision system. It now becomes an acknowledgement of what the digital methods ‘obscure’¹³ rather than what they develop.

Furthermore, the missing metadata and the re-formatting of the latter introduce another layer of uncertainty regarding its exactness, provenance, and certainty. A phenomenon observed in multiple projects that is now being progressively addressed.¹⁴ These issues, brought by the use of computational processes on art historical material, shed light on the important notions of bias and subjectivity. Whether it is the description of the hand pose through keypoints, the computer scientist implementing the machine learning model, the art historian registering data

in an archival system, or the digital humanist processing metadata, every decision represents a bias that has to be made visible and taken into account when addressing potential results of a digital project. As further explained by Frédéric Kaplan “data are never given, but taken and transformed [...] at each step of the data processing, choices are made and biases apply. Understanding these technical choices is crucial to develop new interpretive theories.”¹⁵ In our context, a proper trade-off has to be found between the right amount of cleaning processes and the preservation of the original content in order to make claims and outline tendencies in the dataset. Yet, no ground truth can be established from the results, and more data resolutely does not mean a more objective truth.¹⁶

This deceptive aspect of digital methods not only results in a current difficulty to build more art historical knowledge but it is also a matter of the fundamental assumptions made in both the field of computer science and art history. Indeed, the peculiar context requires some expertise on two complex fields that have different approaches in the treatment of data and the establishment of new knowledge. Therefore, we do not only need a better understanding of the subjective choices that occur in the curation process, but also in the way we perceive and apprehend the field of art history, how computational approaches can or should be integrated to it, and the novelty they may bring.

4. Reconsidering fundamental assumptions

The assumptions made about the practice of art history and the simplification of concepts this entails are one of the main issues that prevent the successful development of a field of digital art history. In our case, the research did not start by questioning the notion of influence and potential artistic intentions in an art historical context, nor the computational approach itself which implies a strong formal approach to art history, nor the actual implications of working with a digital image.

¹³ Rodríguez-Ortega 2019, p. 11.

¹⁵ Kaplan 2015.

¹⁴ Panagiotidou et al. 2023, Daquino et al. 2022.

¹⁶ Boyd / Crawford 2012, p. 667.

Yet these three assumptions impact the way we shape research questions and engage with art historical material in the digital environment. There is an actual need to refine our general understanding of the field of art history and its concerns, as well as the digital material used, in order to better address them in computational perspectives.

First, as we have seen from many digital art history projects,¹⁷ research very often presupposes an artistic influence of one artist on another. They are looking for similar patterns within a dataset, trying to make visual connections as if the images were part of a giant network, without questioning the meaning and implications of such an influence. In art history, the notion raises various positions and arguments. The one of connoisseurship involves the possibility to recognize an artist behind an artwork through the reconstruction of potential external factors that could have led to its creation. These factors involve a practice in a workshop, a regional or time period production, a commission, or the requirements of the art market.¹⁸ Nevertheless, contemporary practice of art history tends to devalue connoisseurship¹⁹ as it involves formal and iconological approaches that do not align with the vision of a 'new art history.'²⁰ Assuming visual connections amongst artworks is assuming a world made of influences, which underrate the artistic intention.²¹ Hence, in order to align with actual art historical concerns, the question whether artists are actually influenced by other artists and their cultural context or if a work of art is the result of an artistic intention should be acknowledged, addressed, and possibly further investigated from a computational perspective.

Another question brought by the digital promise is the actual possibility of offering new developments and methods for the field of art history, knowing that most computational approaches imply formalism, an analysis based on the sole apprehension of the visual

aspect of the artworks. It seems that the field of digital art history is bound to perform some type of pure formal connoisseurship, bringing back to the table issues and concerns from the past century, when the field of art history attempted to move on from these methods. Yet, as questioned by Rodríguez-Ortega on the neo-formalist perspective and the scope of analysis afforded by computational methods, "is it possible to do something different?"²²

The ambition of a computational practice has, in fact, more to offer than a digital mimicking of traditional art historical practice at a large scale, and, in order to shape new narratives, there is a critical need for a better understanding of the computational environment. This awareness starts with what we are attempting to analyze, our object of study, that is, at a primary level, a digital image. As already pointed out a decade ago,²³ in most cases we are treating these images as works of art, when they are in fact digital copies, a non-tangible file made of binary encodings, preserved in the black box of a server. While working with a photographic collection, we actually find ourselves in the lineage of an unresolved problem of objectivity of photography, "the main tools of art historians,"²⁴ and a long tradition of image manipulation for rhetorical convenience. As summarized by Lum on the consideration of a photographic truth: "When an image is needed, it does not matter the means."²⁵

Nevertheless, the digital layer added to these photographs consists of a mathematical transformation, the 'mathesis' of Drucker,²⁶ from which a computational interpretation of the painted hand is performed and a further layer, made of keypoints to express the pose of the hand, is added. The novel definition of the image represents a turning point in the formulation of new methodological propositions in art history. Based on a computational vision of the hand and the creation of features for the comparison of poses, we define the framework

17 Seguin et al. 2016, Shen / Efros / Aubry 2019, Shen et al. 2022, Saleh et al. 2016, Champenois / Joyeux-Prunel 2023.

18 Elsig 2019, p. 12.

19 Elsig 2019, pp. 30–31.

20 Harris 2002, p. 7.

21 Summers 1986, p. 309.

22 Rodríguez-Ortega 2019, p. 10.

23 Drucker 2013, pp. 8–9.

24 Daquino 2019, p. 12.

25 Lum 2020.

26 Johanna Drucker describes *mathesis* as

"knowledge represented in mathematical form, with the assumption that it is an unambiguous representation of thought" (Drucker 2001, p. 141).

of a postural approach in opposition to the formal one. In this context, we shift from aesthetic values of the image to a synthetic and geometric vision of the pose of the hand, where we do not focus on the form anymore, neither essential nor inessential,²⁷ but rather on an interpretation of its sign.²⁸ Although we cannot escape comparative methods that have driven the field of art history since the arrival of photography,²⁹ and which are now an essential aspect of computational approaches,³⁰ it is based on such new conception of the *voir* in the creation of the *savoir*,³¹ along with a proper formulation of the research intention, that new methodologies and research questions can be established regardless of the uncertainties inherent to a digital art history.

27 Beyond the principles established by Heinrich Wölfflin in an attempt to shape a universal approach to art in his *Principles to Art History*, S. Pearl Brilmyer and Filippo Trentin distinguish two types of form as theorized by John Ruskin, the essential and inessential forms. Where the first would consist of “the basic outline or shape of a thing, figural qualities that allow an entity to be identified or recognized for what it is: a mountain, a dog, Jesus,” the inessential form is not fundamental to the recognition of the object but rather determining to the existence of the work of art (Brilmyer / Trentin 2019, p. 485).

28 Here we consider the images from

an interpretative layer, formulated for the purpose of the computational process. It implies a conception of the painting where the latter “ne nous donne pas à voir sa cause formelle, elle nous la donne à interpréter,” as phrased by Georges Didi-Huberman in *Devant l’image* (Didi-Huberman 1990, p. 279).

29 Caraffa 2011, p. 15.

30 Drucker 2012, p. 87.

31 In French, the notion of *voir*, seeing, and *savoir*, knowledge, have the same root and implies that *savoir*, knowledge, derives from *voir*, seeing. Furthermore, the difference between gazing and seeing, *le regard et voir*, in consideration of the *savoir* is, in

French, determinant in the distinction of looking at a painting and seeing a painting. The machine is trained to scan the image, looking at the pixel for a specific piece of information defined as relevant by the *savoir* of the programmer who knows what to see. As explained by Didi-Huberman, regarding the image, “La voir, et non la regarder: car voir sait mieux s’approcher, anticiper ou bien mimer l’acte, supposé souverain, du savoir” (Didi-Huberman 1990, p. 273).

References

- Barasch 1987**
Moshe Barasch, *Giotto and the Language of Gesture*, Cambridge Studies in the History of Art, Cambridge University Press, Cambridge, 1987.
- Bell / Schlecht / Ommer 2013**
Peter Bell, Joseph Schlecht and Björn Ommer, "Nonverbal Communication in Medieval Illustrations Revisited by Computer Vision and Art History," *Visual Resources. An International Journal on images and their uses*, vol. 29, 1–2, (2013), pp. 26–37. Doi: <https://doi.org/10.1080/01973762.2013.761111>.
- Bernasconi 2022**
Valentine Bernasconi, "GAB—Gestures for Artworks Browsing," in *Proceedings of the International Conference on Intelligent User Interfaces*, ACM Digital library 2022, pp. 50–53. Doi: <https://doi.org/10.1145/3490100.3516470>.
- Bernasconi 2024**
Valentine Bernasconi, "Automated Body Pose Estimation on Painting: A Crisis of Uncertainty," 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.0309>.
- Bernasconi / Cetinić / Impett 2023**
Valentine Bernasconi, Eva Cetinić and Leonardo Impett, "A Computational Approach to Hand Pose Recognition in Early Modern Paintings," *Journal of Imaging*, Multidisciplinary Digital Publishing Institute vol. 9, 6, (2023), p. 120. Doi: <https://doi.org/10.3390/jimaging9060120>.
- Boyd / Crawford 2012**
Danah Boyd and Kate Crawford, "Critical Questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon," *Information, Communication & Society*, Routledge vol. 15, 5, (2012), pp. 662–679. Doi: <https://doi.org/10.1080/1369118x.2012.678878>.
- Brandhorst / Posthumus 2017**
Hans Brandhorst and Etienne Posthumus, "Iconclass: A key to collaboration in the digital humanities," in *The Routledge Companion to Medieval Iconography*, ed. Colum Hourihane, Routledge, 2017, pp. 201–218.
- Brilmyer / Trentin 2019**
S.Pearl Brilmyer and Filippo Trentin, "Toward an Inessential Theory of Form: Ruskin, Warburg, Focillon," *Criticism*, vol. 61, 4, (2019), pp. 481–508. URL: <https://muse.jhu.edu/article/754433>. Doi: <https://doi.org/10.13110/criticism.61.4.0481>.
- Cao et al. 2021**
Zhe Cao et al., "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, 1, (2021), pp. 172–186. URL: <https://ieeexplore.ieee.org/document/8765346>. Doi: <https://doi.org/10.1109/tpami.2019.2929257>.
- Caraffa 2011**
Photo archives and the photographic memory of art history, ed. Costanza Caraffa, Italianische Forschungen des Kunsthistorischen Institutes in Florenz Max-Planck-Institutl Mandorli, Bd. 14, Deutscher Kunstverlag, Berlin, 2011. URL: <https://www.khi.fi.it/pdf/publikationen/i-mandorli/Photo-Archives-and-the-Photographic-Memory.pdf>.
- Champenois / Joyeux-Prunel 2023**
Robin Champenois and Béatrice Joyeux-Prunel, "Visual Contagions: extraire et tracer la circulation d'images dans des imprimés illustrés," in *Humanistica 2023*, Association francophone des humanités numériques 2023. URL: <https://hal.science/hal-04108205>.
- Daquino 2019**
Marilena Daquino, *Mining Authoritativeness in Art Historical Photo Archives. Semantic Web Applications for Connoisseurship* Ph.D Thesis, Università di Bologna. Dottorato di ricerca in Culture letterarie e filologiche, 31 Ciclo 2019. Doi: <https://doi.org/10.6092/unibo/amsdottorato/8968>.
- Daquino et al. 2022**
Marilena Daquino et al., "Expressing Without Asserting in the Arts," in *Proceedings of the 18th Italian Research Conference on Digital Libraries*, ed. Giorgio Maria Di Nunzio et al., CEUR 2022. URL: <https://ceur-ws.org/Vol-3160/#short14>.
- Didi-Huberman 1990**
Georges Didi-Huberman, *Devant l'image: Question posée aux fins d'une histoire de l'art*, Les Éditions de Minuit, Paris, 1990.
- Dimova 2020**
Temenuzhka Dimova, *Le langage des mains dans l'art Histoire, significations et usages des chirogrammes picturaux aux XVIIe et XVIIIe siècles*, Théorie de l'art (1400–1800), Brepols, Turnhout, vol. 9, 2020.
- Drucker 2001**
Johanna Drucker, "Digital Ontologies: The Ideality of Form in/and Code Storage—or—Can Graphesis Challenge Mathesis?," *Leonardo*, vol. 34, 2, (2001), pp. 141–145. Doi: <https://doi.org/10.1162/002409401750184708>.
- Drucker 2012**
Johanna Drucker, "Humanistic Theory and Digital Scholarship," in *Debates in the Digital Humanities*, ed. Matthew K. Gold, University of Minnesota Press, Minneapolis, 2012. URL: <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfd1e1/section/ob495250-97af-4046-91ff-98b6eagf83co#cho6>. Doi: <https://doi.org/9780816677948.003.0011>.
- Drucker 2013**
Johanna Drucker, "Is There a 'Digital' Art History?," *Visual Resources. An international journal on images and their uses*, vol. 29, 1–2, (2013), pp. 5–13. Doi: <https://doi.org/10.1080/01973762.2013.761106>.
- Elsig 2019**
Frédéric Elsig, *Connoisseurship et histoire de l'art. Considérations méthodologiques sur la peinture des XVe et XVIe siècles*, Droz, Genève, 2019.
- Gombrich / Huxley 1966**
Ernst Hans Gombrich and Julian Sorell Huxley, "Ritualized gesture and expression in art," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, Royal Society vol. 251, 772, (1966), pp. 393–401. Doi: <https://doi.org/10.1098/rstb.1966.0025>.
- Güler / Neverova / Kokkinos 2018**
Riza Alp Güler, Natalia Neverova and Iasonas Kokkinos, "DensePose: Dense Human Pose Estimation in the Wild," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE 2018, pp. 7297–7306. URL: <https://ieeexplore.ieee.org/document/8578860>. Doi: <https://doi.org/10.1109/cvpr.2018.00762>.

Harris 2002

Jonathan Harris, *The New Art History: A Critical Introduction*, Routledge, Taylor & Francis Group, London; New York, 2002. Doi: <https://doi.org/10.4324/9780203466780>.

Impett 2020

Leonardo Impett, "Analyzing Gesture in Digital Art History," in *The Routledge Companion to Digital Humanities and Art History*, ed. Kathryn Brown, Routledge, New York, 2020, pp. 386–408. Doi: <https://doi.org/10.4324/9780429505188-33>.

Impett / Süssstrunk 2016

Leonardo Impett and Sabine Süssstrunk, "Pose and Pathosformel in Aby Warburg's Bilderatlas," in *Computer Vision – ECCV 2016 Workshops Amsterdam, The Netherlands, October 8–10 and 15–16, 2016. Proceedings, Part I*, ed. Gang Hua and Hervé Jégou, Springer International Publishing 2016, pp. 888–902. Doi: https://doi.org/10.1007/978-3-319-46604-0_61.

Jin et al. 2020

Sheng Jin et al., "Whole-Body Human Pose Estimation in the Wild," in *Computer Vision – ECCV 2020*, ed. Andrea Vedaldi et al., Springer International Publishing 2020, pp. 196–214. Doi: https://doi.org/10.1007/978-3-030-58545-7_12.

Kaplan 2015

Frédéric Kaplan, "A Map for Big Data Research in Digital Humanities," *Frontiers in Digital Humanities*, vol. 2, (2015), Doi: <https://doi.org/10.3389/fdigh.2015.00001>.

Lum 2020

Ken Lum, "From Analog to Digital: A Consideration of Photographic Truth, 2012," in *Everything is Relevant*, Concordia University Press, 2020, pp. 213–223. Doi: <https://doi.org/10.2307/j.ctwvvr2jd.37>.

Madhu et al. 2022

Prathmesh Madhu et al., "Enhancing Human Pose Estimation in Ancient Vase Paintings via Perceptually-grounded Style Transfer Learning," *Journal on Computing and Cultural Heritage*, vol. 16, 1, (2022), pp. 16:1–16:17. Doi: <https://doi.org/10.1145/3569089>.

Panagiotidou et al. 2023

Georgia Panagiotidou et al., "Communicating Uncertainty in Digital Humanities Visualization Research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, 1, (2023), pp. 635–645.

URL: <https://ieeexplore.ieee.org/document/9904442>. Doi: <https://doi.org/10.1109/tvcg.2022.3209436>.

Rodríguez-Ortega 2019

Nuria Rodríguez-Ortega, "Digital Art History: The Questions that Need to Be Asked," *Visual Resources. An International Journal on images and their uses*, vol. 35, 1–2, (2019), pp. 6–20. Doi: <https://doi.org/10.1080/01973762.2019.1553832>.

Saleh et al. 2016

Babak Saleh et al., "Toward automated discovery of artistic influence," *Multimedia Tools and Applications*, vol. 75, 7, (2016), pp. 3565–3591. Doi: <https://doi.org/10.1007/s11042-014-2193-x>.

Seguin et al. 2016

Benoit Seguin et al., "Visual Link Retrieval in a Database of Paintings," in *Computer Vision – ECCV 2016 Workshops Amsterdam, The Netherlands, October 8–10 and 15–16, 2016 Proceedings Part I*, ed. Gang Hua and Hervé Jégou, Springer International Publishing 2016, pp. 753–767. Doi: https://doi.org/10.1007/978-3-319-46604-0_52.

Shen et al. 2022

Xi Shen et al., "Spatially-Consistent Feature Matching and Learning for Heritage Image Analysis," *International Journal of Computer Vision*, vol. 130, 5, (2022), pp. 1325–1339. Doi: <https://doi.org/10.1007/s11263-022-01576-x>.

Shen / Efros / Aubry 2019

Xi Shen, Alexei A. Efros and Mathieu Aubry, "Discovering Visual Patterns in Art Collections With Spatially-Consistent Feature Learning," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE 2019, pp. 9270–9279. URL: <https://ieeexplore.ieee.org/document/8954148>. Doi: <https://doi.org/10.1109/cvpr.2019.00950>.

Summers 1986

David Summers, "Intentions in the History of Art," *New Literary History*, Johns Hopkins University Press vol. 17, 2, (1986), pp. 305–321. Doi: <https://doi.org/10.2307/468898>.

Wittkower 1977

Rudolf Wittkower, "El Greco's Language of Gesture," in *Allegory and the Migration of Symbols The Collected Essays of Rudolf Wittkower*, Thames & Hudson Ltd, London, 1977, pp. 147–158.

Abstract

Piera Riccio & Nuria Oliver

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

Keywords | *Algorithmic content moderation, Social media, Artistic nudity*



This short article focuses on the growing challenge of algorithmic censorship of artistic content on online social platforms, particularly content that exhibits nudity. Contemporary artists heavily rely on social platforms to showcase and give visibility to their work, which is increasingly judged by machine learning algorithms that decide to remove or de-prioritize content if it is deemed inappropriate. While the challenge lies in distinguishing between artistic nudity and explicit material, we address this phenomenon from a techno-feminist perspective, considering different aspects of its cultural, governance, and technological dimensions.

Cite as: Piera Riccio and Nuria Oliver "A Techno-Feminist Perspective on the Algorithmic Censorship of Artistic Nudity" 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. 83–89. Doi: <https://doi.org/10.48431/hsah.0310>.

1. Introduction

Online social platforms have become a crucial tool for contemporary artists to share their art, reach a broader audience, and increase their visibility.¹ With billions of users worldwide, these online environments are subject to community rules including content moderation (deletion, hiding, or de-prioritization of content) to protect vulnerable users from harmful content. The immense volume of user-generated data uploaded to these platforms—in the order of millions of posts every minute—coupled with the challenging psychological working conditions of human content moderators and the impact of the COVID-19 pandemic, has led to the wide adoption of automated content moderation by means of machine learning algorithms,² giving rise to what is referred to as “algorithmic content moderation.” This process typically proceeds as follows:³ Machine learning algorithms are used to automatically detect content that violates the community guidelines to either remove or shadow-ban (de-prioritize) it. In cases where the decision is ambiguous, the algorithms flag the content to human content moderators.

Artistic expressions are not exempt from algorithmic content moderation, frequently leading to contentious case studies where artistic content is removed, banned, or de-prioritized due to e.g., political connotations, copyright violations, or the portrayal of nudity.⁴ Algorithmic censorship directly impacts the creative freedom of artists and leads to economic losses. Thus, the study of this phenomenon extends beyond an

understanding of our cultural ecosystem, as it also raises ethical questions about artistic freedom, the creators’ right to work, and their need for a secure and stable environment.

In this short article, we focus on the subject of algorithmic censorship as it applies to the artistic depiction of nudity. Nudity has been a fundamental subject of human artistic practices for thousands of years. The automatic detection of inappropriate nakedness poses significant challenges, as the distinction between artistic nudity and pornography is not always clear, depends on the context and the artist’s intention, and it may even be non-existent.⁵ Thus, tackling this ambiguity is one of the key challenges that need to be addressed in algorithmic content moderation. Moreover, depictions of nudity are historically and inherently related to notions of gender: Considering them is at the basis of our perspective.

2. Background

Algorithmic censorship of art is an under-studied topic of relevance for different research communities. Recent work has investigated this phenomenon from an artistic perspective, examining the employed creative techniques to bypass content moderation restrictions and their disproportionate impact on feminist artists.⁶ In response to the algorithmic censorship of art, activist movements—such as “Don’t Delete Art”⁷ and “Artists Against Censorship”⁸—have emerged to publicly condemn the matter. Other scholars have studied content moderation and its governance from a sociological and

1 Polaine 2005.

2 Chen 2021.

3 “How does Facebook use artificial intelligence to moderate content?”

URL: <https://www.facebook.com/help/1584908458516247>.

last accessed

10.11.23.

4 Gillespie 2020.

5 Vasilaki 2010, Patridge 2013, Eck 2001.

6 Faust 2017, Riccio et al. 2022.

7 Don’t delete Art. URL: <https://www.dontdelete.art/>.

(last accessed 10.11.23).

8 Artists Against Censorship. URL: <https://www.artistsagainstcensorship.com/> (last accessed 10.11.23).

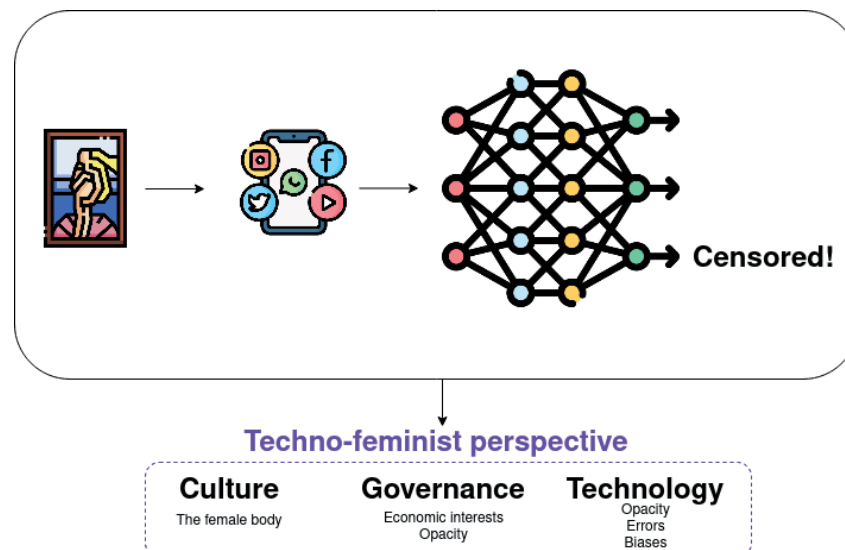


Figure 1. An illustration of the perspective analyzed by this article on the Algorithmic Censorship of Artistic Nudity. Our techno-feminist perspective entails different elements, divided in three dimensions: Culture, Governance, and Technology.

Copyright Nora Al-Badri. Source <https://www.nora-al-badri.de/works-index#neuronal-ancestral-sculptures-series>

a technical perspective,⁹ extending beyond the specific instance of artistic censorship.

the prevalence of algorithmic errors and biases from a technological perspective.

Research and activism are pivotal in enhancing public awareness and emphasizing the cultural, anthropological, and sociological consequences of artistic censorship on social media platforms. This article adopts a techno-feminist perspective, as proposed by Wajcman,¹⁰ framing our discussion of technological development as a socio-technical phenomenon that encapsulates patriarchal values and establishes gender power relationships.¹¹ Through this lens, we shed light on three key dimensions that are implicitly or explicitly associated with gendered concepts in the algorithmic censorship of online nudity. Figure 1 illustrates our framework.

Initially, we offer an insight into how algorithmic content moderation potentially influences the representation of the female body online and its interconnection with other historical periods, operating at the cultural level. Secondly, we examine the influence of private economic interests on online content moderation, operating at the governance level. Thirdly, we address the issue of transparency concerning content moderation practices and decisions, spanning both the governance and technological dimensions. Lastly, we explore

Historically, the concepts of censorship and feminism have been closely linked. Some scholars interpret the imposition of art censorship as a symbol of the war on women's freedom.¹² The 1980s, in particular, witnessed the emergence of various feminist movements actively opposing art censorship.¹³ In this short paper, rather than delving into the established connection between art censorship and feminism in a broad sense, we introduce a perspective that considers the pervasive integration of AI in artistic censorship as a further element to consider in the already intricate relationship between feminism and art censorship. Consequently, we advocate for the adoption of a techno-feminist framework.

3. Analysis

In this section, we delve into the three different elements that compose the framework for our analysis, depicted in Figure 1.

3.1. The Female Body

The way we relate to our bodies and their representation is heavily influenced by a diversity of elements, including the visual arts.¹⁴ While the female body has played a central role in art throughout history, its portrayal has largely

⁹ Gillespie 2020, Blunt et al. 2021, Elkin-Koren 2020, Myers West 2018.

¹⁰ Wajcman 2006.

¹¹ Noble / Tynes 2016.

¹² Jacobsen 1991.

¹³ Rodgerson / Semple 1990.

¹⁴ Gilman 1985, Arnheim 1954.

been shaped by male artists. Thus, women have often been reduced to symbols of male privilege and power rather than recognized as the creators of culture.¹⁵ Consequently, women have been objectified in creative expressions, leading to male-centered interpretations of the history of art. Female artists have been systematically denied proper recognition for their work—e.g., Marietta Robusti (1560–1590)—, de-valued and underestimated by their contemporaries—e.g., Judith Leyster (1609–1660)—, exoticized as exceptions—e.g., Edmonia Lewis (1844–1907)—, constantly forced into a “feminine” lens through their art—e.g., Camille Claudel (1864–1943)—or, conversely associated with masculine adjectives as a way to appreciate and justify their success—e.g., Suzanne Valadon (1865–1938) and Paula Modersohn-Becker (1876–1907). The representations and interpretations of art have always been strongly gendered, with the very ideal of the femininity of the Renaissance woman starting as early as Vasari’s work¹⁶ and continuing later in the history of visual arts.¹⁷

In 1975, Mulvey analyzed the concept of the “male gaze” in Hollywood movies,¹⁸ according to which women are depicted as objects of sexual desire from a heterosexual male perspective. This perspective has become even more pervasive with the rise of social media¹⁹ and the broad use of algorithms. In fact, scholars have argued that the amplification of the male gaze is a result of the gender inequality in the environments where the algorithms are conceived, developed, deployed, and legislated.²⁰ Such amplification is evident in the censorship of female artistic nudity, where female bodies are often perceived as solely objects of pleasure and pornographic content.²¹ Consequently, artists who work with female nudity to challenge the patriarchal sexualization of women face stronger restrictions on their freedom of expression, both because more people report them and because the algorithms are biased.

This systematic misinterpretation and silencing of de-sexualized representations of female bodies is an important phenomenon to reflect upon, study, and address when considering the impact of machine learning algorithms on our cultural and artistic ecosystems.

3.2. Economic Interests

Relying on algorithms for content moderation undoubtedly allows the platforms to obtain significant economic gains in the form of savings,²² as automation allows them to scale-up to levels that would not be possible if they only relied on human moderators.²³ In fact, with billions of users and millions of pieces of content uploaded every day, it is impossible for human moderators to perform this task alone.

Beyond this economic impact, we need to consider that social media and related online platforms are owned by private companies with a business model that largely depends on the monetization of the content that they host. Their ultimate goal is evidently to create an economic gain. Thus, it should not be surprising that social media companies keep whitelists of users for whom content moderation is not applied, or is applied differently.²⁴ Such users are typically high-value members of the network (e.g., influencers), with hundreds of thousands or millions of followers so that their uploaded content generates significant traffic and revenue for the platform owners. This differential treatment has a direct impact on the type of content that is likely to be uploaded as its visibility and popularity will depend on how well it adheres to the mainstream aesthetics accepted on the platform, to the detriment of minority voices and perspectives which are less popular and “economically interesting.”²⁵ Profit-driven content moderation decisions tend to prioritize content that aligns with dominant cultural norms, reinforcing existing power structures and contributing to perpetuating gender

¹⁵ Chadwick 1990.

¹⁶ Vasari 1550.

¹⁷ Duncan 2019.

¹⁸ Mulvey 1975.

¹⁹ Oliver 2017.

²⁰ Sezen 2020.

²¹ Ibrahim 2017, Volkers 2020, Are 2021.

²² “Why Tailored Content Moderation Artificial Intelligence Is Worth Investing In.” URL: <https://besedo.com/knowledge-hub/blog/the-price-of-ai/> (last accessed: 14.11.23).

²³ Gillespie 2020.

²⁴ Wikipedia 2021.

²⁵ Kang / Chen / Kang 2019, You et al. 2017, Sharlow 2015, James 2014.

stereotypes and biases. Such a reinforcement is detrimental to the techno-feminist goal of challenging and dismantling patriarchal structures and amplifying diverse perspectives.

3.3. Opacity

Furthermore, the lack of transparency regarding the use and inner workings of proprietary content moderation algorithms in social platforms hinders the users' understanding of the reasons behind the censorship they experience. Thus, engaging in a public discourse about the legitimacy of this censorship becomes challenging. According to Riccio et al. (2022), content moderation algorithms play the role of an invisible critique in today's online cultural ecosystem as they not only moderate content but also determine what is seen (and not seen), boosted, promoted, or de-prioritized. In essence, algorithms dictate what aesthetic qualities are deemed acceptable or unacceptable within our socio-cultural context by pervasively influencing all the elements of the creative chain,²⁶ with the market acting as the ultimate driving force in the decision-making process. Consequently, social media platforms establish distinct—albeit invisible—boundaries between what can be posted and what cannot. Within such an opaque framework, it is hard to understand what the rules of artistic expression are nowadays. This opacity is both intentional from the side of the social platforms and technical due to the inherent complexity of the deep neural networks that are at the core of content moderation algorithms. Furthermore, the lack of transparency leads to a lack of accountability as there are no clear mechanisms for challenging or appealing content moderation decisions.

3.4. Errors

Examining the consequences of algorithmic censorship demands acknowledging the weaknesses of these algorithms, which are

susceptible to errors and biases. Many of today's social media platforms do not explicitly forbid artistic nudity in their community guidelines.²⁷ However, the discrepancy between these guidelines and the actual experience of censorship by many artists²⁸ suggests that the algorithms might not be sophisticated enough to properly discern pornography from art,²⁹ failing to understand the non-explicit intent and context related to the uploaded images.³⁰ Contextualization is a prominent focus of research in the field of Computer Vision:³¹ Even basic tasks like object detection require explicit cues to contextualize the image.³² Humans, in contrast, rely on their past experiences and diverse memories when engaging with the real world.³³ The ability to contextualize a naked body plays an important role in distinguishing between pornographic and artistic images. As a consequence, detecting porn, nudity, or NSFW (not-safe-for-work) content still represents a technical challenge.³⁴ Existing research on algorithms developed for this purpose often overlooks the nuanced distinction between artistic nudes and pornography.³⁵ Although the Computer Vision literature has explored various techniques for context integration,³⁶ there is a lack of art-centric perspectives in the methods developed to date for automatic porn or adult content detection.

3.5. Biases

In addition, any decision-making process supported or performed by machine learning algorithms requires addressing potential algorithmic biases.³⁷ Algorithmic biases or discrimination refers to the systematic and unfair treatment (discrimination) that can occur when developing, training, or deploying machine learning-based algorithms.³⁸ These biases emerge at different points of the machine learning cycle, from the data used to train them to the design choices made during their development. As a consequence, the biases may be inadvertently embedded in the algorithms and

²⁶ Kulesz 2018.

²⁷ Shapiro 2022.

²⁸ Faust 2017.

²⁹ Riccio et al. 2022.

³⁰ "Does AI have a dirty mind, too?" URL: <https://medium.com/@marekkcichy/does-ai-have-a-dirty-mind-too-6948430e4b2b> (last

accessed 10.11.23).

³¹ Computer Vision is the field within Artificial Intelligence that focuses on enabling computers to interpret and understand visual information from the world.

³² Wang / Zhu 2023.

³³ Vo 2021.

³⁴ Dubettier et al. 2023.

³⁵ Chen 2021, Wang et al. 2018.

³⁶ Lim et al. 2021, Zhang / Tseng / Kreiman 2020, Leng et al. 2021.

³⁷ Buolamwini / Gebru 2018.

³⁸ Kleinberg et al. 2018.

influence their decision-making processes, leading to unequal or unfair treatment of certain individuals, groups or content (e.g., women or portrayal of female bodies). The machine learning algorithms that permeate today's social media platforms thus automate culture through technology, replicating and possibly amplifying human biases.³⁹ In the context of algorithmic censorship, these biases could imply that certain representations of the body are more likely to be accepted than others (exacerbating the economic incentives previously discussed), or that artists representing minorities get censored more often, as Nayyar (2023), Shapiro (2022), Faust (2017) claim.

4. Conclusion

While social media offers a platform for artistic expression to a broad audience, the existing algorithmic content moderation practices create disparities in access for artists. Examining algorithmic censorship and its impact on the artistic depictions of nudity uncovers an intricate interplay of factors.

This phenomenon can be seen as a result of the economic motivations of social media companies, the lack of transparency in algorithmic content moderation, and the inherent limitations of the algorithms, as previously outlined. Furthermore, adopting a techno-feminist lens allows us to delve deeper into the gendered dimensions of algorithmic censorship, revealing how these practices contribute to reinforcing existing power dynamics and societal norms related, but not necessarily limited to, the female body.

This short paper prompts critical reflections on artistic freedom, the ethical considerations surrounding content moderation, and the need to maintain a secure and inclusive digital space for artists. In light of these interconnected issues, we hope to pave the way for innovative solutions that align with the principles of techno-feminism and contribute to a more equitable and just artistic landscape online.

³⁹ Schroeder 2021, Sezen 2020.

References

Are 2021

Carolina Are, "The Shadowban Cycle: An Autoethnography of Pole Dancing, Nudity and Censorship on Instagram," *Feminist Media Studies*, Informa UK Limited vol. 22, 8, (2021), pp. 1–18. Doi: <https://doi.org/10.1080/14680777.2021.1928259>.

Arnheim 1954

Rudolf Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*, University of California Press, 1954.

Blunt et al. 2021

Danielle Blunt et al., "Deplatforming Sex: A Roundtable Conversation," *Porn Studies*, Taylor & Francis vol. 8, 4, (2021), pp. 420–438. Doi: <https://doi.org/10.1080/23268743.2021.2005907>.

Buolamwini / Gebru 2018

Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," in *Proceedings of Machine Learning Research 81:1–15, 2018 Conference on Fairness, Accountability, and Transparency*, ed. Sorelle A. Friedler and Christo Wilson, PMLR 2018, pp. 77–91. URL: <https://proceedings.mlr.press/v81/buolamwini18a.html>.

Chadwick 1990

Whitney Chadwick *Women, Art, and Society*, Thames and Hudson, London, 1990.

Chen 2021

Thomas M. Chen, "Automated Content Classification in Social Media Platforms," in *Securing social networks in cyberspace*, ed. Al-Sakib Khan Pathan, CRC Press, Boca Raton, 2021, pp. 53–71.

Dubettier et al. 2023

Adrien Dubettier et al., "A Comparative Study of Tools for Explicit Content Detection in Images," in *2023 International Conference on Cyberworlds (CW 2023)*, IEEE 2023, pp. 464–471. URL: <https://ieeexplore.ieee.org/document/10337594>. Doi: <https://doi.org/10.1109/cw58918.2023.00077>.

Duncan 2019

Carol Duncan, "Virility and Domination in Early Twentieth-Century Vanguard Painting," in *Feminism and Art History. Questioning the Litany*, ed. Norma Broude and Mary D. Garrard,

Routledge, New York, 2019, pp. 292–313.

Eck 2001

Beth A. Eck, "Nudity and Framing: Classifying Art, Pornography, Information, and Ambiguity," *Sociological Forum*, Springer vol. 16, 4, (2001), pp. 603–632. Doi: <https://doi.org/10.1023/a:1012862311849>.

Elkin-Koren 2020

Niva Elkin-Koren, "Contesting Algorithms: Restoring the Public Interest in Content Filtering by Artificial Intelligence," *Big Data /& Society*, SAGE Publications Sage UK: London, England vol. 7, 2, (2020), Doi: <https://doi.org/10.1177/2053951720932296>.

Faust 2017

Gretchen Faust, "Hair, Blood and the Nipple," in *Digital environments: Ethnographic Perspectives across Global Online and Offline Spaces*, ed. Urte Undine Frömming et al., Transcript Verlag, Bielefeld, 2017, pp. 159–170. URL: <https://doi.org/10.14361/9783839434970-012>.

Gillespie 2020

Tarleton Gillespie, "Content Moderation, AI, and the Question of Scale," *Big Data /& Society*, SAGE Publications Sage UK: London, England vol. 7, 2, (2020), p. 2053951720943234. Doi: <https://doi.org/10.1177/2053951720943234>.

Gilman 1985

Sander L. Gilman, "Black Bodies, White Bodies: Toward an Iconography of Female Sexuality in Late Nineteenth-Century Art, Medicine, and Literature," *Critical inquiry*, University of Chicago Press vol. 12, 1, (1985), pp. 204–242. URL: <https://www.jstor.org/stable/1343468>. Doi: <https://doi.org/10.1086/448327>.

Ibrahim 2017

Yasmin Ibrahim, "Facebook and the Napalm Girl: Reframing the Iconic as Pornographic," *Social Media + Society*, SAGE Publications vol. 3, 4, (2017), p. 205630511774314. Doi: <https://doi.org/10.1177/2056305117743140>.

Jacobsen 1991

Carol Jacobsen, "Redefining Censorship: A Feminist View," *Art Journal*, vol. 50, 4, (1991), pp. 42–55. URL: <https://www.jstor.org/stable/777322?origin=JSTOR-pdf>. Doi: <https://doi.org/10.2307/777322>.

James 2014

Phil James, "8 Reasons Why Social Media Is Decimating Art and

Literature," 2014. URL: <https://qwiklit.com/2014/03/08/8-reasons-why-social-media-is-decimating-art-and-literature/>.

Kang / Chen / Kang 2019

Xin Kang, Wenxin Chen and Jian Kang, "Art in the Age of Social Media: Interaction Behavior Analysis of Instagram Art Accounts," *Informatics*, MDPI AG vol. 6, 4, (2019), p. 52. Doi: <https://doi.org/10.3390/informatics6040052>.

Kleinberg et al. 2018

Jon Kleinberg et al., "Algorithmic Fairness," in *Aea papers and proceedings*, American Economic Association 2018, pp. 22–27. Doi: <https://doi.org/10.1257/pandp.20181018>.

Kulesz 2018

Octavio Kulesz, "Culture, Platforms and Machines: The Impact of Artificial Intelligence on the Diversity of Cultural Expressions," in *Intergovernmental committee for the protection and promotion of the diversity of cultural expressions*, UNESCO 2018. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000380584>.

Leng et al. 2021

Jiaxu Leng et al., "Realize Your Surroundings: Exploiting Context Information for Small Object Detection," *Neurocomputing*, Elsevier vol. 433, (2021), pp. 287–299. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0925231220320051?via%3Di-hub>. Doi: <https://doi.org/10.1016/j.neucom.2020.12.093>.

Lim et al. 2021

Jeong-Seon Lim et al., "Small Object Detection Using Context and Attention," in *2021 international conference on artificial intelligence in information and communication (ICAIIIC)*, IEEE 2021, pp. 181–186. Doi: <https://doi.org/10.1109/icaaic51459.2021.9415217>.

Mulvey 1975

L. Mulvey, "Visual Pleasure and Narrative Cinema," *Screen*, Oxford University Press vol. 16, 3, (1975), pp. 6–18. URL: <https://academic.oup.com/screen/article-abstract/16/3/6/1603296>. Doi: <https://doi.org/10.1093/screen/16.3.6>.

Myers West 2018

Sarah Myers West, "Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms," *New*

Media & Society, SAGE Publications Sage UK: London, England vol. 20, 11, (2018), pp. 4366–4383. Doi: <https://doi.org/10.1177/1461444818773059>

Nayyar 2023

Rhea Nayyar, "On Social Media, Art Censorship Is Alive and Well," 2023. URL: <http://hyperallergic.com/817737/on-social-media-art-censorship-is-alive-and-well/>.

Noble / Tynes 2016

The intersectional Internet: Race, sex, class and culture online, ed. Safiya Umoja Noble and Brendesha M. Tynes, Digital formations, Peter Lang Publishing, Inc., New York, 2016.

Oliver 2017

Kelly Oliver, "The male gaze is more relevant, and more dangerous, than ever," *New Review of Film and Television Studies*, Taylor & Francis vol. 15, 4, (2017), pp. 451–455. Doi: <https://doi.org/10.1080/17400309.2017.1377937>.

Patridge 2013

Stephanie Patridge, "Exclusivism and Evaluation: Art, Erotica and Pornography," in *Pornographic Art and the Aesthetics of Pornography*, ed. Hans Maes, Palgrave Macmillan UK, London, 2013, pp. 43–57. Doi: https://doi.org/10.1057/9781137367938_3.

Polaine 2005

Andrew Polaine, "Lowbrow, High Art: Why Big Fine Art Doesn't Understand Interactivity," in *Refresh! First International Conference on the Histories of Media Art, Science and Technology*, Sept. 28- Oct. 1st 2005, 2005. URL: https://www.polaine.com/images/nodate/apolaine_whybigfineart_refresh_paper.pdf.

Riccio et al. 2022

Piera Riccio et al., "Algorithmic Censorship of Art: A Proposed Research Agenda," in *Proceedings of the 13th International Conference on Computational Creativity*, ed. Maria M. Hedblom et al., ACC 2022, pp. 359–363. URL: <https://computationalcreativity.net/iccc22/accepted-papers/>.

Riccio / Oliver 2024

Piera Riccio and Nuria Oliver, "A Techno-Feminist Perspective on the Algorithmic Censorship of Artistic Nudity," 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.0310>.

Rodgers / Semple 1990

Gillian Rodgers and Linda Semple, "Who Watches the Watchwomen? Feminists Against Censorship," *Feminist Review*, ed. The Feminist Review Collective, 36, (1990), pp. 19–24. Doi: <https://doi.org/10.2307/1395106>

Schroeder 2021

J.E. Schroeder, "Reinscribing Gender: Social Media, Algorithms, Bias," *Journal of Marketing Management*, vol. 37, 3–4, (2021), pp. 376–378. Doi: <https://doi.org/10.1080/0267257x.2020.1832378>

Sezen 2020

Diğdem Sezen, "Machine Gaze on Women: How Everyday Machine-Vision-Technologies See Women in Films," in *Female Agencies and Subjectivities in Film and Television*, ed. Diğdem Sezen et al., Springer International Publishing, Cham, 2020, pp. 271–293. Doi: https://doi.org/10.1007/978-3-030-56100-0_15.

Shapiro 2022

Emma Shapiro, "Censorship on social media not only limits artists' online reach—it can prevent future opportunities, too," 2022. URL: <https://www.theartnewspaper.com/2022/04/18/censorship-on-social-media-not-only-limits-artists-online-reach-it-can-prevent-future-opportunities-too>.

Sharlow 2015

S. Sharlow, "Death of an Artist: How Social Media Is Ruining Creativity," 2015. URL: <https://www.elitedaily.com/life/culture/death-artist-social-media-ruining-creativity/907113>.

Vasari 1550

Giorgio Vasari, *Le Vite de più eccellenti architetti, pittori, scultori italiani, da Cimabue insino a' tempi nostri*, Torrentino, Florence, 1550. URL: https://archive.org/details/gri_vitedepivecco1vasa/page/8/mode/2up.

Vasilaki 2010

Mimi Vasilaki, "Why Some Pornography May Be Art," *Philosophy and Literature*, vol. 34, 1, (2010), pp. 228–233. Doi: <https://doi.org/10.1353/phl.0.0066>.

Vo 2021

Melissa Le-Hoa Vo, "The Meaning and Structure of Scenes," *Vision Research*, vol. 181, (2021), pp. 10–20. URL: <https://www.sciencedirect.com/science/article/pii/S0042698920301796>. Doi: <https://doi.org/10.1016/j.visres.2020.11.003>

Volkers 2020

R. Volkers, "Perverse Media: How Instagram Limits the Potential of Feminist Art," *EUC Student Academic Journal*, 2, (2020). URL: <https://esaj.nl/Roos-Volkers-2>.

Wajcman 2006

Judy Wajcman, "Technocapitalism meets technofeminism: Women and technology in a wireless world," *Labour and Industry: a Journal of the Social and Economic Relations of Work*, Monash University vol. 16, 3, (2006), pp. 7–20. URL: <http://www.buseco.monash.edu.au/mgt/research/acrew/labour-and-industry/index.html>. Doi: <https://doi.org/10.1080/10301763.2006.10669327>.

Wang et al. 2018

Xizi Wang et al., "Adult Image Classification by a Local-Context Aware Network," in *2018 25th IEEE International Conference on Image Processing (ICIP) Proceedings*, IEEE 2018, pp. 2989–2993. Doi: <https://doi.org/10.1109/icip.2018.8451366>.

Wang / Zhu 2023

Xuan Wang and Zhigang Zhu, "Context Understanding in Computer Vision: A Survey," *Computer Vision and Image Understanding*, Elsevier vol. 229, (2023), p. 103646. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1077314223000267>. Doi: <https://doi.org/10.1016/j.cviu.2023.103646>.

Wikipedia 2021

Wikipedia, "2021 Facebook Leak," Wikipedia 2021. URL: https://en.wikipedia.org/wiki/2021_Facebook_leak.

You et al. 2017

Quenzeng You et al., "Cultural Diffusion and Trends in Facebook Photographs," in *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)* 347, ICWSM 2017, pp. 347–356. URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/14902>. Doi: <https://doi.org/10.1609/icwsml.v11i1.14902>.

Zhang / Tseng / Kreiman 2020

Mengmi Zhang, Claire Tseng and Gabriel Kreiman, "Putting Visual Object Recognition in Context," in *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE 2020, pp. 12982–12991. URL: <https://ieeexplore.ieee.org/document/9156547>. Doi: <https://doi.org/10.1109/cvpr42600.2020.01300>

Abstract

Anne-Laure Oberson

Making Sense of AI Generated Images in the
Process of Imagination

Keywords | Knowledge, Imagination, Abstraction,
Concretion, AI generated images, AI



Digital technologies affect our imagination, motivation to discover, and knowledge creation. AI-generated images highlight the need for deeper understanding at the intersection of visualization, information, and aesthetics. This essay examines if these images foster new imagination, referencing Vilém Flusser's analysis. It stresses bridging knowledge gaps in art history, engineering, philosophy, and science, critiquing images with the latest technological insights. It questions traditional representation and ontology, exploring images' evolution from traditional to synthetic, emphasizing their transformative nature.

The text discusses imagination's role in knowledge creation and technical images' impact, challenging traditional logic dichotomies with calculus logic. It scrutinizes AI images' reliance on statistical logic versus imagination in meaning generation, introducing reification to describe AI images as statistical abstractions. This draws parallels to human capacities, cautioning against equating AI with dynamic, living processes.

Cite as: Anne-Laure Oberson "Making Sense of AI Generated Images in the Process of Imagination" 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. 91–97. Doi: <https://doi.org/10.48431/hsah.0311>.

Introduction

In the fields of the humanities, AI has deployed its powerful and radical efficiency in revolutionizing older practices to the extent that new departments have had to be created. Indeed, for any task on which a principle of the grid¹ can be uniformly applied, its tremendous computational power has so attractive as to become indispensable. Yet, at the intersection of specific fields concerned with image analysis, production, and studies, when turning images into concepts, those concepts into discrete data, and this data back into images of another essence, the process must raise some questions. To choose only one—“How do these images make any sense?”—seems to be an important question because images are central to our cognitive processes. And given the greater permeability of various fields of studies and the cross-over from humanities to computer science, new digital departments require both broader and finer knowledge of key elements, amongst which is a critique of image that is up-to-date with the advancements in technology.

In the process of general digitization, the heterogeneity of images has been levelled to the extent that what is under study has become mostly, if not exclusively, dematerialized information (digital data sets). Indeed, Gottfried Boehm had promised to consider the image in its “immense variety” and “diffuse omnipresence” and to forgo the attachment to objects, works of art, and their reproductions, which too often interfered with getting to the root of the ontological question of the image.² Almost thirty years later, this consideration is more often overlooked than taken for granted.

Hence, we have the current necessity to look at images when they do not show or cannot be seen, such as non-predicative, non-discursive images that before showing anything make us see: Transformative images, operational images, generated images. Images that stand beyond the visual as other images have before them: Mental images, archetypal images, neuronal images. Before we exchange, borrow, or adopt terms from different fields, we should ensure to address cross-domain knowledge gaps so that art historians, engineers, philosophers, and scientists have a sufficiently well-informed understanding of their meaning and complexity, share the knowledge of what they are speaking of and dealing with, avoid naïve realism, and stay updated on critical insights “to stretch the notion of image towards some non-image.”³

Indeed, we all call it by the same name—image—yet, we have a very different conception and understanding of what a multifarious thing images truly are.

What Images are

The different meanings we confer to the term image in different situations don’t exclude each other; they add layers of complexity and richness to vital interconnections. According to the media theoretician Vilém Flusser, an image is a message, i.e., it carries meaning (it works), but this meaning has changed with the course of technical and technological evolutions.⁴ It went from signifying a revelation in a pre-historic sacralized space (cavern paintings) to signifying the passage from private to public in the making of history (museum tableaux) to programming behaviors in post-historical,

¹ Whether it is used to achieve proportional accuracy, define modernity in art, control energy consumption, or align pixels in a matrix, the

grid is the ordering structure par excellence; an organizing principle that has become a consistent pattern in digital image analysis.

² Boehm 1995 (Author’s translation).

³ Parikka 2023, p. 87.

⁴ Flusser (1989) 2002.

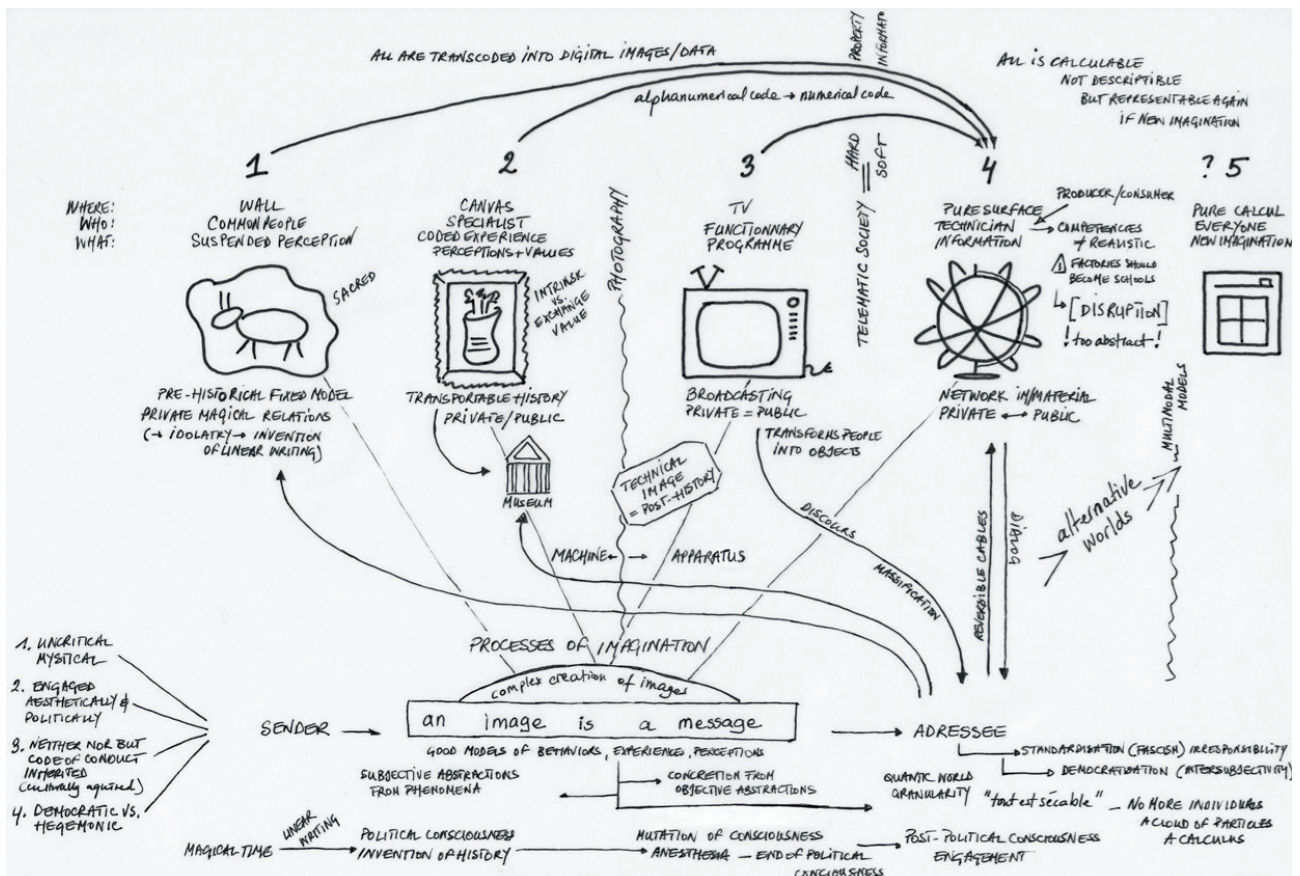


Figure 1. Diagram of Vilém Flusser's four types of images, 2018.

The information comes from his text *Images in the New Media* (1989) 2002, as well as from the French translation by Claude Maillart published in *La Civilisation des médias* 2006. A fifth type of image has been introduced on the right-hand side to start mapping the differences and similarities between AI generated images and original synthetic images, the computer renderings of the 1990s.

post-industrial societies (bedroom televisions), reaching the promise of a fourth message carried by a disembodied surface onto which “many different participants would cooperate to project different meanings on its surface.”⁵ If Flusser had the utopian vision that this new media could “transform people into designers of meaning in a participatory process” and “make it possible to take the political, economic, and social ‘powers’ out of commission,” we must unfortunately acknowledge that our current situation has been taken hostage by all-powerful platforms that have somewhat stolen this dream and enforced an altogether different model. This is demonstrated by Eryk Salvaggio’s proposal of a “Many-to-one to one-to-one” model of transmission, even though he did not acknowledge Flusser’s early fundamental contribution to these mechanisms.⁶

Additionally, besides this shift from traditional to technical to synthetic images, there are some fundamental characteristics of the image—lack of its own essence, its intermediary reality, its quantum-like ubiquity—that must be discussed before moving along.⁷ I will attempt to quickly present these key elements to further understand the role of AI generated images in the process of imagination. Indeed, the exponential increase of algorithmic processes in visual studies, both in terms of their extraordinary capacities to analyze and understand images and in terms of generating these images, requires us to question the ontological nature of these images.

To start, there has been a tremendous amount of critical work done towards getting the image beyond its status as representation and as having an ontology that has been accomplished

5 For a detailed analysis, the reader can refer to my contribution “The fourth meaning of the image: Flusser’s invitation to envision the

world” published in *Flusser Studies*, issue 29, 2020 (Oberson 2020).

6 Salvaggio 2023.

7 I have dedicated a large part of my doctoral thesis *Image as articulation of thought* to explain these characteristics in detail.

by great thinkers such as Georges Didi-Huberman, Vilém Flusser, Emanuele Coccia, Siegfried Zielinski, and Serge Tisseron, among others. For conciseness, I will only quote here from Didi-Huberman:

The great flaw of contemporary image criticism consists in wanting to emit its opinion on an ontology of the image. I would dare to say that, by definition—because the image is first and foremost the image of something else—there can be no ontology of the image. We cannot say: The image, it is this or it is that. We can only say: This image works like this or like that, transforms this or that, transforms itself like this or that.⁸

Whether traditional pictures signify ‘what is’ or ‘what must be,’ being either representations or models, their ontology was always dubious, as “technical images program us as models, and they can do it precisely by what they imagine.”⁹ Non-visual data-based images (also called operational or transactional images) are even further from an ontological essence, as Jussi Parikka suggests when he poetically formulates the repetitive question “What is an image if it is not just an image but also...” in a chapter appropriately titled “What Is Not an Image.”¹⁰ This alternative, ambivalent quality of the image is central. The image has no essence in its own; and even if/when it stands for something else, it does not inherit its ontology as the image is always of something else other “than what it shows.”¹¹ Furthermore, synthetic images (also referred to as computer images) stand beyond simply ‘being,’ as they are “computation of clear and distinct concepts becoming surface” or a bi-dimensional materialization of an abstraction. Hence, as we will see, AI generated images are on a further level of abstraction: They are abstractions of concretized abstraction. If unimaginable thoughts are to be imaginable, a traditional ontology is surely at a loss.¹²

Then, the image always resides in-between. Emanuele Coccia offers one of the simplest and most effective definitions of image: “The image (the sensitive) is nothing but the existence of a thing outside of its proper place. All form and all things capable of existing outside of its proper place become image.”¹³ This characteristic that Coccia traces back to the Middle Ages and names the *metaxu* (a space in the middle) has been identified by many other scholars in various fields of study. The neurologist Antonio Damasio reintroduced the importance of the ‘internal milieu’¹⁴ alongside the visceral and musculoskeletal frames in ensuring the stability that is vital to the proto-self where neuronal images are formed;¹⁵ Gilbert Simondon presented the image as *réalité intermédiaire* (an intermediary reality) to describe images that could be both concrete and abstract, objective and subjective, purely mental or object-like;¹⁶ Carl Gustav Jung and Wolfgang Pauli agreed on the existence of a ‘collective unconscious’ belonging to the *unus mundus* underlying both mind and body, not unrelated to Henry Corbin’s concept of *mundus imaginalis* that includes dimensions going beyond the unconscious into the superconscious.¹⁷

Last but not least, these characteristics resonate precisely with a quantum reality brought forth in Vilém Flusser’s seminal book *Towards a philosophy of photography*, in which he demonstrates the quantum quality of the first technical image.¹⁸ Photography

is neither a fixed medium, nor a medium that fixes anything; rather, it is transient, fluid, undetermined, floating. It does not respond to the Barthian logic of ‘ça a été’ (it has been) but to a constantly renewed ‘c’est’ (it is), which is always in becoming.¹⁹

The quantic quality of a particle is not to be nowhere but in two places simultaneously or

8 Georges Didi-Huberman has explored the resistance of images to an ontology in many extensive publications, from which the complexity and depth of his analysis cannot be summarized. His conversation edited for a series by the French National Audiovisual Institute provides his own synthetic statement: Didi-Huberman 2011 (Author’s translation).

9 Flusser 1984.

10 Parikka 2023, pp. 57–94.

11 Flusser 1984, Tisseron 2010.

12 Salvaggio’s effort to come up with an ontology of AI generated image provides the start of a reflection that should be pursued.

13 Coccia 2010, p. 101 (Author’s translation).

14 A term coined by the French physiologist Claude Bernard in 1865.

15 Damasio 1999.

16 Simondon 2022, p. 13.

17 Corbin (1969) 2016.

18 Flusser (1983) 2000.

19 Oberson 2013.

anywhere in between, just like Coccia's image. Just like archetypes, Carl Jung's "instinctual images" form the collective unconscious and their persistence consists in an alternance of appearing and disappearing. This is interesting because it is also where we find the *a priori* forms of intuition: Forms without content representing the possibilities of a type of perception and action, a *facultas praeformandi*—the capacity to give form.²⁰ And to which echoes Walter Benjamin's dialectical image with its flickering capacity, the very characteristic that allows it to persist, disappearing where challenged and reappearing again somewhere else but not going away.

Technical, synthetic, and generated images have these characteristics: They are transformative, intermittent, and ubiquitous. This clears away any misleading assumptions about the image so that we can look at the role of visual logic in imagination and how it relates to the question of making sense.

How images impact the process of imagination

Already in the late 1980s, Flusser²¹ identified that the linear rules of writing could no longer apply to image criticism when the numeric code had freed itself from the alphabetical and switched to the digital. Counting on fingers at the speed of supercomputers requires a new form of imagination that has been rendered possible by the emergence of synthetic images. His thinking in terms of a logic of calculus²² will help us grasp the possibilities offered by AI generated images.

The long-standing split between reasoning and perceiving in processes of knowledge has been artificially maintained for the necessity of power, when in fact the capacity to imagine via

intuition—*Einbildungskraft*²³—has always relied on a combination of both, from antiquity, to the 18th century and the beginning of the 20th. In an exemplary and seminal analysis, Arthur I. Miller demonstrated the efficacy of visual thinking "as a criterion for selection between alternatives that resists reduction to logic" by exposing and discussing the debates that went on in physics between 1913 and 1927, and drawing the conclusion that

the importance for creative thinking of the domain where art and science merge has been emphasized by great philosopher-scientists of the twentieth century—Bohr, Einstein, Heisenberg and Poincaré [...] for in their research the boundaries between disciplines are often dissolved and they proceed neither deductively through logic nor inductively through the exclusive use of empirical data, but by visual thinking and aesthetics.²⁴

Hence, the old antagonism between the circular logic of images and the linearity of texts is no longer valid and must be set aside once and for all as we are called to endorse a more encompassing view. Accordingly, there is no longer a dialectic of what is inside or outside, of what is or is not, but instead a logic of one extended space that belongs and contributes to our collective memory and impacts how we perceive we think. As Flusser puts it, with the advent of the technical image (camera) and its subsequent synthetic image (computer), we now have an image that is beyond linear text, fully coded, and closer to our neurological mental processes than any other image before. At this point, we have entered the logic of calculus. With the advent of the AI generated image (algorithm) we must consider whether this same logic applies and how this new type of image impacts our thinking processes. If another logic must reside between linearity and circularity, many believe that the statistical probabilistic logic of

²⁰ Jung (1960) 2023, p. 3587.

²¹ Flusser's seminal texts on the topic—"Images in the New Media" (1989) 2002 and "A New Imagination" (1990) 2002—were written in German in 1989 and 1990 respectively and translated in English in the volume *Writings* 2002. I have accessed several unpublished variants from the Flusser

Archiv, Berlin, in order to create the diagrams.

²² Where calculus is the zero-dimensional point or the elementary unit of computerization allowing the logic "whereby humans transform themselves into players who calculate and compute concepts." See Flusser 2015, pp. 95–96. Available online at [https://wiki.flusser.club/doku.](https://wiki.flusser.club/doku.php?id=calculus)

[https://wiki.flusser.club/doku.](https://wiki.flusser.club/doku.php?id=calculus)

²³ Here is a permanent link to a repository page where an ongoing mapping of *Einbildungskraft* is taking place: <https://www.thinkingseeing.com/imagination>.

²⁴ Miller 1978, pp. 77–102.

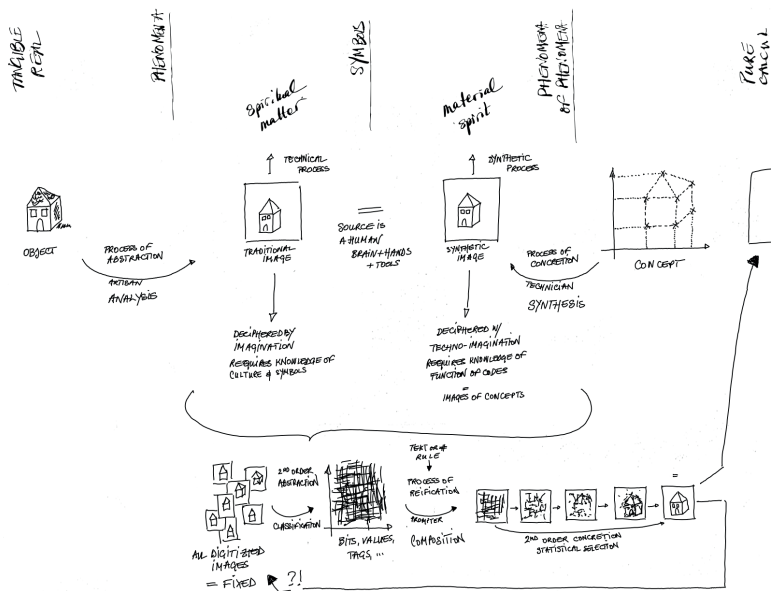


Figure 2. Diagram of Vilém Flusser's process of imagination, 2023.

Vilém Flusser wrote several texts in German and English around the question of imagination, new and second-order imagination. One of them was translated by Erik Eisel as "A New Imagination" in Flusser (1990) 2002, while most remain unpublished and understudied as they are only available in his archive in Berlin. Because his analysis, though specifically addressing earlier synthetic images (Flusser died in 1991), provides an innate and anticipatory contribution to the epistemological questions we are facing today, their study, translation, and comparison will contribute to establishing a finer understanding of AI generated images and their role in imagination processes.

the algorithm may prove to be satisfactory. But is it sufficiently reliant on the role of imagination, and is it contributing to the emergence of a new imagination?

For the mathematician Giuseppe Longo, "Lambda Calculus [...] is a means of great mathematical expressivity, at the heart of Mathematical Logic and of many of the programming styles that created modern computer science." Nevertheless, the madness of the "everything is computational" has created the myth "of the Universe as a Turing Machine," which is being applied "to every possible phenomenon, flattening it onto a universe that is without meaning and made only of formal calculations." He establishes the extremely important distinction between the precision of discrete databases and the great challenge of physical measurement that should not be forgotten.²⁵

Despite the many definitions and applications of the term imagination, let us clearly affirm here that imagination is more than the faculty to conjure up what is absent, even regardless of its veracity. It is a positive faculty, a creative power in the sense given by Castoriadis²⁶ and upon which Vilém Flusser further distinguished between first-order imagination—"the faculty to make and decipher images," the ability to step out of the objective world into our

subjectivity by a symbolic gesture of abstraction—and second-order imagination—"the ability to step backward from imagination into the unsurpassable abstraction," the ability to project our models onto the world by a gesture of concretion, of gathering together points-like dots, rendered possible with the switch from numeric to digital code.

This new imagination, enabled by new technologies and their devices, was supposed to allow us to experiment with images in new ways and deepen our understanding of how they work, independently of any textual context. By extension, it should also allow us to understand the functioning of the mind and its thought processes. From the perspective of the 1990s, Flusser believed that "only when one produces images of calculations instead of facts (it does not matter how 'abstract' the facts) can 'pure aesthetics' (the joy of playing with 'pure forms') find their true expression; only then can *Homo ludens* replace *Homo faber*."²⁷ With hindsight and paraphrasing Flusser, algorithmic processes then are suitably instructed to compute point-like concepts into image-like surfaces, "sufficiently intelligent to compute these concepts, and sufficiently stupid not to try and conceive what they are doing."²⁸

Thus, if traditional images abstract objects and synthetic images concretize concepts, then

²⁵ Longo 2019, pp. 73–94.

²⁶ Castoriadis 1987.

²⁷ Flusser (1990) 2002.

²⁸ Flusser n.d.

AI generated images are both further analyzing traditional and synthetic images one level up into discrete information and re-synthesizing digital data or recombining the values into a surface-like abstract-concretion. In this course, the ‘imaginative’ meanings that were initially given to concepts concretized as images have now been fully emptied. These potentially ‘third-order’ imaginations have become totally meaningless in the process of pure calculation. Indeed, Flusser was envisioning synthetic images that had their source in human conceptual thinking, but the source of AI generated images is statistical and mechanical. They are digital values that are triggered by a text prompt acting merely as a more-or-less sophisticated tool. I would like to propose to describe this process of AI image generation as one of reification, when an abstraction is treated as something concrete and real.

Conclusion

According to Simondon, in the process of invention, intuition, perception, and memory—corresponding to an *a priori*, *a praesenti*, *a posteriori* image—engage reciprocally in a cycle that works because of the spontaneous endogenous movement of anticipation and the heterogeneous plurality of the perceived. These three phases—anticipation, experience, and systematization—constitute a fourth one—invention—in a single process of genesis. AI generated images stand at the threshold of visualization and information, aesthetic and mathematic, and from there they demand that we look at the deeper issues, beyond the purely visible, at how or if they could partake in a process of imagination even though they are not implicated either in endogenous or heterogeneous experiences.

Therefore, I do not believe that AI could work as an exosomatic²⁹ psychic system to draw a parallel with many interpretations of artificial

intelligence as an external extension of human capacities. But could large AI models be considered as an encapsulation of the collective unconscious as it has been suggested metaphorically?³⁰ Several contemporary authors, including Andreas Ervik³¹ and Eryk Salvaggio,³² relate AI’s productions to the collective imaginary, which is a constitutive part that forms what Carl Jung calls the collective unconscious. The temptation to explain the latent space as a form of unconscious is appealing and perhaps even pertinent if we consider its intrinsic potentiality to give form. Further studies will tell.³³ But before drawing inconclusive parallels, it is wise to note the absence of the major structural component that is a being, an organic entity always in movement, in becoming, in transformation: A biological human whose brain doesn’t function as “an *input-output* machine” and whose body gestures contribute to signification.³⁴ Whereas in the latent space, so far, there is only a fixed and finite, if ever large, amount of data. That is not to say that the analogy is not without interest to understand the mechanisms at play in both workings, starting for example with our memory, which is paradoxically based on our capacity to forget—something that is being introduced to models to curb their excellent but problematic capacity to memorize everything and regurgitate it. In the parallels we draw, we should not forego biological materiality and physical multidimensionality for the discrete immateriality and formality of mathematical models of information.

While homogeneous constraints can usefully describe swarm intelligence or crowd behavior, they reduce dynamics to automatisms by excluding any imaginative or creative aspects.³⁵

²⁹ Technological objects have often been compared to exosomatic tools as they are external extensions of the body. See, for example, the camera and the eye, or the book or hard-drive and memory. Alfred Lotka first introduced the idea of “exosomatic evolution”

and Bernard Stiegler drew from it to discuss the impact of technology on the human psyche, society, and economy. For further information, see Stiegler 2021.

³⁰ Cetinić 2022.

³¹ Ervik 2023.

³² Salvaggio 2022.

³³ On the topic, see Hayles 2017.

³⁴ Longo 2023.

³⁵ Sarti / Citti 202

References

- Boehm 1995**
Was ist ein Bild?, ed. Gottfried Boehm, Bild und Text, 2, Fink, München, 1995.
- Castoriadis 1987**
Cornelius Castoriadis, *The Imaginary Institution of Society*, The MIT Press, Cambridge Mass., 1987.
- Cetinić 2022**
Eva Cetinić, "The Myth of Culturally Agnostic AI Models," arXiv 2022. Doi: <https://doi.org/10.48550/arxiv.2211.15271>
- Coccia 2010**
Emanuele Coccia, "Physique du sensible : penser l'image au Moyen-Age," in *Penser l'image*, ed. Emmanuel Alloa, 2010, pp. 95–116.
- Corbin 2016**
Henry Corbin *Creative Imagination in the Sufism of Ibn Arabi*, Princeton University Press, Princeton, N.J., 2016. Doi: <https://doi.org/10.1515/9781400853670>
- Damasio 1999**
Antonio Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Harcourt, Brace, and Co., New York, 1999.
- Didi-Huberman 2011**
Georges Didi-Huberman, "La condition des images : entretien avec Frédéric Lambert et François Nîney," in *L'expérience des images*, INA Editions, Paris, 2011,
- Ervik 2023**
Andreas Ervik, "Generative AI and the Collective Imaginary. The Technology-Guided Social Imagination in AI-Imaginesis," *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft*, vol. 37, 191/2023, (2023), pp. 42–57. URL: <https://image-journal.de/generative-ai-and-the-collective-imaginary/>. Doi: <https://doi.org/10.1453/1614-0885-1-2023-15450>
- Flusser n.d.**
Vilém Flusser, *Imagination*, Vilém Flusser Archiv, Berlin. Manuscript.
- Flusser 1984**
Vilém Flusser, *L'art et l'ordinateur* 1984.
- Flusser 2000**
Vilém Flusser, *Towards a Philosophy of Photography*, Reaktion Books, London, 2000.
- Flusser 2002a**
Vilém Flusser, *Writings*, ed. Andreas Ströhl, Electronic mediations, 6, University of Minnesota Press, Minneapolis; London, 2002.
- Flusser 2002b**
Vilém Flusser, "Images in the New Medias," in *Writings*, ed. Andreas Ströhl, University of Minnesota Press, Minneapolis; London, 2002, pp. 70–74.
- Flusser 2002c**
Vilém Flusser, "A New Imagination," in *Writings*, ed. Andreas Ströhl, University of Minnesota Press, Minneapolis; London, 2002, pp. 110–116.
- Flusser 2006**
Vilém Flusser, *La civilisation des médias*, Circé, Belval, 2006.
- Flusser 2015**
Vilém Flusser, *Flusseriana An Intellectual Toolbox*, ed. Siegfried Zielinski, Peter Weibel and Daniel Irgang, Univocal, Minneapolis, 2015.
- Hayles 2017**
N.Katherine Hayles *Unthought: The Power of the Cognitive Nonconscious*, The University of Chicago Press, Chicago, 2017.
- Jung 2023**
Carl Gustav Jung, *The Collected Works of C. G. Jung*, ed. Gerhard Adler, Michael Fordham and Herbert Read, Routledge, New York, vol. 8, 2023.
- Longo 2019**
Giuseppe Longo, "Letter to Turing," *Theory, Culture & Society*, SAGE Publications Ltd vol. 36, 6, (2019), pp. 73–94. Doi: <https://doi.org/10.1177/0263276418769733>
- Longo 2023**
Giuseppe Longo, *Le cauchemar de Prométhée. Les sciences et leurs limites*, Puf, Paris, 2023.
- Miller 1978**
Arthur I. Miller, "Visualization Lost and Regained: The Genesis of the Quantum Theory in the Period 1913-1927," in *On Aesthetics in Science*, ed. Judith Wechsler, The MIT Press, Cambridge, Mass., London, 1978, pp. 77–102.
- Oberson 2013**
Anne-Laure Oberson, *I see. Do you? Thinking seeing*, Atropos Press, New York, Dresden, 2013.
- Oberson 2020**
Anne-Laure Oberson, "The Fourth Meaning of the Image: Flusser's Invitation to Envision the World," *Flusser Studies*, 29, (2020). URL: <https://www.flusserstudies.net/archive/flusser-studies-29-%E2%80%93-2020-centennial-three-part-issue>.
- Oberson 2024**
Anne-Laure Oberson, "Making Sense of AI Generated Images in the Process of Imagination," 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.0311>
- Parikka 2023**
Jussi Parikka, *Operational Images. From the Visual to the Invisual*, University of Minnesota Press, Minneapolis, London, 2023. URL: <https://manifold.umn.edu/projects/operational-images>.
- Salvaggio 2022**
Eryk Salvaggio, "How to Read an AI Image," 2022. URL: <https://www.cyberneticforests.com/news/how-to-read-an-ai-image>.
- Salvaggio 2023**
Eryk Salvaggio, "Address Not Found (Part 2)," 2023. URL: <https://cyberneticforests.substack.com/p/address-not-found-part-2>.
- Sarti / Citti 2020**
Alessandro Sarti and Giovanna Citti, "A Fistful of Differentials," *La Deleuziana - Online Journal of Philosophy*, vols 11 - Differential Heterogenesis, (2020), pp. 4–18.
- Simondon 2022**
Gilbert Simondon, *Imagination and Invention*, University of Minnesota Press, Minneapolis; London, 2022.
- Stiegler 2021**
Bernard Stiegler, *Bifurcate: There Is No Alternative*, ed. Daniel Ross, Open Humanities Press, 2021. URL: <http://www.openhumanitiespress.org/books/titles/bifurcate/>.
- Tisseron 2010**
Serge Tisseron, *Psychanalyse de l'image Des premiers traits au virtuel*, Pluriel, Fayard, Paris, 2010.