

7 8z4yik26

You Press the Button, They Do the Rest

Moritz von Laufenberg

Chunk 1 Last year, as part of my MA studies in Visual Cultures, I was asked to give a semester-long seminar on a design related AI topic. Coming from a media history research background, I started to consider what I would like to impart onto young students of Design interested in the field of generative technology. This has allowed me to reflect on how my previous, non-AI focused research connects to this new technological shift and what lessons might be drawn from it, and what I see as lacking in the current debates and teachings. My notion has always been that in order to understand present developments in media culture, one must first study and understand the past discourses on our visual history.

Considering my studies in Digital Technology and the History of Art and Design, I have observed a number of factors that I find absent in the ongoing discussions of generational models. Not only in the general public and media but more importantly in the statements of those creating, further developing, and perpetuating this technology. I have observed that there is, on the one hand, (and this is often by design) a public lack of understanding of the technical underlyings of the technology, resulting in either overhype or demonisation, misunderstanding and more often than not, mystification. On the other hand, I see a fundamental lack of cultural knowledge on the creators' side, a blindside towards hundreds of years of philosophical, technical and socio-cultural readings of the media they are drawing from and building upon. The ongoing popularity and usage of these systems has spurred debates about the merits of this technology for human society, its relevance in the realms of art or industry, and its impacts on creativity, labour and politics.

Chunk 2 I contend that in order to understand and properly discuss the fundamental implications of this addition to our visual culture, we must first grasp not only the technological foundations on which these models were built, but also their cultural and historic origins.

III. p. 3, Chunk 2:
Editor's Note
I. p. 111, Chunk 1:
Developing an AI teach-
ing...
II. p. 29, Chunk 5:
Offenbach

Chunk 3 This is paramount especially to those of us who aim to create and perpetuate a technology that is constructed upon – and reframes – our collective visual history.

In the 2022 blog post for the aforementioned release of Stable Diffusion, stability.ai writes: “*This release is the culmination of many hours of collective effort to create a single file that compresses the visual information of humanity into a few gigabytes.*”¹ Besides the obvious issue of conflating images scraped from the web with the entire “*visual information of humanity*” this quote sheds light on the problematic starting point from which these technologies operate. It shows a fundamental lack of knowledge of the collective debates surrounding image production and visual cultures as well as the technological legacies of their own field.

Chunk 4 Returning to the first issue mentioned above, the general public’s lack of understanding of the technical fundamentals of machine learning systems, I find it necessary to recognize that, although highly specialized and complex in its entirety, a certain critical awareness and comprehension of the technical details of AI is nevertheless paramount for anyone interacting with these systems. Before we can suitably discuss *what* these systems are doing with our visual information, and how we may properly interpret their outcomes, we should have a basic understanding of *how* these systems use said visual information. As noted, I don’t expect this to extend to an understanding of code or deeper mathematical principles. What I mean is the understanding that these systems are highly advanced pattern recognition machines, built upon multiple generations of various datasets, often not designed for AI image generation. I mean the understanding that the images these systems have “learnt” from were scraped from the web, filtered by automated systems and classified by text tags that are in most cases not intended for machine learning. I mean the understanding that the images contained within the datasets have been taken without consent and are imprinted with all our human biases, faults, prejudices and sexism as well as being dominated by an English speaking, western culture and by nature of their technical fundamentals being weighted towards supporting the status quo.

How AI systems are created though their datasets has been in the eye of many critical analysts for some time now, and rightfully so. As mentioned before, the various datasets that current AI tools are built on bring with them a host of problems and issues with them that influence every outcome they produce. Even before their rise to mainstream popularity, papers have criticized these datasets for containing “*troublesome and explicit images and text pairs of rape, pornography, malign stereotypes, racist and ethnic slurs, and other extremely problematic content.*”² Additionally, many datasets that are in active use today are closed off and thus obscured for any kind of dissemination at all. It has become obvious that they all, unsurprisingly, contain large amounts of content that has been scraped without consent, is disturbing or even illegal.

Chunk 5 All the while, the flood of images produced and propagated with these image generation tools has also, again unsurprisingly, caused AI generated images to loop back into new learning datasets, the feedback of which is

1: Birhane, A., Prabhu, V. U., & Kahembwe, E. (2021). *Multimodal datasets: Misogyny, pornography, and malignant stereotypes*. arXiv. <<http://arxiv.org/abs/2110.01963>>

3: Metz, R., & Ford, B. (2024). Adobe’s “Ethical” Firefly AI Was Trained on Midjourney Images. *Bloomberg.com*. <<https://www.bloomberg.com/news/articles/2024-04-12/adobe-s-ai-firefly-used-ai-generated-images-from-rivals-for-training>>

still not entirely clear.³ In his article “LAION-5B, Stable Diffusion 1.5, and the Original Sin of Generative AI” published early this year, artist and researcher Eryk Salvaggio points towards a deeper issue that results from a lack of general education on its technical side:

“The way we have framed artificial intelligence since the generative AI boom has been deeply flawed. Rather than understanding AI as an automated form of data analytics, stripped of human supervision, we have seen countless reports on their capacities and outcomes.

4: Salvaggio, E. (2024). LAION-5B, Stable Diffusion 1.5, and the Original Sin of Generative AI \textbar TechPolicy.Press. In *Tech Policy Press*. <https://techpolicy.press/laion5b-stable-diffusion-and-the-original-sin-of-generative-ai>

Chunk 6 *Pivoting our understanding of data collection and algorithms to the frame of “Generative AI” has unnecessarily severed the understanding of this technology, erasing a decade or more of scholarship into algorithmic systems and Big*

Data. This pivot has created a harmful frame shift as policymakers scramble to understand this supposedly “unprecedented” technology.”⁴

This shift in thinking, away from a man-made tool and towards a miraculous entity whose actions are mystified and misunderstood, is exactly the issue I see in the general public discourse. These systems do not “understand” or create meaning, they don’t “learn” information from the internet and give informed answers, they don’t “understand” prompts and logically construct reliable outputs from them.

5: Salvaggio, E. (n.d.). *Thread by @e_salvaggio on Thread Reader App*. Retrieved September 20, 2024, from <https://threadreaderapp.com/thread/1646842263681499136.html>

Chunk 7 *Salvaggio notes that: “An AI image does not say what it means, because it *means* nothing. It is a guess at meaning, made by a machine that cannot comprehend meaning. It does not come from a desire to communicate, but from a design intended to simulate communication...”⁵*

6: Salvaggio, E. (2024). Challenging The Myths of Generative AI \textbar TechPolicy.Press. In *Tech Policy Press*. <https://techpolicy.press/challenging-the-myths-of-generative-ai>

The fact that generative models are all too often seen as deterministic and controllable systems that understand the users prompts and then retrieve factually correct and fitting information, is one of the most concerning myths around AI that is being perpetuated amongst public discourse.

7: Salvaggio, E. (2024). LAION-5B, Stable Diffusion 1.5, and the Original Sin of Generative AI \textbar TechPolicy.Press. In *Tech Policy Press*. <https://techpolicy.press/laion5b-stable-diffusion-and-the-original-sin-of-generative-ai>

The technological truth, namely that AI models respond with statistically likely answers born out of Gaussian noise, is harder to conceptualize and certainly less marketable. We should not forget that most AI companies producing large models are operating first and foremost for profit, not the betterment of humanity, even if they claim otherwise. Salvaggio has recently expanded his thoughts on the production and impacts of these myths produced around generative AI in an essay titled “Challenging The Myths of Generative AI”. In it, he succinctly writes: *“Intelligence myths arise from the reliance on metaphors of thinking in building automated systems. These metaphors – learning, understanding, and dreaming – are helpful shorthand. But intelligence myths rely on hazy connections to human psychology.”⁶*

The way we talk about AI systems matters, and the fact that these concepts are rarely understood or properly taught leads to a general mystification of generative AI that makes informed discussion about its impacts on our visual culture difficult if not impossible. I personally see a stronger education and demystification of these systems to be a necessary baseline when intending to dive deeper into cultural analysis. In the article “LAION-5B, Stable Diffusion 1.5,

and the Original Sin of Generative AI", Salvaggio quotes a paper from researcher Dr. Abeba Birhane, who writes: *"Large-scale AI models can be viewed, in the simplest case, as compressed representations of the large-scale datasets they are trained on. Under this light, it is important to ask what should be compressed within the weights of a neural network and by proxy, what is in a training dataset. [...] Given the wide-scale and pervasive use of such models, it is even more important to question what information is being compressed within them and disseminated to their users."*⁷

Returning to the blog post of stability.ai, one now must wonder if the weight of their actions was realized and considered at all when attempting to compress the *"visual information of humanity"* into a single archive.

Chunk 8 This brings me back to my point of the dire need to study humanity's visual culture, if one wants to stand on its shoulders.

Chunk 9 Two recently published articles jumped out which demonstrate the relevance of my point.

Chunk 10 An interview with Google's Pixel imaging team from August 13th, 2024 and an interview with Samsung's Head of Customer Experience, Patrick Chomet, from January 19th, 2024. In the latter interview, Chomet declares: *"actually, there is no such thing as a real picture. As soon as you have sensors to capture something, you reproduce [what you're seeing], and it doesn't mean anything. There is no real picture."*⁸

8: Hollister, S. (2024). "There is no such thing as a real picture," says Samsung exec. In *The Verge*. <<https://www.theverge.com/2024/2/2/24059955/samsung-no-such-thing-as-real-photo-ai>>

9: Chokkattu, J. (n.d.). For Google's Pixel Camera Team, It's All About the Memories. *Wired*. Retrieved September 20, 2024, from <<https://www.wired.com/story/google-pixel-9-real-tone-pixel-camera-interview/>>

In the former article, Isaac Reynolds, Group Product Manager for the Google Pixel Camera, muses: *"When you define a memory as that there is a fallibility to it: You could have a true and perfect representation of a moment that felt completely fake and completely wrong. What some of these edits do is help you create the moment that is the way you remember it, that's authentic to your memory and to the greater context, but maybe isn't authentic to a particular millisecond."*⁹

Forgoing a proper dissection of the arguments of these articles in too much detail, I will simply juxtapose them with a quote from Susan Sontag's seminal work *"On Photography"*, published forty-seven years prior to these articles in 1977.

Chunk 11 *"In the preface to the second edition (1843) of The Essence of Christianity, Feuerbach observes about 'our era' that it 'prefers the image to the thing, the copy to the original, the representation to the reality, appearance to being' – while being aware of doing just that.*

Chunk 12 *And his premonitory complaint has been transformed in the twentieth century into a widely agreed-on diagnosis: that a society becomes 'modern' when one of its chief activities is producing and consuming images, when images that have extraordinary powers to determine our demands upon reality and are themselves coveted substitutes for firsthand experience become indispensable to the health of the economy, the stability of the polity, and the pursuit of private happiness."*¹⁰

10: Sontag, S. (2008). *On photography* (Reissued). Penguin Books.

Images have always been used as substitutes for our experiences and memories, they have always been questioned and disseminated in terms of authenticity and connection and impact upon reality.

Chunk 13 Sontag's much discussed work is part of a longstanding and evolving history of this discussion, that began even before the invention of the camera.

Chunk 14 The interviewees are trying to say something about a topic as though it were novel, entirely new, yet it has been part of our cultural debate for millennia.

Chunk 15 Generative AI is certainly a part of this discussion and it is a topic that is worth evaluating from many different, interdisciplinary standpoints. However the lack of a baseline knowledge of the discussion up to this point just results in the reiteration of conversations held centuries ago – which in turn hinders us from reading our new technology in a way that allows us to determine its place in this broad and multifaceted exchange and, perhaps, understand its social-cultural implications and mechanisms better. In my opinion this is especially important when the technology being discussed uses, references and in a way reproduces the very history of which knowledge is lacking.

II. p. 214, Chunk 10: KI-Abend
II. p. 178, Chunk 7: Editorial
II. p. 214, Chunk 10: KI-Abend

11: Chokkattu, J. (n.d.). For Google's Pixel Camera Team, It's All About the Memories. *Wired*. Retrieved September 20, 2024, from <<https://www.wired.com/story/google-pixel-9-real-tone-pixel-camera-interview/>>

12: Sontag, S. (2008). *On photography* (Reissued). Penguin Books.

Chunk 16 At the end of one of the articles mentioned, associate professor at MIT Media Lab Ramesh Raskar (half-jokingly) muses about a “non-camera. No sensor, no lens, no flash, just a single button” that generates images via AI on the press of a button, based on GPS location and weather data.¹¹ Besides the fact that I see literally

no difference to the productions of our current image generation model, this quote did bring a fitting similarity to my mind, on which I have based the title of this essay.

In 1888, Kodak released their first consumer camera, marketed with the tagline: “You Press the Button, We Do the Rest”. The release and success of this simple-to-use camera marked the beginning of the age of amateur photography and evermore pushed the medium of photography into critical public discussion. Referring to this sales quote, Susan Sontag wrote: “*Unlike the fine-art objects of pre-democratic eras, photographs don't seem deeply beholden to the intentions of an artist. Rather, they owe their existence to a loose cooperation (quasi-magical, quasi-accidental) between photographer and subject-mediated by an ever simpler and more automated machine...*”¹²

In relation to AI image generators, this quote certainly seems, once again, quite relevant and contemporary. It is worth looking deeper into the past discussions and critical analysis of the photographic medium, as much of it is just as relevant and insightful as the day on which it was written.

13: Sontag, S. (2008). *On photography* (Reissued). Penguin Books.

14: Sontag, S. (2008). *On photography* (Reissued). Penguin Books.

Chunk 17 Since its technological emergence, photography has spurred debates on its effect on art, artists, creativity and even privacy concerns.

When Kodak ushered in the age of simple and affordable photography, its technological background was often misunderstood, multiple exposures were presented as spirit photography and wishes for images of deceased relatives were not rare occurrences. At the dawn of the twentieth century, photography had established its firm place in the world of technology and its effects on our viewing habits and image culture caused, or perhaps even allowed painters to move away from strict realism. Many artists at the turn of the century found themselves rejecting mimicry and imitation in favour of abstraction and Impressionism. Meanwhile, photographers of the same time were equally moving away from mimicking old painterly styles, instead experimenting

with the technological capacities of the camera, finding its limits and embracing its unique possibilities of depicting the world, hoping to firmly establish it as an art form. This venerable discourse was a rich field for theoretical discussion and debate that was carried far into the twentieth century, long after the question of whether or not photography could be art was deemed obsolete and mundane.

We can find fragments in these discussions that seem so contemporary, as if written only yesterday. The sometimes eerie relevance offers us new ways of reading AI discourse, from its impacts and its meaning for art and society and technology. Looking back to the previously discussed interviews, concerning the “realness” of a photograph in the age of digital sensors and filtering, we can continue our juxtaposition with Susan Sontag’s “The Image World”. She writes: “As such, the image is – to use yet another phrase of Walter Benjamin’s – without expression. It doesn’t represent reality. It is a fragment of the real world. It is a thing just like any other – a thing like you and me.”¹³ Later in her text, she elaborates that “the notions of image and reality are complementary. When the notion of reality changes, so does that of the image, and vice versa.”¹⁴

Living in the age of digital image manipulation and filtering, of AI image generation software, of deep fakes and post-truth, we should see that the understanding – the realization – that images influence our lived reality just as much as it influences our images, has been in our collective discourse for over a hundred years. Remembering Eryk Salvaggio’s statement on how generated images are “meaningless”, it becomes clear that AI images – like all images – are a fragment of the real world, without expression. This also goes to show why debates about what can be considered a “real” image miss the point, any image that is put out affects our reality, making them just as real as any other fragment of our reality. Putting this knowledge forth, we can ask ourselves deeper questions on how our reality is shaped by our changing visual culture, and we can draw from a deep well of literature to search for answers to this question. This ocean of cultural knowledge can give us new ways of seeing, new perspectives from different cultural spheres and disciplines. It can also simply allow us to ask the right questions.

Chunk 18 We can attempt to read our past discussion through the lens of our new technology, we can scour them for meaning or probe them for new questions.

Jean Baudrillard’s “Simulacra and Simulation” – for example – is forty years old, but passages of it read as though they were written with uncanny foresight. “Today abstraction is no longer that of the map, the double, the mirror, or the concept.

Chunk 19 *Simulation is no longer that of a territory, a referential being, or a substance. It is the generation by models of a real without origin or reality: a hyperreal.*¹⁵ What lessons of hyperreality can be applied to multimodal machine learning algorithms, what are they, if not “models of a real without origin or reality”? Simulations of simulations, references to references, images reproduced, stripped of meaning?

15: Baudrillard, J. (1994). *Simulacra and simulation*. University of Michigan Press.

Chunk 20 French theorist Guy Debord’s highly influential “The Society of the Spectacle”, written in the late 60s, demonstrates that core aspects of our visual culture, for all their technological progress, have perhaps not changed all that much. Debord’s definition of the spectacle rings frighteningly familiar to those

who critically examine generative AI. *“The images detached from every aspect of life fuse in a common stream in which the unity of this life can no longer be reestablished.*

16: Debord, G., Knabb, K., & Debord, G. (2014). *The society of the spectacle*. Bureau of Public Secrets.

Chunk 21 *Reality considered partially unfolds, in its own general unity, as a pseudo-world apart, an object of mere contemplation. The specialization*

*of images of the world is completed in the world of the autonomous image, where the liar has lied to himself.”*¹⁶ What do we learn when we read machine learning datasets as this “common stream”, or generated images as this “pseudo-world apart”? Keeping in mind stability.ai’s claim of wanting to create an archive of humanities visual history, how do we frame Debord’s claim that *“The spectacle cannot be understood as an abuse of the world of vision, as a product of the techniques of mass dissemination of images.*

Chunk 22 *It is, rather, a Weltanschauung which has become actual, materially translated.*

17: Debord, G., Knabb, K., & Debord, G. (2014). *The society of the spectacle*. Bureau of Public Secrets.

Chunk 23 *It is a world vision which has become objectified.”*¹⁷ Is stability.ai perhaps not creating an archive of our visual history, but instead a manifestation of their own view of what human history is, or perhaps should be?

18: McLuhan, M., Fiore, G., & Agel, J. (2008). *The medium is the message*. Penguin.

Referring to a still valid truth, coined by Marshall McLuhan in the late 1960s we learn that *“Societies have always been shaped more by the nature of the media by which men communicate than by the content of the communication.”*¹⁸

Accordingly, studying the nature of the media that shape our society becomes evermore paramount as our lives become increasingly saturated by its presence.

Chunk 24 The technical capabilities need to be understood, but most importantly, so must the social effects on all aspects of our society.

Chunk 25 When technology inevitably touches our lives, aspects such as our communication, our labour, our art or our education react and change, as they have done in the past in the wake of every new media technology. Our past reciprocal actions with these media, in art, philosophy, literature or all other forms of cultural output can give us valuable glimpses as to where our current journey may lead us, and whether or not we wish to go there.

It must be said, though, for all the parallels in the discussions of our past modes of image production and lessons to be drawn from them, there do exist some inescapable differences in certain critical aspects. Contrary to photography, AI technologies are, more often than not, closed source black boxes that only a select few have access to. Digital technology itself operates in a “hidden” realm of sorts, its inner workings are obscured from end users, there are no gears to be seen, and recently almost no buttons to press, everything is enclosed in an infinitely complex machine, masked by a glass screen and a smooth interface.

Returning to 1888 and Eastman Kodak taglining: “You Press the Button, We Do the Rest”, I feel the use of “we” implies a certain amount of inclusiveness and collaboration, something the photographer and Kodak have achieved together.

Chunk 26 I realize the “we” stems from the fact that the advertising comes straight from the mouth of the maker, not from a third party, but I nonetheless

found it more fitting to use a more detached “they” in my title instead. The fact that prompted images are sourced from unknowable datasets that were scraped without consent and colluded with concealed code, the fact that prompts are inserted without knowledge and internal weights adjusted based on conceived assumptions on what the user wants to see, affords – to me – a wording that hints at a difference between then and now.

Chunk 27 Using the same tagline might imply that nothing has changed at all, but the reality of commercial AI image generators is that they show us only what AI companies believe we want to see.

In this moment it also becomes pertinent to remember that – regardless of what AI companies want – generative AI can only produce what has already been. Being visually relegated strictly to the past, they have only processed images that already were and are biased towards yesterday.

Chunk 28 Their outputs are predictive, but the predictions are based on a statistically calculated status quo, stripped of context. Niche, fringe and radical elements will be overshadowed by the infinitely larger “popular demand” by design. Mainstream visual representations, perpetuated by large corporations, stock images and sanitized visuals, as well as floods of pornography and spam rule the statistical latent space and overexpose the already overexposed. Ultimately, to quote media theorist Roland Meyer, current models of image generation AI are “*structurally conservative, even nostalgic. Everything it produces is based on interpolating existing data.*”

Chunk 29 *In a way, it’s a backward prediction: It makes plausible guesses on what could have been, based on images from the past.*¹⁹

19: Meyer, R. (2024). Roland Meyer (@bildoperationen.bsky.social). In *BLuesky Social*.
<<https://bsky.app/profile/bildoperationen.bsky.social/post/3k1kzfdvfvf2m>>

In light of this I find it especially important, necessary even, to study and understand our relationship to past media theory. The models we are attempting to research are fundamentally and inseparably intertwined with our sociocultural visual history, by their nature of being solely based on the media of the past. Generative AI stands on the shoulders of millennia of visual culture, it should be no less required that all who work on, or with, AI systems understand this foundation. Remembering the past and viewing it through the lens of the present can afford new insights also of the past, the two being similarly interwoven.

Chunk 30 New forms of media can allow us to reinterpret fundamental aspects of old media, the same way painters once realized that realistic mimicry was perhaps not the highest plane of existence for their craft and searched for a realm beyond the scope of photographic realism. John Berger, in his seminal work “Ways of Seeing” reminds us that “*The past is never there waiting to be discovered, to be recognized for exactly what it is.*”

Chunk 31 *History always constitutes the relation between a present and its past. [...] The past is not for living in; it is a well of conclusions from which we draw in order to act.*²⁰

20: Berger, J. (2008). *Ways of seeing*. Penguin.

Read in an appropriate context, AI systems might allow us to draw new conclusions about our past visual history. I want to mention at this point that although I am critical of generative AI, I do believe it has great potential that is worth pursuing.

Chunk 32 I don’t want to give the impression that I dismiss the very impressive technical achievements that constitute their existence.

Chunk 33 I do believe that in their current form they are used haphazardly and overly mystified. In the Gartner Hype Cycle, we are now probably somewhere at the late stage of peak expectations, as critique and disillusionment are starting to slowly become more apparent in the public sphere. It is no secret that new media often begin their formative years mimicking past media, and where generative AI will go once this phase has been played out remains to be seen – but I do believe it holds potential for truly novel ways of seeing.

In 2009, writer Jörg Heiser pondered how *“Boris Groys was right in arguing that acceptance of innovation depends on cultural archiving – one can only distinguish and appreciate the new in relation to the old.*

21: Heiser, J. (n.d.). *Torture and Remedy: The End of -isms and the Beginning Hegemony of the Impure - Journal #11*. Retrieved September 21, 2024, from <https://www.e-flux.com/journal/11/61346/torture-and-remedy-the-end-of-isms-and-the-beginning-hegemony-of-the-impure/>

Chunk 34 *But what if that archive becomes so vast that it can't be held in check, if it extends beyond any single human being's capacity?”²¹*

its sheer size? As remnants of our reality, they pose a fragmented mirror into our visual culture and their computational power could allow us to find patterns and connections previously beyond the scope of human visual capability.

22: Steyerl, H. (2023). *Mean Images*. *New Left Review*, 140/141, 82-97.

Chunk 35 Naturally, all the technological limitations and biases these systems incur must then be considered with great care before grand conclusions are drawn. However, simply affording a new vantage point might be a place to start for said novel way of seeing. We can see some of these new readings emerging and taking necessary steps into discerning aspects of this technology that are distinct from previous media.

23: Steyerl, H. (2023). *Mean Images*. *New Left Review*, 140/141, 82-97.

24: Impett, L., & Offert, F. (2022). *There Is a Digital Art History*. *Visual Resources*, 38(2), 186-209. <https://doi.org/10.1080/01973762.2024.2362466>

25: Impett, L., & Offert, F. (2022). *There Is a Digital Art History*. *Visual Resources*, 38(2), 186-209. <https://doi.org/10.1080/01973762.2024.2362466>

For example, artist Hito Steyerl recently analysed an AI generated portrait of herself as an *“approximation of how society, through a filter of average internet garbage, sees me.”*, reading it as a *“mean image; a rendition of correlated averages – or: different shades of mean.”²²* She looks into the technological background of the image creation machine, connecting its classification of human faces to the early works of eugenicists, resulting in a concept of *“neural networks that bend existing social relations to converge towards a highly ideological ‘optimum’ through all sorts of market-related weights and parameters.”²³*

Very recently Fabian Offert and Leonardo Impett offered an answer to the question *“Is there a digital art history?”* in the light of *“large-scale, transformer-based vision models.”²⁴* They proposed that *“in reading a corpus of visual culture through a neural network, we are always also doing the reverse.”²⁵* Viewing classical art through the eyes of the computer vision algorithm CLIP, they equally investigate the algorithm they use as well the corpus of images they use it on.

Chunk 36 They conclude that *“we have to accept that the scope of the field needs to expand.*

26: Impett, L., & Offert, F. (2022). *There Is a Digital Art History*. *Visual Resources*, 38(2), 186-209. <https://doi.org/10.1080/01973762.2024.2362466>

Chunk 37 *Models – and their idiosyncratic ways of seeing the world – are our responsibility now, and*

any art-historical study harnessing the power of contemporary machine learning must necessarily, at least in part, also be a study of contemporary machine learning.”²⁶

These examples highlight how artists, scientists and writers are using generative AI to examine the medium itself and deconstruct the tangled web of dependencies and interactions it has with our history, our society and our other media. If we employ this type of technology to read our visual culture – and considering its current trajectory it seems very much like this is becoming more a question of when, not if – we must always remember that we are then just as much reading the technology through the lens of visual culture.

Chunk 38 Ultimately, this is the primary insight that I wanted to impart on the students of my seminar I mentioned at the beginning of this text. Understanding and employing key lessons of our cultural-historic debates on our visual culture are a necessary tool in understanding the defining cultural questions about generative AI. I believe that strong hybrid interplay, of media as well as disciplines, is necessary for us to delve into new frontiers. Considering the current trends of generative AI, it seems it will be here to stay, and we will have to deal with its presence and impact. But I believe there is a lot to gain, ultimately, through integration and understanding of any new media, especially when it is allowed to interact with a wide variety of other media and disciplines.

Chunk 39 The lessons we will learn will nonetheless require a certain amount of knowledge of the interior workings of this technology, as

well as an understanding of the cultural field it is constructed upon. As is the case with every new media, it affords us a unique way of seeing our world, but requires thought-out and critical reflection, otherwise we will just continue to keep asking the same questions and miss the bigger picture. Lack of cultural and historic knowledge, in any field for that matter, leads us into falling for that age-old trap of repeating the past. But when we do understand our technology, when we penetrate the initial spectacle that it presents to us, we can find new contexts, new uses and new meanings in it.

27: Franklin, U. M. (2004). *The real world of technology* (Revised edition). Anansi.

“The web of technology can indeed be woven differently, but even to discuss such intentional changes of pattern requires an examination of the features of the current pattern and an understanding of the origins and the purpose of the present design.” – Ursula Franklin²⁷

This article is separately published as ↗ 10.25358/openscience-11835.