

Roland Meyer

### The New Value of the Archive: AI Image Generation and the Visual Economy of 'Style'

2023

<https://doi.org/10.25969/mediarep/22314>

Veröffentlichungsversion / published version

Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Meyer, Roland: The New Value of the Archive: AI Image Generation and the Visual Economy of 'Style'. In: *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft*. Generative Imagery: Towards a 'New Paradigm' of Machine Learning-Based Image Production, Jg. 19 (2023), Nr. 1, S. 100–111. DOI: <https://doi.org/10.25969/mediarep/22314>.

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under a Deposit License (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual, and limited right for using this document. This document is solely intended for your personal, non-commercial use. All copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute, or otherwise use the document in public.

By using this particular document, you accept the conditions of use stated above.

Roland Meyer

## The New Value of the Archive: AI Image Generation and the Visual Economy of ‘Style’

*Abstract:* Text-to-image generators such as DALL·E 2, Midjourney, or Stable Diffusion promise to produce any image on command, thus transforming mere ekphrasis into an operational means of production. Yet, despite their seeming magical control over the results of image generation, prompts should not be understood as instructions to be carried out, but rather as generative search commands that direct AI models to specific regions within the stochastic spaces of possible images. In order to analyze this relationship between the prompt and the image, a productive comparison can be made with stock photography. Both stock photography databases and text-image generators rely on text descriptions of visual content, but while stock photography searches can only find what has already been produced and described, prompts are used to find what exists only as a latent possibility. This fundamentally changes the way value is ascribed to individual images. AI image generation fosters the emergence of a new networked model of visual economy, one that does not rely on closed, indexed image archives as monetizable assets, but rather conceives of the entire web as a freely available resource that can be mined at scale. Whereas in the older model each image has a precisely determinable value, what DALL·E, Midjourney, and Stable Diffusion monetize is not the individual image itself, but the patterns that emerge from the aggregation and analysis of large ensembles of images. And maybe the most central category for accessing these models, the essay argues, has become a transformed, de-hierarchized, and inclusive notion of ‘style’: for these models, everything, individual artistic modes of expression, the visual stereotypes of commercial genres, as well as the specific look of older technical media like film or photography, becomes a recognizable and marketable ‘style’, a repeatable visual pattern extracted from the digitally mobilized images of the past.

## The Question of Value

“Why is DALL-E scam?” asked artist David O’Reilly in July 2022 in a much-discussed Instagram post, and his answer was straightforward: “It rips off the past generation for the current one and charges them money for it” (O’REILLY 2022: n.pag.). In O’Reilly’s view, AI models such as DALL-E, which draw on vast quantities of photographs, illustrations, and other visual content scraped from online sources, exploit human creativity without giving anything back to the creators, or even asking them for permission. For O’Reilly, generative AI thus ultimately amounts to little more than algorithmically refined plagiarism: “Because it’s a black box, passing off DALL-E images as one’s own work is always going to be akin to plagiarism” (O’REILLY 2022: n.pag.). At the time when this was written, such fundamental criticism, which now seems almost commonplace, was hardly heard on social media. Since its launch in April 2022, DALL-E had generated fascination, even enthusiasm, not least thanks to OpenAI’s clever marketing campaign. Initially, DALL-E 2, as it was then still called, was only available to an exclusive circle of test users, ostensibly to prevent abuse. This circle was gradually expanded, but far too slowly for many of those on the waiting list. The ‘chosen few’, in turn, rewarded the exclusive access granted to them not only with their usage data but often also by starting to share their AI-generated images on Instagram, Twitter, or Facebook. Beta testers became influencers: a perfect hype machine.

And while the initial hype lasted, critical voices seemed sparse, focusing mainly on the issue of algorithmic bias – an issue that Open AI itself addressed in its “Risks and Limitations” statement back in April 2022 (OPENAI 2022a). DALL-E, for example, provided mostly male-coded images for prompts such as “CEO”, and almost exclusively female-coded ones for “assistant”. In both cases – and many others – the faces shown were predominantly *white*. This apparent lack of diversity may have been one reason why the images on DALL-E 2’s website and official Instagram account so often featured cats in space, skateboarding teddy bears, and similarly cute and supposedly innocent subjects. By July 2022, however, OpenAI had made some improvements, albeit only at the level of the text interface: In the background and without the user’s knowledge, the software now regularly mixes in keywords such as “woman” or “black” to increase the diversity of its results; or, as Fabian Offert and Thao Phan (2002: 2) put it: by “literally putting words in the user’s mouth” OpenAI “did not fix the model, but the user”.

But that wasn’t the point of O’Reilly’s criticism which focused on producers rather than on production and was a direct response to OpenAI’s announcement that it was now entering the ‘official’ beta testing phase (cf. OPENAI 2022b). This not only meant that up to a million new users were invited to try out the software, but also the introduction of a payment model. From now on, only 15 prompts per month (instead of the previous 50 per day) would be free, with

OpenAI charging for the rest. But who was the actual producer of these digitally generated images, which now cost about thirteen cents per prompt? What about those whose work the algorithm was trained on? One does not have to share O'Reilly's view that AI-based image generation is a form of plagiarism to recognize that concerns are raised here that go far beyond the idle (but very popular) debate about whether AI models such as DALL·E, Midjourney, or Stable Diffusion can create something like 'art' or even replace human artists. Rather, his intervention points to a question that seems central to understanding AI image generation as a 'new paradigm of image production' (cf. WILDE 2023): What is the value of a single image under conditions of mass digital availability of vast virtual image archives? In other words, what does image production mean when almost every conceivable image already seems to exist as a statistical possibility in a latent image space fed by images from the past? Rather than tackle these big and ultimately hard-to-answer questions head-on, I'd like to reflect in the following on two smaller, related questions that I think might shed some light on how models like DALL·E, Midjourney, and Stable Diffusion might transform our visual economy: What is a prompt, and what does 'style' mean today?

## What is a Prompt?

"Start with a detailed description", it says on the DALL·E interface, just above the text box where you can enter your prompt. In the 'new paradigm of image production', linguistic codes in the form of highly specific verbal descriptions seem to take on the role of a means of production, and the image produced is presented as a visual interpretation of a previous verbalization: at the same time as the effect and the result of a verbal prompt. Hannes Bajohr (2022) has aptly addressed prompts as a form of "operative ekphrasis", using the classical Greek term for a literary description of an image. Paradoxically, however, as a form of ekphrasis, prompts become operative only insofar as they must be understood as more than mere descriptions: They do not describe what already exists, even if only in the imagination, but are meant to produce what they describe (and what did not exist before their description). In this respect, prompts seem to resemble commands, instructions, or even lines of code – operative forms of language that also aim not just to represent pre-existing perceptions or concepts, but to produce real effects. Unlike lines of code in a programming language, however, prompts do not function as unambiguous commands: They do not follow a standardized syntax, nor are they interpreted according to transparent protocols. Most importantly, they do not produce predictable and repeatable results. Rather, and this seems to be true of all diffusion models to date, one can never predict what specific image a particular prompt will produce, since minimal

changes in the prompt will lead to visually completely different results, and even the exact repetition of a formula will conjure up ever novel, though in some respects similar images. Indeed, this “unpredictability of the results” may very well be, as Andreas Ervik argues in his contribution to this issue, “[p]art of the intrigue” (ERVIK 2023: 50).

Thus, at least from the user’s point of view, the process of image generation with text-to-image models such as DALL·E resembles a search query rather than a production process: You type a few words into a text box and four images appear that may have some relationship to what you’ve written but are far from an exact realization of the parameters you’ve specified. It is perhaps no coincidence that DALL·E’s interface design, with its clean, white, reduced looks, seems to mimic that of Google’s search engine. In a sense, DALL·E’s prompts function as search queries, directing the model to a particular region within the latent space of possible images, a region that correlates in some way with the verbalized, semantic concepts indicated in your prompt. And this search process in the latent image space can be quite time consuming, as the example of the June 2022 issue of the American magazine *Cosmopolitan* shows. For the cover of their so-called “A.I. issue”, the editors of *Cosmo* (who were also enthusiastic participants in the hype machine) wanted DALL·E to produce an image of a female astronaut on Mars. But getting the software to do exactly what they wanted was no easy task: Sometimes the astronaut didn’t look strong enough, sometimes not feminine enough (cf. CHENG 2022). Contrary to what the cover would later claim, the final image was not “created in 20 seconds” but took hours of extensive ‘prompt engineering’, the iterative optimization of text input based on trial and error. The length of the formula they eventually arrived at gives an idea of the complicated process of finding it: “Wide angle shot from below of an astronaut with an athletic female body walking with momentum towards the camera in an infinite universe on Mars, Synthwave Digital Art” (LIU 2022: n.pag.).

Here, the ‘detailed description’ is not so much a single starting point that immediately triggers the production of an image, but rather the end point of an iterative process of adjusting expectations and effects, gradually refining parameters and thereby steering the model towards the intended results. What is new about the ‘new paradigm of image production’, then, is not exactly the primacy of language. Indeed, image production as a form of visual interpretation of prior verbalization has a long history: Baroque emblematics or the pictorial programmes of Christian iconography, for example, were also based on the earlier verbalization of visual content, on descriptions as instructions for the artists who had to interpret them. In the new paradigm, however, the relationship between description and image seems to be less one of instruction and interpretation than one of navigation and matching: Verbal description does not

determine what is to be produced, but functions as a means of narrowing down selections in a space of possibilities not yet realized.

To understand this specific relationship between text and image, a productive comparison might be provided with stock photography. As Matthias Bruhn (2003) and others have shown, the value of stock images is measured by their archival accessibility and retrievability, which presupposes their prior keywording and indexing. An image that cannot be found in an agency's database, or at least not under the appropriate keyword, appears worthless, regardless of its aesthetic quality. The history of stock photography is therefore above all a media history of image retrieval systems. When the first commercial image agencies, such as the Bettmann Archive in the 1930s, turned the recycling of previously published images into a business model, the core of this model, as Estelle Blaschke (2016) has pointed out, was the storage medium of index cards. Such cards, modeled on library index cards, allowed images and metadata, visual and textual information, to be combined on a single physical data carrier, making thousands of reproducible and licensable images available to publishers, photo editors, advertising agencies, and other potential users.

With the advent of early relational database systems in the 1970s and 1980s, a decoupling of visual image and textual information took place. Mirco Melone (2018: 51-71) has analyzed how early digitization changed the function of press image archives, transforming them from mere repositories into valuable assets. With digital databases, information about press photographs was, for the first time, systematically recorded in standardized metadata, making it possible to search for individual images by photographer or location, as well as by subject, motif, or keyword. This was a prerequisite for the stock photography business, as newspaper image archives now became a commercial resource for publishers. Initially, however, this only applied to newly produced photographs, as the vast quantities of historical photographs stored in archives were only gradually being digitally indexed and made accessible. As Bruhn (2003: 9) has noted, bureaucratic management and the commercial exploitation of visibility go hand in hand: Turning a mere collection of images into an economic asset required archival logistics of image retrieval, and these logistics ultimately defined the value of images as commodities.

While stock and press photo databases only allow to search for images that already exist and have been indexed, text-to-image generation prompts allow to 'search' for images that don't exist yet and therefore have never been indexed – blurring the lines between production and re-production, search and generation. Rather than being optimized for expected and likely queries, as is the case with many stock photo services today, text-to-image models such as DALL-E, Midjourney, or Stable Diffusion open up possibility spaces for unlikely and unanticipated search commands. In particular, they allow us to formulate

search queries that do not need to be matched by any prior image, not even in our imagination. When formulating a prompt, words can be combined counterfactually, even meaninglessly or purely randomly. In fact, text-image models like DALL-E may surprise you rather than give you exactly what you are looking for, and perhaps the best way to be surprised is to formulate queries that do not match anything already found in the vast virtual image archives on which the software has been trained – to ask DALL-E for a self-portrait, for example. Rather than a logistics of image retrieval that transforms vast archived collections of images into valuable assets, what we have with AI image generation is a logistics of accessing and navigating vast latent spaces of possible images, made possible by, but by no means limited to, already archived images. In a sense, the individual image produced by these models is not just an element of an archive, but rather its product, a contingent outcome that recombines, synthesizes, and interpolates what has already been produced and described.

Such aspects of combinatorics and contingency, especially in the way images and descriptions are matched (and more often than not also mis-matched), link DALL-E to the historical Surrealism alluded to by its name, a portmanteau of (Salvador) Dali and (Pixar's) *Wall-E*. As Sven Spieker (2008: 85-103) has pointed out, the early Surrealists were fascinated by the idea of the unconscious as a kind of linguistically structured archive. In order to reveal the latent structures behind unconscious phenomena such as dreams, the Surrealist group around André Breton used office media such as index cards and filing cabinets, which provided a technical means of disrupting the logic of the everyday. The recombination of words, letters, and other linguistic elements, as well as the re-mixing and re-filing of documents, allowed contingency and chance to produce an “order of disorder” (SPIEKER 2008: 98) which was based on the combinatorial, structural, and relational logic of the archive. Many AI-generated images, especially those made with DALL-E, look like a strange blend of Surrealism and stock photography, maybe because they conflate a linguistically structured combinatorial ‘dream logic’ with a visual conventionality fueled by commercial image archives. In a sense, they realize what Fredric Jameson once claimed about experimental video art: a “surrealism without the unconscious” (1991: 67). Indeed, the cultural logic of postmodernism in general, which, in Jameson’s words, “ceaselessly reshuffles the fragments of preexisting texts, the building blocks of older cultural and social production” (JAMESON 1991: 96), now seems to have become the technical logic of automated image generation.

The infrastructural precondition for this never-ending reshuffling of cultural fragments is the existence of vast amounts of images found online, already annotated and described, on which AI models such as DALL-E, Midjourney, and Stable Diffusion can be trained. In other words, what makes these models operational is the fact that, in today’s platform-based visual economy, digital images are always

already surrounded by clouds of textual information and are therefore related to semantic concepts in multiple ways. Text-to-image generation thus presupposes extensive semantic pre-processing of digital image cultures, often the product of crowdsourced “ghost work” (GRAY/SURI 2019: ix) by underpaid click-workers. In this respect, the latent spaces of AI image generation are unthinkable without the emergence of what Adrian MacKenzie and Anna Munster (2019: 3) have called “image ensembles”, huge aggregations not only of images but of images that have been formatted, labeled, enriched with metadata, and thus made “platform-ready” (MACKENZIE/MUNSTER 2019: 5). In fact, with text-to-image generation, this semantic preprocessing of digital images almost comes full circle, as prompts can be understood as metadata descriptions attached to an image even before it is generated.

But first and foremost, prompts are generative search queries for exploring and exploiting latent image spaces: A huge virtual archive of possible images is organized and made navigable based on semantic concepts. The contingent combinability of semantic units thus becomes the operative principle of a generative search: a search process that produces what it is looking for within the limits of statistical possibility. Whereas, in the earlier database logic of stock photography, pre-existing images were stored and indexed as stable and individual units, forming a kind of asset or “image capital” (BLASCHKE/LINKE 2022), now vast archives of text-image pairs have become not only a training ground for machine learning, but also a multidimensional data manifold capable of generating never-before-seen images. More than just an asset, the archive thus becomes a veritable resource of image production. And this, I will argue in the concluding paragraphs of this essay, fundamentally changes the way in which value is ascribed to images. But for this, let me first turn to my second question: What does ‘style’ mean today?

## The New Meaning of Style

“An Impressionist oil painting of sunflowers in a purple vase...” This is the suggested prompt you can read in light grey letters in DALL-E’s main text field, just before you enter your own prompt. This pre-formulated, generic prompt serves as a kind of example and inspiration, and also gives hints about the basic, though not binding, ‘grammar’ of prompts: a combination of terms denoting style (“impressionist”), medium (“oil painting”), and subject or motif (“sunflowers in a purple vase”), not necessarily in that order. I’ll come back to the question of medium, but for now, let’s focus on the category of style. Formulating prompts allows considering subject and style, iconography and form, as separate parameters: Historical as well as contemporary, collective as well as individual forms of



representation can seemingly be detached at will from their time and place of origin and the work of their authors. It is not least for this reason that O'Reilly and others speak of plagiarism.

More importantly, the logic of the prompt radically expands and de-hierarchizes the notion of style: Style can refer to the classical art historical sense of an epochal style or the individual style of a canonized artist, but it can also refer to the aesthetic qualities of certain products of popular culture or the visual appearance associated with specific genres and media formats. *The DALL·E 2 Prompt Book*, a popular online tutorial on how to write better prompts (DALL·ERY GALLERY 2022), aptly illustrates this expansion of the concept of style by suggesting that the words “in the style of...” be combined with the names of individual painters, photographers, and illustrators, as well as with those of popular cartoons and TV series such as *South Park* or *The Simpsons*. But the category of style, at least according to the *Prompt Book*, also includes generic illustration styles such as “botanical illustration”, “political cartoon”, and “IKEA manual”, specific artistic techniques and media such as “airbrush” and “vector art”, and many more (cf. DALL·ERY GALLERY 2022).

In other words, in models like DALL·E, the individual brushstrokes of Van Gogh or Vermeer and the recognizable look of “steampunk” or “synthwave” seem to be almost interchangeable, transferable, and even, at least to some extent, combinable parameters within an extended category of ‘style’. As examples such as “airbrush”, “cartoon” or “digital art” show, this notion of style cannot be clearly separated from the category of medium neither. In the logic of the prompt, “in the style of Vermeer” or “1970s Polaroid” both function as modifiers indicating a certain ‘look’ that affects not only certain elements within the image but the image as a whole. Everything becomes a ‘style’, and while, in name, all these different ‘styles’ are still associated with people, media, genres, techniques, formats, places, or historical periods, in the production logic of the AI model they are nothing more than typical visual patterns extracted from a latent space of possible images accessed through generative (and often iterative) search queries.

Thus ‘style’ ceases to be a historical category and becomes a pattern of visual information to be extracted and monetized. As Jens Schröter (2022) has pointed out, this tendency has already been described to some extent by Hal Foster in his essay “The Archive without Museums” (1996). Foster distinguishes here between the discipline of art history, which relied on photographic reproductions to “abstract a wide range of *objects* into a system of *style*” (1996: 97, original emphases), and the (then) new discourse of visual culture, which, he suggests, relies on information technologies “to transform a wide range of *mediums* into a system of *image-text* – a database of digital terms, an archive without museums” (FOSTER 1996: 97, original emphases). The main difference here is between a system of

styles, which also organized the classical art museum, and a system of image-text, which organizes the digital archive. In the logic of the museum, styles had a life of their own and their story could be told through exemplary masterpieces. In the archive and its digital derivatives, style becomes a search term for accessing a manifold of visual data. And while the museum necessarily excluded everything that did not fit into its narrative and its pre-stabilized categories, the archive can accommodate all kinds of images as information, without boundaries or hierarchies.

In the case of DALL·E, Midjourney, and Stable Diffusion, the latent space that forms a kind of virtual image archive also includes an infinite number of images that (almost) look like ordinary photographs but are not photographs at all. For these models, the 'photographic' seems to be just another 'style', an aesthetic, a certain 'look', not a privileged mode of indexical access to the world. And this 'photorealistic style', I would argue, simulates *visual* rather than *optical* aspects of the photographic. For unlike, say, game engines, architectural renderings, or Hollywood CGI effects, AI image generation does not use a three-dimensional model of a physical reality calculated according to optical laws and the rules of perspective but recombines and synthesizes visual surface textures and 'looks'. The world it shows is basically flat, as it does not consist of bodies and objects, not even virtual ones, but of visual patterns that have been transformed into digital information.

As experienced 'prompt engineers' soon discovered, particularly 'photorealistic' effects can be achieved if the prompt already contains technical information referring to photographic equipment, such as lenses and shutter speeds (cf. MERZMENSCH 2022). Again, however, unlike for the parameters of virtual cameras in video game engines and CGI programs (cf. SCHRÖTER 2003), technical specifications such as "wide angle lens" or "Sigma 24mm f/8" in a text-to-image prompt do not feed into an *optical* simulation of the photographic apparatus. Rather, they are merely typical keywords and attributes that, in the logic of the model, correlate with recurring *visual* qualities of large quantities of images – not unlike generic quality statements such as "perfect" or "prize-winning photograph". We are thus dealing here, as in many other cases of networked visual culture, with computational images that are not based on prior models of a physical reality, but on the posterior statistical analysis of large collections of two-dimensional images and their descriptions.

What models like DALL·E, Midjourney, and Stable Diffusion thus show us are not images of the world, but images of images – indeed, ultimately images about images, filtered through language. To see this as a mere affront to human creativity, or even as a scam, may miss the point. Rather, such AI models mark a crucial stage in the progressive exploitation of virtual image archives as a productive data resource. The archive of semantically encoded and digitally

mobilized images of the past thus becomes a seemingly inexhaustible source of visual patterns that can be extracted, varied, and transformed at will, across time, and beyond established hierarchies of cultural value. This process goes far beyond the field of AI image generation and is linked to two tendencies that seem to characterize our current visual economy and culture in general: First, operating with digital images today means navigating the virtual image archives of big data. In a networked digital culture, images are no longer isolated artifacts, but elements within “virtually unlimited populations of images” (JOSELIT 2013: 13), already semantically predefined and pre-processed, enriched with non-visual information that significantly determines their accessibility and thus also their value. And secondly, a concept such as ‘style’, broadly understood as a nameable and repeatable form of visual aesthetics, a ‘vibe’, ‘mood’, or ‘look’ is now becoming an algorithmically exploitable resource capable of generating infinite variants of new images. As a pattern that can be extracted from large aggregations of digitally mobilized visual content, and thus detached from the individual image, its author, its medium, and its conditions of production, ‘style’ becomes a source of value. This may or may not be ethically problematic, but it is undoubtedly bad news for individual creators and for industries that still depend on licensing individual creations.

It may come as no surprise, then, that at the time of writing this, Getty Images, one of the world’s largest stock image agencies, is suing the company behind Stable Diffusion, Stability AI, for copyright infringement. In fact, this is not the only court case to consider whether the use of copyrighted visual content to train AI models is a practice of ‘fair use’ or rather a form of plagiarism (cf. VINCENT 2023). While it remains to be seen how the courts will decide, the case itself seems telling: Getty and Stability AI essentially represent two very different definitions of the value of images. While Getty stands for an older system of closed image archives as monetizable assets, where licenses are sold for individual uses of images, and which historically goes back to the Bettman Archive (whose licenses are now part of the Getty portfolio), under the ‘new paradigm of image production’ we can see the emergence of a networked model of image monetization which understands the entire web as a freely available resource that can be mined at scale. And while in Getty’s business model, each image has a precisely determinable value, for DALL·E, Midjourney, and Stable Diffusion the single image doesn’t matter much. The commodity they sell is less the individual image artifact itself, but the patterns derived from aggregating and analyzing vast image ensembles. Getty’s lawsuit, then, seems to be the attempt of a major player of an older economy to stake a claim to future markets, to at least be still recognized as a player, albeit a minor one, in this new visual economy. Whatever the outcome, one thing seems certain: Even if AI companies are required to license the images they use for training, creators will receive only a tiny share.

## Bibliography

- BAJOHR, HANNES: Operative Ekphrasis: Meaning, Image, Language in Artificial Neural Networks. Unpublished lecture, conference “*Künstliche Intelligenz – Intelligente Kunst? Mensch-Maschine-Interaktion und kreative Praxis*”, TU Braunschweig, October 8, 2022
- BLASCHKE, ESTELLE: *Banking on Images: The Bettmann Archive and Corbis*. Leipzig [Spector] 2016
- BLASCHKE, ESTELLE; LINKE, ARMIN (eds.): *Image Capital*. Essen [Folkwang Museum] 2022. <https://image-capital.com/intro/> [accessed February 16, 2023]
- BRUHN, MATTHIAS: *Bildwirtschaft: Verwaltung und Verwertung der Sichtbarkeit*. Weimar [VDG] 2003
- CHENG, KAREN X. (@karenxcheng): Created the First Ever AI Cover for Cosmopolitan Magazine! Video on *YouTube*. June 22, 2022. <https://www.youtube.com/watch?v=8fthDHDshvg> [accessed February 16, 2023]
- DALL-ERY GALL-ERY (ed.): *The DALL-E 2 Prompt Book*. In: *Dall-ery gall-ery: Ressources for Creative DALL-E Users*. July 14, 2022. <https://dallery.gallery/the-dalle-2-prompt-book/> [accessed February 2, 2023]
- ERVIK, ANDREAS: Generative AI and the Collective Imaginary: The Technology-Guided Social Imagination in AI-Imaginesis. In: *Generative Imagery: Towards a 'New Paradigm' of Machine Learning-Based Image Production, special-themed issue of IMAGE: The Interdisciplinary Journal of Image Sciences*, 37(1), 2023, pp. 42-57
- FOSTER, HAL: The Archive without Museums. In: *October*, 77, 1996, pp. 97-119
- GRAY, MARY L.; SIDDHART SURI: *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston [Houghton Mifflin Harcourt] 2019
- JAMESON, FREDRIC: *Postmodernism, or, The Cultural Logic of Late Capitalism*. London [Verso] 1991
- JOSELIT, DAVID: *After Art*. Princeton [Princeton University Press] 2013
- LIU, GLORIA: The World's Smartest Artificial Intelligence Just Made its First Magazine Cover. In: *Cosmopolitan*. June 21, 2022. <https://www.cosmopolitan.com/lifestyle/a40314356/dall-e-2-artificial-intelligence-cover/> [accessed February 16, 2023]
- MACKENZIE, ADRIAN; ANNA MUNSTER: Platform Seeing: Image Ensembles and their Invisibilities. In: *Theory, Culture & Society*, 36(5), 2019, pp. 3-22
- MELONE, MIRCO: *Zwischen Bilderlast und Bilderschatz: Pressefotografie und Bildarchive im Zeitalter der Digitalisierung*. Paderborn [Fink] 2018
- MERZMENSCH: Prompt Design for DALL-E: Photorealism – Emulating Reality. In: *Medium*. June 9, 2022. <https://medium.com/merzazine/prompt-design-for-dall-e-photorealismemulating-reality-6f478df6f186> [accessed February 16, 2023]
- O'REILLY, DAVID: Why is DALL-E a Scam? Post on *Instagram*. July 22, 2022. [https://www.instagram.com/p/CgSqRxhPF\\_X/](https://www.instagram.com/p/CgSqRxhPF_X/) [accessed February 16, 2023]

- OFFERT, FABIAN; THAO PHAN: A Sign That Spells: DALL-E 2, Invisual Images and the Racial Politics of Feature Space. *arXiv:2211.06323*. October 26, 2022. <https://arxiv.org/abs/2211.06323> [accessed February 16, 2023]
- OPENAI: DALL-E 2 Preview – Risks and Limitations. In: *GitHub*. April 11, 2022. <https://github.com/openai/dalle-2-preview/blob/main/system-card.md> [accessed February 16, 2023]
- OPENAI: DALL-E now Available in Beta. In: *OpenAI Blog*. July 20, 2022. <https://openai.com/blog/dall-e-now-available-in-beta> [accessed March 5, 2023]
- SCHRÖTER, JENS: Virtuelle Kamera: Zum Fortbestand fotografischer Medien in computergenerierten Bildern. In: *Fotogeschichte* 88, 2003, pp. 3-16
- SCHRÖTER, JENS: AI, Automation, Creativity, Cognitive Labor. Unpublished lecture, conference “Künstliche Intelligenz – Intelligente Kunst? Mensch-Maschine-Interaktion und kreative Praxis”, TU Braunschweig, October 8, 2022
- SPIEKER, SVEN: *The Big Archive: Art from Bureaucracy*. Cambridge, MA [MIT Press] 2008
- VINCENT, JAMES: Getty Images is Suing the Creators of AI Art Tool Stable Diffusion for Scraping its Content. In: *The Verge*. January 17, 2023. <https://www.theverge.com/2023/1/17/23558516/aiart-copyright-stable-diffusion-getty-images-lawsuit> [accessed February 16, 2023]
- WILDE, LUKAS R.A.: Generative Imagery as Media Form and Research Field: Introduction to a New Paradigm. In: *Generative Imagery: Towards a 'New Paradigm' of Machine Learning-Based Image Production, special-themed issue of IMAGE: The Interdisciplinary Journal of Image Sciences*, 37(1), 2023, pp. 6-33