

Lukas R.A. Wilde

Generative Imagery as Media Form and Research Field: Introduction to a New Paradigm

Abstract: This introduction to the collection “Generative Imagery: Towards a ‘New Paradigm’ of Machine Learning-Based Image Production” discusses whether – or to what respect – generative imagery represents a new paradigm for image production; and if that constitutes even a novel media form and an emerging research field. Specifically, it asks what a humanities approach to machine learning-based image generation could look like and which questions disciplines like media studies will be tasked to ask in the future. The essay first focuses on continuities and connections rather than on alleged radical shifts in media history. It then highlights some salient differences of generative imagery – not only in contrast to photography or painting but specifically to earlier forms of computer-generated imagery. Postulating a ‘new paradigm’ will thus be based 1) on generative imagery’s emergent or stochastic features, 2) on two inter-related, but often competing entanglements of immediacy-oriented and hypermediacy-oriented forms of realisms, and 3) on a new text-image-relation built on the approximation of ‘natural’, meaning here *human* rather than machine code-based language. The survey closes with some reflections about the conditions under which to address this imagery as a distinct media (form), instead of ‘merely’ as a new technology. The proposal it makes is to address generative imagery as a form of *mediation* within evolving dispositifs, assemblages, or socio-technological configurations of image generation that reconfigure the distribution of agency and subject positions within contemporary media cultures – especially between human and non-human (technological as well as institutional) actors. Of special importance to identify any (cultural) distinctness of generative imagery will thus be a praxeological perceptive on the establishment, attribution, and negotiation of cultural ‘protocols’ (conventionalized practices and typical use cases), within already existing media forms as well as across and beyond them.

Introduction

The emergence of machine learning-based platforms has been a prominent and increasingly prevalent topic in both popular as well as specialized academic discussions for many years now (cf. for a survey NILSSON 2010; SUDMANN 2018a; MITCHELL 2019). Around the middle of the year 2022, these emerging technologies left the spheres of R&D departments, computer science labs, and our speculative imagination. Generative platforms started to pervade the everyday life of people around the globe. Beginning with text-to-image technologies such as DALL·E, Stable Diffusion, or Midjourney (flanked by a range of other competitors such as Imagen, Wombo Dream, DeepDream, or Leonardo AI), and succeeded by further evolving and increasingly easier to-access prompt-to-text-applications like ChatGPT, Bing, or Bard. Discussions about the imminent threats, potentials, and transformations of media and communication now permeate news media, popular culture, and academic discourses. Other forms of machine learning technologies are developing steadily too, with text-to-music, text-to-video, text-to-code, or even text-to-3D rapidly progressing. Certainly, machine learning-based image generation technologies – commonly referred to as ‘AI imagery’ or ‘generative imagery’ – are only a small part of these developments. Their history was long in the making long before the summer of 2022 (cf. MILLER 2019: 59-122; BAJOHR 2021). The successive stages of technological developments in the area of generative imagery have been historized (cf. OFFERT 2022) as a transition from classification to generation (2012–2015), over five years of GAN development (generative adversarial networks, 2015–2020), leading up to the currently popular diffusion models (2020–present), whose ‘multimodal’ deep learning through CLIP (contrastive language-image pre-training) and GLIDE (guided language-to-image diffusion for generation and editing) combines and consolidates techniques from NLP (natural language processing) and “computer vision” (DOBSON 2023). Despite this gradual progress and the fact that the actual deep learning-“media revolution” (SUDMANN 2018b: 66; my translation) has happened a while ago – or rather: has been happening for a long time now – the summer of 2022 introduced a moment of radical shift in the public awareness, mainly due to the fact that generative imagery since left the confinement and control of companies, research labs, or specialized artistic experiments, becoming available to the general public. This also marked the beginning of what Fabian Offert (2022: n.pag.) called the “Photoshop era” of such image synthesis. It is now feasible to use generative models as an everyday tool to create highly realistic images from a rough sketch, adding AI-based modifications layer by layer. Stability AI’s open-source application Stable Diffusion, for instance, is characterized by a modular architecture that allows working with more and more fine-tuned extensions such as OpenPose Editor or ControlNet (cf. ZHANG/AGRAWALA 2023) and through the

exchange of individual, pre-trained models through the collaborative hosting and exchange platform GitHub. As of early March 2023, there are already Stable Diffusion plugins available for Adobe Photoshop and other graphics programs (cf. ALFARAJ 2023), integrating generative imagery seamlessly into established practices of digital image production and editing.¹¹

After an initial rush of public interest in this imagery around July to October of 2022, prompt-to-text platforms seem to attract not only much more press coverage at present (March 2023) – at times excited, worried, or increasingly annoyed. They also seem to necessitate more ‘emergency meetings’ in universities and other institutions where decisions need to be made quickly on how to deal with the impacts of ChatGPT and the like on all aspects of social, cultural, and political life. In many other ways, too, earlier prompt-to-image platforms appear more harmless to existing regulations. As Hannes Bajohr (2023) remarks in his contribution to this collection, nobody would (and, to my best knowledge, nobody *did*) speak of DALL·E, Stable Diffusion, or Midjourney as having any sort of consciousness or personality – let alone a range of *alternate* personalities ‘discovered’ in ChatGPT or Bing (cf. TANGERMANN 2023; VINCENT 2023). For AI chatbots simulating direct communicative interactions, this is currently discussed daily (even if arguably in some frame of suspension of disbelief, make-believe, or role-play, as René Walter, 2023, has argued). Generative imagery still seems to retain a much more salient *instrumental* role, discussions of alleged ‘autonomy’ or ‘creativity’ restricted to the interpretation of prompts and the subsequent production of results, not the interaction or communication with human users (via images) itself. This might partly be owed to present interface design limitations: None of the currently available generative imagery platforms retain memories between input prompts, which is a mere technical limitation at this point. Certainly, as both prompt-to-image as well as prompt-to-text technologies make their APIs interfaces available (cf. BROCKMAN et al. 2023), a dialogue and memory-based image platform is probably not too far away (enabling hypothetical commands like “combine the last three results, and then respond with another picture representing a next moment in time”).¹² Arguably, however, it would still be the verbal interaction through chats prompts that could generate the uncanny impression of a ‘responding agent’ once again, not the immediate ‘communication’ through image generation, for the simple reasons that this core

1 Only on March 21, Adobe even unveiled their own generative AI, “Firefly”, advertised as *not* drawing on proprietary material of earlier artists that did not agree to this (cf. ADOBE 2023) – which should change a lot of things argued for within the essays in the present collection. Given the speed of current developments, it will be harder and harder to write texts that are still somewhat up to date, it seems (cf. WILDE 2023).

2 Actually, also mere days before the manuscript for this publication was finalized, OpenAI not only announced that a later version of GPT-4 would be multimodal (cf. OPENAI 2023), Microsoft also published a press release that Bing would soon entail DALL·E to do, under the name “Bing Image”, exactly what was merely imagined here (cf. MICROSOFT 2023).

function – producing novel images at rapid speed in seconds – simply has *no* equivalent in earlier human (or even human-machine-augmented) communication and thus runs contrary to all communicative intuition.^[3]

These cursory thoughts are, in any case, about the only remarks about prompt-to-text platforms provided within the present collection of essays – with the exception of Bajohr who dives more deeply into the ‘artificial semantics’ of large-language-models behind both prompt-to-image and prompt-to-text-platforms. The following thirteen essays instead offer a range of humanities-based perspectives on the ‘discourse event’ that started the AI discussion back in July–October 2022. Limiting our interest to AI-generated pictorial representations and image forms, the overarching question for our workshop “DALL·E, Midjourney, Stable Diffusion: Responses from Media Studies towards a ‘New Paradigm’ of Image Production” (University of Tübingen, February 13/14, 2023) seemed ambitious enough: Does the availability of generative imagery as an everyday resource represent a moment of media change in contemporary image and media history, perhaps as consequential as the transition from mechanically to photochemically produced pictures or even as the emergence of mechanical reproduction before? In October 2022, when Klaus Sachs-Hombach and I published the Call for Participation asking these questions, answers seemed uncertain at best. As every responsible scholar would, we hence put the ‘new paradigm’ of image production into single quotation marks. Half a year later now, it seems less complacent to do without them confidently. This certainly demands some reasoning. In the present introduction to our collection, I would like to provide a few parameters and coordinates for the ‘latent space’ of media studies and picture theory discourse, if this metaphorical expression is allowed, that the following essays might be situated in. Their proposed perspectives are based on a range of fields across the humanities, of which media studies is just one. Indeed, the urgent concerns and questions posed by generative imagery are going to be of paramount importance for all disciplines working with and on images, pictoriality, and visual or multimodal communication. What media studies – or the conceptual and analytical departing point of *mediation* – could offer for these discussions, perhaps, is a framework connecting and interrelating communicative-semiotic, material-technological, and cultural-institutional concerns and perspectives.

First, I want to set out from a perspective focusing on continuities and connections rather than on radical shifts in media history. Secondly, I do want to

3 An interesting point of comparison might be found in the narratological observation that verbal texts usually generate the impression of an anthropomorphic narrator or of a personalized voice (perhaps even distinct from the actual author), while this is not necessarily true for the pictures of films or comic books: “Written narrative text is perceived as analog to the process of verbal narration, it is (in Fludernik’s 1996 terminology) ‘naturalized’. Comics, as well as films, have, regarding their visual components, no equivalent in mundane, everyday communication” (SCHÜWER 2008: 389; my translation).

highlight some salient differences of generative imagery possibly constituting a ‘new paradigm’ – not only with regard to photography or painting but specifically in contradistinction to earlier forms of computer-generated imagery or ‘machine vision’. Finally, if we understand generative imagery as an emerging, distinct field of research in the humanities, can we identify some of the key concerns within this paradigm? My introduction closes with a few reflections about the conditions under which to address these new image technologies as a distinct media (form). The proposal I want to make is this: addressing generative imagery as a (partially novel) form of *mediation* asks how these developing dispositifs, assemblages, or socio-technological configurations of picture generation reconfigure the distribution of agency and subject positions within contemporary media cultures, especially between human and non-human (technological as well as institutional) actors.

Continuities and Connections?

Generative image platforms produce pictorial artifacts without the indexical relations of photography to light waves or of painting to brush strokes. As Eryk Salvaggio (2023a) argues most convincingly in his present contribution, they instead recombine and perhaps also reveal aspects of underlying pictorial datasets as well as of the human decisions behind their classification and organization. Still, we might ask skeptically: what is genuinely new about that, really? The abandonment of referential reality (of an indexical relationship to physical reality), is hardly new for digital pictures and has been established through CGI for decades (cf. MITCHELL 1992; RICHTER 2008; GOOSKENS 2011). The partial autonomy of a ‘non-human apparatus’ generating pictures ‘automatically’ might even constitute one of the points of departure of media theory with the emergence of photography over a hundred years ago (cf., for instance, BENJAMIN 2007 [1935]). Generative imagery is then remarkable perhaps not in quality but in quantity, speed, and availability as platforms like DALL-E, Midjourney, or Stable Diffusion can generate, through rapid feedback loops, an infinite number of pictures in all possible stylistic variations at incredible speed even for laypersons. All the resulting *individual* pictures then seem so arbitrary and ephemeral that they hardly seem to deserve deepened individual attention or analysis. This, however, makes generative imagery perhaps an especially suited topic for media studies and media theory interested less in individual artifacts (or ‘imagetexts’) than in the structural impact of media technologies on culture and society in general.

The lasting consequences of this moment of media transformation on social, political, and cultural practices, conventions, and institutions are certainly

far from decided or determined at this point in time. What can be stated with some confidence, however, is that the speed of recent developments has been surprising for most observers. For the time being, our institutions and laws are hopelessly lagging to regulate some very old (and some newly emerging) questions. As Jay D. Bolter (2023) points out in his present contribution, high on this list of urgent concerns are certainly questions of authorship (plagiarism vs. fair use) under these new technological decisions. One possible task, specifically for media studies, could then be to highlight continuities and connections a) between generative imagery and earlier forms of “machine vision” (cf. GALLOWAY 2021; RETTBERG 2022; 2023, as well as DOBSON 2023), b) between the present and earlier moments of media transformation and media change, as well as perhaps c) between practices and uses of pictures that have either proven resilient to such changes or are resurfacing. A respective praxeological perspective might go way back, indeed. Lev Manovich, who inspired our discussions as early as in July 2022 in a series of Facebook ‘micro essays’ (for lack of a better term), described something that he coined “the return of the classical ‘art of the copy’” (MANOVICH 2022: n.pag.). His observation was that art historical storytelling, focusing on individual, outstanding pictures, largely ignored the hundreds and thousands of similar copies and variations that were actually produced in studios and workshops – in favor of a highly selective (and thus finally ideological) ‘slice of history’ in museums today. The production of pictures has then, maybe, always been dominated by practices of imitating, copying, and slightly varying existing patterns of visual representation. We are all the more excited to have an opening essay by Manovich (2023a) in the present collection that draws especially on his perspectives and experiences as an artist and practitioner.

Praxeological questions might reveal many more such connections, the most saliently one the notion of “remix” and “remix culture” that Bolter (2023) and Lamerichs (2023) discuss in more detail. Not only audio remix (in hip-hop) has been established for decades, but also “the somewhat younger video remix, which involves the editing and often complex layering of a series of video clips together” (BOLTER 2023: 199). Comparisons to older, ‘analog’ media and image technologies can also reveal interesting analogies with regard to their ‘statistical’ nature as Jens Schröter (2023) discusses in his essay on Francis Galton’s composite photography portraits and Sigmund Freud’s fascination with them. To Freud, superimposed composite images corresponded to the generalized visual condensation of dreams through the subconsciousness – or at least our recollection of dreams. Can AI-generated imagery thus be seen as a contemporary, mediated form of a collective “statistical unconscious” (SCHRÖTER 2023: 111)? Roland Meyer (2023c), in turn, discusses a more immediate media-historical connection between generative imagery and stock photography and press image archives on the one hand and recent digital search engines on the other. Meyer traces

how the Bettmann Archive in the 1930s created a new form of image valorization by collecting and ‘assembling’ pictures together with metadata on physical data carriers like index cards. The mediality of both image forms – in physical archives as well as especially in generative platforms – is thus determined by their valorization and commodification which in turn rest on a “media history of image retrieval systems” (MEYER 2023c: 103). Another analogy could be found with respect to fan cultures and fan practices, currently certainly the sociocultural context where generative imagery is exploited, tested, and negotiated most viciously. Nicolle Lamerichs (2023) discusses in her survey of these developments to what extent generative platforms could be considered a form of ‘transformative fan fiction’ even on a technological level, albeit one that is deeply entangled in platform economies and respective data-driven business models that have been evolving rapidly for about 10 years now. A different form of continuity is then again pointed out by Pamela Scorzin in her survey of artistic practices that include not only the newest iterations of machine learning-based technologies. Technologically distinct phenomena such as humanoid robots on media stages, avatar design in the metaverse, or partly algorithmic created music videos are employed to represent similar questions or recurring topics like artistic authorship or mediated body representations. Manovich (2023a) likewise points out such connections with regard to Ivan Sutherland’s computer program Sketchpad (1961-1962) that finished half-drawn circles or rectangles; within “cultural perception” (!) this “was undoubtedly ‘AI’ already” (MANOVICH 2023a: 33).

As important a task as it will be to describe generative imagery on the level of social practices – and thus in terms of continuities and connections rather than in dramatic ‘turns’ – there *are* many perspectives that focus on mostly new aspects of mediation between human and non-human agents. Many of the contributors in the present collection still turn to well-known protagonists of media studies and media theory to pinpoint what, exactly, distinguishes generative imagery from photography as much as from ‘analog’ picture forms before them. These readings at the same time create and challenge notions of continuity in media history. We will thus once again visit the thoughts of authors like Theodor Adorno (OFFERT 2023) and Walter Benjamin (ERVIK 2023), John Austin and Ludwig Wittgenstein (FEYERSINGER et al. 2023), Roland Barthes (ERVIK 2023; OFFERT 2023; SALVAGGIO 2023a; SCHRÖTER 2023) and Susann Sonntag (MICHOS 2023), Stuart Hall (SALVAGGIO 2023) and Fredric Jameson (MEYER 2023c), or Marshall McLuhan (ERVIK 2023) and Sybille Krämer (OFFERT 2023), to name just a few. Our contributors thus explore what their thoughts could highlight about generative imagery and the (dis)continuities within this most recent chapter of media history that we are, for better or worse, a part of. To this list of authors, many more names could be added and certainly *will* be added in the future. For my part, for instance, I cannot stop thinking about Villem Flusser’s notion of the “techno

imaginary” (FLUSSER 2006 [1983]: 88) or the “technical image” (FLUSSER 2011 [1985]: 10); ideas that seemed so fascinating and strange decades ago, but which seem to capture so perfectly the ‘platform ready’ formats, labels, and metadata of this new pictoriality, and the latent ‘bounded space’ of pictorial possibilities (cf. SALVAGGIO 2022b for a similar reading). “The difference between traditional and technical images, then, would be this: the first are observations of objects, the second computations of concepts” (FLUSSER 2011 [1985]: 10). It will be up for debate whether such media theoretical thoughts – developed in this case on and about photography, not AI imagery, to be sure – can still contribute to our understanding of these emerging image technologies.

Categorical Differences to ‘Analogue’ Imagery and Earlier ‘Machine Vision’?

If there is indeed a categorical difference of generative imagery, our task goes beyond highlighting continuities and connections. Half a year after Manovich’s first note about the “return of the art of the copy” he remarked in a new post, with respect to new generative platforms in general, that “another new media is emerging in front of our eyes” (MANOVICH 2023b: n. pag., cf. 2023a). Could we indeed conceptualize generative imagery as such a new media form, perhaps comparable to photography, film, radio, or computer games? Or, more modestly, could we at least uphold that AI imagery constitutes this new paradigm of image production under discussion? A few common strands running through the contributions in this collection indeed indicate such a shift. They might help us to identify and conceptualize salient categorical differences to earlier forms of imagery. I want to focus on three, specifically: 1) generative imagery’s emergent or stochastic features, 2) two interrelated, but often competing entanglements of immediacy-oriented and hypermediacy-oriented forms of realisms, and 3) a new text-image-relation built on the approximation of ‘natural’, meaning here *human* rather than machine code-based language.

First, the most obvious point to be made here is that generative imagery has *emergent* features: the ‘decisions’ of the respective platforms are neither reducible to the programmers, nor to a stable code. Technologically, the more fundamental distinction here is related to the difference between symbolic vs. subsymbolic AI, or between atomistic vs. holistic operating principles, as Bajohr (2021: 25) has reconstructed in a useful survey. Artificial neural networks do “not contain any explicit knowledge”. “[A] neural network does not follow the paradigm of logical deduction or explicitly stated rules that are executed sequentially; rather, it operates by statistical induction, and it is the system as a whole that does the computing” (BAJOHR 2021: 26). One of the consequences from that is that a user

can produce potentially infinite variations of imagery through the same prompt used multiple times while the exact workings of the algorithms remain as much a black box phenomenon to them as to the developers themselves. Alternative terms proposed for generative imagery are thus *stochastic*, *statistic*, or *probabilistic* images (cf. SCHRÖTER 2023).

For a humanities-based approach, it is also important to note that such technological aspects of probabilistic image production are not necessarily visible with the resulting artifacts – especially not if and when they are further distributed and recontextualized from the platforms where they originated. Within a DALL·E, Stable Diffusion, or Midjourney output interface, we can immediately see that every individual picture is only one prompt result out of a range of perhaps four or more alternatives. The algorithmic ‘blackbox’ is part of their mediality. As with many other technologies before them, we can recognize it especially when it is not functioning ‘properly’. For generative imagery, this collapse of transparency has accumulated a range of recognizable markers, the most prominent one probably a wrong number of fingers, as Amanda Wasielewski (2023) discusses in detail in her present contribution. An especially revealing meme circulating on Facebook, Reddit, and Twitter in February 2023, jokingly presented the synthetic prop of a ‘sixth finger’ attachable to a “criminal’s” hand (cf. fig. 1). If photographed, the caption mocked, the picture would be mistaken for an AI image and thus become “inadmissible as evidence”. The widely shared meme thus reverts the intermedial relationship that we easily mistake generative imagery for photographic one these days. The ‘glitch’ of the sixth finger thus functions as a (humorous) intermedial index, highlighting a salient difference between both media and image forms that is normally invisible. Crucially, however, two different *forms of realism* are interwoven or interlaced here, and this points to a second categorical difference of generative imagery to earlier picture media.

Not only can generative imagery masquerade a non-existing person for an existing one, but their representations as an (absent) media form – such as photography. Frequently, it is not the ‘content’ of a DALL·E, Stable Diffusion, or Midjourney picture that is mistaken for a mediated ‘slice of reality’, but its mode of representation itself. Generative imagery, as Jay D. Bolter (2023) elaborates in his present contribution, does indeed continuously simulate or remediate earlier media and image technologies and techniques by creating not only simulations of ‘photos’ but also of ‘oil paintings’, or other established media and image forms like line drawing, woodcuts, comic book covers, graffiti, medical imagery, as well as earlier computer graphics. A media analytical perspective that I am currently developing together with Jan-Noël Thon would thus focus on two connected, sometimes interlaced, but often competing forms of realism. Evolving theoretical conceptualizations and popular notions of realism have been central

to media history and theory, especially with regard to digital media (cf. WANG/DOUBE 2011; GIRALT 2017; MIHAILOVA 2019). Digital media forms not only perpetuate and simulate conceptualizations of realism that are connected to previous ‘analogue’ media forms but reconfigure them into new forms, which sometimes highlight, sometimes hide their digital mediality. This is obviously far from new, either: More than 20 years ago, Jay David Bolter and Richard Grusin (2002) characterized the continuous “remediation” of older media forms into newer ones, especially within digital media landscapes, as a continuous dialectic between the logic of *immediacy* and *hypermediacy*. Generative imagery now arguably employs this dialectic in a perhaps new, media-specific – or at least recognizable – fashion to reconfigure the relationship between human knowledge and communication and what is perceived as physical and social reality.

Ring-Finger-Ring



Figure 1: A widely shared meme circulating on Facebook, Reddit, and Twitter in February 2023, Dan 2023

On the one hand, many contemporary sources indeed express a growing unease that something fundamental is about to change with regard to the human relationship to reality, going back perhaps to Phillip Wang’s online exhibition “This Person Does Not Exist” from 2018 (showcasing a series of portraits created entirely by machine learning).⁴ As many of our contributors (especially Ervik) address, the popular resource of the *DALL·E 2 Prompt Book*, too, opens its introduction with the statement that “nothing you are about to see is real”. All the images shown are “photos that are not real photos”, “paintings that are not real paintings and people, places and things that do not exist” (*DALL·ERY GALL·ERY* 2022: 2). A headline of a 2020 *New York Times* article on generative imagery already suggested that these images were “designed to deceive” (*HILL/WHITE* 2020: n.pag.). Such problems attributed to digital imagery are arguably further complicated by the post-truth discourses surrounding ‘deep fake’

4 <https://thispersondoesnotexist.com/> [accessed March 10, 2023].

technologies (cf. DAGAR/VISHWAKARMA 2022). Bolter and Grusin described immediacy somewhat differently as the appearance of “a transparent interface [...], one that erases itself, so that the user would no longer be aware of confronting a medium” (BOLTER/GRUSIN 2002: 318). One of the oldest aspirations of (digital) media – but still highly relevant today – is indeed a specific form of immediacy typically achieved through photorealism or visual verisimilitude. A key term here is “perceptual realism” which was introduced by Steven Prince (1996) to describe the aesthetic appearance of realism without the concept of indexicality. This is obviously what is at stake here when generative imagery creates digital artifacts that are increasingly able to pass as photographs. In February 2023, for example, the artist Jos Avery ‘came clean’ and ‘confessed’ to his 26,000 Instagram followers that a series of photographic ‘portraits’ he had published on his account were in fact generated by Midjourney and then edited with Photoshop (cf. fig. 2). To his account, he first wanted to fool the public intentionally, then reconsidered in order to reveal the AI production as indeed a new sort of artistic technique (cf. EDWARDS 2023). As the wide press coverage surrounding Avery’s confession indicates, generative imagery seems in fact able to achieve a level of immediacy that can become a problem. This is certainly a matter of honesty or transparency about the process itself, but also a matter of (un)conventionalized degrees of digital manipulation. We expect photographs like Avery’s to be digitally edited through software such as Photoshop without specific notice, so *some* sort of digital mediation is acceptable while others are not – if it is not made transparent.



Figure 2: Jos Avery’s ‘photographic portraits’, revealed to be created through Midjourney, from Edwards 2023^{5]}

5 Cf. Jos Avery’s Instagram-profile <https://www.instagram.com/p/CiirUY8O3Bu/?hl=de> [accessed March 23, 2023]

On the other hand, most of the textual prompts presented within popular resources such as the *DALL-E 2 Prompt Book* focus on earlier image and media techniques, styles, and technologies that do *not* strive for immediacy-oriented realism. We could thus speak of a hypermediacy-oriented realism or a stylistic realism. Hypermediacy “strives to make the viewer acknowledge the medium as a medium and indeed delight in that acknowledgment” (BOLTER/GRUSIN 2002: 335). This acknowledgment is further highlighted by the fact that many picture posts generated through DALL-E, Stable Diffusion, or Midjourney and shared via social media platforms like Facebook, Twitter, and Instagram often advertise their AI generation, either by revealing and discussing the linguistic input prompts or by concealing them like a well-protected, enigmatic ‘magic spell’ (cf. FEYERSINGER et al. 2023). All these use cases highlight the specific part of their mediality related to their AI production. It could even be argued, as Meyer (2023c) does in his contribution, that the immediacy-oriented realism associated with photography has become nothing but one among countless ‘styles’ within an overarching paradigm of hypermediacy-oriented realism. Meyer elaborates on the huge ramifications this has on the notion of pictorial *style* in general which, under this new paradigm, entails a radical expansion and de-hierarchization: “Style can refer to the classical art historical sense of an epochal style or the individual style of a canonized creator, 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” (MEYER 2023c: 106). ‘Style’ now entails people, media, genres, techniques, formats, places, and historical periods, all turned into visual patterns ready to be reproduced and mixed. All visual and formal aspects of a picture can become such a ‘style’ now on all levels of abstraction – and “the entire web [...] a freely available resource that can be mined at scale” (MEYER 2023c: 99).

A specific interrelation of and a conceptual distinction between immediacy-oriented realism and hypermediacy-oriented realism might nevertheless remind us that the remediation of styles is far from ‘evenly distributed’ across communicative contexts. Fabian Offert’s (2023) contribution highlights that differences in immediacy-oriented realism and hypermediacy-oriented realisms might even constitute a novel sort of syntax vs. semantics of generative imagery. Generative imagery should not only be *criticized* for its underlying biases, ideologies, and stereotypes (cf. SALVAGGIO 2023a) but can also be used as a new, technology-guided *access* to the collective cultural imaginary, as Ervik (2023) already suggests. Offert employs DALL-E to produce striking evidence for the fundamental mediatedness of (parts of) our cultural imagination – especially where it concerns terms and concepts connected to historicity: Prompts like “fascism”, he shows, will almost inevitably be remediated in early Kodachrome aesthetics by DALL-E, even if not explicitly demanded. “And it turns out that

it is hard to get rid of, too [...]. There exists, in other words, a strong default in models like DALL·E that conjoins historical periods and historical media and thus produces a (visual) world in which fascism can simply not return because it is safely confined to a black-and-white media prison” (OFFERT 2023: 120). A specific preference for hypermediacy-oriented realism will thus not be up to the individual users (or programmers, for that matter), but engrained in our cultural imaginary – and within the way technological models like CLIP currently work. Whether generative imagery can thus also serve as a powerful tool to reveal and expose this implicit, ideological ‘remediation grammar’ of the cultural imagination or whether these technologies merely perpetuate and reinforce them (for instance through additional filter and censoring mechanisms, as Offert observes), will remain open for discussion.

All of this seems to embed generative imagery deeply into the history and evolution of earlier forms of computer-generated imagery (CGI). In fact, however, many contributions in the present collection point out how different DALL·E, Stable Diffusion, Midjourney, and the images they produce are from earlier computer-generated graphics. Ervik captures this with reference to Alexander Galloway’s “gnostic” view of a 3D CGI simulation, “promising immediate knowledge of all things at all times from all places” (GALLOWAY 2021: 59). Generative imagery, in contrast, offers something else entirely, since even an image generated from the prompt “3D render” does *not* rely on such a model and neither does the platform generate or work with one. The path from linguistic prompt to a flat surface output leads not through simulated 3D space, but through a multi-dimensional latent space of linguistic categories. The results are fundamentally flat surface appearances of visual, not optical patterns. As Meyer (2023c) again points out, even parameters of technical specification (such as “wide angle lens” or “Sigma 24mm f/8”) do not feed into an optical simulation of a photographic apparatus – they function as mere keywords correlating with recurring visual patterns, entirely like generic quality statements such as “perfect” or “award-winning”. In other words, all generative imagery is modeled entirely after and intended for *human* language users. They rely on verbalized semantics to navigate the space of all potential images in recursive iterations (“narrowing down selections in a space of possibilities not yet realized”, MEYER 2023c: 103). Humans also remain paramount for the production of generative imagery at the moment which is based on the still mostly manual labor of indexing, captioning, and ‘cleaning’ the visual data (cf. WILLIAMS et al. 2022). Importantly, prompt-to-image generation is only one aspect of generative imagery and there is also image-to-image generation or techniques like ‘outpainting’ that do not necessarily require linguistic input. Nevertheless, the generation relies on the multi-dimensional vector space of NLP (natural language processing) modeled after human language use. In other words, the current working mechanisms of

generative platforms seems to turn language prompts and verbalized semantics always back into “signs close to perception” (cf. SACHS-HOMBACH 2011) – an emphatically human perception, because this is what the language models are built from and after. In practical uses of generative imagery, this is not a one-way street from text to image, however: Erwin Feyersinger, Lukas Kohmann, and Michael Pelzer (2023) point out in their contribution how DALL·E, Stable Diffusion, or Midjourney can also be used as tools to work on the conceptual level, “to brainstorm, prototype, and refine visual ideas as well as conceptual and stylistic approaches to a given topic or idea” (FEYERSINGER et al. 2023: 143). All of this seems to indicate that generative imagery occupies a rather novel multimodal position continuously oscillating *between* linguistic and pictorial forms of expressions – both, however, firmly revolving around the approximation (and, sometimes, a surprising subversion) of *human* semantics as well as human aesthetics.

An Emerging Field of Research for an Emerging Media Form?

All of this only points to the fact that, despite the prevalent notion of a supposed ‘AI autonomy’, many of the problems and questions surrounding generative imagery that emerged in the second half of 2022 are eminently centered around human and social concerns. These include, but are not limited to, the ‘invisible’ labor of workers especially from the Global South responsible for identifying, cropping, indexing, and labeling images for minimum wages (cf. GRAY/SURI 2019, or for generative AI WILLIAMS et al. 2022), ‘cleaning’ the data by classifying examples of violence, hate speech, or sexual abuse (cf. PERRIGO 2023), as well as supplying private data themselves (cf. EDWARDS 2022). Despite all precautions, the available samples on which generative platforms draw have been shown to contain misogyny, pornography, and harmful stereotypes as well as countless examples of violent, racist, and sexist imagery and text description biases, especially with respect to Black, Asian, or otherwise marginalized women (cf. BIRHANE et al. 2021; OFFERT/PHAN 2022). AI-generated imagery is already used to generate ‘hyper-realistic’ police sketches of suspects (cf. XIANG 2023). The datafication of embedded social, racial, and gender biases perpetuates them in a framework of perceptual realism that hides its constructedness within an “illusion of ‘neutral’ and unbiased technologies which is still prevalent in the discourse around these tools” (SALVAGGIO 2023a: 96). In contradistinction, the perhaps most visible controversies and concerns surrounding generative imagery are centered around plagiarism and the theft of intellectual property (cf. MAZZONE/ELGAMMAL 2019; SOMEPALLI et al. 2022), as well as the exploitation of the labor of artists whose works the algorithms are trained on (cf. BENZINE 2022). While fan cultures have by and large celebrated the emerging possibilities to produce

creative artworks and remix existing styles into new image forms, huge parts of the artistic community have adopted an openly dismissive stance towards generative imagery (cf. DORSEN 2022). As is perhaps hardly surprising, there are also countless platforms for generative pornography on the web⁶ and the use of AI-based imagery for political propaganda is exploding. Politicians of Germany's far right AfD party, for instance, posted imagery of alleged refugees on Facebook with hateful, manic facial expressions (cf. fig. 3). Despite the obvious lack of quality or care within these fakes – the wrong number of fingers, once again – countless readers in the comments reply with agitated remarks (e.g., “Omg, how they even look 🤡👎”, “All this hatred in their faces!”, both quoted from KLEINWÄCHTER 2023, translations mine). Scrolling through accounts (like Norbert Kleinwächter's quoted here), one currently finds generative imagery in almost every new post – although, interestingly, not too often aiming for an immediacy-oriented realism like in figure 3, but more often hypermediacy-oriented (highly ‘stylized’).



Figure 3: “No to even more refugees”: Generative imagery from Germany's extreme right as hate mongering propoganda, Kleinwächter 2023

Several scholars thus argue for the urgent need for ethical and political discussions surrounding generative technologies that are built on enormous amounts of visual data and meta-data (cf. MATZNER 2018; ASHOK et al. 2022; KIESLICH et al. 2022). For the humanities, it will become ever more important to follow up

6 I am not posting the websites here. They can be easily found through a Google search, however.

on these technological developments and to generate an expanding understanding of the distribution of mediated and mediating agency between human and non-human (technological as well as institutional) actors: “In generating images, agency is shared between the prompting user, the platform holders, and the AI. Users write prompts that trigger and steer the diffusion process of AI towards actualizing the possibilities of the latent space. Platform holders can both exclude certain terms and add others without user knowledge” (ERVIK 2023: 52). Salvaggio (2023a) reconstructs in detail how *some* of the parameters limiting or directing user agency are obvious and remain visible (restriction like content policies preventing certain prompts, cf. also OFFERT 2023), others are not – as when words are covertly added into user prompts to diversify image results (cf. OFFERT/PHAN 2022). What media studies could offer here is addressing generative imagery not as a distinct technology, but as a (partially novel) form of *mediation* in a communicative-semiotic, material-technological as well as social-institutional sense. As Richard Grusin put it: “[M]ediation should be understood not as standing between preformed subjects, objects, actants, or entities but as the process, action, or event that generates or provides the conditions for the emergence of subjects and objects, for the individuation of entities within the world” (GRUSIN 2015: 129; cf. MITCHELL/HANSEN 2010; KEMBER/ZYLINSKA 2012).

Bolter presents some perspectives for addressing generative imagery as a “medium” in this sense: “[N]ot just the prompt itself but the whole process of creating the model and producing images would constitute the medium” (BOLTER 2023: 199); it would thus also entail “the database, model, and algorithms behind systems like DALL-E 2 [as] constituents of a new medium” (199) just as “a step-by-step process by a team of programmers and an anonymous crowd of image taggers” (199). In media studies, terms like ‘assemblages’, ‘networks’, or ‘dispositifs’ have been proposed for such interconnected configurations (cf. JUNG et al. 2021), “heterogeneous totalit[ies] that potentially include everything imaginable, whether linguistic or non-linguistic: discourses, institutions, buildings, laws, policing measures, philosophical tenets, etc. The dispositif itself is the network that can be created between these elements” (AGAMBEN 2008: 9, my translation). Respective approaches to mediated and mediating agency have first been developed within actor-network-theory, science and technology studies (STS), and interface studies. In recent years, they have been further developed into a refined theoretical project that is pursued under the header of “actor-media-theory” (cf. SCHÜTTPELZ 2013; KRIEGER/BELLIGER 2014; SPÖHRER/OCHSNER 2017). From this perspective, images can no longer be understood as distinct (material *or* digital) artifacts, but instead appear as networked interfaces between human and non-human actors (including platforms, databases, and corporations) within heterogeneous dispositifs, assemblages, or socio-technological configurations (cf. MACKENZIE/MUNSTER 2019). Feyersinger’s, Kohmann’s,

and Pelzer’s (2023: 143) perspectives on generative imagery as “an accelerated form of externalized visual thinking” seem of special importance here as they conceptualize DALL·E, Stable Diffusion, or Midjourney not from resulting pictorial artifacts, but from affordances provided in iterative interactions. Lamerichs also remarks that “AI art is not an outcome but a process or a performance. It is best understood as the interplay of different agencies and a way of collaborating” (LAMERICHS 2023: 154). Understanding generative imagery not merely as a technology, but as a media form, would then not at all depend on technological, but on cultural developments and praxeological questions. Or, as Jens Schröter (2008; 2011) put it: ‘Media’ are always discursively ‘singled out’ out of technical procedures, institutions, programs, formal strategies, author figures, practices, etc. according to specific strategic purposes. The “arch-intermedial network” (SCHRÖTER 2008: 579, my translation), the discursive “intermedial field” (SCHRÖTER 2011: 1) remains especially visible when there are no conventionalized practices, no established use cases, no “cultural protocols” (GITELMAN 2008: 5) in place yet – which is arguably where we are with generative imagery in the spring of 2023. It will thus be important to trace and map how generative imagery is conceptualized, attributed, negotiated, and commodified in specific sociocultural contexts, such as art, fan culture, news media, the sciences, etc. Concerning protocols of typical usage as “normative rules and default conditions, which gather and adhere like a nebulous array around a [...] nucleus” (GITELMAN 2008: 7), two developments seem equally possible at present, and probably they are not mutually exclusive.



Figure 4: Kris Kashtanova’s Midjourney-Comic *Zarya of the Dawn*, Kashtanova 2022

On the one hand, it stands to reason that the recognizable image or media forms (and aesthetics) remediated by generative imagery (from oil paintings over photographic portraits to drawn fan artworks) might carry with them and thus

recontextualize (but also transform) the cultural protocols and conventionalized practices of typical production, distribution, and reception, as well as the ascribed assumptions about their cultural values. The question is where, when and by whom – in what situations – are these image media usually employed; and will generative imagery ‘fill’ these spaces, if only through their economic accessibility? Will generative imagery thus be integrated or ‘absorbed’ within other media forms such as films, television shows, comic books, or video games or ‘stand out’ as another (marked) intermedial reference? Who will abstain from using them in which socio-cultural contexts? That the use cases and concerns surrounding generative imagery are and will be entirely different ones across socio-cultural fields and discourse strands is something Konstantinos Michos (2023) reminds us of in his contribution: for academic research and science communication, the ‘blackboxes’ of image generation and the stochastic nature of their results can generate serious concerns where absolute precision is of eminent importance. All of these are fundamentally praxeological issues, some of which are pointed out by Feyersinger, Kohmann, and Pelzer. At the current moment, partially AI-generated works like Jason M. Allen’s award-winning artwork “Théâtre D’opéra Spatial” (created through Midjourney, cf. ROOSE 2022), comic books like Kris Kashtanova’s *Zarya of the Dawn* (with images created through Midjourney, cf. KASHTANOVA 2022; FOLEY 2022; cf. fig. 4), a Netflix animated film like *The Dog and the Boy* by director Ryotaro Makihara (with background images by an undisclosed generative platform, cf. DEIKOVA 2023) or Boris Eldagsen’s Sony World Photography Awards 2023-winning ‘synthetical’ image *The Electrician* (created with two undisclosed generative platform, cf. EL DAGSEN 2023) attract wide press coverage and generate heated controversies precisely for the fact *that* generative imagery is employed within these established media forms and that they are thus not (yet) seamlessly integrated into more conventionalized forms of imagery and their uses.

On the other hand, generative imagery might accumulate protocols, practices, conventions, and finally institutions of ‘their own’, for instance by providing the distinct media practice of ‘prompt engineering’ or of the still contested social role of an ‘AI artist’ (cf., for instance, DONNELLY 2023). Scorzin points to many artistic practices and experiments in which the ‘generativity’ of the imagery is key to any artistic statement or provocation; from her observations, one could, perhaps, almost speak of an emerging tradition. Even beyond the confines of the ‘art world’, however, generative imagery *could* be recognized as a distinct media form – just as photography can be art, without exhausting itself in that function. For this question, it is also extremely relevant whether or not a recognizable ‘AI aesthetics’ is emerging across and despite the range of all possible stylistic remediations. Roland Meyer just diagnosed a “midjourneyfication” (MEYER 2023a: n.pag.) of DALL·E’s newest March 2023 update, addressing a specific, strongly conventionalized style that the artist Nils Pooker (2023: n.pag.)

described as a “fluffy glamour glow” after his beta test (cf. fig. 5). This aesthetics, alongside a recognizable color scheme (“Teal and Orange”), would not be strictly technologically determined, but become increasingly dominant due to a complicated concoction of recurring user preferences, commercial restraints, and most importantly the relevance of the amateur art exchange platform DeviantArt for the underlying training dataset. If this is true, then generative imagery is already consolidating into a distinct node within the intermedial field, ready to accumulate conventionalized practices, cultural values, and sociocultural roles together with its conventionalized aesthetics. During our first workshop in February 2023, it certainly felt as if we were witnessing the ‘Vaudeville days’ of generative imagery, comparable perhaps to the early days of cinema when institutions, studios, and professional roles – the protocols of production, distribution, and reception – where not yet established. And certainly, the companies responsible for generative platforms up to this point are still mostly startups – even OpenAI (DALL-E) has not even a thousand employees at this point. The technological and most certainly also socio-cultural developments will continue to progress rapidly now that ‘big players’ like Microsoft, Google, or Meta are about to enter into the generative AI business. The Vaudeville days might be over soon.



Figure 5: “Fluffy Glowing Cute Teal and Orange Vibe” as an increasingly conventionalized ‘AI aesthetics’, generated by Meyer (2023b) with Midjourney, March 2023

Perhaps, however, cinema is the wrong analogy, to begin with. An alternative comparison to conceptualize generative imagery might be provided by *animation* which retained a much more complicated and tense relationship to media theories and popular conceptions of mediality. Animation was never fully accepted as a ‘distinct’ media form but often misunderstood as a filmic genre among others. Currently, animation is increasingly recognized as a completely trans-medial technique (which we also find in video games or in digital interfaces, for instance) or even as an umbrella term for *cinematically* generated illusions of

movement in which ‘live action’ would then just be one specific form of animation (cf. MANOVICH 2001). Generative imagery could retain such a medial ambiguity just as well, and perhaps its complicated entanglement of immediacy-oriented and hypermediacy-oriented realisms make it especially suited for that. Only time and future media history will tell. Once again, it will be important for the humanities to trace how generative imagery is conceptualized, attributed, negotiated, and commodified in different sociocultural contexts, perhaps understood as discourse strands. At the moment, the most prominent ‘use cases’ can certainly be found within fan cultures, attributing special importance to research from fan studies represented by Nicolle Lamerichs’ (2023) survey in the present collection. If media studies want to provide a critical framework for these ongoing discussions – whatever that might look like – it seems clear to me that this must include both a deepened knowledge about the technological workings behind the ‘interface blackboxes’ (how CLIP and GLIDE *actually work*, for instance, cf. especially BAJOHR 2023; SALVAGGIO 2023a), just as a critical reflection of the emerging cultural, social, and economic uses – the practices and conventions that transform technologies into media forms – which might be evolving at a much quicker speed now than in earlier moments of media transformations (cf. WILDE 2023). In any case, this certainly requires a joint effort from and between experts from all disciplines in the humanities concerned with pictures, pictoriality, and visual communication – from media studies and communication studies to art history, design, multimodal linguistics, media sociology, and media anthropology, to name just a few. No less importantly, though, it will require a dialogue with the technical and social sciences, specifically with colleagues from science and technology studies (STS) and computer sciences.

Not surprisingly, the emergence of a “Critical AI Studies” (ROBERGE/CASTELLE 2021) is already discussed as “a field in formation” (RALEY/RHEE 2023: 188). Generative imagery constitutes only a small part of these developments, and, certainly, current multimodal distinctions will grow together rapidly: While the first version of ChatGPT was strictly limited to textual inputs and outputs, the new iteration 4 can interpret pictures. As prompt-to-text technologies make their APIs interfaces available (cf. BROCKMAN et al. 2023), the multimodality of AI platforms will progress rapidly, too. Nevertheless, the cultural distinctness of AI imagery as a media form will hardly depend on such technological factors. In our ‘postdigital’ media ecologies, all media differences could be said to be mere interface effects – based on the same digital infrastructures and hardware – for decades already (cf. HOOKWAY 2014). Again, far from everything seems new in that respect. ‘Critical AI Studies’ might thus develop coexistently with a more specialized field interested in this new paradigm of imagery. Research into this also calls for collaboration with artists, computer designers, and other practitioners. Most importantly, it will be crucial to create an inclusive and diverse

exchange of research and perspectives, especially for concerns and emerging technologies that are dispersed globally across languages and cultures. It is all the more unfortunate that the group of scholars represented in this collection is, despite our best intentions as organizers and editors, overwhelmingly male and especially *white*. The idea for our gathering started as a small, local workshop and we were overwhelmed by the large number of registered online participants from every continent around the globe. This cannot serve as an excuse for the actual line-up of presenters and authors presented here, though, so we certainly need to do better. This will not only be important for future workshops, conferences, and publications on generative imagery, but also with respect to our bibliographies if we do not want to write the history of yet another medium as a male, eurocentric one. Probably there will be many opportunities to do so. It seems likely that generative imagery is going to stay.

Acknowledgments

Many of the present ideas and observations have been developed together with Jan-Noël Thon. I would also like to thank Tolmie McRae, Marcel Lemmes, and Erwin Feyersinger for important remarks and corrections.

Bibliography

- ADOBE: Adobe Unveils Firefly, a Family of new Creative Generative AI. In: *Adobe News*. March 23, 2023. <https://news.adobe.com/news/news-details/2023/Adobe-Unveils-Firefly-a-Family-of-new-Creative-Generative-AI/default.aspx> [accessed March 23, 2023]
- AGAMBEN, GIORGIO: *Was ist ein Dispositiv?* Translated by Andreas Hiepko. Zürich [Diaphanes] 2008
- ALFARAJ, ABDULLAH: Auto-Photoshop-StableDiffusion-Plugin. In: *GitHub*. 2023. <https://github.com/AbdullahAlfaraj/Auto-Photoshop-StableDiffusion-Plugin> [accessed March 23, 2023]
- ASHOK, MONA; ROHIT MADAN; ANTON JOHA; UTHAYASANKAR SIVARAJAH: Ethical Framework for Artificial Intelligence and Digital Technologies. In: *International Journal of Information Management*, 62, February 2022. <https://doi.org/10.1016/j.ijinfomgt.2021.102433> [accessed April 30, 2023]
- BAJOHR, HANNES: The Gestalt of AI: Beyond the Holism-Atomism Divide. In: *Interface Critique*, 3, 2021, pp. 13-35
- BAJOHR, HANNES: Dumb Meaning: Machine Learning and Artificial Semantics. 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. 58-70

- BENJAMIN, WALTER:** *The Work of Art in the Age of its Technological Reproducibility*. Second Version. Translated by Edmund Jephcott, Rodney Livingstone, Howard Eiland, and others. In: *The Work of Art in the Age of its Technological Reproducibility, and Other Writings on Media*. Cambridge, MA [Belknap Press] 2008 [1935], pp. 19-55
- BENZINE, VITTORIA:** 'A.I. Should Exclude Living Artists from its Database,' Says One Painter whose Works were Used to Fuel Image Generators. In: *Artnet*. September 20, 2022. <https://news.artnet.com/art-world/a-i-should-exclude-living-artists-from-its-database-says-one-painter-whose-works-were-used-to-fuel-image-generators-2178352> [accessed March 23, 2023]
- BIRHANE, ABEBA; VINAY UDAY PRABHU; EMMANUEL KAHEMBWE:** Multimodal Datasets: Misogyny, Pornography, and Malignant Stereotypes. *arXiv:2110.01963*. October 5, 2021. <https://arxiv.org/abs/2110.01963> [accessed March 23, 2023]
- BOLTER, JAY D.:** AI Generative Art as Algorithmic Remediation. 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. 195-207
- BOLTER, JAY D.; RICHARD GRUSIN:** *Remediation: Understanding New Media*. Cambridge, MA [MIT Press] 2002
- BROCKMAN, GREG; ATTY ELETI; ELIE GEORGES; JOANNE JANG; LOGAN KILPATRICK; RACHEL LIM; LUKE MILLER; MICHELLE POKRASS:** Introducing ChatGPT and Whisper APIs. In: *OpenAI Blog*. March 1, 2023. <https://openai.com/blog/introducing-chatgpt-and-whisper-apis> [accessed March 23, 2023]
- DAGAR, DEEPAK; DINESH KUMAR VISHWAKARMA:** A Literature Review and Perspectives in Deepfakes: Generation, Detection, and Applications. In: *International Journal of Multimedia Information Retrieval*, 11, 2022, pp. 219-289
- DAN (@bristowbailey):** Criminals will Start Wearing Extra Prosthetic Fingers ... Tweet on *Twitter*. February 13, 2023. <https://twitter.com/bristowbailey/status/1625165718340640769?s=20> [accessed March 23, 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 March 23, 2023]
- DEIKOVA, MASCHA:** Netflix Uses AI to Generate Anime Short Film – Reactions Follow. In: *CineD*. February 6, 2023. <https://www.cined.com/netflix-uses-AI-to-generate-anime-short-film-reactions-follow> [accessed March 23, 2023]
- DOBSON, JAMES E.:** *The Birth of Computer Vision*. Minneapolis [University of Minnesota Press] 2023
- DONNELLY, MATT:** WME Signs AI Artist Claire Silver. In: *Variety*, March 6, 2023. <https://variety.com/2023/digital/news/wme-signs-AI-artist-claire-silver-louvre-1235544502/> [accessed March 23, 2023]

- DORSEN, ANNIE: AI is Plundering the Imagination and Replacing it with a Slot Machine. In: *Bulletin of the Atomic Scientist*. October 27, 2022. <https://thebulletin.org/2022/10/AI-is-plundering-the-imagination-and-replacing-it-with-a-slot-machine/> [accessed March 23, 2023]
- EDWARDS, BENJ: Artist Finds Private Medical Record Photos in Popular AI Training Data Set. In: *Ars Technica*. September 21, 2022. <https://arstechnica.com/information-technology/2022/09/artist-finds-private-medical-record-photos-in-popular-AI-training-data-set> [accessed March 23, 2023]
- EDWARDS, BENJ: Viral Instagram Photographer has a Confession: His Photos are AI-Generated. In: *Ars Technica*. February 21, 2023. <https://arstechnica.com/information-technology/2023/02/viral-instagram-photographer-has-a-confession-his-photos-are-AI-generated/> [accessed March 23, 2023]
- ELDAGSEN, BORIS: Sony World Photography Awards 2023. In: *Boris Eldagsen*. March 14, 2023. <https://www.eldagsen.com/sony-world-photography-awards-2023/> [accessed March 23, 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
- FEYERSINGER, ERWIN; LUKAS KOHMANN; MICHAEL PELZER: Fuzzy Ingenuity: Creative Potentials and Mechanics of Fuzziness in Processes of Image Creation with Text-to-Image Generators. 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. 135-149
- FLUDERNIK, MONIKA: *Towards a 'Natural' Narratology*. London [Routledge] 1996
- FLUSSER, VILÉM: *Towards a Philosophy of Photography*. Translated by Anthony Mathews. London [Reaktion Books] 2006 [1983]
- FLUSSER, VILÉM: *Into the Universe of Technical Images*. Translated by Nancy Ann Roth. Minneapolis [University of Minnesota Press] 2011 [1985]
- FOLEY, JOSEPH: The First Copyrighted AI Art Looks Uncannily like Zendaya. In: *Creative Bloq*. October 4, 2022. <https://www.creativebloq.com/news/AI-art-copyright> [accessed March 23, 2023]
- GALLOWAY, ALEXANDER: *Uncomputable: Play and Politics in the Long Digital Age*. London [Verso] 2021
- GIRALT, GABRIEL F.: The Interchangeability of vfx and Live Action and its Implications for Realism. In: *Journal of Film and Video*, 69(1), 2017, pp. 3-17
- GOOSKENS, GEERT: The Digital Challenge: Photographic Realism Revisited. *Proceedings of the European Society for Aesthetics*, 3, 2011, pp. 115-125
- GRAY, MARY L.; SIDDHART SURI: *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston [Houghton Mifflin Harcourt] 2019
- GRUSIN, RICHARD: Radical Mediation. In: *Critical Inquiry*, 42(1), 2015, pp. 124-148

- HILL, KASHMIR; JEREMY WHITE: Designed to Deceive: Do these People Look Real to You? In: *The New York Times*. November 21, 2020. <https://www.nytimes.com/interactive/2020/11/21/science/artificialintelligence-fake-people-faces.html> [accessed March 23, 2023]
- HOOKEY, BRANDEN: *Interface*. Cambridge, MA [MIT Press] 2014
- JUNG, BERENIKE; KLAUS SACHS-HOMBACH; LUKAS R.A. WILDE: Agency postdigital: Verteilte Handlungsmächte in medienwissenschaftlichen Forschungsfeldern. In: BERENIKE JUNG; KLAUS SACHS-HOMBACH; LUKAS R.A. WILDE (eds.): *Agency postdigital: Verteilte Handlungsmächte in medienwissenschaftlichen Forschungsfeldern*. Cologne [Herbert von Halem] 2021, pp. 7-41
- KASHTANOVA, KRIS: English Version of my Graphic Novel Zarya of the Dawn. Post on *Instagram*. September 23, 2022. <https://www.instagram.com/p/C1rUY8O3Bu/?hl=de> [accessed March 23, 2023]
- KEMBER, SARAH; JOANNA ZYLINSKA: *Life after New Media: Mediation as a Vital Process*. Cambridge, MA [MIT Press] 2012
- KIESLICH, KIMON; BIRTE KELLER; CHRISTOPHER STARKE: Artificial Intelligence Ethics by Design: Evaluating Public Perception on the Importance of Ethical Design Principles of Artificial Intelligence. In: *Big Data & Society*, 9(1), 2022, pp. 1-15
- KLEINWÄCHER, NORBERT: Nein zu noch mehr Flüchtlingen!. Post on *Facebook*. March 21, 2023. <https://www.facebook.com/norbert.kleinwaechter/photos/a.375576792808512/1871862683179908> [accessed March 26, 2023]
- KRIEGER, DAVID J.; ANDRÉA BELLIGER: *Interpreting Networks: Hermeneutics, Actor-Network Theory and New Media*. Bielefeld [transcript] 2014
- LAMERICHS, NICOLLE: Generative AI and the Next Stage of Fan Art. 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. 150-164
- MACKENZIE, ADRIAN; ANNA MUNSTER: Platform Seeing: Image Ensembles and their Invisibilities. In: *Theory, Culture & Society*, 36(5), 2019, pp. 3-22
- MANOVICH, LEV: *The Language of New Media*. Cambridge, MA [MIT Press] 2001
- MANOVICH, LEV: Note on AI Image Synthesis and Return of the Classical Art – ‘Art of a Copy’. Post on *Facebook*. July 20, 2022. <https://www.facebook.com/lev.manovich/posts/pfbido2wXx3qiqherA585LWXnjbRhdzAhGa7vYcqA6r89GsSD5o38VtF6GDn1F1u1Qm83SWl> [accessed March 23, 2023]
- MANOVICH, LEV: AI Image Media through the Lens of Art and Media History. 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), 2023a, pp. 34-41

- MANOVICH, LEV: AI Video Research is Making Quick Progress. Post on *Facebook*. February 7, 2023b. <https://www.facebook.com/lev.manovich/posts/pfbid02YFZrtyBYCkAAGDsno7sPiUoz8AflDgfsfzutmL28WyzCiXfd66EMbUcqXsrd6cXiDl> [accessed March 23, 2023]
- MATZNER, TOBIAS: Grasping the Ethics and Politics of Algorithms. In: ANN RUDINOW SAETNAN; INGRID SCHNEIDER; NICOLA GREEN (eds.): *The Politics of Big Data: Big Data, Big Brother?* London [Routledge] 2018, pp. 39-45
- MAZZONE, MARIAN; AHMED ELGAMMAL: Art, Creativity, and the Potential of Artificial Intelligence. In: *Arts*, 8(1), 2019, 26. <https://www.mdpi.com/2076-0752/8/1/26> [accessed March 23, 2023]
- MEYER, ROLAND: Es schimmert, es glüht, es funkelt – Zur Ästhetik der KI-Bilder. In: *54 Books*. March 20, 2023a. https://www.54books.de/es-schimmert-es-gluieht-es-funkelt-zur-aesthetik-der-ki-bilder/?fbclid=IwAR37Ff7wD8aJTcl0xzDRgxPjKDARACgxVJVdi_OMhrYg-IwIvUX5tVGERc [accessed March 23, 2023]
- MEYER, ROLAND: PPS: May I Introduce: “Fluffy Glowing Cute Teal and Orange Vibe”. Tweet on *Twitter*. March 7, 2023b. <https://twitter.com/bildoperationen/status/1633165036259536922?s=20> [accessed March 23, 2023]
- MEYER, ROLAND: The New Value of the Archive: AI Image Generation and the Visual Economy of ‘Style’. 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), 2023c, pp. 100-111
- MICHOS, KONSTANTINOS: AI in Scientific Imaging: Drawing on Astronomy and Nanotechnology to Illustrate Emerging Concerns About Generative Knowledge. 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. 165-178
- MICROSOFT: Create Images with your Words – Bing Image Creator Comes to the New Bing. In: *Microsoft Blog*. March 21, 2023. <https://blogs.microsoft.com/blog/2023/03/21/create-images-with-your-words-bing-image-creator-comes-to-the-new-bing/> [accessed March 23, 2023]
- MIHAILOVA, MIHAELA: Realism and Animation. In: DOBSON, NICHOLA; ANNABELLE HONESS ROE; AMY RATELLE; CAROLINE RUDELL (eds.): *The Animation Studies Reader*. New York [Bloomsbury Academic] 2019, pp. 47-57
- MILLER, ARTHUR I.: *The Artist in the Machine: The World of AI-Powered Creativity*. Cambridge, MA [MIT Press] 2019
- MITCHELL, MELANIE: *Artificial Intelligence: A Guide for Thinking Human*. New York [Farrar, Straus and Giroux] 2019
- MITCHELL, WILLIAM J.T.: *The Reconfigured Eye: Visual Truth in the Post-Photographic Era*. Cambridge, MA [MIT Press] 1992

- MITCHELL, WILLIAM J.T.; MARK B.N. HANSEN: Introduction. In: WILLIAM J.T. MITCHELL; MARK B.N. HANSEN (eds.): *Critical Terms for Media Studies*. Chicago [Chicago UP] 2010, vii–xxii
- NILSSON, NILS J.: *The Quest for Artificial Intelligence: A History of Ideas and Achievement*. Cambridge [Cambridge University Press] 2010
- OFFERT, FABIAN: Ten Years of Image Synthesis. In: *Zentralwerkstatt*. November 10, 2022. <https://zentralwerkstatt.org/blog/ten-years-of-image-synthesis> [accessed March 23, 2023]
- OFFERT, FABIAN: On the Concept of History (in Foundation Models). 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. 121-134
- OFFERT, FABIA; 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 March 23, 2023]
- OPENAI: GPT-4. In: *OpenAI*. March 14, 2023. <https://openai.com/research/gpt-4> [accessed March 23, 2023]
- PERRIGO, BILLY: Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. In: *Time*. January 18, 2023. <https://time.com/6247678/openai-chatgpt-kenya-workers> [accessed March 23, 2023]
- POOKER, NILS (@pookerman): Die “Midjourneyfizierung” als KI-Trend.... Tweet on *Twitter*. March 12, 2023. <https://twitter.com/pookerman/status/1634954416762814470?s=20> [accessed March 23, 2023]
- PRINCE, STEVEN: True Lies: Perceptual Realism, Digital Images, and Film Theory. In: *Film Quarterly*, 29(3), 1996, pp. 27-37
- RALEY, RITA; JENNIFER RHEE: Critical AI: A Field in Formation. Advance Publication. In: *American Literature*, 95(2), 2023. pp. 185-204. <https://read.dukeupress.edu/american-literature/article/doi/10.1215/00029831-10575021/344223/Critical-AI-A-Field-in-Formation> [accessed March 23, 2023]
- RETTBERG, JILL WALKER: DALL-E and Human-AI Assemblages. In: *jilltxt.net*. June 23, 2022, <https://jilltxt.net/dall-e-and-human-AI-assemblages/> [accessed March 23, 2023]
- RETTBERG, JILL WALKER: *Machine Vision: How Algorithms Are Changing the Way we See the World*. Newark [Polity Press] 2023/forthcoming.
- RICHTER, SEBASTIAN: *Digitaler Realismus: Zwischen Computeranimation und Live-Action*. Bielefeld [Transcript] 2008
- ROBERGE, JONATHAN; MICHAEL CASTELLE (eds.): *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*. Cham [Palgrave Macmillan] 2021
- ROOSE, KEVIN: An A. I.-Generated Picture Won an Art Prize: Artists aren’t Happy. In: *The New York Times*. September 2, 2022. <https://www.nytimes.com>

- com/2022/09/02/technology/AI-artificialintelligence-artists.html [accessed March 23, 2023]
- SACHS-HOMBACH, KLAUS: Theories of Image: Five Tentative Theses. In: JAMES ELKIN; MAJA NAEF (eds.): *What Is an Image? The Stone Art Theory Institutes* Vol. 2. University Park [Pennsylvania State University Press] 2011, pp. 229-232
- SALVAGGIO, ERYK: How to Read an AI Image: Toward a Media Studies Methodology for the Analysis of Synthetic Images. 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), 2023a, pp. 83-99
- SALVAGGIO, ERYK: The Most Generated Barn in America. In: *Cybernetic Forests*. January 8, 2023b. <https://cyberneticforests.substack.com/p/the-most-generated-barn-in-america> [accessed March 23, 2023]
- SCHRÖTER, JENS: Das ur-intermediale Netzwerk und die (Neu-)Erfindung des Mediums im (digitalen) Modernismus: Ein Versuch. In: JOACHIM PAECH; JENS SCHRÖTER (eds.): *Intermedialität – analog/digital: Theorien, Methoden, Analyse*. München [Fink] 2008, pp. 579-601
- SCHRÖTER, JENS: Discourses and Models of Intermediality. In: *CLCWeb: Comparative Literature and Culture*, 13(3), 2011. <http://docs.lib.purdue.edu/clcweb/vol13/iss3/3> [accessed March 23, 2023]
- SCHRÖTER, JENS: The AI Image, the Dream, and the Statistical Unconscious. 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. 112-120
- SCHÜTTPELZ, ERHARD: Elemente einer Akteur-Medien-Theorie. In: TRISTAN THIELMAN; ERHARD SCHÜTTPELZ (eds.): *Akteur-Medien-Theorie*. Bielefeld [transcript] 2013, pp. 9-70
- SCORZIN, PAMELA C.: AI Body Images and the Meta-Human: On the Rise of AI-generated Avatars for Mixed Realities and the Metaverse. 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. 179-194
- SOMEPELLI, GOWTHAMI; VASU SINGLA; MICAH GOLDBLUM; JONAS GEIPING; TOM GOLDSTEIN: Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models. *arXiv:2212.03860*. December 7, 2022. <https://arxiv.org/abs/2212.03860> [accessed March 23, 2023]
- SPÖHRER, MARKUS; BEATE OCHSNER (eds): *Applying the Actor-Network Theory in Media Studies*. Hershey [IGI Global] 2017
- SCHÜWER, MARTIN: *Wie Comics erzählen: Grundriss einer intermedialen Erzähltheorie der grafischen Literatur*. Trier [Wissenschaftlicher Verlag Trier] 2008

- SUDMANN, ANDREAS: On the Media-Political Dimension of Artificial Intelligence: Deep Learning as a Black Box and OpenAI. In: *Digital Culture & Society*, 4(1), 2018a, pp. 181-200.
- SUDMANN, ANDREAS: Szenarien des Postdigitalen: Deep Learning als MedienRevolution. In: CHRISTOPH ENGEMAN; ANDREAS SUDMANN (eds.): *Machine Learning: Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*. Bielefeld [transcript] 2018b, pp. 66-68
- TANGERMANN, VICTOR: Microsoft's Bing AI Is Leaking Maniac Alternate Personalities Named "Venom" and "Fury". In: *Futurism*. December 15, 2023. <https://futurism.com/microsofts-bing-AI-leaking-maniac-alternate-personalities> [accessed March 23, 2023]
- VINCENT, JAMES: Microsoft's Bing is an Emotionally Manipulative Liar, and People Love it. In: *The Verge*. February 15, 2023. <https://www.theverge.com/2023/2/15/23599072/microsoft-AI-bing-personality-conversations-spy-employees-webcams> [accessed March 23, 2023]
- WALTER, RENÉ: Suspension of Disbelief (in Sentient AI). In: *Good Internet*. February 19, 2023. <https://goodinternet.substack.com/p/suspension-of-disbelief-in-sentient> [accessed March 23, 2023]
- WANG, NORMAN; WENDY DOUBE: How Real Is Reality? A Perceptually Motivated System for Quantifying Visual Realism in Digital Images. In: *2011 International Conference on Multimedia and Signal Processing*, 2011, pp. 141-149
- WASIELEWSKI, AMANDA: "Midjourney Can't Count": Questions of Representation and Meaning for Text-to-Image Generators. 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. 71-82
- WILDE, LUKAS R.A.: AI-Bilder und Plattform-Memes: Post-digital, post-artifiziell, post-faktisch? In: *Tübinale: Das studentische Kurzfilmfestival Tübingens*. April 11, 2023. <https://www.tuebinale.de/aibilder-und-plattformen> [accessed April 26, 2023]
- WILLIAMS, ADRIENNA; MILAGROS MICELI; TIMNIT GEBRU: The Exploited Labor Behind Artificial Intelligence. In: *Noema*. October 13, 2022. <https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence> [accessed March 23, 2023]
- XIANG, CHLOE: Developers Created AI to Generate Police Sketches: Experts Are Horrified. In: *Vice*. February 7, 2023. <https://www.vice.com/en/article/qjk745/AI-police-sketches> [accessed March 23, 2023]
- ZHANG, LVMIN; MANEESH AGRAWALA: Adding Conditional Control to Text-to-Image Diffusion Models. *arXiv:2302.05543v1*. February 10, 2023. <https://arxiv.org/abs/2302.05543> [accessed March 23, 2023]