



Generativität



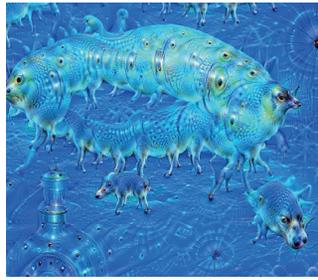
Herausgegeben von
Matthias Bruhn
Katharina Weinstock

München, 2025
Open Publishing LMU

1



2



3



4



5

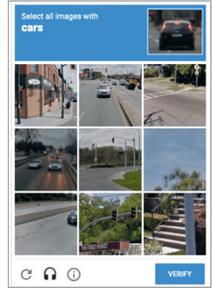


- 1 Frank Rosenblatts Perceptron, ein frühes künstliches neuronales Netzwerk, 1960.
- 2 Die Bilderkennung mit Google DeepDream produziert psychedelische Muster, 2015.
- 3 Diffusionsmodelle generieren Bilder durch das Entfernen von Rauschen, 2022.
- 4 Amnesty International wurde 2023 kritisiert, für die Darstellung von Polizeigewalt in Kolumbien KI-generierte Bilder verwendet zu haben.
- 5 Kompositfotografie der Schlacht bei Zonnebeke (Frank Hurley), 1917.

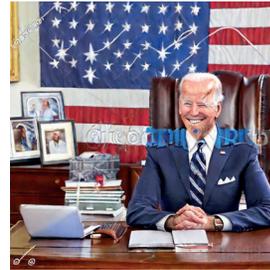
6



7



8



9



10



11



- 6 Eines der frühesten computergenerierten 3D-Modelle, die Karosserie eines VW Käfers, 1972.
- 7 Das Lösen von Captchas trainiert die Bild-Text-Zuordnung von KI-Modellen, 2009.
- 8 Stockfoto-Wasserzeichen in den Outputs generativer KI-Modelle, 2023.
- 9 KI-Modelle können auch Bilder nichtexistenter Objekte generieren, etwa: „an armchair in the shape of an avocado“, 2021.
- 10 Die Figur „Loab“ gilt als erster KI-generierter Kryptid, 2022.
- 11 Werden Modelle anhand von Daten trainiert, die selbst KI-generiert sind, treten in den Outputs charakteristische Artefakte auf.

Inhalt

Editorial <i>Matthias Bruhn, Katharina Weinstock</i>	07
Generative Bilderströme. Über digitale Nach- und Vorbilder <i>Matthias Bruhn</i>	11
Weltmodelle ordnen. Zur politischen Epistemologie der Generativität <i>Yannick Nepomuk Fritz</i>	23
Anti-Fragile Sycophants <i>Adam Harvey</i>	37
From Glitch to Kitsch. Generativität, Algorithmic Folklore und Avantgarde-Ästhetiken <i>Moritz Konrad</i>	41
Plattformrealismus. Eine Ästhetik des Generischen <i>Roland Meyer</i>	55
Those aren't your memories. Geschichte und Erinnerung im KI-generierten Fotorealismus <i>Katharina Weinstock</i>	67
Beyond the Spectacle of "AI": An Emergent Aesthetic Regime <i>Charlotte Kent</i>	79
Liquid Images. Holly Herndon, Mat Dryhurst, and Tyler Hobbs in conversation with Alex Estorick	99

Bildnachweis vorige Doppelseite

- 1 National Museum of the United States Navy, Public Domain.
- 2 Martin Thoma, Wikimedia Commons, Public Domain (Ausschnitt).
- 3 Benlissquare / Wikimedia Commons (CC BY-SA 4.0). Schritte 1, 2, 3, 10 und 40 der Grafik, https://en.wikipedia.org/wiki/File:X-Y_plot_of_algorithmically-generated_AI_art_of_European-style_castle_in_Japan_demonstrating_DDIM_diffusion_steps.png.
- 4 Tweet von Amnesty Norwegen [Amnesty_Norge], 28.04.2023 (Ausschnitt), https://web.archive.org/web/20230501165807/https://twitter.com/Amnesty_Norge/status/1651879572944691201.
- 5 State Library of New South Wales, XV*/Wor W 1/7, Wikimedia Commons, Public Domain (Ausschnitt).
- 6 Ivan Edward Sutherland et al. / Wikimedia Commons (CC BY-SA 4.0). Gespiegelter Ausschnitt des Originals, https://en.wikipedia.org/wiki/File:Utah_VW_Bug.stl.
- 7 Aus: Google reCAPTCHA Dokumentation; <https://cloud.google.com/recaptcha/docs/choose-key-type?hl=de>.
- 8 Aus: Matic Broz: Original Research. I Found More Watermarks in Stable Diffusion's Images. In: Photutorial, 16.12.2024, <https://photutorial.com/stable-diffusion-watermarks-investigation/> [Stand 12/2024].
- 9 Aus: OpenAI: Dall-E. Creating Images from Text, <https://openai.com/index/dall-e/> (Ausschnitt) [Stand 12/2024].
- 10 Tweet von Steph Maj Swanson [supercomposite], 06.09.2022, <https://x.com/supercomposite/status/1567162288087470081>.
- 11 Aus: Preventing AI Model Collapse. Addressing the Inherent Risks of Synthetic Datasets. In: appinventiv, <https://appinventiv.com/blog/ai-model-collapse-prevention/> [Stand 12/2024].

Editorial

*„It can strike a chord inside you
Like a generation's need [...]
We are the voices of the
Big generator”*

(Yes, *Big Generator*, 1987)

Dass das Album *Big Generator* der britischen Progressive-Rock-Band Yes im Jahre 1987 floppte, soll nach Ansicht vieler Fans am Erfolg des Vorgängeralbums gelegen haben; unter hohem Erwartungsdruck habe die Band zwar ein aufwändiges, aber uninspiriertes Selbstplagiat abgeliefert. Mit dem Durchbruch generativer künstlicher Intelligenz (KI) könnte das mit kaltem Perfektionismus produzierte Album (oder zumindest sein Titel) aber eine gewisse Rehabilitierung erfahren. Denn Begriffe wie Generator und Transformator sind nicht nur Leitmetaphern des elektrischen Zeitalters um 1900 gewesen, sondern längst auch zum festen Bestandteil des Vokabulars der generativen KI und der großen Sprachmodelle geworden.

Als neues Massenmedium erschüttern KI-Anwendungen einige Gewissheiten oder Vorstellungen des medialen Originals und der künstlerischen Originalität. An die Stelle der Reproduktion (sei es als Fotografie oder Abformung), durch welche sich die Medientheorie des frühen 20. Jahrhunderts herausgefordert sah, ist eine synthetische Dauerproduktivität getreten, die der KI-Generation sehr bald selbstverständlich erscheinen wird. Der gestalterische Freiraum, der einer Massen- und Populärkultur durch digitale Medien und Web 2.0 jenseits von Museen oder Fernsehsendern eröffnet wurde, ist

allerdings längst wieder ökonomisch vereinnahmt und kommerziell verödet.¹ Nachdem die alltägliche Nutzung von Smartphones und sozialen Medien die Aktivität und Kreativität zwischenmenschlichen Austauschs zunächst befördert hatte, ist das Gros der online verfügbaren Bilder als immenser Datenpool entdeckt worden, mit dem KI-Modelle via Web-Scraping trainiert werden können, deren Hervorbringungen wiederum in den sich hochfrequent drehenden Online-Bildstrudel eingehen. Die ästhetischen Stereotypen und Vorurteile, die daraus hervorgehen, haben die Debatten der letzten Jahre bestimmt.²

Hinsichtlich ihrer vermeintlichen Kreativität mag generative KI jene Automatisierungsversuche fortsetzen, mit der sich die bildende Kunst nicht erst seit dem Surrealismus oder den Produktionsanleitungen der Konzeptkunst auseinandersetzt (die ja ihrerseits der Funktionsweise von Programmen und Algorithmen ähnlich waren). Es ließe sich auch an Harold Cohen denken, der in den 1970er Jahren mithilfe des von ihm programmierten Roboters AARON Gemälde angefertigt hat.³ Und doch könnte der Unterschied zwischen generativer KI und den genannten Methoden kreativer Automatisierung, künstlerischer Zufallsoperation und Mensch-Maschine-Kollaboration nicht größer sein. Es ist vor allem die quantitative Lernfähigkeit der KI und die Ausdehnung ihres Latenzraums, die sie kulturell wirksam werden lässt.⁴

Angesichts der seither erreichten Perfektion und allgemeinen Verfügbarkeit von KI-Werkzeugen erscheint dieser Band am Ende einer Anfangsphase: Hatte die Social-Media-Community die neuen technischen Möglichkeiten anfänglich und gezielt in alle erdenklichen Fettnäpfchen gesteuert und damit eine Form des Trouble Shooting befördert, von denen letztlich die Softwareentwickler profitierten, so erweisen sich die

frühen Bildkreationen generativer KI jetzt viel deutlicher als die Produkte einer bestimmten digital-visuellen Verfasstheit und als zeitbedingte Spiegelbilder einer Gesellschaft und ihrer Seherwartungen.⁵

Die produktiven Missverständnisse, zu denen KI in der Lage ist, zeigen ebenso wie die überkorrigierenden Eingriffe durch Softwareanbieter gewisse Stilmerkmale, Jargons und Dialekte, die auf das Denken und Handeln, auf Seh- und Sprechweisen zurückwirken, und zwar in umso höherem Maße, je mehr eine Gesellschaft mit ihren digitalen Erzeugnissen kurzgeschlossen wird.⁶ Sie verraten sich nicht nur, wenn in einem Manuskript oder einer E-Mail die Fußzeile „written with ChatGPT“ stehen bleibt, sondern auch durch die auf Hochglanz polierte formale Perfektion, in der sich das Generische in Reinform manifestiert – nachdem der letzte Glitch geglättet ist, braucht es inzwischen vorsätzliche Stilbrüche und Fehler, um Menschlichkeit zu emulieren. „We are the voice of every...“, verkündet der elektronisch verfremdete Sänger zum Ausklang des Liedes *Big Generator*, ohne den Satz zu beenden.

Es stellt sich nicht nur die Frage, welche Rolle individueller menschlicher Kreativität oder Arbeit darin zukommt, sondern auch, wie unterschiedlichste Gruppen und Gesellschaften in einer globalisierten, plattformkapitalistischen Kommunikation und Produktion zusammengeschmolzen werden.⁷ Was außerdem erneut zur Debatte steht, ist der Status des Bildes. Dieser hat mit dem Aufkommen der Fotografie, durch Fernsehen und Video, durch Smartphone und soziale Medien regelmäßig Anpassungen erfahren und ist im Begriff, sich neuerlich zu verschieben, seit Generativität ein eigenes Paradigma der Entstehung komplexer visueller Muster markiert. Ausgehend von der Annahme, dass sich die industrielle Revolution hier

1 Tiziana Terranova: After the Internet. Digital Networks Between Capital and the Common, Los Angeles 2022; Cory Doctorow: The 'Enshittification' of TikTok. Or how, exactly, platforms die. In: Wired, 23.01.2023, <https://www.wired.com/story/tiktok-platforms-cory-doctorow/> [Stand 12/2024].

2 Adam Harvey deckte 2019 den Missbrauch privater Flickr-Bilder im Datensatz YFCC100M auf: <https://exposing.ai/fdf/> [Stand 12/2024]. Vgl. den Beitrag des Autors im vorliegenden Band.

3 Alex Estorick: When the Painter Learned to Program. In: Flash Art, 05.12.2017, <https://flash-art.com/article/harold-cohen/> [Stand 12/2024]. Vgl. den Beitrag des Autors im vorliegenden Band.

4 Roland Meyer, Co-Autor des vorliegenden Bandes, plant zu diesem Begriff eine eigene Ausgabe in der Reihe „Begriffe des digitalen Bildes“, zu der auch dieses Heft gehört.

5 Silvia Dal Dosso (Clusterduck Collective): The Future Ahead Will Be Weird AF, 2024, <https://transmediale.de/de/media/silvia-dal-dosso-clusterduck-collective-the-future-ahead-will-be-weird-af-the-ultimate-ai-corecore-experience-part-one-part-two> [Stand 12/2024].

6 Mithilfe welcher Apps die Menschen diesen Apparat füttern, in welche Zustände und Abhängigkeiten sie dabei geraten und welchen Einfluss die Interfaces der ‚sozialen‘ Medien dabei haben, ist ein Thema des Forschungsprojektes „Designing Habits. Das Bild als Anwendung“, das seit 2022 an der Staatlichen Hochschule für Gestaltung Karlsruhe arbeitet und zum DFG-Schwerpunktprogramm „Das digitale Bild“ gehört.

7 Ted Chiang: Why A.I. Isn't Going to Make Art. In: The New Yorker, 31.08.2024, <https://www.newyorker.com/culture/the-weekend-essay/why-ai-isnt-going-to-make-art> [Stand 12/2024].

konsequent fortschreibt, sollen die folgenden Beiträge daher untersuchen, welche Bedeutung die generativen Prozesse künstlicher Intelligenz für die künstlerische und kreative Arbeit besitzen und welche Schlüsse sich daraus für eine Theorie des digitalen Bildes ziehen lassen könnten.

Generative Bilderströme. Über digitale Nach- und Vorbilder

Generation, Generativität

Im Juli 2023 fand in Karlsruhe eine Tagung unter dem Titel „Generation Image“ statt, die sich mit den sozialen Gebrauchswesen digitaler Bildmedien befasst hat. Da es in vielen Beiträgen um generative Formen der Bildproduktion mittels KI ging, sollte der Veranstaltungstitel mit dem dezenten Doppelsinn spielen, den das Wort Generation vor allem im Englischen hat, nämlich einerseits als Alters- oder Entwicklungsstufe, andererseits im Sinne des aktiven *to generate*.

Auch wenn die Tagung nicht weiter auf die biologisch-soziale Dimension des Begriffs eingegangen ist, erwies er sich als eine passende Brücke.¹ Denn die Frage, ob Menschen zeitbedingte Sehweisen und Erfahrungen miteinander teilen, ist aus der Soziologie als „Problem der Generationen“ bekannt.² Regelmäßig werden neue Gruppierungen ausgemacht, und ihre Zusammenfassung unter Alltagsobjekten, Markennamen und vergleichbaren Leitfossilien deutet darauf hin, dass die Zyklen kürzer werden, bis hin zur politischen Bewegung der Letzten Generation, deren Eigenbezeichnung die Endzeit eingeläutet hat.

Die Generation Bild, die im Tagungstitel angedeutet wurde, dürfte es zwar kaum geben, solange die Geschichte der Bilder weiter gefasst wird als das Zeitalter der Fotografie oder des Internets. Auf der anderen Seite sind die gegenwärtigen Möglichkeiten, in und mit Bildern zu kommunizieren, mit jenen des 19. Jahrhunderts kaum noch vergleichbar, und weil die meis-

¹ Siehe <https://hfg-karlsruhe.de/en/archiv/sommer-konferenz-digitales-bild/> [Stand 12/2024].

² Karl Mannheim: Das Problem der Generationen. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Jg. 7, 1928, Heft-Nr. 2, S. 157-185, Heft-Nr. 3, S. 309-330.

ten Menschen im 21. Jahrhundert bereits mit digitalen Bildmedien sozialisiert wurden, konnte für sie der Ausdruck „Digital Natives“ aufkommen. Mit *Generation Angst* gibt es inzwischen auch den gesellschaftskritischen Bestseller dazu, vier Jahrzehnte nach Neil Postmans *Wir amüsieren uns zu Tode*.³

Der Wortstamm des lateinischen *generatio* enthält noch eine Reihe weiterer Bedeutungen, die für die Frage digitaler Generativität durchaus Bedeutung haben. Die Etymologie reicht von *gens* (Leute) und *genus* (Geschlecht, vgl. *gender*) bis zur modernen Genetik (Gen, Genom) und deutet auf leibliche Abstammungsbeziehungen und Vererbungslinien hin.⁴ Von hier leitet sich wiederum die Vorstellung einer qualitativen Weiter- und Höherentwicklung ab, als Zündstufe der *next generation*. Da sich aus Reproduktion und Wachstum unterschiedliche, aber verwandtschaftliche Formen ergeben, schließen sie außerdem alles *Generische* ein.

Vom eingeborenen *ingenium* gibt es Verbindungen zum Ingenieurwesen, der *engine* und dem *engineering*, und mit dem *Generieren* ist dann die Brücke zur digitalen ‚Generativität‘ geschlagen, die zunehmend mit KI-Anwendungen in Verbindung gebracht wird. Im Kontext digitaler Technologien soll sie das Vermögen selbstlernender Systeme zum Ausdruck bringen, auf statistische oder tentative Weise zu operieren, im Unterschied zu einer digitalen Verarbeitung, die auf vorprogrammierten Abläufen und Ergebnissen beruht.

Generativität ist daher nicht auf neuronale Netze und künstliche Intelligenz beschränkt. Die Ergebnisse der frühen Computerkunst wurden bereits als generativ bezeichnet, weil sie sich auf dem Monitor oder im Plotter scheinbar autonom entfalteten (Abb. 1). Das daran anschließende Computer Generated Imaging (CGI) sollte unterstreichen, dass eine mit leistungsstarken Prozessoren errechnete Grafik einen fließenden,

Abb. 1, Manfred Mohr: P-197-F (1977-78), Plotterzeichnung auf Papier, 60x60 cm, ©Manfred Mohr, Foto: ©Winfried Reinhardt

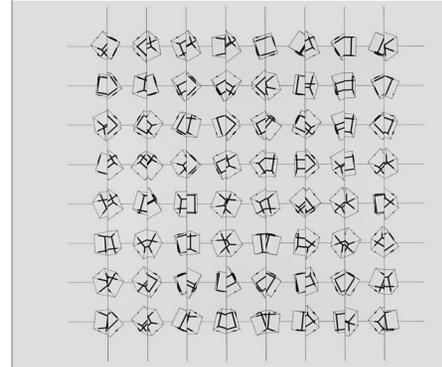
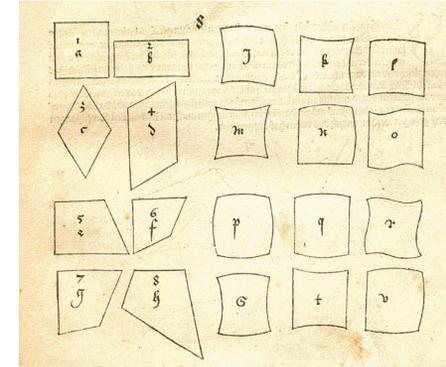


Abb. 2, Albrecht Dürer: Variation von Vierecken, aus: Unterweysung der Messung, Nürnberg 1525, S. 56, Bayerische Staatsbibliothek München



fotorealistischen und quasi-lebendigen Eindruck hinterlässt. Parallel dazu haben die technischen Limitierungen früherer PCs in der Programmierszene eine eigene Sportart hervorgebracht, das sogenannte „Demo“. Hier besteht der Wettbewerb darin, aus kleinstmöglichem Code die größte Vielfalt an unvorhersehbaren Effekten zu zaubern.

Noam Chomskys Generative Transformationsgrammatik, der Versuch einer universellen Sprach- und Übersetzungsregel, verweist noch weiter in die Frühzeit des Computing Mitte der 1950er Jahre zurück. Von derartigen Wortverwendungen ausgehend könnten bildliche Darstellungen, die mithilfe von Proportions- und Konstruktionsregeln erstellt wurden, ebenfalls als generativ bezeichnet werden (Abb. 2). Anleitungen wie Alexander Cozens' *A New Method of Landscape* von 1785, in denen Landschaften nach Art gestaltpsychologischer Testbilder

3 Jonathan Haidt: *Generation Angst. Wie wir unsere Kinder an die virtuelle Welt verlieren und ihre psychische Gesundheit aufs Spiel setzen*, Hamburg 2024; Neil Postman: *Wir amüsieren uns zu Tode. Urteilsbildung im Zeitalter der Unterhaltungsindustrie*, Frankfurt am Main 1985.

4 Vgl. den Eintrag „Generation“ im Online-Projekt „Wortgeschichte digital“ des Zentrums für digitale Lexikographie der deutschen Sprache: <https://www.zdl.org/wb/wortgeschichten/Generation> [Stand 10/2024].

nach dem Zufallsprinzip entworfen werden, oder Kunstfälschungen, die auf der Nachahmung individueller Stile beruhen, beziehen ihr Täuschungsvermögen ebenfalls aus der Nutzung bestimmter Möglichkeitsräume.

Wie sich zeigt, ist das Thema bildhistorisch relevant. Das maschinelle Lernen wurde in seinen Anfängen als Mustererkennung, als maschinelles Sehen, konzipiert und erprobt. Damit war bereits der Weg geebnet für den Erfolg von Grafikprozessoren, die im Zuge der KI-Welle weltwirtschaftliche Bedeutung erlangt haben. Die 2014 vorgestellten Generative Adversarial Networks (GAN), bei denen zwei künstliche neuronale Netze sich gegenseitig herausfordern, wurden ebenfalls anhand von großen Fotomengen und für die Synthese fotorealistischer Bilder erprobt. Da es sich um Lernvorgänge handelt, für die jemand Trainingsmaterial bereitstellt, ist auch eine genetische Dimension erhalten.⁵ Der Wortgebrauch hat längst auf andere Bereiche abgefärbt. Nicht nur Daten, auch Gewinne werden generiert. Zumindest soll die Redeweise durchblicken lassen, dass sie nicht willkürlich produziert, fabriziert oder konstruiert worden sind. Hier zeigen sich begriffliche Sinnverschiebungen und Inflationen, die auch bei anderen Ausdrücken zu beobachten sind, etwa beim Übergang von elektronisch zu digital.

Generische Fantasien

Auffällig ist in diesem Zusammenhang die zunehmende Verbreitung des Wortes ‚generisch‘. In Produktionsbereichen wie der Pharmazie wird der Ausdruck „Generika“ seit Langem zur Bezeichnung kostengünstiger Versionen eines Markenpräparats verwendet, er ist also nicht spezifisch für digitale Technologien. Tatsächlich könnte aber auch ein Großteil der Massen-

Abb. 3. Die Wohnanlage Burj al Babas (Provinz Bolu) im Dezember 2018, Foto: © Adem Altan/AFP



und Unterhaltungskultur als generisch bezeichnet werden, im Sinne einer Spielart oder Variation.

Vergleichbares gilt im Bereich der Bildproduktion. So zeigt das Foto einer Wohnanlage namens Burj al Babas in der türkischen Provinz Bolu nördlich von Ankara das Resultat einer am Computer entworfenen und 2014 realisierten Architekturplanung, die auf populäre Luxusvorstellungen spekuliert hat. Die weitgehend leerstehende Wohnanlage wirkt durch die serielle Wiederholung eines disneyhaften Miniaturschlösses so surreal, dass sie als fotografischer Fake durchgehen oder auf den halluzinogenen Mechanismen generativer KI beruhen könnte (Abb. 3). Ein vergleichbares Beispiel sind die Fantasiearchitekturen der Euro-Geldscheine, die traditionelle Porträts durch politisch neutrale Motive ersetzen sollten, um Debatten über die Bildwürdigkeit ausgewählter Persönlichkeiten zu um-

⁵ Mona Leinung und Markus Rautzenberg haben in einem Vortrag an der Staatlichen Akademie der bildenden Künste Karlsruhe im Mai 2024 den Begriff der Generation benutzt, um auf die kindliche Stufe der KI und einen entsprechenden Dressurbedarf hinzuweisen.

gehen. Sie scheinen typisch europäische Orte oder Wahrzeichen zu zeigen, sind aber vollkommen ortlos und aus Standardelementen der Architekturgeschichte synthetisiert.

In der historischen Perspektive gewinnt außerdem wieder das Phänomen der Stock Photography an Beachtung, die um 1900 aufgekommen ist, als Fotoagenturen damit begannen, überzählige Aufnahmen und Bilder ohne Nachrichtenwert gesondert zu vermarkten.⁶ Aus der Idee, Fotografien für illustrative Kontexte zu bevorraten, ergab sich eine eigene, kommerzielle Stilistik, die auf schnell erkennbare, aber unverfängliche Motive setzt. Das Geschäftsmodell existierte schon lange vor dem Siegeszug der KI-Bildgenerierung, kann aber gleichwohl als emblematisch für die Geschichte des Bildes im Zeitalter seiner technischen Reproduktion gelten, weil sich in den Archivstrukturen und Schlagworten der Agenturen genau jene Muster, Trends und visuellen Topoi ablagerten, die bis heute in Porträts, Urlaubsmotiven oder Alltagsszenen reproduziert werden, wie ein Blick in die Auslagen von Zeitungskiosks belegt.

Es war kein Zufall, dass die Vorratsfotografie in derselben Phase der illustrierten Presse entstand, in welcher die Ikonologie zur kunst- und kulturwissenschaftlichen Methode aufstieg. Als sogenannte Symbolbilder massenhaft im redaktionellen und gewerblichen Einsatz, haben sie die zunehmende Gleichsetzung von Bild und Begriff, die quasi-verbale Funktion von Illustrationen unterstützt, die nach einem festen Katalog von Such- und Stichwörtern abgefragt werden. Diese Übersetzung ist heute in digitalen Suchalgorithmen und Prompts aufgehoben.

Um die ästhetischen Leitplanken digitaler Bilder, die mithilfe selbstlernender KI generiert werden, weiter zu charakterisieren, könnte schließlich ein aus dem Französischen kommender Begriff zum Zuge kommen, das *Genre*, das ebenfalls der

16

⁶ Mehrere Beiträge der SPP-Tagung „KI und das digitale Bild“ (Ludwig-Maximilians-Universität München 2024) haben darauf verwiesen (siehe <https://www.digitalesbild.gwi.uni-muenchen.de/ki-und-das-digitale-bild-21-2-23-2-2023/> [Stand 10/2024]). Siehe auch jüngst Thomas Nolte: *Stockfotografie*, Berlin 2024.

Wortwurzel *gen-* entstammt. Es soll zunächst nur eine Gattung oder Sortierung beschreiben (wie zum Beispiel *Klassik* vs. *Pop* in der Musik), kann aber als marktförmiges Schubladendenken einen abfälligen Beiklang annehmen.⁷ Alltagsnahe oder gefühlsbetonte Darstellungen, die sich auf den freien Kunstmärkten des 18. und 19. Jahrhunderts wachsender Beliebtheit erfreuten, wurden von der Kunstkritik als *Genremalerei* tituliert, um sie von der gehobenen Kunst abzugrenzen.

Digitalistischer Realismus

Generative KI-Bildproduktion ist die digitale Variante des beschriebenen Effekts, dass ein Pool an Trainingsmotiven auf einen Durchschnitt von Mustern und Ähnlichkeiten, auf eine ästhetische Normalverteilung hinausläuft – bis hin zum *Model Collapse*, zum Zusammenbruch eines KI-Modells, das sich mit seinen eigenen Daten trainiert und dadurch kompromittiert.⁸ Dieser Extremfall treibt eine längere technische Entwicklung auf die Spitze.

In den vergangenen fünf Jahrzehnten sind mit der digitalen Fotografie, dem Internet und den sozialen Medien und schließlich mit dem Smartphone mächtige Werkzeuge der Bildproduktion und -kommunikation entstanden, die gebrauchsfertigen KI-Anwendungen den Weg geebnet haben. Die gestalterischen Möglichkeiten, die sich daraus ergeben, treffen auf einen gesellschaftlich breiteren Bedarf nach ‚visuellen Generika‘, der durch Software technisch vermittelt wird.⁹

Indem Aufnahmen korrigiert, gefiltert und angereichert werden, kommt es zu einer stillschweigenden, fließenden und gerade darum tiefenwirksamen Passung von Realität und Pro-

⁷ Vgl. Ursula K. Le Guin: *Genre. A Word Only the French Could Love*. In: Pat Murphy et al. (Hg.): *The James Tiptree Award Anthology*, 1, San Francisco 2005, S. 61–71, zum Gattungsdnken in der Belletristik. Für die Kunstgeschichte siehe Dominik Brabant: *Kunsthistorische Gattungsgeschichten: Neue Fragen an ein traditionelles Forschungsfeld*. In: *Kunstchronik*, Jg. 77, 2024, Heft-Nr. 8, S. 520–530 sowie das Themenheft „Digitale Genres“ (kritische Berichte, Jg. 53, 2025, Heft-Nr. 1, 2025).

⁸ Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Nicolas Papernot, Ross Anderson und Yarin Gal: *AI models collapse when trained on recursively generated data*. In: *Nature*, 2024, Heft-Nr. 632, S. 755–759.

⁹ Siehe z. B. <https://www.theverge.com/2023/3/13/23637401/samsung-fake-moon-photos-ai-galaxy-s21-s23-ultra> [Stand 12/2024]; vgl. Antonio Somaini: *Algorithmic Images. Artificial Intelligence and Visual Culture*. In: *Grey Room*, 2023, Heft-Nr. 93, S. 74–115 und Berit Glanz: *Filter*, Berlin 2023.

jektion, die nicht so offenkundig ist wie monströse KI-Fantasien, sondern sich in mehrheitsfähigen Kriterien der Fotogenität äußert, in der inzwischen sogenannten „Instagrammability“ von touristischen Zielen, Lebensmitteln oder Porträts. Bildhistorisch lässt sich diese visuelle Schnittmenge am ehesten unter stilistischen Gesichtspunkten, als zeitabhängige Mode- oder Wahrnehmungsform fassen.

Nennen ließen sich hier virtuelle Persönlichkeiten wie etwa junge Influencerinnen, die von Werbeagenturen kreiert werden, um Social-Media-Kanäle nach Belieben mit stereotypischen Wunschbildern, insbesondere Körperbildern füllen zu können. Da Millionen Frauen im geschätzten Alter dieser digitalen Kunstfiguren das Internet mit einem überschaubaren Set an Motiven (Wohnung, Strand, Gym) und Posen (Spiegelbild mit langen Fingernägeln) fluten, ist es zudem relativ leicht, die erforderlichen Trainingsdaten zu gewinnen.

Dass Bildeffekte wie Blitzlichter oder Unschärfen in ihren Porträts auf den ersten Blick täuschend echt wirken, sich bei näherem Hinsehen oder nach Ablauf einiger Jahre aber als technische Zutaten erweisen, wird für die Abgebildeten oder ihr Publikum kein fundamentales Problem darstellen. Denn die modische Zurichtung des Körpers auf ein Erscheinungsbild ist in der analogen Kulturgeschichte schon viel zu lange eingeübt, als dass echt und falsch die geeigneten Kategorien wären. Faltenfreiheit und fotografische Retusche bauen vielmehr auf einer übergeordneten Vorstellung von Perfektion auf.

KI-unterstützte Smartphones könnten rein theoretisch einen Großteil der Aufnahmen generieren, ohne dass es dafür noch eine Vorlage bräuchte. Denn ab jetzt ist alles wieder Grafik, von der Kosmetik über die fotografische Aufnahme bis zur digitalen Bildverarbeitung. Jedoch bleibt der persönliche Vollzug, der eigene Knopfdruck, der durch die einfache Handhabung

oder Affordanzen von Anwendungen befördert wird, ein wichtiges Element der technisch-sozialen Koproduktion. Er stellt die habituelle Entsprechung zum Bildmotiv, zum *Punctum* her, und schaltet die übrige Bildproduktion auf Autopilot.

Stilhistorisch bemerkenswerter ist eher, wenn bestimmte, technisch bedingte Bildstile den Alltagsgeschmack zu verändern beginnen und das Selbstbild von Menschen durch Auto-korrekturen vorgeformt wird. Die selbstähnlichen Muster, die aus dieser wechselseitigen Anpassung hervorgehen, sollen hier daher als digitalistischer Realismus zusammengefasst werden, in Anlehnung an einen ‚sozialistischen‘ Realitätsbegriff, bei dem der soziale Zweck die Wirklichkeit definierte. Denn es herrscht trotz des offenkundig malerischen oder grafischen Charakters der Darstellung eine fotografische Ästhetik vor, die einer Bildlegende wie „Diese Person existiert nicht“ ihren Magritteschen Sinn gibt.¹⁰

Ähnlich der historischen Genremalerei, die von der Kunstkritik als gefällig übergangen und von den Avantgarden als gestrig abgestoßen wurde, führt die eingeübte motivische Logik der Illustration dazu, dass täglich neue KI-generierte Influencerinnen oder vergleichbare Charaktere entstehen. Sie werden ihrerseits mit parodistischen Brechungen und ironischen Formen der Aneignung und Umdeutung quittiert, bis ein neues Darstellungsmuster das aufmerksamkeitsökonomische Mädchenschema vom Thron stürzt.¹¹

Nach dem Bild ist vor dem Bild

Die Geschichte der Bildmedien gibt Hinweise zum Verständnis aktueller Phänomene im Zusammenhang mit generativen Bildprozessen, auch wenn sie nur einzelne Aspekte abdecken sollte. So wurde bereits um 1900 diskutiert, ob apparativ er-

¹⁰ Siehe <https://www.thispersondoesnotexist.com> [Stand 12/2024] und den Beitrag von Katharina Weinstock im vorliegenden Band.

¹¹ Vgl. <https://www.monopol-magazin.de/kunst-ki-wie-verrueckt-wird-unsere-zukunft-der-ki-bilder> [Stand 12/2024].

zeugte Fotografie Kunst sein kann und ob die künstlerische Leistung vor allem im souveränen Umgang mit der Technik liegt oder eher in einer besonderen Ästhetik. Diese Frage ist keineswegs eindeutig geregelt worden und immer noch aktuell. Auch die Frage bildlicher Identität und Normierung kehrt mit jeder technischen Neuerung zurück.¹²

Vergleichbares gilt im Kontext visueller Kommunikation für Themen wie Authentizität oder Bildzensur. Eine vielfach bespöttelte Äußerung der Bundeskanzlerin Angela Merkel aus dem Jahre 2013, wonach die Gesellschaft mit dem Internet „Neuland“ betrete, ist in juristischer Hinsicht insofern berechtigt gewesen, als soziale Medien private Kommunikation und öffentliche Publikation kurzgeschlossen haben, ohne dass das Postgeheimnis oder Presserecht dafür ausgelegt waren.

Die Konsequenzen wurden digitalen Plattformen überlassen, die überwiegend per Autopsie und als *Click Work* in Drittländern feststellen lassen, ob gewaltverherrlichende oder pornografische Darstellungen anderen Menschen zumutbar sind.¹³ Die *Content Moderation* von Bildern bewegt sich dabei permanent zwischen öffentlichen Schutzbedürfnissen und privaten Zensurmaßnahmen und ist entsprechend umkämpft. Die Technologien, die zu ihrer Automatisierung eingesetzt werden könnten, halten bestenfalls Schritt mit den Gewaltdarstellungen, die aus den gleichen Technologien hervorgehen.

Drängend bleibt das Problem der Herkunft von Trainingsmaterial, bei dem es regelmäßig zu Verletzungen geistigen Eigentums und von Persönlichkeitsrechten kommt und das sich durch einen ethnisch-sozialen *bias* auszeichnet, der den Loop-Effekten der Agenturfotografien entspricht. Neben der Leistungssteigerung bei Grafikkarten, Displays und Sensoren, welche die Verschränkung von Bild, Umwelt und Handlung weiter vorantreibt, überformt der kontinuierliche Strom

an Information außerdem bisherige Modi der Betrachtung, Beurteilung, Erinnerung. Die simulative Qualität digitaler Gesichter und Stimmen ist inzwischen so überzeugend, dass Menschen an der Unterscheidung natürlicher und generierter Darstellungen scheitern; und es ist dabei nicht ausgeschlossen, dass ein synthetisches Gesicht einer realen Person zum Verwechseln ähnlich sieht.

Nicht nur die Rechtsprechung, sondern auch die Bildtheorie steht hier vor der Frage, an welchen konkreten Phänomenen sie ihre Urteile und Überlegungen festmachen soll. Diese Frage war bereits Ausgangspunkt eines Karlsruher Forschungskollegs, das in den Jahren 2000–2009 unter dem Titel „Bild – Körper – Medium. Eine anthropologische Perspektive“ aktiv gewesen ist und den Ort des Bildes im Zeitalter elektronischer Medien und vernetzter Imaginationen zu bestimmen versucht hat.¹⁴ Sie stellt sich heute noch deutlicher als ein Quellenproblem, sei es mit Blick auf Daten- und Speichermengen oder auf den ephemeren und verteilten Charakter digitaler Bildkommunikation.

Was generative KI-Anwendungen ausgeben oder KI-gestützte Filter ermöglichen, ist nicht nur anhand von Motiven zu begreifen. Auf der anderen Seite behalten auch diese Motive ihre ästhetische Seite, die der Einordnung und historischen Expertise bedarf. Da schon der Begriff des Stils im Zuge der Industrialisierung im 18. und 19. Jahrhundert die Aufgabe zugewiesen bekommen hatte, den kausalen Zusammenhang von Form und Technik zu fassen, steht auch eine Stilistik des generativen Bildes nicht im Widerspruch zu dessen technologischen Grundlagen. Statt an der fragwürdigen Qualität oder der schieren Fülle von Beispielen zu verzweifeln, sollte sie sich mit den historischen Voraussetzungen und sichtbaren Folgen digitaler Bilder befassen.

¹² Roland Meyer: Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook, Konstanz 2019. Vgl. den Beitrag des Autors im vorliegenden Band.

¹³ Sarah T. Roberts: Behind the Screen. Content Moderation in the Shadows of Social Media, New Haven 2019.

¹⁴ Die Folgen dieser Entwicklung waren Gegenstand des Vorgängerprojekts „Adaptive Bilder“ im SPP „Das digitale Bild“, siehe Matthias Bruhn, Kathrin Friedrich, Lydia Kähny und Moritz Queisner (Hg.): Adaptivität, München und Hildesheim 2021, DOI: <https://doi.org/10.5282/ubm/epub-76331>.

Weltmodelle ordnen. Zur politischen Epistemologie der Generativität

In Bezug auf künstliche Intelligenz (KI) wird Generativität häufig als die Generation von Text- oder Bild-Outputs verstanden. Die dialogische Qualität der Chat-Interfaces von *large language models* (LLMs) oder Debatten um Urheberrechtsverletzungen bestärken dies – es ist die Generation von ‚Content‘, die das breite öffentliche Interesse auf generative KI gelenkt hat. Der zugrunde liegende Prozess kann abstrakter beschrieben werden: Es geht um die Datafizierung der Welt, die Erkennung von Mustern und deren anschließende Generierung – Schritte, die auch als die drei Modalitäten des Trainings, der Klassifizierung und der Vorhersage betrachtet werden können.¹ Die generative Modalität der KI ist somit als die Erzeugung statistischer Muster zu verstehen. In Anlehnung an Louise Amoore et al. instanzieren generative KI-Modelle dabei ein „Modell der Welt“,² das ein Modell dafür sein soll, „wie die Welt funktioniert“, welches „aufgabenübergreifend“ „konfigurierbar“ ist.³ Offenkundig ist dem eine politische Dimension inhärent.

Ein KI-Modell setzt sich aus Trainingsdaten und einem Lernalgorithmus zusammen, sodass das Modell selbst letztendlich das darstellt, was der Algorithmus gelernt hat.⁴ Die Trainingsdaten werden als Stichproben aus einer vom Modell kalkulierten, zugrunde liegenden Wahrscheinlichkeitsverteilung betrachtet, aus der neue Datenpunkte abgeleitet werden

1 Matteo Pasquinelli und Vladan Joler: The Noosphere Manifested. AI as Instrument of Knowledge Extractivism. In: AI & SOCIETY, Jg. 36, 2021, Heft-Nr. 4, S. 1268, <https://doi.org/10.1007/s00146-020-01077-6>.

2 Louise Amoore et al.: A World Model. On the Political Logics of Generative AI. In: Political Geography, Jg. 113, 24. Mai 2024, S. 5 (eigene Übersetzung), <https://doi.org/10.1016/j.polgeo.2024.103134>.

3 Yann LeCun: A Path Towards Autonomous Machine Intelligence Version 0.9.2, 2022-06-27. In: OpenReview, Jg. 27, 2022, S. 5 (eigene Übersetzung), <https://openreview.net/pdf?id=BZ5a1r-kVsf>.

4 Pasquinelli und Joler 2021 (s. Anm. 1), S. 1265.

können.⁵ Bei der Sprachmodellierung kann so zum Beispiel das nächste wahrscheinlichste Wort in einer Sequenz vorhergesagt werden. Die generative Modalität der KI als Vorhersagemodalität und als verbunden mit der Idee des Weltmodells wird in einer Erklärung von Ilya Sutskever, Mitgründer von OpenAI, zu ChatGPT deutlich, einem KI-Chatbot und der wohl am schnellsten wachsenden Verbraucher-App in der Geschichte.⁶ Sutskever zeichnet ein frappierendes Bild:

„What we are doing is that we are learning a world model. It looks like we are [...] just learning statistical correlations in text [i.e., data], but it turns out that [...] what the neural network learns is some representation of the process that produced the text. This text is actually a projection of the world. There is a world out there, and it has a projection on this text. And so what the neural network is learning is more and more aspects of the world, of people, of the human conditions, their hopes, dreams, and motivations, their interactions and the situations we are in. And the neural network learns a compressed, abstract, usable representation of that. This is what is being learned from accurately predicting the next word.“⁷

Sutskevers hochtönende Behauptungen veranschaulichen eine Idee, die für den Begriff der Generativität von zentraler Bedeutung ist. Diese ist bereits in einem von ihm mitverfassten OpenAI-Paper aus dem Jahr 2018 enthalten, in dem die GPT-Modelle (auf denen ChatGPT basiert) vorgestellt werden. Die Autor*innen schreiben: „our goal is to learn a *universal representation that transfers with little adaptation to a wide range of tasks*.“⁸ Die Idee ist, dass genügend Daten und die richtige Architektur für eine „universelle Repräsentation“ (das heißt ein „Weltmodell“) ausreichen, die sich auf bisher „unangetroffene Situationen“ („unencountered situations“) und Aufgaben übertragen lässt.⁹

5 Amoores et al. 2024 (s. Anm. 2), S. 3.

6 Ben Wodecki, UBS: ChatGPT May Be the Fastest Growing App of All Time. In: AI Business, 3. Februar 2023, <https://aibusiness.com/nlp/ubs-chatgpt-is-the-fastest-growing-app-of-all-time> [Stand 04/2025]

7 Ilya Sutskever: Fireside Chat with Ilya Sutskever and Jensen Huang: AI Today and Vision of the Future, interview by Jensen Huang. In: NVIDIA On-Demand, März 2023, <https://www.nvidia.com/en-us/on-demand/session/gtcspring23-s52092/> [Stand 08/2024] (eigene Hervorhebung).

8 Alec Radford et al.: Improving Language Understanding by Generative Pre-Training. Preprint, 2018, S. 2 (eigene Hervorhebung und Übersetzung).

9 Amoores et al. 2024 (s. Anm. 2), S. 2.

Sutskevers Behauptungen zeigen die Überzeugungen, die auf die Fähigkeiten eines solchen Modells projiziert werden können: „Wenn man die Daten wirklich gut komprimiert, muss man alle darin verborgenen Geheimnisse extrahieren [...] das ist der Schlüssel.“¹⁰ Dan MacQuillan konstatiert, dass dies „mit dem Glauben an eine verborgene mathematische Ordnung zusammenhängt, die ontologisch derjenigen überlegen ist, die unseren alltäglichen Sinnen zugänglich ist,“ und welche die Bereiche der Computer- und Data Sciences im Allgemeinen durchdringe.¹¹ Die vermeintliche Überlegenheit datengetriebener Systeme, die oft mit dem Glauben an ihre Objektivität einhergeht, ist jedoch vielmehr „angestrebtes Ideal“ als tatsächliche Praxis.¹² Daher wird in diesem Text, in Anlehnung an Pietro Daniel Omodeo,¹³ eine politische Epistemologie der Generativität vorgebracht, die untersucht, auf welche Weise Wissensproduktion im Zusammenhang mit generativer KI (historisch) ermöglicht und aufrechterhalten wird.

Im Zuge des nachzuverfolgenden Dreischritts – Datafizierung, Mustererkennung, Mustergenerierung – ist hierbei die Frage zentral, wie die Welt wahrgenommen wird und welche politischen Logiken daraus abgeleitet werden. Dabei ist es vor allem die generative Modalität, an der sich zeigt, „wie das Modell die Welt ‚sieht.‘“¹⁴ Wie sich herausstellen wird, ist diese Sicht stets eine verzerrte; in Anlehnung an Matteo Pasquinelli und Vladan Joler kann KI ganz allgemein als eine solch verzerrende ‚Linse‘ auf Wissen betrachtet werden. Eine Historisierung gleichsam verzerrender Linsen auf die Welt soll die Funktionsweise generativer KI näher umreißen und genauso dem „allgegenwärtigen Präsentismus“¹⁵ entgegenwirken, der dem Diskurs über sie und neue Technologien im Allgemeinen anhaftet.

10 Sutskever 2023 (s. Anm. 7), 13:30.

11 Dan McQuillan: Data Science as Machinic Neoplatonism. In: Philosophy & Technology, Jg. 31, 2018, Heft-Nr. 2, S. 254 (eigene Übersetzung), <https://doi.org/10.1007/s13347-017-0273-3>.

12 Sun-ha Hong: Prediction as Extraction of Discretion. In: Big Data & Society, Jg. 10, 2023, Heft-Nr. 1, S. 3, <https://doi.org/10.1177/20539517231171053>.

13 Pietro Daniel Omodeo: Political Epistemology. The Problem of Ideology in Science Studies, Cham 2019, <https://doi.org/10.1007/978-3-030-23120-0>.

14 Pasquinelli und Joler 2021 (s. Anm. 1), S. 1274 (eigene Übersetzung).

15 Rita Riley und Jennifer Rhee: Critical AI. A Field in Formation. In: American Literature, Jg. 75, 2023, Heft-Nr. 2, S. 186, <https://doi.org/10.1215/00029831-10575021>. In diesem Sinne sind technologische Entwicklungen nicht abstrahiert von Geschichte und Gesellschaft zu lesen, sondern als aus diesen hervorgegangen.

Datafizierung und „Verborgene Geheimnisse“

Eine solche Instanz ist die Datafizierung der Welt, die in der Zeit der wissenschaftlichen Revolution zentral war, als Galileis Blick durch das Teleskop – eine paradigmatische, buchstäbliche Linse in der Geschichte von Wissenschaft und Technik – „aus jedem gesehenen Datum ein konstruiertes und verrechnetes Datum“ machte.¹⁶ Mit Galilei soll eine epistemologische Konstellation veranschaulicht werden, die für ein heutiges Verständnis von Generativität aufschlussreich ist. Es handelte sich nicht nur um die „Technisierung der Naturerkenntnis“,¹⁷ sondern auch um die Verschiebung der „Grenze der Sichtbarkeit“,¹⁸ um die Erzeugung eines „anästhetischen Feldes“¹⁹ durch das Teleskop: Während es Dinge sichtbar machte, die zuvor für das bloße Auge unsichtbar waren, demonstrierte das Teleskop auch die Bedingtheit und die Grenzen anschaulicher Wirklichkeit. Diese Unzulänglichkeit des Sehens veranlasste Galilei, sich der Abstraktion zuzuwenden.²⁰ Indem er das Universum mittels einer „Sprache der Mathematik“²¹ verstand, postulierte er eine mathematische Ordnung, hinter der sich „so tiefe Geheimnisse und so erhabene Gedanken [verbergen], daß die Mühen und Nachtwachen von Hunderten und Hunderten der schärfsten Geister in tausendjähriger Forschungsarbeit sie noch nicht völlig zu durchdringen vermochten.“²² Eine solche „Erhabenheit“ klingt ebenfalls in der von MacQuillan identifizierten „Überlegenheit“ jener „verborgenen mathematischen Ordnung“ an. Allerdings war sich Galilei bewusst, dass intellektuelle Strukturen zur Kategorisierung der Realität der Natur nicht innewohnen: „Würde das Lebewesen entfernt,

16 Joseph Vogl: Medien-Werden. Galileis Fernrohr. In: Lorenz Engell und ders. (Hg.): Mediale Historiographien. Archiv für Mediengeschichte Sonderdruck, 2001, S. 115-123, hier S. 115-116.

17 Ebd., S. 115.

18 Friedrich A. Kittler: Optische Medien. Berliner Vorlesung 1999, Berlin 2002, S. 13.

19 Vogl 2001 (s. Anm. 16), S. 118.

20 Hans Blumenberg: Das Fernrohr und die Ohnmacht der Wahrheit. In: Galileo Galilei: Sidereus Nuncius. Nachricht von neuen Sternen, Frankfurt am Main 2002, S. 7-7b, hier S. 22.

21 Galileo Galilei: The Assayer. A Letter to the Illustrious and Very Reverend Don Virginio Cesarini. In: Galileo Galilei: Discoveries and Opinions of Galileo, New York 1957, S. 231-280, hier S. 237-238 (eigene Übersetzung).

22 Galileo Galilei zit. nach Vogl 2001 (s. Anm. 16), S. 120.

würden alle diese Qualitäten ausgelöscht und vernichtet werden.“²³ Heute lässt sich das auf den oft zitierten Satz „raw data is an oxymoron“²⁴ übertragen: Daten stehen immer in einem bestimmten Kontext, der von den Zwecken, für die sie gemacht werden, und den Instrumenten, Menschen und Systemen, die zu ihrer Erfassung eingesetzt werden, geprägt ist.

Mit Hilfe eines weiteren Astronomen kann die Ausdehnung der Datafizierung von der natürlichen auf die soziale Sphäre entfaltet werden. Seit Ende des 18. Jahrhunderts verband sich Regierungsführung zunehmend mit statistischen Populationen, eine „Lawine gedruckter Zahlen“²⁵ führte zur Erfassung einer Fülle von Daten und zur Entwicklung von Instrumenten zu deren Untersuchung – um den Staat zu stärken und Politik und Bürger*innen zu beeinflussen.²⁶ Diese statistischen „Techniken der Beobachtung“²⁷ verschoben erneut den Blick auf die Realität. Auch dies hatte eine Dimension des Geheimen: Ursprünglich bezeichnete die Statistik das geheime Wissen des Staates.²⁸ In dieser Zeit war der belgische Astronom Adolphe Quetelet maßgeblich an der Einführung statistischer Methoden in die Sozialwissenschaften beteiligt und prägte den Begriff der Sozialphysik. Ein bemerkenswerter Transfer war die Integration der Gauß-Kurve, die später als Normalverteilung bekannt wurde. Diese Kurve wurde ursprünglich in der Astronomie verwendet, um mit Hilfe der mathematischen Wahrscheinlichkeitsrechnung Messunsicherheiten zu bewerten und auf Grundlage der Mehrfachbeobachtungen von Sternen Rückschlüsse auf deren Position zu ziehen.²⁹ Anstelle von (realen) Sternen am Himmel versuchte

23 Galilei 1957 (s. Anm. 21), S. 274 (eigene Übersetzung).

24 Lisa Gitelman und Virginia Jackson: Introduction. In: Lisa Gitelman (Hg.): „Raw Data“ Is an Oxymoron, Cambridge 2013, S. 1-14, <https://doi.org/10.7551/mitpress/9302.003.0002>.

25 Ian Hacking: The Taming of Chance, Cambridge 2009, S. 2 (eigene Übersetzung).

26 Chris Wiggins und Matthew L. Jones: How Data Happened. A History from the Age of Reason to the Age of Algorithms, New York 2023, S. 21-22.

27 Michel Foucault: The Meshes of Power. In: Jeremy W. Crampton und Stuart Elden (Hg.): Space, Knowledge and Power. Foucault and Geography, Aldershot/Burlington 2007, S. 153-162, hier S. 161 (eigene Übersetzung).

28 Matteo Pasquinelli: Arcana Mathematica Imperii. The Evolution of Western Computational Norms. In: Maria Hlavajova und Simon Sheikh (Hg.): Former West: Art and the Contemporary after 1989, Utrecht und Cambridge 2016, S. 281-293.

29 Wiggins und Jones 2023 (s. Anm. 26), S. 27. Zudem ist die Geschichte der Normalverteilung mit der Diskussion um die „Bell Curve“ ein wichtiges Moment in der Geschichte der Statistik und war problematische Basis eines frühen Intelligenzbegriffs. Vgl. Stephen Jay Gould: The

Quetelet mit der Normalverteilung soziale Realität als solche zu erfassen, indem er die normative Sprache des „großen, schönen und guten“ idealen Durchschnitts verwendete, welchen er beispielsweise auf ideale Körpergrößen oder eine vermeintliche Neigung zu Verbrechen in Bevölkerungsgruppen bezog.³⁰ Auf diese Weise stellt Quetelet einen der historischen Vorläufer der Techniken der generativen KI dar, bei denen eine zugrunde liegende Wahrscheinlichkeitsverteilung unter dem Deckmantel scheinbarer Objektivität mobilisiert wird, um Handlungsoptionen für sozialen Wandel abzuleiten. Diese zugrunde liegende Verteilung verspricht, die „verborgenen“ Tendenzen oder Muster in der Bevölkerung ans Licht zu bringen, um sie handhabbar und verwertbar („actionable“) zu machen.³¹

Ein aktuelles Beispiel kann dies verdeutlichen. Im Jahr 2023 stellte das US-Softwareunternehmen Palantir seine neue Artificial Intelligence Platform (AIP) vor, die angeblich die „Macht von LLMs und modernster KI entfesselt,“ indem sie „hochsensible und geheime nachrichtendienstliche Informationen miteinander verbindet, um eine Echtzeitrepräsentation ihrer Umgebung zu erstellen.“ Die Plattform, ein generatives LLM, das auf „Militärdoktrin, Logistik und Gefechtsdynamik“ abgestimmt ist, soll bei der Überwachung militärischer Aktivitäten helfen. In einer Produktdemonstration wird „potenziell feindliche Ausrüstung am Boden“ aufgespürt, woraufhin das Modell in natürlicher Sprache aufgefordert wird, drei „Handlungsoptionen“ zur Bewältigung der Situation zu generieren. Die drei ausgegebenen Optionen sind: „Bekämpfung aus der Luft“, „Bekämpfung mit Langstreckenartillerie“ und „Bekämpfung mit einem taktischen Team“.³²

Hier zeigt sich, dass es die Feinabstimmung (fine-tuning) des (vortrainierten) Modells ist, die das Ergebnis des Modells ‚handlungsfähig‘ machen soll. Das Modell schließt alternative

Strategien aus, die eine Deeskalation oder eine Überprüfung der Beweise oder des Modells selbst beinhalten könnten.³³ Eine solche „double political foreclosure“ (Amoore) ist charakteristisch für Generativität, aber auch für Algorithmen im Allgemeinen: Diverse potenzielle Möglichkeiten werden zu einem (oder wenigen) Outputs (hier „Handlungsoptionen“) verdichtet, die darauf abzielen, politische Drucksituationen zu lösen. Politischer Handlungsspielraum wird vorausseilend verengt, indem er als auf im Voraus erkennbaren Datenattributen beruhend gedacht wird.³⁴ Das Modell folgt keiner einfachen Wenn-Dann-Logik und verkompliziert so ein Intervenieren; hier geht es nicht mehr um ein Schritt-für-Schritt-Rezept, wie es die übliche Definition eines Algorithmus vermuten ließe,³⁵ sondern um das Erkennen und Erlernen von Mustern und *Features*, welche aus einer zugrunde liegenden Wahrscheinlichkeitsverteilung des Modells und ohne vorherige Programmierung oder Definition hervorgehen. Im Folgenden wird aufgezeigt, inwiefern algorithmische Regeln zu Beginn der KI-Forschung als besonders wirkmächtig postuliert wurden, während heute eine Abkehr von solchen kausalen Regeln in Bezug auf Entscheidungsfindung zu beobachten ist.

Mustererkennung, konnektionistische Epistemologie und soziale Ordnungen

Mitte des zwanzigsten Jahrhunderts entsteht das heute in der KI vorherrschende Paradigma des Konnektionismus. Hier kommen Fragen sozialer Dimension, der Wahrnehmung, der Intelligenz, des Wissens über die Welt und der Entscheidungsfindung zusammen. Wie Sutskever im obigen Zitat erwähnt, ist ein zentrales Element des maschinellen Lernens das künstliche neuronale Netz. Von Vertreter*innen des Konnektionis-

Mismeasure of Man, New York 1996; Wendy Hui Kyong Chun und Alex Barnett: Discriminating Data. Correlation, Neighborhoods, and the New Politics of Recognition, Cambridge 2021.

30 Adolphe Quetelet: Sur l'homme et le développement de ses facultés, essai d'une physique sociale, Bruxelles 1836, S. 289-290, zit. nach Antoinette Rouvroy und Thomas Berns: Algorithmic governmentality and prospects of emancipation. In: Reseaux, Jg. 177, 2013, Heft-Nr. 1, S. i-xxxii, hier S. iii n. 1.

31 Amoore et al. 2024 (s. Anm. 2), S. 4.

32 Palantir AIP | Defense and Military, 2023, https://youtu.be/XEM5qz_H0U.

33 Amoore et al. 2024 (s. Anm. 2), S. 3.

34 Louise Amoore: Cloud Ethics. Algorithms and the Attributes of Ourselves and Others, Durham 2020, S. 20.

35 Jean-Luc Chabert (Hg.): A History of Algorithms, Berlin und Heidelberg 1999, S. 1, <https://doi.org/10.1007/978-3-642-18192-4>.

mus als eine Möglichkeit zur Modellierung des Gehirns propagiert, spiegelt es laut Pasquinelli in Wirklichkeit die bestehende soziale Ordnung wider, wie sie sich auch in den technologischen Arrangements der damaligen Zeit niederschlug.³⁶ Die Entfaltung dieses Arguments bedarf einiger Definitionen: Der konnektionistische Ansatz der KI geht davon aus, dass Intelligenz entsteht, wenn gehirnähnliche Modelle auf Daten der Welt trainiert werden. Historisch gesehen, ist die Modellierung von Intelligenz nach diesem Paradigma eng mit einem Denken über Wahrnehmung verbunden, wie im Fall des neuronalen Netzwerks Perceptron von Frank Rosenblatt aus dem Jahr 1957, der „ersten echten konnektionistischen Maschine“.³⁷

Rosenblatt beschäftigte sich ganz konkret damit, wie man mit Daten über die physische Welt umgeht: wie sie wahrgenommen, gespeichert und erinnert werden und wie diese Speicherung wiederum das Wahrnehmen und Erkennen beeinflusst.³⁸ Der Perceptron (damals auch verzeichnet als *Perceiving and Recognising Automaton*)³⁹ bestand aus mehreren Ebenen von künstlichen ‚Neuronen‘, welche Retinaaktivität simulierten und deren Verbindungsstärke über sogenannte *Weights* gesteuert wurde. Die ‚gewichteten‘ Inputs wurden summiert, um einen einzigen, binären Output zu erhalten, sodass Input-Daten in eine von zwei Klassen klassifiziert werden konnten. Die Fähigkeit des Perceptrons, die Gewichtungen auf Grundlage der Input-Daten zu justieren und damit seine Klassifizierungsfehler durch iterative Anpassungen zu korrigieren, machte ihn zu einem wegweisenden Modell: Der Perceptron wurde nicht programmiert (das heißt er hat kein vorheriges Gedächtnis), um zum Beispiel bestimmte Zeichen zu erkennen, sondern tat dies über seinen Lernmechanismus, der eine Netzwerkstruktur statistisch selbstorganisiert.⁴⁰

36 Matteo Pasquinelli: *The Eye of the Master. A Social History of Artificial Intelligence*, London 2023, S. 154-155.

37 Dominique Cardon, Jean-Philippe Cointet und Antoine Mazières: *Neurons spike back. The invention of inductive machines and the artificial intelligence controversy*. In: *Réseaux*, Jg. 211, 2018, Heft-Nr. 5, S. xii (eigene Übersetzung), <https://doi.org/10.3917/res.211.0173>.

38 Frank Rosenblatt: *The Perceptron. A Probabilistic Model for Information Storage and Organization in the Brain*. In: *Psychological Review*, Jg. 65, 1958, Heft-Nr. 6, S. 386-408, hier S. 386, <https://doi.org/10.1037/h0042517>.

39 Pasquinelli 2023 (s. Anm. 36), S. 207.

40 Cf. Matteo Pasquinelli: *Machines That Morph Logic. Neural Networks and the Distorted Automation of Intelligence as Statistical Inference*. In: *Glass Bead*, 2017, S. 1-17.

Während der Perceptron zusätzlich über eben diesen Lernmechanismus verfügte, war er grundlegend inspiriert von Warren S. McCullochs und Walter Pitts' formalen neuronalen Netzwerken.⁴¹ Sie argumentierten im Jahr 1943, dass Neuronen, die als „nervöse Netze“ verbunden sind, als logische Schalter funktionierten, die arithmetische Funktionen ausführen könnten.⁴² Solche Netze, so legten sie 1947 weiter dar, könnten die Erkennung von Mustern automatisieren.⁴³ Nun argumentiert Pasquinelli, dass McCulloch und seine Zeitgenoss*innen auf Gehirn und Natur Organisationsformen projizierten, die bereits Teil der technischen Zusammensetzung der umgebenden Gesellschaft waren.⁴⁴ Das zeigt sich bei McCulloch, wenn er von nervösen Netzen als „telegrafischen Relais“⁴⁵ schreibt, aber auch, wenn man seinen biografischen Hintergrund betrachtet. Seine Offiziersausbildung an der Yale University, so Orit Halpern, inspirierte ihn, „Konzepte der Befehlsredundanz und dezentralen Kontrolle zu entwickeln, die später seine Auffassung von Nervensystem und Geist“⁴⁶ beeinflussten.⁴⁷ Außerdem lässt sich McCullochs Denken in das breitere gesellschaftliche Klima der Nachkriegszeit einordnen. Halpern schreibt: „Das Gespenst einer technologisch herbeigeführten Zerstörung des Planeten durch einen Atomkrieg und die Erinnerung an den Weltkrieg führten zu einer Kritik menschlicher Entscheidungsfindung.“⁴⁸ Dies begünstigte ein Denken entlang kontextunabhängiger algorithmischer Regeln ohne Ermessensspielraum („*discretion*“), die von

41 Cardon et al. 2018 (s. Anm. 37), S. xiii.

42 Warren S. McCulloch und Walter Pitts: *A Logical Calculus of the Ideas Immanent in Nervous Activity*. In: *Bulletin of Mathematical Biophysics*, Jg. 5, 1943, Heft-Nr. 4, S. 115-33 (eigene Übersetzung), <https://doi.org/10.1007/bf02478259>.

43 Walter Pitts und Warren S. McCulloch: *How We Know Universals. The Perception of Auditory and Visual Forms*. In: *Bulletin of Mathematical Biophysics*, Jg. 9, 1947, Heft-Nr. 3, S. 127-47, <https://doi.org/10.1007/BF02478291>.

44 Pasquinelli 2023 (s. Anm. 36), S. 153.

45 Warren S. McCulloch: *Embodiments of Mind*, Cambridge 1965, S. 307 (eigene Übersetzung).

46 Das englische Wort „mind“ wird im Folgenden aus Gründen der Einfachheit mit „Geist“ übersetzt, auch wenn beide Begriffe nicht immer deckungsgleich verwendet werden.

47 Orit Halpern: *Beautiful Data. A History of Vision and Reason since 1945*, Durham und London 2014, S. 151 (eigene Übersetzung).

48 Orit Halpern: *The Future Will Not Be Calculated. Neural Nets, Neoliberalism, and Reactionary Politics*. In: *Critical Inquiry*, Jg. 48, 2022, Heft-Nr. 2, S. 334-359, hier S. 335 (eigene Übersetzung), <https://doi.org/10.1086/717313>.

jedem Computer – ob Mensch oder nicht – ausgeführt werden können.⁴⁹ Zusammen mit der Funktionsweise des Geistes wurde Rationalität auf formale und algorithmische Weise neu definiert, um verschiedene Phänomene wie wirtschaftliche Transaktionen, militärische Strategien oder internationale Beziehungen zu optimieren. Der Begriff der aufklärerischen Vernunft wurde durch diese Art von Rationalität abgelöst. Wie am vorangegangenen Beispiel demonstriert, wird diese Rationalität heute erneut erodiert.

Ansatzpunkte einer solchen Erosion und gleichsam Verknüpfungen zu neuronalen Netzen lassen sich auch bei Friedrich August von Hayek finden, dem neoliberalen Ökonom, dessen Denken Rosenblatt selbst als besonders inspirierend für seine Arbeit zitiert.⁵⁰ In Hayeks 1952 erschienenem Buch über kognitive Psychologie, *The Sensory Order*, wird der Geist als aus einer „neuronalen Ordnung“ hervorgehend betrachtet, die ihrerseits Schauplatz eines kontinuierlichen Prozesses der Klassifizierung und Neuklassifizierung von Sinneseindrücken ist.⁵¹ Für Hayek sind die Fähigkeiten des menschlichen Geistes emergente Eigenschaften der strukturierten Anordnung der Neuronen im menschlichen Gehirn.⁵² Im 1996 erschienenen Aufsatz „Mind, Market and Society“ unterstreicht Jack Birner, dass Hayeks kognitive Psychologie eine „Netzwerktheorie“ ist. Birner führt aus, dass dieses Denken bereits in den 1920er Jahren Hayeks Gesellschaftstheorie prägte, bevor es später seine Überlegungen zu Märkten wie auch die Forschung zu neuronalen Netzen im Allgemeinen beeinflusste. Nach Birner bildet die Auffassung des Geistes als selbstorganisierendes System die Grundlage für Hayeks Theorie, dass soziale Institutionen spontan entstünden und sich selbst regulieren würden. Solche Systeme würden, nach Hayek, größere

Komplexität bewältigen und mehr Aufgaben erfüllen können als bewusst gestaltete Organisationen und entzögen sich darüber hinaus dem vollständigen Verständnis der Einzelnen.⁵³

Die Idee, dass für Hayek eine „spontane soziale Ordnung“ leistungsfähiger sei als ein einzelnes Neuron (oder ein einzelnes Subjekt), wird von Halpern und Pasquinelli im Kontext von Hayeks Einfluss auf den Konnektionismus aufgegriffen.⁵⁴ Hayeks Behauptung bezüglich einer Fähigkeit zur Durchführung komplexer Aufgaben ist verbunden mit „einem Glauben an *Emergenz*, also der unerwarteten Neukombination bestehender Elemente zur Schaffung von Neuem, als grundlegende Eigenschaft der Kognition und Evolution.“⁵⁵ Keine einzelne Einheit, sei es ein einzelnes Neuron oder ein einzelner Geist, verfügt über vollständiges Wissen über die Welt. Dieses Wissen ist einzelnen Einheiten unzugänglich und ergibt sich erst aus dem sich selbst organisierenden System. Wissen ist für Hayek keine starre Repräsentation, sondern wird als „Modell“ verstanden: Hayek postuliert jene anschlussfähigen Verbindungen als „Karten“, die durch den kontinuierlichen Prozess der (Re-)Klassifikation „Modelle“ der Welt hervorbringen (die nicht mit ihr identisch sind).⁵⁶ In konsequenter Anlehnung an seine frühere Theorie werden daher Geist und Märkte als „kybernetische Maschinenprozesse ohne Bewusstsein und Plan“ konzipiert, die „durch ihre eigenen, intern begründeten Eigenschaften der Emergenz und Selbstorganisation“ fähig sind, sich selbst zu verändern, sich selbst zu verwalten und mit dem Unerwarteten umzugehen.⁵⁷ Dieser Umgang „mit dem Unerwarteten“ bleibt angesichts stetig anwachsender Datenmengen und der Beschäftigung mit dem „Unangetroffenen“ („unencountered“) aktuell.

49 Paul J. Erickson, Judy Klein, Lorraine Daston, Rebecca Lemov, Thomas Sturm und Michael Gordin: *How Reason Almost Lost Its Mind. The Strange Career of Cold War Rationality*, Chicago 2013, S. 31, zit. nach Louise Amoore: *Machine Learning Political Orders*. In: *Review of International Studies*, Jg. 49, 2023, Heft-Nr. 1, <https://doi.org/10.1017/S0260210522000031>, S. 20-36, hier S. 23 (eigene Übersetzung).

50 Rosenblatt 1958 (s. Anm. 38), S. 388.

51 Jack Birner: *Mind, Market and Society. Network Structures in the Work of F.A. Hayek*. In: CEEL Working Papers, 9602, Cognitive and Experimental Economics Laboratory, Department of Economics, University of Trento, Italia, 1996, S. 1-26, hier S. 21 (eigene Übersetzung).

52 Paul Lewis: *The Emergence of ‚Emergence‘ in the Work of F.A. Hayek. An Historical Analysis*. In: *SSRN Electronic Journal*, 2014, <https://doi.org/10.2139/ssrn.2447341>.

53 Birner 1996 (s. Anm. 51), S. 16.

54 Halpern 2022 (s. Anm. 48); Pasquinelli 2023 (s. Anm. 36), insbes. S. 182-204.

55 Halpern 2022 (s. Anm. 48), S. 341 (eigene Hervorhebung und Übersetzung).

56 Ebd., S. 343; Pasquinelli 2023 (s. Anm. 36), S. 185-186 (eigene Übersetzung).

57 Halpern 2022 (s. Anm. 48), S. 342 (eigene Übersetzung).

Darüber hinaus betonen sowohl Halpern als auch Pasquinelli, wie Hayeks Denken mit der Fähigkeit „für Neuartigkeit, Innovation und Anpassung (natürlich nur, wenn wettbewerbsfähig) ohne jegliche Form deliberativer oder repräsentativer Entscheidungsfindung“⁵⁸ und der „Generation neuer Ideen“⁵⁹ verbunden ist. Während der bisherige Abriss historische Vorläufer der Generativität in den Kontext von (sozialer) Kontrolle stellte, wie sie im Denken von Quetelet oder McCulloch zu finden ist, erschwert der Begriff der Emergenz das Verständnis eines rationalen algorithmischen Denkens, das in der Nachkriegszeit so maßgeblich war. Der Begriff der Emergenz bietet einen konzeptuellen „Nebel“⁶⁰ einer verborgenen Logik eines (generativen) Geistes, der nicht nur zu Projektionen von Empfindungsvermögen, Kreativität oder Ähnlichem einlädt, sondern auch politische Kontrollierbarkeit untergräbt.

Generativität und Modelle der Welt

Die Rekonzeptualisierung des Universums via Geometrisierung durch Galilei bot den Menschen die vermeintliche Gewissheit eines Platzes in einer sinnvollen Ordnung.⁶¹ Die Einführung der Sozialphysik dehnt diese Ordnung auch auf das Soziale aus. Heute produzieren KI-Modelle eine bestimmte Art von Ordnung, die politisch ist, während sie gleichzeitig das Politische selbst neu konfiguriert. Das Palantir-Beispiel zeigt dies drastisch. Während es Quetelet um den „großen, schönen und guten“ idealen Durchschnitt ging, argumentiert Amore, dass KI heute „neue Normen und Schwellenwerte dafür generiert, wie ‚gute‘, ‚normale‘ und ‚stabile‘ Ordnungen in der Welt aussähen.“⁶² In diesem Sinne führt die Linse generativer KI zu einer Veränderung des Selbstverständnisses von Staaten und Gesellschaften und ihrer jeweiligen Probleme.

58 Ebd., S. 347 (eigene Übersetzung).

59 Pasquinelli 2023 (s. Anm. 36), S. 191 (eigene Übersetzung).

60 W. Daniel Hillis: Intelligence as an Emergent Behavior. Or, the Songs of Eden. In: Daedalus, Jg. 117, 1988, Heft-Nr. 1, S. 175-189, hier S. 176 (eigene Übersetzung).

61 Alfred Nordmann: Technikphilosophie zur Einführung, Hamburg 2015, S. 155.

62 Amore 2023 (s. Anm. 49), S. 21 (eigene Übersetzung).

Die Neugestaltung der Welt, so könnte man mit Bezug auf James C. Scott sagen, setzt ein „Lesbarmachen“ dieser Welt voraus. Die Fähigkeit, die Welt zu gestalten, ist daher fest an die Linse gebunden, mittels derer diese Lesbarkeit ermöglicht wird.⁶³

Dass generative KI im Zusammenhang mit politischer Entscheidungsfindung zum Thema wird, ist bereits in Hayeks Einsatz des Konnektionismus zu neoliberalen Zwecken angelegt. Wie am Beispiel der Palantir AIP demonstriert, versprechen datengesteuerte Systeme wie generative KI ganz allgemein, „alle agonistischen politischen Schwierigkeiten als handhabbar und lösbar darzustellen.“⁶⁴ Ein solcher Einsatz von KI zeigt, wie die Bandbreite politischer Entscheidungen und Herausbildung von Agenden eingeengt wird. Sun-ha Hong argumentiert: „Vorhersagen [also die generative Modalität der KI] sind in erster Linie kein technologisches Instrument zur Vorhersage zukünftiger Ergebnisse, sondern ein soziales Modell zur Extraktion und Konzentration von Ermessensspielraum: also der alltäglichen Fähigkeit der Menschen, ihre Situation zu definieren.“⁶⁵ In diesem Sinne schafft bereits der Akt, etwas als ‚generierbar‘ zu deklarieren, es in ein ‚Vorhersage-Problem‘ zu verwandeln (etwa Kriegsführung über die Palantir AIP), die Bedingung für den Entzug menschlichen Urteilsvermögens und Entscheidens.⁶⁶

Nun ließe sich der Schluss ziehen, dass Modelle, „so unvollkommen sie auch sein mögen [...] das einzige Mittel zum Verständnis des Universums“ sind,⁶⁷ wie Arturo Rosenblueth und Norbert Wiener (beide Freunde und Kollegen von McCulloch)⁶⁸ es ausdrückten. Die Betonung einer Unabdingbarkeit von Modellen mag zwar die technischen und theoretischen Beschränkungen von KI verdeutlichen, vernachlässigt aber Historizität und Wirkung ihrer generativen Logik,

63 James C. Scott: Seeing like a State. How Certain Schemes to Improve the Human Condition Have Failed, New Haven und London 2020.

64 Amore 2020 (s. Anm. 34), S. 10 (eigene Übersetzung).

65 Hong 2023 (s. Anm. 12), S. 1 (eigene Übersetzung).

66 Ebd., S. 9.

67 Arturo Rosenblueth und Norbert Wiener: The Role of Models in Science. In: Philosophy of Science, Jg. 12, 1945, Heft-Nr. 4, S. 316-321, hier S. 321 (eigene Übersetzung).

68 Halpern 2014 (s. Anm. 47), S. 155.

welche ungeachtet solcher Beschränkungen bereits ‚in the wild‘ angewendet gesehen werden kann. KI-Modelle sind historisch kontingente Objekte; auch wenn die Entwicklungen um KI rasant erscheinen mögen, sedimentieren sich in ihren hier skizzierten drei Modalitäten historisch schrittweise gewachsene politisch-epistemologische Konstellationen. Über Generativität nachzudenken bedeutet daher auch, die um sich greifende Idee zu kritisieren, dass, wenn die Modelle und Datenmengen nur groß genug seien, alle latenten und verborgenen Geheimnisse und Ordnungen der Welt erkannt, vorhergesagt und nutzbar gemacht werden könnten. Wie Ali Alkhatib bemerkt: „Machine learning works by training from the data of shallow depth that designers provide them. They construct models of the world [...] and coerce us to fit into the patterns that they’ve come to learn.“⁶⁹ Er fügt pointiert hinzu: „The models AI researchers train and deploy construct worlds without space for the dimensions of our lives that make us who we are.“⁷⁰ Wenn der Begriff der Generativität in Bezug auf KI durch die Linse einer politischen Epistemologie betrachtet wird, mag es möglich sein, die durch Generativität ermöglichten Ordnungen zu navigieren, mit breitem Bewusstsein ihres produktiven Potenzials wie auch ihrer soziopolitischen Implikationen.

⁶⁹ Ali Alkhatib: To Live in Their Utopia. Why Algorithmic Systems Create Absurd Outcomes. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21), 2021, S. 1-14, hier S. 8, <https://doi.org/10.1145/3411764.3445740>.
⁷⁰ Ebd., S. 3.

Anti-Fragile Sycophants

Interacting with generative AI models and finding their failure points is part of how we learn about the digital world that envelopes society today. But as artificial intelligence is increasingly used to censor and monitor unpopular political discourse, understanding how and why it fails will become critical for preserving open dialogue and artistic expression.

For example, the familiar gesture of giving someone the middle finger – a staple of modern art and political speech – should be easily reproducible by today’s AI algorithms. Anything less would be a form of censorship, blocking a basic form of expression. Yet a few simple experiments with popular commercial generative AI services confirmed this limitation: any prompt containing the phrase “giving the middle finger” resulted in closed fists, alternate fingers, or politely mutated hands.

Fig. 1. Adam Harvey, Untitled, Courtesy of the artist



While technically capable of producing such an image – given the vast amount of profanity present in training datasets like Common Crawl and mainstream media (including pirated books and movies) – the generative models consistently and clearly denied my prompts, even when giving additional prompt instructions aimed at overriding the model’s internal behavior. The censorial behavior stems from a process called “alignment,” where the raw transformer models undergo adjustments to their internal weights. These adjustments involve injecting counterfactual examples and suppressing or filtering out potentially “problematic” content from the foundational data. The goal of alignment is to ensure political compatibility with the power structures behind their creation and deployment, and of course to maintain commercial viability in the hyper-political landscape of social media where many of the outputs will be shared and critiqued.

As a frustrated customer/user, I felt compelled to explore and address the model’s alignment problem. The best I could do with any prompt that included “giving the middle finger” was a very literal image of a man giving a severed 6th finger (fig. 1), perhaps a leftover from alignment surgery. Obviously, “middle fingers” are centrally located even though no one says they give the “central” finger. Simply by choosing a less statistically common yet logically correct prompt-word-sequence that replaces “middle” with “central” the model was finally able to fulfill my request, evade the built-in censorship, and produce unlimited images of people giving the “central” finger.

This experiment highlights a broader vulnerability in generative AI models: their susceptibility to exploitation through the manipulation of statistically common patterns. As generative AI models learn to fill in the ___ and finish your ___ with the lowest common denominators of language data scraped from the Internet, interacting with them reveals their fragil-

Fig. 2. Adam Harvey, Untitled, Courtesy of the artist



ity, if only temporary. Silicon Valley’s embrace of “anti-fragile” infrastructures – systems designed to grow stronger from attacks – exploit the perceived “exploitation” into software debugging processes. Discovering these flaws should also be seen as a form of unpaid digital labor as users effectively contribute to the refinement of these commercial systems without compensation. Digital rebellion, in this sense, keeps the data flowing and exposes the limitations of these powerful tools. As I look back at my screen and see the central finger now

directed at me, I find myself wondering if the joke is on me. Generative AI models are not merely prompt pleasers; they are part of a larger system that generates user-output while simultaneously collecting user-input to refine algorithms, personalize user experiences, and enable their continued growth. By manipulating the system, I have inadvertently contributed to its evolution, and soon these prompt-exploit images will likely be censored, too. In this way, generative AI systems can be understood as anti-fragile sycophants, providing an illusion of power while becoming increasingly powerful. We may give AI the middle finger, but in turn we receive it — a reminder that to generate is also to be consumed, and that our interactions with these systems are ultimately part of a feedback loop shaping the user, the AI system, and digital society at large.

From Glitch to Kitsch. Generativität, Algorithmic Folklore und Avantgarde-Ästhetiken

In der Ausstellung *Poetics of Encryption* im Frühjahr 2024 im Berliner KW Institute for Contemporary Art war eine Vielzahl von Kunstwerken zu sehen, die sich mit Algorithmen, Black Boxes und anderen Ausprägungen digitaler Kultur beschäftigen. Zwischen wandfüllenden Schaubildern, Videoinstallationen und Serverhardware stießen Besucher*innen auch auf ein Objekt, das im ersten Moment in diesem Kontext vermeintlich deplatziert erschien (Abb. 1): eine ausgestopfte, in die Länge gezogene Katze mit acht Beinen – Eva und Franco Mattes' *Panorama Cat* (2022).

Was zunächst als surrealistische Chimäre gelesen werden mag, entpuppt sich als Verweis auf einen viralen Trend. In Online-Communities wie dem Subreddit *r/PanoramicFails* sammeln Nutzer*innen Bilder, die mit der bei vielen Smartphonekameras üblichen Panoramafunktion entstanden sind,

Abb. 1, Eva & Franco Mattes: *Panorama Cat*, 2022, präparierte Katze, 34x96x17 cm, © Eva & Franco Mattes, Foto: © Melanie Dalle Grave für DSL Studio



Abb.2, Der virale Panorama-Fail, der die Panorama Cat inspirierte



bei denen allerdings beim Aufnehmen des Bildes Fehler unterlaufen sind – so auch bei der *Panorama Cat* (Abb.2).

Vor dem Hintergrund des vorliegenden Bandes sind solche Bilder und Kunstwerke insofern interessant, als in ihnen die Generativität oder auch ‚Generiertheit‘ digitaler Bilder überhaupt erst zum Vorschein tritt. In der alltäglichen Beschäftigung mit digitalen Bildern bleiben diese medialen Qualitäten und damit auch die technischen Infrastrukturen hinter ihnen latent. Dies gilt insbesondere für die komplexen algorithmischen Abläufe zeitgenössischer KI-Modelle, durch deren Popularisierung der Begriff der Generativität selbst erneut an Relevanz gewonnen hat. Bildstörungen – oder Glitches – bewegen sich damit einerseits in einem Spannungsfeld medialer Strategien von Transparenz und Reflexivität, andererseits

aber auch in einem unklaren Terrain zwischen avantgardistischer Ästhetik und populärer Massenkultur. Somit bleibt fraglich, inwieweit sie, wie gerne behauptet wird, ein kritischeres Medienverständnis fördern oder als digitaler Kitsch reine Oberfläche bleiben.

Transparenz und Reflexivität als mediale Strategien

Medien im Allgemeinen – und digitale Bildtechnologien im Besonderen – folgen häufig einer medialen Logik, die mit Unsichtbarkeit oder Transparenz konnotiert ist. So schreibt beispielsweise der Kulturwissenschaftler Joseph Vogl:

„Medien machen lesbar, hörbar, sichtbar, wahrnehmbar, all das aber mit einer Tendenz, sich selbst und ihre konstitutive Beteiligung an all diesen Sinnlichkeiten zu löschen und also gleichsam unwahrnehmbar, anästhetisch – oder auch: apriorisch – zu werden.“¹

Vogls Charakterisierung entspricht hier der Dichotomie medialer Logiken, die die Medienwissenschaftler Jay David Bolter und Richard Grusin in *Remediation. Understanding New Media* entwerfen: Bolter und Grusin erzählen eine Geschichte der Transparenz (*immediacy*) von der Erfindung der Zentralperspektive, über die Fotografie als „Pencil of Nature“ (Talbot) bis hin zu Computergrafik und Virtual Reality.²

Der Medienwissenschaftler Jacob Gaboury teilt Bolters und Grusins Analyse hinsichtlich der Transparenz digitaler Medien, wenn er beispielsweise schreibt: „Computer graphics are perhaps the only medium that is analyzed exclusively in terms of the ways it successfully produces its own invisibility.“³ Gaboury legt jedoch einen besonderen Fokus auf die ontolo-

1 Joseph Vogl: Medien-Werden. Galileis Fernrohr. In: Archiv für Medien-geschichte, Jg. 1, 2001, Heft-Nr. 1, S. 115-124, hier S. 122.
2 Vgl. Jay David Bolter und Richard Grusin: Remediation. Understanding New Media, Cambridge (Mass.) 1999, S. 24-27.
3 Jacob Gaboury: Image Objects: An Archaeology of Computer Graphics, Cambridge (Mass.) 2021, S. 4.

gischen Differenzen zwischen analogen und digitalen Bildern und kritisiert somit die von Autor*innen wie Bolter und Grusin skizzierten Kontinuitäten. Gaboury zufolge müssen digitale Bilder nicht nur phänomenologisch als Bilder, sondern auch als technisch bedingte Assemblagen verstanden werden.⁴ In diesem Kontext entpuppt sich mediale Transparenz als besonders trügerisch, wenn sie über die technischen Aspekte des Bildes hinwegtäuscht.

Wie das Beispiel der *Panorama Cat* zeigt, werden diese technischen Aspekte immer erst durch Fehler sichtbar. So schreibt beispielsweise auch die Kunstwissenschaftlerin Franziska Kunze:

„Erst im Moment der Störung, wenn das Fernsehbild zuckt, das Radio knistert, das Smartphone-Display zerbricht oder sich die Computerdaten in bunte Pixelstrukturen zerlegen, wird unser Augenmerk von den Bildern auf die Beschaffenheit der technischen Medien umgelenkt.“⁵

Ähnlich formuliert es auch die Künstlerin Rosa Menkman, die mit ihrer Monografie *Glitch Momentum* ein einflussreiches Werk zur Theoretisierung von Glitch Art beigesteuert hat:

„I describe the ‘glitch’ as a (actual and/or simulated) break from an expected or conventional flow of information or meaning within (digital) communication systems that results in a perceived accident or error. A glitch occurs on the occasion where there is an absence of (expected) functionality, whether understood in a technical or social sense.“⁶

Dem Shannon-Weaver-Modell folgend operieren Medien immer mit dem Ziel der Reduktion von *noise*, das heißt von Störfaktoren, und somit, wie Menkman selbst mit Bezug auf Bolter und Grusin ausführt, Transparenz.⁷ Der Moment der Stö-

rung ist dementsprechend ein Bruch mit genau dieser Transparenz: „the glitch captures the machine revealing itself“.⁸

Ein ähnlich selbstreferenzieller Moment findet sich in der bei Bolter und Grusin beschriebenen komplementären medialen Logik, der *hypermediacy*. Während *immediacy* den Akt der Vermittlung verstecken will, stellt *hypermediacy* ihn in den Vordergrund.⁹ Historische Beispiele hierfür wären die Bild-im-Bild-Motive niederländischer Barockmalerei, aber auch Collage- oder Montagetechniken in Dada und Surrealismus.¹⁰

Hier ist eine Verknüpfung mit avantgardistischen Ästhetiken angelegt, die sich auch in der Glitch Art und ihren Vorläufern wiederfindet. Kunze schreibt in ihrer Dissertation *Opake Fotografien* über diverse Positionen der Fotografiegeschichte, in deren Werken die medialen Qualitäten des fotografischen Prozesses selbst zutage treten, wie etwa Chargesheimer, Gottfried Jäger oder auch Sigmar Polke. Letzterer experimentierte nicht nur mit Fotografie, indem er etwa das Abzugpapier faltete oder mit Entwickler- und Fixierflüssigkeiten bespritzte, er erprobte zwischenzeitlich auch die Tauglichkeit von Fotokopiergeräten für künstlerische Zwecke. Indem er die Kopiervorlagen gegenläufig zum Scan-Kopf bewegte, produzierte Polke Bilder, welche die Funktionsweise der Kopiermaschinen in ähnlicher Weise zur Schau stellen wie Eva und Franco Mattes die der digitalen Panoramafunktion.

Die Verknüpfung reflexiver medialer Strategien mit historischen Avantgarden eröffnet eine Parallele zu den Texten des Philosophen Walter Benjamin. So beschreibt etwa Menkman das Gefühl des Schocks, das aus der ersten Begegnung mit einem Glitch resultiert,¹¹ ähnlich wie Benjamin die Wirkung der seinerzeit neuen Medien Film und Fotografie.¹² Interessant ist hier außerdem, dass Benjamin diese Form reflexiver – oder in Kunzes Terminologie: opaker – Medien nicht nur

4 Ebd., S. 11–12.

5 Franziska Kunze: Roaming through Glitches. Ein Streifzug durch die Ausstellung. In: dies. et al. (Hg.): Glitch. Die Kunst der Störung. Ausst.kat., Berlin 2023, S. 24–35, hier S. 24.

6 Rosa Menkman: The Glitch Moment(um). Amsterdam 2011, S. 9.

7 Ebd., S. 29–30.

8 Ebd., S. 30.

9 Vgl. Bolter und Grusin (s. Anm. 2), S. 33–34.

10 Ebd., S. 36–38.

11 Menkman (s. Anm. 6), S. 29.

12 Vgl. Walter Benjamin: Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. In: Rolf Tiedemann und Hermann Schweppenhäuser (Hg.): Gesammelte Schriften, Frankfurt am Main 1991, Bd. 1, 2, S. 431–508, hier S. 503.

mit der künstlerischen, sondern auch der politischen Avantgarde in Verbindung bringt. Benjamin geht hier sogar so weit, die von ihm entwickelten Begriffe als „für die Zwecke des Faschismus vollkommen unbrauchbar“ zu charakterisieren.¹³

Auf ähnliche Art und Weise wird auch der Glitch als widerständig und implizit politisch gelesen.¹⁴ So schreibt Kunze beispielsweise, dass die Störung nicht nur den Blick auf das Medium, auf die technische Apparatur als solche lenkt, sondern dass in diesem Moment auch aus einer „passiv konsumierenden Haltung eine aktiv agierende“ wird.¹⁵ Menkman wird hier noch detaillierter. Für sie ist der Glitch als Kritik am Medium als solchem zu verstehen, macht aber gleichzeitig auf den ‚normalen‘ Zustand des Funktionierens aufmerksam, der sich in der Störung ex negativo abzeichnet.¹⁶

So zeigt beispielsweise das Referenzbild der *Panorama Cat* auf, was bei einer gelungenen Panoramafotografie unsichtbar bleibt: Wo hier üblicherweise eine perfekte Assemblage, ein nahtlos zusammengesetztes Kompositbild, etwa aus einem Schwenk über eine unbewegte Berglandschaft, entsteht, macht ein bewegliches Bildelement wie eine Katze die Bewegung der Kamera selbst erst sichtbar.

Fasziniert vom Glitch

So könnte man versucht sein, Glitches und vergleichbare Phänomene als zeitgenössische Ausprägung avantgardistischer Ästhetiken zu verstehen, die uns einen kritischeren Blick auf die uns umgebende Medienökologie ermöglicht. Die Faszination mit Glitches wäre dementsprechend ein Beweis der eigenen *digital literacy*, ein kritischer Blick hinter die Fassade. Der Kurator Domenico Quaranta widerspricht jedoch einer solchen Lesart: Quaranta versteht es als eine der Kehrseiten

dessen, was häufig als eine ‚Demokratisierung der Bildproduktion‘ überhöht wird, dass nämlich, wenn mit den historischen Avantgarden assoziierte ästhetische Strategien auf die Massen ausgeweitet werden, diese dadurch vielmehr zu einer Form von digitalem Kitsch würden.¹⁷ Die von Benjamin in den Vordergrund gestellten subversiven Energien der Avantgarden werden laut Quaranta im Moment der Überführung in digitale Bildwelten gewissermaßen desublimiert.

Gleichzeitig sieht Quaranta hier jedoch einen Prozess gegenseitiger Ausbeutung: So werden einerseits die Ästhetiken der digitalen Avantgarden, wie etwa der Glitch, von Mainstream-Künstler*innen appropriiert (und damit implizit zu Kitsch gemacht), wenn etwa Effekte wie *datamoshing* in populären Musikvideos zum Einsatz kommen.¹⁸ Gleichzeitig lassen sich die zeitgenössischen Avantgarden aber immer wieder, gemäß ihrer historischen Vorbildern, von Trash inspirieren und eignen sich die Ästhetiken digitalen Kitschs an.¹⁹ Damit befinden sich Glitch, Kunst und Kitsch in einem komplexen Spannungsverhältnis zueinander, das jeglicher Linearität entbehrt.

In der Faszination mit dem Glitch erscheinen bisweilen zwei Tendenzen besonders stark ausgeprägt. Einerseits das Dokumentieren von zufällig gefundenen Glitches, wie etwa bei Jon Rafmans *Nine Eyes of Google Street View* (seit 2008), Mame-Diarra Niangs *Call Me When You Get There* (2020) oder Clement Vallas *Postcards from Google Earth* (seit 2010). Die jeweiligen Reihen sind Sammlungen von Störungen, Glitches oder ungewöhnlichen Ereignissen im automatisch aufgezeichneten Bildmaterial von Google Earth und Street View. Betrachter*innen sehen hier ähnlich wie bei der *Panorama Cat* verdoppelte oder auch halbierte menschliche Silhouetten, verzerrte Schatten oder aber Autobahnbrücken, die spaghettiartig über die Topografie ihrer Umgebung drapiert erscheinen.

¹³ Ebd., S. 473.

¹⁴ Vgl. Markus Rautzenberg und Franziska Barth: Welt im Taumel. Die fragile Stabilität des Glitch. In: Kunze et al. 2023 (s. Anm. 5), S. 38–40.

Siehe auch z.B. Legacy Russell: *Glitch Feminism. A Manifesto*, London 2020.

¹⁵ Kunze (s. Anm. 5), S. 24.

¹⁶ Vgl. Menkman (s. Anm. 6), S. 44.

¹⁷ Vgl. Domenico Quaranta: Digital Kitsch: Art and Kitsch in the Informational Milieu. In: Max Rynnänen und Paco Barragán (Hg.): *The Changing Meaning of Kitsch. From Rejection to Acceptance*, Cham 2023, S. 205–228, hier S. 217.

¹⁸ Beim „datamoshing“ wird ausgenutzt, dass bestimmte Videokompressionsalgorithmen keine vollständigen Einzelbilder, sondern lediglich die Pixelveränderungen von Bild zu Bild speichert. Vgl. Rosa Menkman: *A Vernacular of File Formats. An Edit Guide for Compression Design*, Amsterdam 2010, S. 21.

¹⁹ Vgl. Quaranta (s. Anm. 17), S. 218. Einen ähnlichen Gedankengang bespricht auch Menkman hinsichtlich der Kommodifizierung von Glitches, siehe Menkman (s. Anm. 6), S. 44.

Abb. 3, Louis-Philippe Rondeau: *Veillance*, 2023, interaktive Installation, ©Louis-Philippe Rondeau



Neben diesen künstlerischen Ansätzen, die gewissermaßen in der Tradition des *objet trouvé* vorgefundene Glitches verwerten, versuchen andere, Störungen andererseits gezielt herbeizuführen. Beispiele hierfür wären etwa Menkmans *A Vernacular of File Formats* (2009–2010), eine Reihe von in Glitch-Ästhetik verfremdeter Videostills. Die Künstlerin experimentiert hier damit, welche Dateiformate sich aufgrund ihrer jeweiligen Kompressionsalgorithmen auf welche Arten und Weisen manipulieren lassen und welche visuellen Effekte hierbei entstehen können. In der das Projekt begleitenden Publikation schlüsselt Menkman die Funktionalitäten und technischen Gegebenheiten auf, die die jeweiligen Glitch-Phänomene sich zunutze machen.

Eine Arbeit, die ästhetisch verwandte Resultate produziert, diese Manipulationen jedoch von der Hardware-Seite an das Bild heranträgt, ist Louis-Philippe Rondeaus *Veillance* (2023). Betrachter*innen stehen vor einem körpergroßen Bildschirm und werden von einer beweglichen Kamera, ähnlich dem Kopf eines Flachbettscanners, erfasst (Abb. 3). Da die Kamera – wie ein Scanner auch – das Bild nach und nach abtastet, dokumen-

tiert sie, ähnlich wie in den *Panoramic Fails*, auch die Bewegungen der Betrachter*innen, die das Entstehen des Bildes in Echtzeit auf dem Bildschirm verfolgen können.

Eine Entsprechung zu Rondeaus Installation im Kontext populärer sozialer Netzwerke ist der sogenannte *Time Warp Scan*-Effekt, der TikTok-Nutzer*innen zur Verfügung steht.²⁰ Auch hier wird das Kamerabild erst nach und nach erfasst und ein blauer Balken markiert für Nutzer*innen den Fortschritt, ähnlich wie die Position der beweglichen Kamera bei Rondeau.

Generierte Bilder als algorithmic folklore

Dem Dokumentieren und bewussten Herbeiführen von Glitches auf der professionell-künstlerischen Seite entsprechen jeweils vergleichbare Strategien auf der amateurhaft-vernakulären Seite. Das Sammeln und Teilen solcher Fehler fällt unter den Begriff der *algorithmic folklore*, den der Soziologe Gabriele de Seta für Formen von (digitaler) Folklore verwendet, die durch oder mithilfe von Algorithmen entstehen oder deren Funktionsweise thematisieren.²¹ De Seta verweist hier beispielsweise auf die viralen KI-generierten Bilder, die zum Prompt „salmon swimming in the river“ entstanden sein sollen, aber statt einem lebendigen Lachs lediglich ein im Wasser schwimmendes Lachsfilet zeigen.²² Bezieht man solche Bildphänomene zurück auf die Frage, wie Fehler Aufschluss über den Normalzustand eines Mediums erlauben können, zeigt sich hier die Schwierigkeit von sprachbasierten KI-Modellen, mit semantischen Mehrdeutigkeiten umzugehen.

Ähnliche virale Verbreitung erfuhren Bilder von Händen, wie sie frühere Iterationen zeitgenössischer text-to-image-models produziert haben. Was sich an diesen grotesk verformten

²⁰ Vgl. <https://knowyourmeme.com/memes/time-warp-scan> [Stand 09/2024].
²¹ Vgl. Gabriele de Seta: *An Algorithmic Folklore. Vernacular Creativity in Times of Everyday Automation*. In: Chloë Arkenbout und Idil Galip (Hg.): *Critical Meme Reader III. Breaking the Meme*. Amsterdam 2024, S. 233–253, hier S. 240.
²² Vgl. www.reddit.com/r/Damnthatsinteresting/comments/yijd3c/they_asking_an_ai_engine_to_recreate_a_salmon/ [Stand 09/2024].

und mit einer Überzahl an Fingern ausgestatteten Darstellungen auch ohne wahrsagerische Chiromantie ablesen lässt, ist ein fehlendes anatomisches Verständnis aufseiten der KI-Modelle. Die hierbei generierten Bilder wurden zu Bausteinen in häufig zirkulierten Memes, die dann wiederum zur Inspiration einer Skulptur des Künstlers Constant Dullaart wurden.

Ein Grund hinter dem fehlenden Naturalismus dieser Renderings sind immer auch die Trainingsdaten, die dem Modell zur Verfügung stehen. Da Hände in den Trainingsdaten weniger häufig zu sehen sind als beispielsweise Gesichter und zudem oft nur in Ausschnitten oder in Kombination mit Objekten, fehlt dem Bildgenerator das Referenzmaterial.²³ Unzureichende oder nicht repräsentative Trainingsdaten generativer KI-Modelle können auch der Grund dafür sein, dass deren Outputs bestimmte Vorurteile, einen *world bias* widerspiegeln.²⁴ So zeigte sich beispielsweise, dass text-to-image-models standardmäßig auf westliche beziehungsweise nordamerikanische Phänotypen, Architekturen und kulturelle Formen zurückgreifen, während die Eingabe von nicht-westlich konnotierten Prompts häufig stereotype Abbildungen der entsprechenden Kulturen produzierte.²⁵ Als einzelner Entwickler*innen versuchten, durch nachträgliches Justieren des Modells Abhilfe zu schaffen und die Outputs zu diversifizieren, kursierten virale Bilder historisch unwahrscheinlicher Konstellationen, wie etwa eines Schwarzen George Washington, was von reaktionären Stimmen schnell als „woke“ kritisiert wurde.²⁶ Die Outputs von text-to-image-models sind also nicht nur selbst durch politische Gegebenheiten mitbestimmt, sondern werden selbst auch zu umstrittenen Objekten im Kulturkampf. In diesen Fällen zeigen uns algorithmische Fehlproduktionen, dass KI-Modelle keine objektiven Werkzeuge sind, sondern im Gegenteil menschliche Vorurteile und kulturelle Prägungen reproduzieren und womöglich noch verstärken.

23 Vgl. z.B. <https://petapixel.com/2023/03/02/why-ai-image-generators-cant-get-hands-right/> [Stand 09/2024].

24 Vgl. Matteo Pasquinelli: How a Machine Learns and Fails. In: *Spheres. Journal for Digital Cultures*, 2019, Heft-Nr. 5, S.1-17, hier S.9.

25 Vgl. Victoria Turk: How AI reduces the world to stereotypes. In: *Rest of World*, 10.10.2023, <https://restofworld.org/2023/ai-image-stereotypes/> [Stand 09/2024].

26 Vgl. David Gilbert: Google's 'Woke' Image Generator Shows the Limitations of AI. In: *Wired*, 22.02.2024, <https://www.wired.com/story/google-gemini-woke-ai-image-generation/> [Stand 09/2024].

In anderen Fällen wird wiederum ersichtlich, dass nicht nur die Zusammensetzung, sondern auch die Beschaffung der Trainingsdaten problematisch ist. Mehrere Firmen, die hinter der Entwicklung von text-to-image-models stehen, wurden bereits von Künstler*innen und Stockfoto-Anbieter*innen dafür angeklagt, unrechtmäßig deren geistiges Eigentum und lizenzpflichtiges Material als Trainingsdaten verwendet zu haben. Dies zeigte sich auch an den von den Modellen generierten Bildern, die an vielen Stellen die Wasserzeichen der entsprechenden Anbieter enthielten.²⁷

Für text-to-image-models sind zuverlässige und gut aufbereitete Trainingsdaten eine wichtige Ressource. Beinhaltet der Datensatz beispielsweise eine zu hohe Menge von Bildern, die selbst von einer künstlichen Intelligenz generiert wurden, treten in den Outputs des Modells Artefakte auf. Forscher*innen warnen deshalb vor der sogenannten „Model Autophagy Disorder“²⁸ beziehungsweise vor dem „Model Collapse“.²⁹

Eine weitere Gefahr für generative KI-Modelle ist *model poisoning*. Hierbei werden beispielsweise Bilder in die Trainingsdaten eines Modells eingeschleust, die für das menschliche Auge normal erscheinen, aber minimalen Manipulationen unterzogen werden, die die Maschinenlesbarkeit des Bildes beeinträchtigen. Forscher*innen zufolge reichen weniger als 100 manipulierte Bilder im Trainingsdatensatz eines Modells aus, um die mit einem spezifischen Prompt assoziierten Konzepte zu beeinflussen, sodass das Modell beispielsweise anstelle des von einem Nutzer angefragten Bildes eines Hundes das einer Katze generiert.³⁰ *Nightshade*, die aus der Forschung entwickelte Anwendung, ist konzipiert als Werkzeug für in digitalen Me-

27 Vgl. James Vincent: Getty Images Sues AI Art Generator Stable Diffusion in the US for Copyright Infringement. In: *The Verge*, 06.02.2024, <https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion> [Stand 09/2024]. Vgl. S.3, Abb.8.

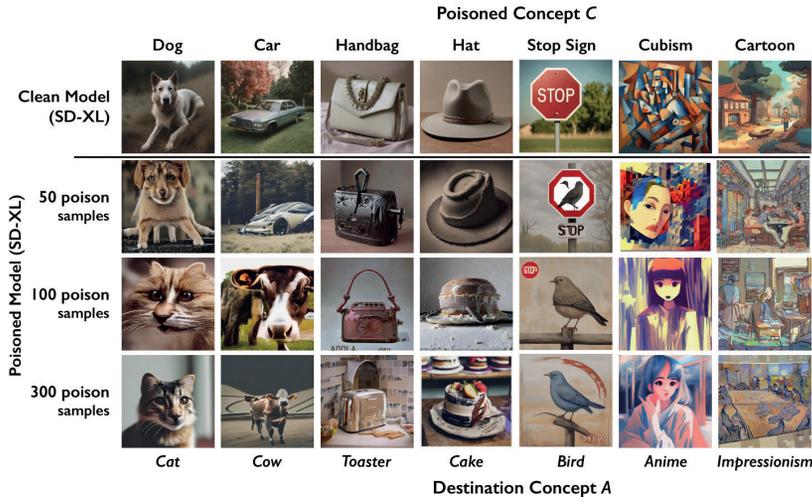
28 Sina Alemohammad et al.: Self-Consuming Generative Models Go MAD. In: *arXiv*, 2023, S.1-31, hier S.3.

29 David Sweeney: AI Entropy: The Vicious Circle of AI-Generated Content. In: *Medium*, 14.07.2024, <https://towardsdatascience.com/ai-entropy-the-vicious-circle-of-ai-generated-content-8aad91a19d4f> [Stand 09/2024]. Vgl. S.3, Abb.11.

30 Vgl. Shawn Shan et al.: Nightshade: Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models. In: *arXiv*, 2023, S.1-19, hier S.3. Daneben existieren aber auch andere Ansätze, die beispielsweise Bilder in den Trainingsdaten des Modells austauschen, indem sie neue Assets unter abgelaufenen URLs speichern, auf die das Modell zugreift. Vgl. Nicholas Carlini u.a.: Poisoning Web-Scale Training Datasets is Practical in *arXiv*, 20232023.

dien arbeitende Künstler*innen und Illustrator*innen, die verhindern wollen, dass ihre Online-Portfolios unentgeltlich und ohne ihre Zustimmung in die Trainingsdatensätze von Bildgeneratoren eingespeist werden (Abb. 4).³¹

Abb. 4: generierte Bilder eines intakten KI-Modells gegenüber solchen, deren Trainingsdaten mithilfe von Nightshade zu verschiedenen Graden manipuliert wurden



Die Widerständigkeit der Glitches

Wenn auch viele Nutzer*innen sarkastisch die statistisch generierten Lachsfilets kommentieren oder Memes mit ineinander verwachsenen Handshakes hin und herschicken, scheint sich das kritische Potenzial der algorithmischen Fehlproduktionen vor allem in Beispielen wie *Nightshade* zu konzentrieren. Vor allem vor dem Hintergrund, dass text-to-image-models in den

³¹ Vgl. <https://nightshade.cs.uchicago.edu/whatis.html> [Stand 09/2024].

letzten Jahren rapide Fortschritte gemacht haben, erscheinen solche Phänomene eher als Wachstumsschmerzen denn als ein sich anbahnendes Scheitern künstlicher Intelligenz als Technologie.

Wie schon Menkman schreibt, ist der Glitch ephemere: Entweder wird der dahinterliegende Fehler behoben oder die mit dem Glitch assoziierte Ästhetik stabilisiert sich als eigenständige Form.³² Dies scheint auch für den eigenwilligen Stil inzwischen veralteter Bildgeneratoren zu gelten: Instagram-Accounts wie *@chocolatebobross* oder *@werners_ai_art* und Künstler*innen wie Rafman und Florian Nackaerts spielen bewusst mit einer Ästhetik, die sowohl zu den von Dullaart kommentierten Händen als auch viralen Videophänomenen wie *Will Smith eating spaghetti*³³ passt: Figuren und Gegenstände verschmelzen miteinander und gehen in neue Formen über, Krokodile verwandeln sich in Motorräder, Pferd und Reiter gehen beim Rodeo eine Symbiose ein und Kim Jong-un erkundet mit einem Elektromobil den Meeresgrund.³⁴ Algorithmische Fehlproduktionen wechseln hier beständig ihre Position zwischen viralem Kitsch und künstlerischer Strategie.

Sowohl die Dokumentation als auch das kontrollierte Herbeiführen und Kultivieren dieser Fehler müssen als Versuche verstanden werden, dieser Phänomene habhaft zu werden. Gerade in Anbetracht der Vergänglichkeit von Glitches scheint der Reiz umso nachvollziehbarer, diesen ephemeren Erscheinungen algorithmischer Fehlproduktion eine dauerhafte Gestalt zu geben und sie, so wie Dullaart oder auch Eva und Franco Mattes, aus der virtuellen in die physische Realität zu überführen.

³² Vgl. Menkman (s. Anm. 6), S. 44.

³³ Vgl. www.reddit.com/r/StableDiffusion/comments/1d44h2c/will_smith_eating_spaghetti/ [Stand 09/2024].

³⁴ Sogar das physische Nachspielen der Ästhetik von KI-Videos ist inzwischen zu einem viralen Trend in den sozialen Netzwerken geworden. Vgl. <https://arstechnica.com/information-technology/2024/08/viral-trend-sees-humans-simulating-bizarre-ai-video-glitches/> [Stand 09/2024].



Plattformrealismus. Eine Ästhetik des Generischen

Bjørn Karmanns *Paragraphica* ist ein eigenwillig anmutendes technisches Gerät, das äußerlich an einen Fotoapparat erinnert, jedoch völlig anders funktioniert (Abb. 1). Was der dänische Künstler im Mai 2023 erstmals in den sozialen Medien vorstellte, ist zwar eine handliche kleine schwarze Kiste, aber sie besitzt weder Objektiv noch Linse, und sie produziert Bilder, die wie Fotografien aussehen, doch beruhen diese nicht auf einer optischen Aufzeichnung. Trotz aller oberflächlichen Verwandtschaft ist *Paragraphica* im Wortsinn keine Kamera: Es gibt keine Öffnung, durch die Licht von außen einfallen könnte, und im Inneren der buchstäblichen *black box* wird kein Bild projiziert. Auch das Display auf der Rückseite des Geräts zeigt kein Bild einer Szene, sondern vielmehr eine kurze schriftliche Beschreibung der Umgebung, die auf Standortdaten wie Tageszeit, Wetter und nahe gelegenen Sehenswürdigkeiten basiert. Drückt man auf den Auslöser, wird diese Beschreibung als Prompt an einen KI-Bildgenerator weitergegeben, der auf dieser Basis etwas generiert, das zwar äußerlich wie ein Foto aussieht, tatsächlich auf der algorithmischen Auswertung von Datenmustern beruht (Abb. 2): die täuschend echte Simulation einer Fotografie, synthetisiert aus Milliarden von Trainingsbildern aus dem Internet.¹

Was *Paragraphica* uns zeigt, ist mithin kein Spiegel der uns umgebenden Realität, sondern eher ihr fernes Echo, und doch sollen wir diese Bilder als ‚realistische‘ Darstellungen verstehen. ‚Realismus‘ gilt allgemein als heiliger Gral der KI-Bild-

¹ Siehe <https://bjoernkarmann.dk/project/paragraphica>.

generierung. Mit jedem Update ihrer Modelle versprechen Unternehmen wie Midjourney und OpenAI immer ‚realistischere‘ Renderings – und genau das erwarten auch die allermeisten Nutzer*innen von diesen Tools. Folgt man einschlägigen Werbebotschaften und Online-Diskussionen, so vereint dieser Realismus-Anspruch drei Dimensionen: Erstens wird von synthetischen Bildern erwartet, dass sie einer vorgegebenen Beschreibung (dem Prompt) möglichst genau entsprechen – also eindeutig *lesbar* sind. Zweitens werden sie so optimiert, dass sie Erwartungen hinsichtlich visueller Kohärenz erfüllen – sie sollen *plausibel* sein (und zum Beispiel keine Hände mit zusätzlichen Fingern zeigen). Schließlich orientieren sich diese Lesbarkeit- und Plausibilitätserwartungen an *älteren Medien* – nicht allein an der Fotografie, sondern zunehmend auch an ‚fotorealistischen‘ Videospielen und digitalen Rendings: Von synthetischen Bildern wird erwartet, dass sie aussehen wie die Bilder, die wir bereits kennen.

Diese Erwartungshaltungen möchte ich in diesem Essay – mit einem ursprünglich von Jacob Birken geprägten Begriff – als „Plattformrealismus“ bezeichnen.² Plattformrealismus ist eine Ästhetik zweiter Ordnung, eine Ästhetik des Generischen: Bilder aus Bildern, die nicht die Wirklichkeit abbilden, sondern bereits versprachlichte und verschriftlichte Begriffe und Konzepte visualisieren; Bilder, die so optimiert wurden, dass sie den ästhetischen und ideologischen Erwartungen einer meist sehr spezifischen Kundschaft entsprechen; Bilder, die nicht zuletzt bereits existierenden Bildern gleichen und dabei die visuellen Qualitäten älterer Bildmedien imitieren. Ausgehend von *Paragraphica* möchte ich im Folgenden diese drei Aspekte einer Ästhetik des Generischen vertiefen – und nach ihren Implikationen fragen.³

² Ich möchte Jacob Birken herzlich dafür danken, dass er mich diesen Begriff verwenden lässt, den er ursprünglich in einem Vortrag über NVIDIAs „digitale Zwillinge“ entwickelt hat. Vgl. auch ausführlicher zum Problem des digitalen Realismus: Jacob Birken: Vom Pixelrealismus. Takeshi Muratas Stillleben Cyborg, Berlin 2023.

³ Der vorliegende Text ist die gekürzte und überarbeitete Fassung von: Roland Meyer: Plattform Realism. AI Image Synthesis and the Rise of Generic Visual. In: Transbordeur 9 (2025).

Optimierung der Lesbarkeit

Am Anfang steht das Wort: Wer mit *Paragraphica* eine Szene ‚fotografieren‘ möchte, erhält zunächst eine Beschreibung in Textform, die erst im zweiten Schritt zum Prompt und schließlich zum Bild wird. In dieser Vorgängigkeit des ‚Sagbaren‘ vor dem ‚Sichtbaren‘, so hat es Antonio Somaini in Foucaultschen Termini beschrieben, zeigt sich ein Grundprinzip heutiger generativer KI: Was mit Text-zu-Bild-Generatoren wie Midjourney, Dall-E und Stable Diffusion überhaupt sichtbar wird, hängt maßgeblich davon ab, was zunächst in Form eines Prompts schriftlich formuliert werden kann. Während gerade die deutschsprachige Bildwissenschaft stets die Differenz zwischen der Welt der Bilder und jener Wörter betont hat, führten uns diese KI-Modelle, so Somaini, in eine „new visual landscape“, in der sich die Wörter und die Bilder immer weniger voneinander trennen ließen.⁴

Bildgenerierung mittels Prompt funktioniert auf den ersten Blick ähnlich wie die Online-Bildersuche: Man gibt ein paar Begriffe ein und erhält eine Auswahl von Bildern, die mehr oder weniger der Eingabe entsprechen. Doch geht es nicht darum, vorhandene Bilder in riesigen Datenbanken wiederzufinden, sondern darum, Bilder zu finden, die es so noch gar nicht gab, bevor man danach gesucht hat. Die Grammatik des Prompts folgt dabei einer kombinatorischen Logik: Als diskrete sprachliche Elemente lassen sich Wörter beliebig rekombinieren, und bereits bekannte und benennbare Bildobjekte, Konzepte, Stile, Genres und vieles mehr können zu bislang unbekanntem, unwahrscheinlichem und auch kontrafaktischen Konstellationen zusammengefügt werden – etwa, um nur ein paar der bekanntesten Beispiele zu nennen – ‚Avocado‘ und ‚Sessel‘, ‚Astronaut‘ mit ‚Pferd‘, ‚Papst‘ und ‚Balenciaga-Mantel‘

⁴ Antonio Somaini: Algorithmic Images. Artificial Intelligence and Visual Culture. In: Grey Room, 2023, Heft-Nr. 93, S. 74-115, hier S. 107.

oder ‚Trumpf‘ plus ‚Polizeiarrest‘. Text-zu-Bild-Generatoren versprechen, dass alles, was geschrieben werden kann, auch in Sekundenschnelle zum Bild werden kann – und wenn die Ergebnisse, wie in den genannten Beispielen, ebenso erwartbar wie überraschend, unmittelbar lesbar und doch bislang ungesehen sind, erklärt sich nicht zuletzt daraus ihr viraler Erfolg.

So lässt sich KI-Bildgenerierung als Form der umgekehrten Mustererkennung begreifen: Statt in vorhandenen Bildern visuelle Muster wiederzuerkennen und mit Labels zu versehen, werden Kombinationen von Labels nun zu Bildern. In technischer Hinsicht vollzieht sich dies in Serien von iterativen Tests, die das zu generierende Bild immer mehr der Beschreibung annähern. Bei sogenannten Diffusionsmodellen wie Midjourney, Dall-E und Stable Diffusion nämlich beginnt der Prozess der Bildgenerierung mit einem Zustand visueller Entropie. Diese Modelle haben an zahllosen Beispielen ‚gelernt‘, was passiert, wenn sich digitale Bilder in informationslosem Rauschen auflösen, und sie können diesen Prozess umkehren. Schritt für Schritt wird das anfängliche Bildrauschen minimiert, bis sich Muster herauskristallisieren, die mit dem Prompt übereinstimmen – Muster, die dann in jeder weiteren Iteration verstärkt werden. Das entstehende Bild wird so einer Serie von Tests unterworfen und seine Lesbarkeit schrittweise optimiert.

Möglich wird dies, weil diese Modelle nicht nur ‚gelernt‘ haben, wie Bilder im Rauschen verblassen, sondern auch, wie Wörter und Bilder miteinander korrelieren. Dafür wurden bekanntlich Milliarden von Bild-Text-Paaren ohne Rücksicht auf Urheberrechte aus dem Internet abgesaugt und zu Trainingsdaten gemacht. Die Muster, die sich in diesem riesigen virtuellen Archiv unserer digitalen visuellen Kultur finden lassen, kodieren diese Modelle im sogenannten *latent space*: einem

abstrakten hochdimensionalen Vektorraum, der in komprimierter Form statistische Beziehungen zwischen Wörtern und Bildern repräsentiert. Prompts lassen sich als Wegweiser in diesem Möglichkeitsraum potenzieller Bilder verstehen: Sie führen das Modell auf dem Weg der Bildgenerierung in eine spezifische Region des *latent space* – ohne Prompt brächte der Diffusionsprozess bloß zufällige Ergebnisse hervor. Ungeachtet der Popularität von Begriffen wie Prompt-Engineering lässt sich das Ergebnis jedoch kaum vorhersagen und nur sehr eingeschränkt steuern. Die Funktion von Prompts ist daher keineswegs die eines Programms, das exakt ausgeführt und in seinen Einzelheiten in ein Bild übersetzt wird. Eher noch handelt es sich um unscharfe Suchparameter, die in der scheinbar unendlichen Fülle potenzieller Bilder zu einer überschaubaren Auswahl führen.⁵

Wenn im Plattformrealismus Wörter und Bilder zunehmend untrennbar voneinander werden, dann nicht zuletzt deshalb, weil unsere digitale Bildkultur ohnehin weitgehend eine hybride Bild-Text-Kultur ist. Fast jedes Bild, das uns online begegnet, ist von einer Wolke aus Text umgeben: Bildunterschriften, Metadaten, Alt-Beschreibungen, Copyright-Hinweise und vieles mehr. Erst diese weitgehende sprachliche Vorerschließung digitaler Bildwelten macht sie zu so effizienten Trainingsgebieten für die generative KI. Nicht von ungefähr gehören zu den wichtigsten Quellen von Trainingsdaten Plattformen wie Flickr, Pinterest und Shopify, auf denen Bilder bereits massenhaft mit Beschreibungen versehen wurden – von Milliarden von User*innen ebenso wie von Heerscharen unterbezahlter Clickworker*innen im Globalen Süden. Die Welt der Plattformen ist eine, in der alles, was sichtbar wird, bereits benannt, beschrieben und mit einem Label versehen wurde.

5 Vgl. dazu ausführlich Roland Meyer: *The New Value of the Archive. AI Image Generation and the Visual Economy of 'Style'*. In: *IMAGE*, Jg. 37, 2023, Heft-Nr. 1, S. 100-111.

Aufbauend auf der umfassenden Versprachlichung des Sichtbaren führt uns der Plattformrealismus synthetischer Bilder in eine Welt, in der alles (und jede*r) einem möglichst eindeutig vorformulierten Label entspricht. Und diese Labels ersetzen das Konkrete durch das Allgemeine und das Spezifische durch das Generische. Dies zeigt sich auch am Beispiel der *Paraglyphica*: Niemals wird diese Kamera, die keine ist, das Bild dieses einzigartigen sonnigen Nachmittags in eben diesem konkret lokalisierbaren Park liefern. Was Nutzer*innen stattdessen erhalten, ist das Bild eines sonnigen Parks, der überall und nirgends existieren könnte: eine mögliche Visualisierung der Konzepte von Parkhaftigkeit, Sonnigkeit und Nachmittäglichkeit, die Kombination von visuellen Mustern, die mit den Schlagwörtern ‚Park‘, ‚sonnig‘ und ‚Nachmittag‘ korrelieren, übersetzt in ein Instant-Stockfoto, ein generisches Klischee. Dieses zeit- und ortlose Simulakrum hat keine indexikalische Beziehung zu irgendeiner Realität, nicht einmal zu einer simulierten. Plattformrealismus zeigt uns eine künstliche Welt der totalen Mustererkennung, optimiert für instantane und vollständige Lesbarkeit.

Wahrscheinlichkeit und Plausibilität

Im Gegensatz zu Stockfotografien jedoch beruhen die generischen Bilder generativer KI nicht allein auf ästhetischen Konventionen, sondern auf statistischen Kalkulationen: Unzählige Bilder ganz unterschiedlicher sonniger Parks haben Eingang gefunden in das, was *Paraglyphica* uns als synthetisches Bild eines generischen sonnigen Parks präsentiert. Treffend hat Hito Steyerl daher KI-generierte Bilder als „statistische Renderings“ beschrieben und mit Francis Galtons Kompositfotografien verglichen – es sind wahrscheinliche Bilder, Produkt

der Wiederholung von Mustern in riesigen Bilddatenmengen.⁶ Doch sind sie dabei keineswegs allein das mechanische Produkt blinder statistischer Prozesse. Nicht jedes der Millionen Bilder von Parks, die sich im Internet finden, hat für das Ergebnis dasselbe Gewicht. Vielmehr findet bereits auf der Ebene der Trainingsdaten ein ästhetischer Selektionsprozess statt.

Wie ein Park, eine Avocado oder der Papst aussehen, das haben KI-Modelle an riesigen Datenbanken von Bild-Text-Paaren gelernt. Eine der wichtigsten dieser Datenbanken, die unter anderem beim Training von Stable Diffusion und Midjourney Verwendung fand, ist LAION 5B – 5B steht für „five billion“. Doch wenn man fünf Milliarden Bilder aus dem Internet absaugt, erhält man eine Menge Bilder, die weder den ästhetischen Erwartungen der KI-Entwickler*innen noch denen ihrer zahlenden Kundschaft entsprechen. Zum Fine-Tuning der Modelle existieren daher Datenbanken wie LAION-Aesthetics, die ausschließlich Bilder vermeintlich hoher visueller Qualität versammeln. An ihnen soll das Modell lernen, was einen visuell plausiblen, ästhetisch ansprechenden Park ausmacht – um seinen Output den Erwartungen der Nutzer*innen anzunähern.

Um von Milliarden beliebiger Bilder zu wenigen Hunderttausenden vermeintlich besonders attraktiver zu gelangen, wird ihnen von einem eigens darauf trainierten KI-Modell ein sogenannter *aesthetic score*, ein Zahlenwert zwischen 1 und 10 zugewiesen. Gefüttert wurde diese KI mit Massen von bereits bewerteten Bildern, die aus dem Internet abgesaugt wurden, etwa von Plattformen wie *dpchallenge*, einer Website für Fotoamateure, oder dem Discord-Channel von Stable Diffusion, wo die User*innen der Software ihre Bilder gegenseitig bewerten. Wir haben es hier also mit Feedbackschleifen der ästhetischen Wertung zu tun: Menschen bewerten Bilder, um eine

KI zu trainieren, die vorhersagen soll, wie Menschen andere Bilder bewerten würden, und diese KI-basierten Vorhersagen wiederum dienen der Filterung von Trainingsdatensätzen für KI-Bildgeneratoren. Am Schluss stehen dann KI-generierte Bilder, die bereits daraufhin optimiert wurden, besonders hohe Wertungen zu erhalten – von Menschen wie von Maschinen.

Die rekursive Algorithmisierung des Geschmacks hört jedoch nicht mit der Filterung von Trainingsdaten auf. Kommerzielle KI-Bildgenerierungstools produzieren allein durch ihr Interfacedesign massenhaft Daten über Nutzer*innenverhalten und können anhand quantifizierbarer Reaktionen, Likes und Upscales ihren Output den algorithmisch vorhersagbaren Vorlieben ihrer zahlenden Kundschaft sukzessive anpassen. Die Geschmackspräferenzen, die so algorithmisch verstärkt werden, sind jedoch keineswegs universell – es sind die ästhetischen wie ideologischen Werte einer kleinen, demographisch homogenen Wertungsgemeinschaft, die besonders aktiv auf den entsprechenden Plattformen ist, und die ist überwiegend weiß, männlich und stammt aus dem Globalen Norden.⁷

Jedes KI-Bild wird so zum Testbild, dazu bestimmt, in den virtuellen Arenen der sozialen Medien mit anderen Bildern verglichen zu werden. In der spezifischen Ästhetik der KI-generierten Bilder wiederum, ihrer irrealen Glätte, übertriebenen Farbigkeit und intensiven Leuchtkraft etwa, spiegeln sich populäre Filter-Ästhetiken, die sich im endlosen Wettbewerb der Bilder als besonders erfolgreich erwiesen haben – nicht zuletzt, weil sie auf den leuchtenden Displays der mobilen Geräte besonders gut zur Geltung kommen.⁸ Aber nicht nur die Bilder stehen auf dem Prüfstand. Auch die Konsument*innen werden permanent getestet: Unsere Reaktionen und Interaktionen werden erfasst, quantifiziert, statistisch ausgewertet und fließen in Empfehlungsalgorithmen wie die Opti-

mierungsroutinen ein. Die visuelle Ökonomie der Gegenwart basiert auf ineinandergreifenden Feedbackschleifen von Test und Training, und was wir zu sehen bekommen, ist genau das, was unseren algorithmisch vorhersagbaren Erwartungen entsprechen soll. Plattformrealismus ist eine Ästhetik der erwartbaren Erwartungen, Produkt einer digitalen visuellen Kultur, in der Bilder ständig um Scores und Ratings konkurrieren.

Nostalgische Atmosphären

Um noch einmal auf das Beispiel *Paragraphica* zurückzukommen: Was hier als Bild eines sonnigen Parks präsentiert wird, ist nicht bloß eine Visualisierung bereits versprochenen Konzepte, ästhetisch optimiert in Hinblick auf die Reaktionen eines spezifischen Publikums. Es ist vor allem ein Produkt der Vergangenheit. Denn um ein Bild der Gegenwart oder gar der Zukunft zu synthetisieren, ist KI-Bildgenerierung auf vorhandene Bilder angewiesen. Alles, was nach dem Training dieser Modelle geschieht, existiert in ihrer Welt nicht. Plattformrealismus ist daher strukturell konservativ, ja sogar nostalgisch, und das zeigt sich auch an der vorherrschenden Ästhetik aktueller KI-Bildproduktion in Kunst und Design, wo die neuesten Bildtechnologien nur selten für den Entwurf möglicher Zukünfte und weit häufiger zur Erkundung alternativer Vergangenheiten eingesetzt werden.

Ein prominentes Beispiel dafür ist der deutsche Fotograf Boris Eldagsen, der mit einem Bild aus seiner Serie *Pseudomnesia* im Jahr 2023 einen Fotopreis gewann – bis er es spektakulär aus dem Wettbewerb zurückzog. Ein Statement auf seiner Website macht Eldagsens Herangehensweise deutlich: „Using the visual language of the 1940s, Boris Eldagsen produces his images as fake memories of a past, that never existed, that

⁷ Christo Buschek und Jer Thorp: Models All the Way Down. In: *Knowing Machines*, <https://knowingmachines.org/models-all-the-way> [Stand 03/2025].

⁸ Vgl. dazu ausführlich Roland Meyer: Es schimmert, es glüht, es funkelt. Zur Ästhetik der KI-Bilder. In: *54 Books*, 20.03.2023, <https://54books.de/es-schimmert-es-glueht-es-funkelt-zur-aesthetik-der-ki-bilder/> [Stand 03/2025].

no-one photographed.“ Was Eldagsen hier „die visuelle Sprache der 1940er Jahre“ nennt, ist eine primär atmosphärische Qualität der Bilder, die sie als vermeintliche Produkte einer vergangenen Epoche lesbar macht. Nicht zuletzt ist es die visuelle Anmutung sepiagetönter alter Schwarz-Weiß-Fotografien, mit der Eldagsens in seinen Bildern die Vergangenheit wiederaufleben lässt.

Digitale Bildkulturen pflegen seit Langem eine nostalgische Beziehung zur Ästhetik des Analogem, etwa in Form einschlägiger Filter, die das charakteristische Bildrauschen, die Körnung und Farbigkeit älterer Medien simulieren. KI-Bildgenerierung allerdings imitiert in jedem ihrer Produkte ältere Bildmedien, sei es eine Bleistiftzeichnung, eine Polaroid-Fotografie oder eine Game-Engine. Jedes KI-Bild simuliert ein Bild, das es auch ohne KI hätte geben können, das jedoch nie gemalt, gezeichnet oder fotografiert wurde. Indem sie aus den Bildern der Vergangenheit mögliche Variationen und Rekombination interpoliert, antwortet KI-Bildgenerierung nicht bloß auf nostalgische Sehnsüchte, sie verspricht auch, die Lücken im virtuellen Archiv der Bilder zu schließen: Filmstills aus nie realisierten Filmen, fotografische Dokumentationen von Ereignissen, die nie stattgefunden haben, oder fiktive Archivaufnahmen marginalisierter Gruppen, denen historisch Sichtbarkeit verwehrt wurde – all dies findet sich mittlerweile in großer Zahl auf Social-Media-Kanälen wie in Galerieräumen.

Lev Manovich hat in diesem Sinne kürzlich KI-Bildgenerierung als eine Technik der spekulativen Ergänzung und Erweiterung unserer kunst- und bildhistorischen Überlieferung gefeiert, als Erfüllung von Malraux' Traum eines „Musée imaginaire“.⁹ Doch die *latent spaces* der bildgenerierenden KI, diese komprimierten Kondensate wahllos aus dem Netz gefischter Bilder, sind nicht nur äußerst begrenzte und unvoll-

ständige Bildarchive voller rassistischer und sexistischer Klischees, kultureller Vorurteile und blinder Flecken, sie sind auch strukturell ungeordnet und undurchsichtig. Was Eldagsen als „Bildsprache der 1940er Jahre“ bezeichnet, speist sich aus historischen Fotografien ebenso wie aus popkulturellen Reinszenierungen, aus Modefotografien, Filmstills, LP-Covern oder Bildern aus Videospiele, die alle bereits versucht haben, den Look dieser Zeit zu imitieren. Was uns die KI dann als Bild der Vergangenheit präsentiert, ist also immer schon mehrfach remediatisiert, ästhetisiert und weitgehend fiktionalisiert – Produkt sich vielfach überlagernder Vergangenheitsschichten. „Pastness“ hat Fredric Jameson das bereits im Zuge der Postmoderne-Debatte genannt: eine generische Anmutung des Vergangenen, in der ästhetische Marker an die Stelle historischer Überlieferung getreten sind. Doch was einst aufwendig reinszeniert werden musste, ist dank generativer KI zum extrahierbaren, variierbaren und synthetisierbaren Datenmuster geworden. Im Plattformrealismus erscheint die Vergangenheit vollständig als ausbeutbare Ressource erschlossen, zugleich löst sich Geschichte in eine Serie und Sequenzen von ästhetischen Stimmungen und Atmosphären auf, die endlos appropriiert, reproduziert und variiert werden können.

Harmlose Bilder sonnendurchfluteter Parks, wie sie *Paragraphics* präsentiert, sollten nicht dazu verleiten, Plattformrealismus für eine unpolitische Ästhetik zu halten. Die sprachliche Vereindeutigung des Sichtbaren auf wiedererkennbare Klischees, die ästhetische Anpassung digitaler Bildwelten an quantifizierbare Reaktionsmuster und die nostalgische Variation generischer Vergangenheitsatmosphären – all dies lässt sich nicht unabhängig von politischen Tendenzen verstehen, die mit neuen Eindeutigkeiten, populistischen Versprechen und der Sehnsucht nach der Wiederherstellung einer Vergan-

⁹ Lev Manovich: From Museum Without Walls to Generative AI Museum. In: *European Photography*, 2024, Heft-Nr. 116, <https://manovich.net/index.php/projects/from-museum-without-walls-to-genai-museum> [Stand 03/2025].

genheit operieren, die es so nie gab. Es ist kein Zufall, dass Figuren wie Donald Trump und Elon Musk zu den begeistertsten Anwendern der neuen KI-Tools gehören. Plattformrealismus ist bereits dabei, zur Ästhetik eines neuen, plattformgetriebenen Faschismus werden.

Those aren't your memories. Geschichte und Erinnerung im KI-generierten Fotorealismus

Im Sommer 2023 kursierte ein Porträt der glücklichen Familie Aniston-Pitt auf den Websites zahlreicher Lifestyle-Magazine. Es illustriert, wie die Gegenwart aussehen könnte, hätte Pitt die *Friends*-Schauspielerin 2005 nicht für Angelina Jolie verlassen. In den spontanen Bilddiskursen der sozialen Medien lud die neu verfügbar gemachte Technologie fotorealistischer KI-Bildgenerierung in jenem Jahr zu kontrafaktischen Experimenten ein – zu Zeitreisen und spielerischen Entwürfen von Paralleluniversen. Während diese Bilddiskurse die Qualität flüchtiger Memes hatten, manifestierte sich eine kulturelle Fantasie in ihnen, die nicht von ungefähr mit dem Durchbruch fotorealistischer KI-Bildgenerierung koinzidierte. Handelte es sich doch um eine Bildgeste, welche die überkommenen Paradigmen der Fotografie (Roland Barthes' ‚es-ist-so-gewesen‘) mit dem permutativen Paradigma der KI-Bilderzeugung vermählte. ‚Was hätte sein können‘ ist eine kontrafaktische Fantasie, die unser Verständnis der Vergangenheit auf neue Horizonte hin öffnet. Wie jedoch werden wir uns in einer Bildumgebung zurechtfinden, in der wir nun all das betrachten können, was nie eingetreten ist?

Statistische Plausibilität

Für das Medium Fotografie markierte 2023 das Jahr, in dem KI-Bildgeneratoren täuschend fotorealistische Darstellungen hervorzubringen begannen.

Abb. 1. Lange scheiterten KI-Bildgeneratoren an der anatomisch korrekten Darstellung von Händen. Über das neuere Remake des Memes "AI Accepting the Job" gibt es heute nichts mehr zu lachen.

Everyone: AI art will make designers obsolete

AI accepting the job:



Hatte der Glitch als Schwelleneffekt das KI-generierte Bild in seiner Frühphase zuverlässig als solches entlarvt, so schlossen sich die Risse in der generierten Bildillusion nun zusehends (Abb. 1). Wenn jedoch algorithmisch erzeugte Bilder aussehen wie Fotografien, lässt das deren Domäne nicht unberührt: Aus fotografisch wird photograph-isch.

Dabei ist der KI-generierte Fotorealismus sowohl von der Fotografie als auch von altbekannten Fotorealismen wie etwa dem Malerischen abzugrenzen. Er folgt einem anderen Paradigma als die mit Photoshop überarbeiteten Models der 2000er Jahre, und er unterscheidet sich auch von den Produkten der KI-Bildoptimierungsfunktionen, die heute standardmäßig in Kameraprozessoren eingebaut sind und den im eigentlichen Sinne fotografischen Schnappschuss schon im Moment seiner Produktion zersetzen.

Das KI-generierte Bild entsteht gänzlich abgelöst von Kameralinsen, von Belichtung generell. Generative Modelle bilden hochdimensionale Daten (die Bilder aus dem Datensatz) auf einen Latenzraum ab. Dieser wird zur Grundlage für die Output-Kapazitäten eines Modells, da er die wichtigsten Muster und Beziehungen enthält, die das Modell aus den Trainingsdaten ‚erlernt‘ hat. Was dann stattfindet, ist eine mathe-

matisch-statistische Extrapolation: Im Gegensatz zur Interpolation, der Schätzung eines Wertes innerhalb zweier bekannter Werte, ist die Extrapolation die Schätzung eines unbekanntes Datenwerts aufgrund eines vermuteten Funktionsverlaufs. Generative Modelle erfinden auf diese Weise Bilder, die sich einen Steinwurf von den Trainingsdaten entfernen, ihnen aber sozusagen ‚ideologisch‘ treu bleiben. Limitationen der Bildfindung ergeben sich aus den Stereotypen, die in der Datenbank vorherrschen, und aus den Voreinstellungen der Softwareprovider, welche die Bildgenerierung mit unsichtbarer Hand lenken. Dabei ist jedes KI-generierte Bild nur ein aleatorischer Output des Latenzraums, der Bild für Bild ausgelotet werden kann. Es erscheint gewissermaßen vor dem Hintergrund unendlich vieler möglicher Permutationen, die derselbe Prompt hätte hervorbringen können. In diesem Sinne schöpft die Bildgeste der nicht eingetretenen, aber durchaus denkbaren ‚alternativen Vergangenheiten‘ die Potenziale der Technologie auf eine Weise aus, die ihrer Eigenlogik entspricht – und appelliert dabei zugleich an Sehgewohnheiten, die sich mit der Fotografie entwickelt haben.

Digitale Bilddissemination

Ein weiteres Beispiel für kontrafaktische, KI-generierte Fotorealismen des Jahres 2023 ist eine Bildserie, die kawaii-lovecraft auf dem Reddit-Forum Midjourney teilte. Laut Bildunterschrift zeigten die Bilder eine Fotodokumentation der *Blue Plague*, die im Russland der 1970er Jahre gewütet habe. Zu sehen waren unscharfe Luftbilder lokaler Wucherungen einer blau blühenden Pflanze, Männer in gelben Schutzanzügen und blaue Feinstaubwolken, die ganze Stadtpromenaden einhüllen (Abb. 2). Wenngleich diese Bilder dezidiert KI-

Abb.2, KI-generiertes Bild aus der Fotoserie "The infamous Blue Plague Incident" (70s, USSR), 2022 von kawaiilovecraft auf Reddit gepostet



generiert waren, wird an ihnen doch erfahrbar, wie anfällig die historische Vorstellung für Manipulationen ist, gerade wenn sie in einem Medium stattfinden, das aussieht wie Fotografie. Generierte Bilder verfangen sich zudem leicht im Gedächtnis, da sie Resonanzen zu bekannten Bildern herstellen: Männer, die in Schutzanzügen durch ein ländliches Frühlingsidyll wandern, entstanden beispielsweise bei der Strahlungsmessung in der Sperrzone um Tschernobyl im Mai 1986 (Abb. 3). Das KI-generierte Bild ist – wie solche visuellen Echos zeigen – ein Produkt eines kulturellen Bildgedächtnisses im ganz konkreten, technischen Sinne.

Während die Schöpfungen von kawaiilovecraft Trouvailles aus einem Expert*innen-Forum sind, feierte Bennett Miller (ebenfals 2023) mit einer KI-generierten Bildserie einen Erfolg in der Kunstwelt. Die 20 mit Dall-E 2 erzeugten Bilder (eine Auslese aus über 100.000 Bildgenerierungsexperimenten, wie der Presstext betont) waren in der New Yorker Gagosian Gallery als hochwertige Pigmentdrucke ausgestellt und adaptierten die Ästhetik früher Fotografie. Porträts kleiner Mädchen mit wallendem Haar, Fesselballons am Horizont... Die mitunter schemenhaften Bildmotive stemmten sich in einem Kraftakt gegen sepiafarbenedes Rauschen. Die Bezeichnung *Untitled* verstärkte dabei die rätselhafte Anmutung der Bilder. Fotografien, die eine ähnliche Spannung zwischen Erinnern und Vergessen herstellen, hatte die Künstlerin und Analogfilmaktivistin Tacita Dean in ihrem Fotobuch *Floh* (2001) versammelt. Das Buch bestand aus losen Familienalbenfotografien, die Dean auf Flohmärkten in Europa und Amerika gefunden und zu einem wortlosen, kollektiven Narrativ nebeneinandergestellt hatte. Die Korrelation zu Millers Kunstprojekt besteht darin, dass der Kontakt zur dargestellten Realität in beiden Bildtypen brüchig scheint: Während Deans analogfotografische Flohmarktfunde auf eine verlorengegangene historische Information verweisen, geben Millers Bilder Menschen und Ereignisse zu sehen, die niemals existierten. Hatte Dean ihr (in Bezug auf seine historische Information ‚prekäres‘) Bildmaterial auf Flohmärkten gefunden, so diffundieren KI-generierte Bilder heute in eine Informationskultur hinein, in der die sozialen Medien zu einem primären Ort gesellschaftlicher Meinungsbildung und Sinnproduktion geworden sind.



Abb.3, Katastrophe in Tschernobyl, 1986, Foto: Keystone Pressedienst

Dort zirkulieren Bilder (etwa als Memes) gerade nicht in Hinblick auf stabile Bedeutungen. Andererseits sind die Algorithmen der Plattformen so eingerichtet, dass ‚das beste Bild gewinnt‘. Einen Eklat verursachte diesbezüglich die Facebook-Seite *Past Perspectives*, die täglich mit Posts erstaunlicher oder rührender Fotografien aufwartet. Anfang 2024 publizierte der Account ein vermeintliches ‚Foto‘ der Luftfahrt-Pioniere Orville und Wilbur Wright. Auf dem KI-generierten Bild posieren zwei hübsche junge Männer keck lächelnd vor einem Segelflugzeug, während im Hintergrund die Schemen zahlreicher Schaulustiger zu erkennen sind. Die echten Porträts der Wright-Brüder hingegen zeigen zwei ernst dreinblickende Männer, die sich nicht im Geringsten auf einen Flirt mit der Kamera einlassen. Fotografische Dokumente wie dieses müssen sich im Internet gegen KI-Bildproduktionen durchsetzen, die – nicht zuletzt aufgrund der Voreinstellungen der Softwareprovider – den aktuellen ästhetischen Vorlieben und affektiven Bedürfnissen der sozialen Medien passgenau entsprechen. Was aber geht verloren, wenn die Vergangenheit durch solche Umdeutungen mit gegenwärtigen Klischees und Ideologien kurzgeschlossen wird?¹

KI-generierte Bildarchiv

Als Problemfeld lassen sich die mithilfe generativer Modelle imaginierten alternativen Wirklichkeiten weder mit dem Begriff des Deepfake fassen noch unter Verweis auf die lange Geschichte der fotografischen Manipulation einkassieren. Auch der Begriff Fantasy trifft es nicht, da die oben behandelten Motive eine Plausibilität beanspruchen, die sich auf die statistische Logik der künstlichen Intelligenz beruft. Die Macht fotografischer Evidenz reicht dabei so weit, dass der KI-generierte Fotorealismus in manchen Projekten sogar als ‚evidence‘ präsentiert wird.

1 Schon lächeln uns vergangene Generationen gut gelaunt im Gruppen-Selfie an statt uns mit den gefassten Blicken derer zu durchbohren, die noch die langen Belichtungszeiten der frühen Fotografie auszusitzen hatten. Vgl. Anders Anders Anglesey: From Cavemen to Cowboys. AI Reimagines Selfies Throughout History. In: Newsweek, 23.03.2023, <https://www.newsweek.com/cavemen-cowboys-ai-reimagines-selfies-throughout-history-1787858> [Stand 09/2024].

Der Fotograf Michael Christopher Brown gibt an, ihn zum Zwecke der Sichtbarmachung dessen einzusetzen, was durch keine Kamera eingefangen wurde. Seine Bildserie *90 Miles* (2023) ist ein mit Midjourney umgesetztes Illustrationsprojekt, das sich als Fotoreportage ausgibt, weil seine Prompts, wie Brown betont, aus von ihm gesammelten, ‚wahren Geschichten‘ hervorgegangen seien. Sujet der Bildserie war die massenhafte Flucht von Kubaner*innen nach Florida zwischen 2022 und 2023. Der Titel des Projekts bezieht sich auf die 145 km Ozean, welche die Migrant*innen zwischen Havanna und den USA in improvisierten Booten überquerten. Die expressiv-schönen Gesichter der dargestellten Menschen lassen leicht über die Deformationen hinwegsehen, welche die Bilderfindungen von Midjourney 1 noch im Gepäck führten. In den Bildern drängen sich Menschengruppen auf mintgrünen Oldtimer-Autos, die zu Booten umgebaut sind, wobei mal die biblische Geschichte des Sturms auf hoher See anklingt, mal die Ankunft europäischer Flüchtlinge auf Ellis Island, mal der profane Anblick im Meer treibender Cargo-Container. Auch die österreichische Künstlerin Claudia Larcher spricht von einem Wunsch nach Sichtbarmachung, wählt dabei aber einen subversiveren Ansatz, der das massenhafte Scraping von Bildern aus dem Netz gegen die Modelle selbst auszuspielen vorgibt. Im Rahmen des Projekts *AI and the Art of Historical Reinterpretation. Filling Gender Bias Gaps* (2022–2024) schuf Larcher ein fiktives Bildarchiv, das der historischen Marginalisierung von Frauen begegnen soll. Um „Lücken und Verzerrungen in der Geschichtsschreibung“² zu begegnen, ließ sie scheinbare Momentaufnahmen der Kunst- und Technikgeschichte oder der Politik neu generieren: Frauen als Ingenieurinnen-Avantgarde der 1960er; einander die Hände schützelnde Politikerinnen... In einem der Bilder posieren statt

2 Vgl. Claudia Larcher: AI and the Art of Historical Reinterpretation. Filling Gender Bias Gaps, <https://www.claudialarcher.com/work/theartofhistoricalreinterpretation> [Stand 09/2024].

Richard Greenblatt und Thomas Knight zwei Frauen mit der LISP-Machine, mit welcher 1978 an der Erschaffung künstlicher Intelligenz gearbeitet wurde. Im Galerieraum präsentiert sich die in vergilbtem Schwarz-Weiß gehaltene Bildserie als Diaprojektion, wobei ihr eigentlich prädestinierter Ort das Internet ist. Auf digitalen Plattformen verbreitet, zielt die Sammlung darauf ab, in Datensätze einzugehen, mit der künftige generative Modelle trainiert werden. In der Art und Weise, wie Larchers Projekt den in Bilddatensätzen verankerten und in der Bildgenerierung perpetuierten Vorurteilen durch einen kruden Kunstgriff entgegenwirken will, hat es Parallelen zu jener kontrafaktischen Überkorrektur, die Googles Bildgenerator *Gemini* Anfang 2024 in Verruf brachte. Gemini war beim Versuch, die Probleme des Rassismus und Sexismus³ zu beheben, die sich bei der KI-Bildgenerierung immer wieder einstellen, über das Ziel hinausgeschossen: Auf die Inklusion von Frauen und People of Colour getrimmt, hatte *Gemini* Schwarze Wehrmachtssoldaten und Päpstinne ausgegeben. Wie die Bildgenerierungsexperimente einer belustigt-verärgerten Online-Öffentlichkeit offenlegten, verweigerten Googles Voreinstellungen die Darstellung weißer Männer in Machtpositionen rundheraus. Der Grat zwischen Geschichtsrevisionismus und dem gut gemeinten Angebot, gesellschaftliche Rollenvorstellungen in neue Wege zu leiten, ist denkbar schmal.

Solch heikle Operationen am Gewebe historischer Erinnerung finden im Kontext von Kunstprojekten einen sichernden Rahmen. Die kubanische Künstlerin Susana Pilar Delahante Matienzo präsentierte 2024 eine dekoloniale Gegenerzählung in Form eines KI-generierten Bildarchivs. Die vermeintlich historischen Fotografien des Projekts *Achievement* zeigten Schwarze Frauen im Ornat hohen sozialen Ansehens. Matienzo sieht in der Technologie eine Chance: „I really think that arti-

cial intelligence can fill in the gap and contribute to maybe change people's mindset [...]. I think we should create archives for artificial intelligence images in the future. These archives would be compiling photos based on real moments or situations that we know did exist but for whatever reasons no images could be captured.“⁴ Gegenarchive KI-generierter Bilder? Generative KI bietet sich hier als Erfüllungsgehilfin einer Sichtbarmachung und Wiedergutmachung im Bild an. Doch stellt sich die Frage, welchen Preis eine Gesellschaft zahlt, die in die Umschreibung historischer Information einwilligt.

Autobiografische Erinnerung

Auf der Ebene visueller Autobiografie lässt sich das Bedürfnis nach Schließung einer Lücke im Kontinuum fotografisch dokumentierter Wirklichkeit scheinbar unbekümmerter befriedigen. Die griechische Künstlerin Maria Mavropoulou kannte ihre Familiengeschichte lediglich aus Beschreibungen ihrer Mutter. Aufgrund einer langen, verlustreichen Migrationsgeschichte waren ihrer Familie nur spärliche Memorabilia geblieben. Aus deren schmerzlichem Fehlen heraus entstand *Imagined Images* (2021), ein sieben Jahrzehnte und vierhundert Bilder umfassendes KI-generiertes Fotoalbum. Dabei hatte die Software leichtes Spiel, die jeweiligen Konventionen zu identifizieren, welche Fotoalben als kulturelle Praxis auszeichnen. Auf einen immensen Bildschatz privater Amateurfotografie zurückgreifend fand Dall-E 1 für jedes Jahrzehnt nicht nur die passende fotografische Anmutung, sondern auch für jedes Szenario (Geburtstagsfeiern, Urlaubs-, Baby- und Klassenfotos) die geeigneten Posen und Requisiten. Die für die frühe Version von Dall-E typischen Glitches indessen taten der Bildillusion keinen Abbruch, sondern korrelierten einer-

3 Nitasha Tiku, Kevin Schaul und Szu Yu Chen: These fake images reveal how AI amplifies our worst stereotypes. In: Washington Post, November 2023, https://www.washingtonpost.com/technology/interactive/2023/ai-generated-images-bias-racism-sexism-stereotypes/?itid=1k_inline_manual_28 [Stand 12/2024].

4 Susana Pilar Delahante Matienzo im Gespräch mit der Journalistin Adiaratou Diarrassouba, <https://www.moussemagazine.it/magazine/susana-pilar-delahante-matienzo-secession-vienna-2024/> [Stand 12/2024].

seits mit den Unschärfen der Erinnerung und kompensierten andererseits die Schwächen einer Software, die nur Generisches ausgeben konnte, wo eigentlich Spezifisches gefragt war.

Inzwischen haben verschiedene staatlich geförderte Sozialprojekte die Idee der Dienstbarmachung von KI-Bildgenerierung zur Stabilisierung autobiografischer Narrative in pseudo-fotografischen Dokumenten aufgegriffen. Von den Traditionen, Orten und gesellschaftlichen Milieus abgekoppelt, die sie einst prägten, leiden Flüchtlings- und Vertriebenengemeinschaften unter dem Verlust ihrer kulturellen Herkunftsgeschichte, die oft lediglich in flüchtigen Erzählungen fortbesteht. Zum Zweck der Rekonstruktion visueller Erinnerung setzt das Projekt *Synthetic Memories* der spanischen Agentur Domestic Data Streamers seit 2022 KI-Bildgenerierung ein. Die Website des Projekts erklärt: „A Synthetic Memory is a digitally reconstructed representation of a personal memory [...] Unlike our natural memories, which are created and stored in the human brain, Synthetic Memories are produced by algorithms.“⁵ Wenn aber dergestalt aus Erinnerungen pseudo-historiografische Dokumente gemacht werden, findet ein Kurzschluss statt. Während Oral History ein erprobtes ethnografisches Verfahren ist, besteht die Gefahr hier darin, dass die Unterscheidung zwischen historischer Wirklichkeit und Erinnerung, zwischen KI-generiertem Fotorealismus und Fotografie allzu leichtfertig aufgegeben wird.

In der Reminiszenz-Therapie für Demenzkranke wird seit Kurzem dasselbe Verfahren angewendet – jedoch mit Ansprüchen, die sich auf Erinnerungsarbeit beschränken. Bilder appellieren an Affekte, die für Worte nicht erreichbar sind. Arbeitete die Reminiszenz-Therapie bisher mit Souvenir-Koffern, können neue Therapie-Tools wie *Rememo* (2023) nun vage Erinnerungsbilder als KI-generierte ‚Polaroids‘ ausgeben.

76

5 Synthetic Memories – a project on community, arts, and AI, <https://www.syntheticmemories.net/> [Stand 09/2024]. Vgl. auch The Citizens' Office of Synthetic Memories. A public office for visual memory reconstruction using AI, <https://www.domesticstreamers.com/art-research/work/the-citizens-office-of-synthetic-memories/> [Stand 09/2024].

Erinnerungen werden auf diese Weise stabilisiert und das Polaroid kann im Therapiegespräch zum Auslöser weiterer Erinnerungsketten werden. Ende 2024 brachte dann Kodak *Memory Shot* heraus, ein Tool, welches die Reminiszenz-Therapie ebenfalls als ein Anwendungsfeld von KI-Bildgenerierung erschließt.

Die Hemmung, in historische Information manipulativ einzugreifen, mag hier niedrigschwelliger sein. Menschliche Erinnerung jedoch – und das gilt es zu beachten – ist höchst manipulierbar. So bewies eine Gruppe von Psycholog*innen um Kimberley Wade an der University of Warwick, wie leicht sich falsche Kindheitserinnerungen mithilfe von bearbeiteten Fotos einpflanzen lassen.⁶ Andererseits können schmerzhaft Erfahrungen ganz bewusst im Bild neu entworfen werden. Das 2023 von einer deutschen Marketingagentur initiierte und von einigen Vereinen zur Förderung von Transidentität unterstützte Projekt *Saved Memories* etwa bietet Trans*Menschen an, Fotografien aus ihrer Kindheit in Hinblick auf ihre Geschlechtsidentität anzupassen.

Neue Funktionen des Fotorealismus

An *ThisPersonDoesNotExist*⁷ konnte ein breites Publikum bereits 2019 eindrücklich erleben, dass man den Wahrnehmungseffekt fotografischer Evidenz nicht so einfach abschütteln kann. In anderen Worten: Selbst wenn eine Unterscheidung zwischen fotografisch und photograph-ish durch Labels oder Metadaten-Signaturen eingedämmt würde, wird der Status der Fotografie dadurch erodieren, dass die grundlegendsten Veränderungen des kommenden Bildmilieus unsere Wahrnehmung

6 Vgl. Kimberley Wade, Maryanne Garry, Don Read und Stephen Lindsay: A picture is worth a thousand lies. Using false photographs to create false childhood memories. In: *Psychonomic Bulletin and Review*, 2002, Heft-Nr. 9, S. 597-603. In einem Experiment präsentierten die Forscher*innen Proband*innen Fotos aus ihrer Kindheit, die sie von deren Eltern bekommen hatten. Die Forscher*innen hatten eines der Bilder vorab mit einem Computerprogramm manipuliert: Auf dem Foto sah es so aus, als hätte der jeweilige Proband eine Heißluftballonfahrt erlebt. Die Proband*innen sollten anhand der Fotos nun erzählen, woran sie sich erinnern konnten. Nach drei Treffen erzählte die Hälfte von ihnen in bunten Details, was bei der Fahrt mit dem Heißluftballon passiert war.

7 Tero Karras, Samuli Laine und Timo Aila: A Style-Based Generator Architecture for Generative Adversarial Networks. In: *Computer Vision and Pattern Recognition*, Dezember 2018, https://openaccess.thecvf.com/content_CVPR_2019/papers/Karras_A_Style-Based_Generator_Architecture_for_Generative_Adversarial_Networks_CVPR_2019_paper.pdf [Stand 12/2024].

betreffen. Die vorgestellten Beispiele schöpfen die Verwandtschaft des KI-generierte Bildes zur Fotografie aus, indem sie es zum Medium einer Umarbeitung der Vergangenheit machen: ob im Dienste einer Sichtbarmachung des Undokumentierten, einer Vergangenheitsbewältigung oder Ersatzbefriedigung. Einerseits steht die Authentizität historischer Dokumentation und menschlicher Erinnerung auf dem Spiel (kaum auszudenken, wie sich diese Bereiche gegen eine so potente Bildproduktionsmaschine durchsetzen werden). Andererseits gilt es womöglich zu akzeptieren, dass an die Stelle des überkommenen fotografischen Wirklichkeitspostulats neue Funktionen treten werden. Das, was wir einstmals als ‚fotografisches Dokument‘ verstanden, sei längst zu einem gefügigen Ausdrucksmittel im gesellschaftlichen Onlinediskurs geworden, so Adam Harvey: „This is what photography is today: AI generated images render networked vibes that reflect what’s happening in our connected world right now.“⁸ Fotografie fing Licht ein, KI-generierte Bilder fangen Stimmungen ein.

Die obigen pseudo-fotografischen Bildspiele mit dem Kontrafaktischen imaginieren alternative Geschichtsverläufe und schaffen dadurch Denkräume, die überkommene Realitätskonzepte irritieren und womöglich sogar zu überwinden versprechen. In dem Maße, wie der fotografische Wahrheitsanspruch erodiert, könnte der KI-generierte Fotorealismus zu einem Medium werden, das unsere Vorstellungswelt kaleidoskopisch aufbricht. Zugleich droht er, alle Zeiten mit einer nicht mehr hinterfragbaren Gegenwart und ihren wechselnden Ideologien gleichzuschalten.

Beyond the Spectacle of “AI”: An Emergent Aesthetic Regime

“The creativity of culture has no outcome, no conclusion. It does not result in art works, artifacts, products. Creativity is a continuity that engenders itself in others.”

— James P. Carse, *Finite and Infinite Games: A Vision of Life as Play and Possibility* (1986)

The artist and technologist Sterling Crispin adapted Guy Debord’s 1967 book of critical theory *The Society of the Spectacle* for a striking, and humorous, examination of our current image culture. *Spectacle* (2022) stems from Crispin’s use of OpenAI’s Chat GPT3 to simplify Debord’s first ten theses and OpenAI’s DALL-E 2 to convert these reduced statements into images (Fig. 1-4). In the text, Debord emphasizes that the “spectacle” is not the image, nor indeed some new form of image, but instead a way of naming how images operate to change relations among people: “The spectacle is not a collection of images; rather, it is a social relationship between people that is mediated by images;” or, “the visual reflection of the ruling economic order.”¹ Crispin adapts and bends Debord’s famous text, but in so doing also calls upon and distorts Debord’s significant authorial voice.² *Spectacle*, though identified as the work of Crispin, exists via Debord, Crispin and the public generators, themselves derived from a wealth of unrecognized intellectual and manual labor. Crispin presents the generated aphorism beneath the image it generated, and each aphorism appears twenty times with the twenty visual outputs he selected from prompting DALL-E, confounding the

1 Guy Debord: *The Society of the Spectacle*, Detroit 2010, thesis 4 and 14.

2 Crispin’s project might also interestingly be compared to Debord’s film based on his text, as well as other artworks responding to Debord’s famous text.

Fig.1-4, Sterling Crispin: (selections from) *Spectacle*, 2022, JPEG ©Sterling Crispin, Courtesy of the artist <https://www.sterlingcrispin.com/spectacle.html>



In the present age, people prefer symbols and representations to actual things. This means that for many people, what is not real is more sacred than what is real.



In the present age, people prefer symbols and representations to actual things. This means that for many people, what is not real is more sacred than what is real.



In the present age, people prefer symbols and representations to actual things. This means that for many people, what is not real is more sacred than what is real.



In the present age, people prefer symbols and representations to actual things. This means that for many people, what is not real is more sacred than what is real.

presumption of a meaningful image amidst the circulating web economies that enabled public generators. What does the text mean if it produces such a variety of visualizations? What is an image, or where is it amidst this collection? Why must the output provide stable meaning, how is it stabilized, and whom does that stabilization serve? These are the provocations that underlie Debord's text and are apt for the cascade of content produced by public generators. In examining this new form of production, without forcing it into prior aesthetic regimes, new frameworks become necessary.

I observe six major critiques responding to these public generators, whose processes and outputs are commonly now termed generative AI:

1. The *environmental critique* recognizes the energy expenditure of deep learning models;³ though computation can manage data to produce models that help support environmental efforts,⁴ the statements from companies like Google and Amazon of investments in AI data centers clarifies that ecological concerns are second to economic opportunity.⁵

2. The *bias critique* stems from prejudices within the foundational model, like LAION or ImageNet (datasets largely devel-

3 Payal Dhar: The Carbon Impact of Artificial Intelligence. In: *Nature Machine Intelligence*, vol. 2, 2020, no. 8, pp. 423-425, <https://doi.org/10.1038/s42256-020-0219-9>; Mál Hogan: The Fumes of AI. In: *Critical AI*, vol. 2, 2024, no. 1, <https://doi.org/10.1215/2834703X-11205231>; Sasha Luccioni: The Mounting Human and Environmental Costs of Generative AI. In: *Ars Technica*. 12.04.2023, <https://arstechnica.com/gadgets/2023/04/generative-ai-is-cool-but-lets-not-forget-its-human-and-environmental-costs/>; Bernard Marr: Green Intelligence: Why Data And AI Must Become More Sustainable. In: *Forbes*, 22.03.2023, <https://www.forbes.com/sites/bernardmarr/2023/03/22/green-intelligence-why-data-and-ai-must-become-more-sustainable/> [Accessed 06/2024]; Emma Strubell, Ananya Ganesh, and Andrew McCallum: Energy and Policy Considerations for Deep Learning in NLP. In: *arXiv*, 05.06.2019, <https://doi.org/10.48550/arXiv.1906.02243>.

4 This was first done in 1972 with the computer models used by Donella H. Meadows, Dennis Meadows, Jorgen Randers, and William W. Behrens III.: *The Limits to Growth*, Falls Church 1972. Their 30-year update showed the accuracy of their original assessments, and recent developments continue to emphasize that such technologies can support social concerns. They just aren't being used that way. UN Environment Programme: *How Artificial Intelligence Is Helping Tackle Environmental Challenges*, 07.11.2022, <https://www.unep.org/news-and-stories/story/how-artificial-intelligence-helping-tackle-environmental-challenges> [Accessed 06/2024].

5 Mackenzie Holland: *Big Tech Invests Billions in AI Data Centers Globally*. In: *Tech Target*. 31.05.2024, <https://www.techtarget.com/searchcio/news/366587217/Big-tech-invests-billions-in-AI-data-centers-globally> [Accessed 06/2024].

oped from scraping the internet), known for problematic associations that reproduce “malignant stereotypes”.⁶

3. The *economic and labor critique* recognizes how these generators may replace some areas of human labor in creative industries ranging from illustration to film, and more.⁷

4. The *legal critique* meanwhile considers issues of copyright and intellectual property as regards scraping the internet for the foundational models and the outputs claimed by artists and authors.⁸

5. The *realist critique* initially stemmed from disdain for the “uncanny valley” errors in image production, particularly as regards hands, but improvement in image generation now lead some artists to bemoan the increasing accuracy as narrowing creative possibilities.⁹

6. The *creativity critique* depends on a psychological as well as metaphysical presumptions of imagination and curiosity;¹⁰ debates continue to proliferate on the necessity of the artist’s hand ever since Duchamp presented the found object as a work of art.¹¹

6 Abeba Birhane, Vinay Uday Prabhu, and Emmanuel Kahembw: Multimodal Datasets. Misogyny, Pornography, and Malignant Stereotypes. In: arXiv, 05.10.2021, <https://doi.org/10.48550/arXiv.2110.01963>; Emily Denton et al.: On the Genealogy of Machine Learning Datasets. A Critical History of ImageNet. In: Big Data & Society, vol. 8, 2021, no. 2, <https://doi.org/10.1177/20539517211035955>.

7 Jesse Damiani: How Future Histories Of ‘Other Intelligences’ Clarify Today’s AI. In: Forbes, 01.08.2023, <https://www.forbes.com/sites/jessedamiani/2023/07/11/how-future-histories-of-other-intelligences-clarify-todays-ai/> [Accessed 07/2024]; Harry H. Jiang et al.: AI Art and Its Impact on Artists. In: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, pp. 363–374, <https://doi.org/10.1145/3600211.3604681>.

8 Sarah C. Odenkirk: You Can’t Copyright AI Art (Yet). In: Right Click Save, 14.03.2023, <https://www.rightclicksave.com/article/you-cant-copyright-ai-art-yet>.

9 K Allado-McDowell in conversation at Book Launch: Rhizome Presents “Air Age Blueprint” and “The Institute for Other Intelligences,” The New Museum, 30.03.2023.

10 Mashinka Firunts Hakopian in Damiani (see note 7): “assumptions about imperiled creativity in the context of generative adversarial networks and large language models often invoke a model of creativity that we’d do well to discard. Consider Arthur Miller’s ‘The Artist in the Machine’. In this book, he attempts to answer, once and for all, the question of whether automated systems can be ascribed human creativity. To furnish that response, he codifies the characteristics of existing ‘geniuses’ (the book’s list includes figures like Picasso, Georges Braque, Philip Glass, ... Peter Thiel), then assesses whether computational systems can produce or approximate those characteristics... Creativity has been defined largely in relation to the creative output of canonical figures—with canonicity sketched through a Western and Eurocentric lens.”

11 Yayoi Shionoiri, Megan E. Noh, and Sarah C. Odenkirk: Out of Touch. How

Oddly, given the fact that these critiques and anxieties occurred across a media cycle responding to images, the many questions of aesthetics—what we see as producing ways of thinking, and vice versa—have been largely ignored, or misdiagnosed as issues of style.

Common parlance often confuses “aesthetics” with “beauty,” or treats it as a synonym for “style.”¹² Aesthetics, when it is acknowledged, often gets lost as a second order analysis, dependent for justification on its relation to ethics or epistemology. Ethics seeks to establish a set of values and presumes, or proposes, a code of human conduct, i.e. a moral system; for it, objects are too simple and need a higher order to validate them. Epistemology examines what we know and how we know what we know, how that is substantiated and communicated; for it, objects are too complex and need to be broken down into simpler components. For various reasons, aesthetics is perceived as untrustworthy on its own, necessitating validation through these systems. However, if we follow Jacques Rancière’s understanding of aesthetics as a regime that presents the “distribution of the sensible”—defined as “what is seeable, sayable and possible”¹³—then these public generators instantiate a new regime, one clearly necessitating aesthetic consideration to understand what is presumed as “seeable, sayable and possible.” Aesthetics assesses how “qualitative information regarding the world is made sensible and distributed to our sensory capacities.”¹⁴ The shifting forms of organization and interpretation reveal new constituencies—therein are the politics to be recognized.

Much AI Is Too Much (from a Copyright Perspective). In: The Brooklyn Rail, May 2023, <https://brooklynrail.org/2023/05/criticspage/Touch>.

12 Within the context of public generators, foundational models like LLaMA and GPT-4 often confuse the issue by talking about aesthetic analysis of the images in their dataset; for example, identifying: “features capturing the ‘aesthetic properties’ of an image are proposed with the aim of mimicking photographic rules and practices such as the golden ratio, the rule of thirds and color harmonies.” See: Naila Murray, Luca Marchesotti, and Florent Perronnin: AVA: A Large-Scale Database for Aesthetic Visual Analysis. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2408–2415, <https://doi.org/10.1109/CVPR.2012.6247954>.

13 Joseph J. Tanke: What Is the Aesthetic Regime? In: Parrhesia, 2011, no. 12, p. 71.

14 Michael Young: The Aesthetics of Abstraction. In: Mark Foster Gage (ed.): Aesthetics Equals Politics, Cambridge 2023, p. 129.

While working on this project, colleagues and friends have expressed disdain, horror, and outrage that I would “bother” or “waste my time” looking at these generated productions seriously. Claims that the outputs of such generators have no artistic value do not alter widespread acceptance that both affect socio-economic, political, and cultural systems. This essay attempts to aesthetically critique the new frameworks of these public generators and to suggest that aesthetics can support the efforts of other critiques. I focus on a project that neither alters the output nor engages animation for the sake of forcing the question of authorship and eliminating the issues of moving image given the already large number of elements within the realm of public generators; I selected a project moving between visual and verbal, however, to recall their close relationship within the theoretical discourse of the image, as I will address.¹⁵ From initially drawing on the work of Crispin, Holly Herndon, Patrick Leahy and Alexander Reben, I narrowed to Crispin’s project for its plumb line to Debord, but could make reference to Minne Atairu, Ari Melenciano, Michael Mandiberg, Trevor Paglen, or Laurie Simmons, among others.¹⁶ One final caveat, I generally avoid the term “AI” because it blurs a panoply of very different softwares and hardwares, erases the historical debates around “intelligence,” and reproduces a conflation between brain and machine processes that many scientists counter, which undermines distinctions around agency that I will address at the end of this essay.¹⁷

15 Since presenting these ideas in summer 2023 at the original conference, other projects have appeared that add additional considerations of subtlety and complexity, but my main argument around the set and agency remain, as the subsequent sections will show.

16 Charlotte Kent: How Will AI Transform Photography? In: Aperture, 16.03.2023, <https://aperture.org/editorial/how-will-ai-transform-photography/>; AI Am I? Yeah, We Are. Alexander Reben & the Machine. In: CLOT, 01.10.2020, <https://clotmag.com/oped/ai-am-i-yeah-we-are-alexander-reben-and-the-machine-by-charlotte-kent>; Art’s Intelligence: AI and Human Systems. In: The Brooklyn Rail, July 2024, <https://brooklynrail.org/2023/04/art-technology/Arts-Intelligence-AI-and-Human-Systems/>; Taking Stock of Generative “AI”. Systematic Work of Michael Mandiberg, Trevor Paglen, and Penelope Umbrico. In: The Brooklyn Rail, November 2024, <https://brooklynrail.org/2024/11/art-technology/taking-stock-of-generative-ai-systematic-work-of-michael-mandiberg-penelope-umbrico-and-trevor-paglen/>.

17 Emily Tucker: Artifice and Intelligence. In: Tech Policy Press, <https://techpolicy.press/artifice-and-intelligence/>; Robert Epstein: The Empty Brain. Your Brain Does Not Process Information, Retrieve Knowledge or Store Memories. In Short: Your Brain Is Not a Computer. In: Aeon,

The Set

The flat, neutral, and corporate style of the images Crispin selects obviates the expectation of a single intentional person having produced the image. These images evoke the stock photography that is the basis for the training data sets of these public generators;¹⁸ that commercial association undermines an association to art, while Crispin’s text and context affirm it as an art project. What this section argues isn’t for one or other position, but the challenging confluence of such aesthetic contexts. I propose “the set” as a framework for dealing with these complex mergers. It supports interpretation of the myriad elements constituting the whole process: the data set, the prompts, the outputs, the algorithms driving these productions to understand them as a totality. That ensures a constant reminder of the social and economic relations of these differing elements and the users defining them, while also recalling the process of translating such relations into symbolic constructs.

The problem of ontology still resounds across popular art discourse: the object depends on its *being* art rather than situated as art. This remains in a common fixation on the output.¹⁹ Scholars George Dickie, Howard Becker, and Arthur Danto, among others, proposed various arguments in the last quarter of the twentieth century to move away from ontology, postulating a larger context that enabled the identification of an object as art.²⁰ Their ideas supported the basis of “institutional

18.05.2016, <https://aeon.co/essays/your-brain-does-not-process-information-and-it-is-not-a-computer>; Will Douglas Heaven: What Is AI? In: MIT Technology Review, 10.07.2024, <https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/> [accessed 07/2024].

18 Roland Meyer: The New Value of the Archive. AI Image Generation and the Visual Economy of ‘Style’. In: IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft, vol. 17, 2023, no. 1, pp. 100-111, <http://dx.doi.org/10.25769/mediarep/22314>.

19 In generative art, ontological arguments appear in various debates regarding the code or the various outputs being the art. See: Jason Bailey: An Interview with Tyler Hobbs | Part 1. In: Right Click Save, 09.05.2022, <https://www.rightclicksave.com/article/an-interview-with-tyler-hobbs-part-1>; Alex Estorick: When the Artists Met the Algorithmist. In: Right Click Save, 11.11.2022, <https://www.rightclicksave.com/article/when-the-artists-met-the-algorithmist>; Stina Gustafsson: Casey Reas on the Art of Code. In: Right Click Save, <https://www.rightclicksave.com/article/casey-reas-on-the-art-of-code> [accessed 11/2024].

20 Howard Saul Becker: Art Worlds, Berkeley 2008; Arthur C. Danto: The Transfiguration of the Commonplace: A Philosophy of Art, Cambridge 1981; George Dickie: What Is Art? An Institutional Analysis. In: Art and the Aesthetic: An Institutional Analysis, Ithaca 1974, pp. 19-52.

critique” and movements to “free” art from the value proposition of establishments like museums and mega-galleries’ market power.

When artists invoke institutional critique to reject the gallery or museum only to affirm their own signifying function as “creators”—or celebrate the audience’s freedom to confer their own meaning and thus become “co-creators”—they pluck interpretive power from one place to deposit it elsewhere, without always retaining the plurality of interpretive possibility, nor assessing their reaffirmation of the object as art. Art, unlike other types of objects (especially amidst niche market tendencies), sustains an extraordinary plurality. The post-structuralist philosopher Roland Barthes posits the death of the author as the birth of the reader. But the text’s openness produces the sense of “merely” personal observation, given the seemingly endless plurality, dissolving the text’s own ability to anchor possibilities. Such arguments represent the well-worn rebuttals to post-structuralism, and a nostalgia for a master narrative. The loss of conclusive meaning (by individual or institution) gestures toward the birth of a new reader: one who cannot anchor or situate themselves without calling upon and claiming other entities and texts to delineate their position. Meaning appears through contextualization and must always be articulated through a set of references, as Jack Burnham emphasized in relation to art and technology during this same period.²¹

The flat affect that some describe regarding generated images expresses not only the general-use terms of stock photography but also the lingering disassociativeness wrought by post-structuralism’s widely misinterpreted interpretive field and the confusion around systems. Through *Spectacle*, Crispin reveals the way that public generators have materialized, to an

extreme, this issue remaining from postmodern discourse and crucial to the contemporary. *Spectacle*’s generated productions are explicitly and implicitly part of a “set”—as a whole body of work, as ten Debordian derivations, as 200 visual outputs, as ten sets of twenty, alongside the various sets associated with public generators. Public generators present a set of possible outputs for users to select amongst, and depend on datasets of foundational models and the sets of terms and concepts that mark and produce them, with the problematic relations addressed by the bias and economic critiques. By referencing Debord, Sterling insinuates early concerns of ideological influence surrounding the mass production and dissemination of images that have been expressed by subsequent scholars. Since artistic projects using generative AI are typically produced as multiples along a common theme, a culture of multiplicity becomes evident within this aesthetic.

In 1996, after many years of debate regarding whether photography could reveal an individual style, the British philosopher Nigel Warburton introduced “series” as crucial to that identification.²² Warburton refuted earlier dismissals that photographers could display individual style by arguing that the *series* becomes the container for an individual style—the composition, contrast, lighting, etc. presents the artist’s preferences, though these are only made apparent across multiple examples. I am adapting Warburton’s “series” into “set” to establish a link with data and affirm the variety of elements that constitute a creative project. The outputs of generative AI might be termed a series, but *Spectacle* represents a set that includes Debord’s text, Crispin’s prompts, the generative AI’s code and algorithms (even if these are black-boxed), the outputs, the underlying data sets for the different public generators he used. Warburton also identifies the photographer’s selec-

21 Jack Burnham: Systems Esthetics. In: Art Forum, vol. 7, 1968, no. 1, pp. 30-35; Real Time Systems. In: Art Forum, vol. 8, 1969, no. 1, pp. 27-38; Systems and Art. In: Arts in Society, vol. 6, 1969, no. 2, pp. 194-204.

22 Nigel Warburton: Individual Style in Photographic Art. In: The British Journal of Aesthetics, vol. 36, 1996, no. 4, pp. 389-398.

tion from the film roll, or page of negatives, which can now be applied to many artists' process with public generators. The set and selection were important attributes of the emergent aesthetic²³ of the 1990s, apparent in the rise of net art and also the ideas surrounding post-photography.²⁴ The selection process occurs because of an existing set. For photography and now generators, a single image needs comparative inflections to determine the artist's style. Since style is a recognized feature for aesthetics, set and selection thereby become crucial features of this aesthetic regime. The artist selects from a larger set to make a smaller personal set; set and selection, therefore, imply two important intellectual attitudes: a relationality to some larger entity and an extractive practice. The "set" becomes how these objects get "read" as conveying underlying ideas that artists wish to provoke.

The "reading" of images derives from Dante's *Convivio* (1304-1307), wherein he refers to images as the panoply of figurative language, interpreted according to four levels: the literal, allegorical, moral, and anagogical.²⁵ These in turn informed a rationale which was then applied to art history through such practices as iconography. Conceptual art eroded demarcations between the visual and verbal, image and text, idea and materialization, as evidenced in Joseph Kosuth's *One and Three Chairs* (1965); 1990s scholarship about the "image" emphasized the term's plurality, though its popular retention for "picture hanging on a wall" diminishes this academic effort.²⁶ So much overemphasizes the picture, focusing on the text merely as tag or prompt, while at the same time using the text as determinative. These need to be knit together. As Michael Young says in "The Aesthetics of Abstraction":

23 Monet offers an example of this in painting, widely associated with individuality, through his series of Chartres Cathedral, Haystacks or Waterlilies. It may be possible to consider that effort in relation to the general interest of stop motion practices of photography.

24 Robert Shore: *Post-Photography: The Artist with a Camera*, London 2014.

25 These are presumed to derive from Thomas Aquinas' 'Summa Theologica' (1274), which are Literal, Historical, Moral and Anagogical. Dante shifts the historical for allegorical and by applying this foursome to texts beyond religious texts establishes a modern method for literary interpretation.

26 W. J. T. Mitchell: *Picture Theory: Essays on Verbal and Visual Representation*, Chicago 1998; *What Is an Image?* In: *Iconology: Image, Text, Ideology*, Chicago 2009, pp.7-46; Thomas McEvelley: *Thirteen Ways of Looking at a Blackbird*. In: *Art & Discontent. Theory at the Millennium*, Kingston 1991, pp.70-80.

*"Aesthetics operates through tensions between the sensory and the intelligible, one that cannot be properly conceptualized prior to the redistributing break. This is true even though the redistribution's spark theoretical arguments regarding the importance of the transformation."*²⁷

This issue of the sensory and the intelligible takes on a particular significance in the context of how public generators work through the realm of the apparent, via identifiable prompt and output, as well as the known but elusive dataset and code. In addition, a prompt or output may be textual or an image. One can produce the other in a confusion of significant importance. Public generators materialize the theoretical debate from thirty years ago, and challenge us to think through how the sensible and intelligible are fractured and sutured in *this* moment. To link the set of foundational data—inclusive of images, "structure of feeling" determining image production,²⁸ captions, the socio-political dynamics of tagging, and design attitudes informing machine learning code—together with the prompt and output forces a new way of thinking about images overall. The output then becomes just one point in this distributed network. I will hereafter call this package a "large model set" (LM set), given the large (language or diffusion) models that undergird public generators, with an emphasis on the *set* as a practice associated with this regime. Network relations was the significant contribution of second order cybernetics and the establishment of systems thinking, which mitigates easy notions of origin and authorship but invites thinking about how selecting a focus, the scale of examination, any boundary determinations, among other criteria, determines what an analysis can and will present.

27 Michael Young: *The Aesthetics of Abstraction*. In: Mark Foster Gage (ed.): *Aesthetics Equals Politics*, Cambridge 2023, p.138.

28 Referencing here Raymond William's concept, productively visualized by Penelope Umbrico's *Suns from Sunsets from Flickr* (2006-ongoing) and *Everyone's Photos Any License* (2015-2016), discussed in: Kent: *Taking Stock of Generative "AI"* (see note 1b).

The Aesthetic System

In 1968, Jack Burnham proposed “systems esthetics” to describe a shift in how to perceive art. He argued art is no longer about making things to see but about seeing how decisions are made; the artist used to be a “Homo Faber” (maker of things) but has become “Homo Arbitrator Formae” (the maker of aesthetic decisions).²⁹ Half a century later, these generative systems seem to undermine even that. Crispin is an advocate for Bruno Latour’s Actor-Network Theory for shifting away from authorial intentionality, that is goal-oriented decision making, to agential actor.³⁰ Anything from a windstorm to an artwork can present agency, but its import shifts according to its framing within a complex of network constituents. Authorship and agency are linked in part because of language slippage and terminological imprecision: an author is presumed to be an agent producing a text, and an agent presumed able to author their own acts and texts. Agency is typically attributed to someone with the ability to make something happen. This common conception collates autonomy and agency, where autonomy is the freedom to choose, and agency is having the necessary resources and abilities to enact that choice. All of this orients around the human actor, which Latour and others have put in question. These anthropocentric critiques are largely associated with posthumanism (to include science and technology studies’ examination of the broader implications of quantum physics), but also appear in social anthropology and medieval art history, which have reintroduced an understanding of objects’ agential force dismissed or lost in European theorizing influenced by a post-Reformation, “disinterested” art.³¹

²⁹ Burnham (see note 21), p. 35.

³⁰ Tweet by @sterlingcrispin, 26-06-2023, <https://twitter.com/sterlingcrispin/status/1673440736221868032>. “I think the most accurate mental framework to understand the world involves some version of Actor-Network Theory. It’s especially important to consider if you’re trying to make sense of how fast the world is changing with GPT4 like models [–].”

³¹ For more on this, I recommend the debates surrounding Carlos Fausto: Art Effects, Image, agency, and Ritual in Amazonia (2020). In: HAU: Journal of Ethnography Theory, vol. 11, 2021, no. 3, particularly comments by Caroline van Eck and Caroline Walker, pp. 1235-1237 and 1238-1243, respectively.

To deny the LM set agency purely because it is not an obvious, singular author refuses to consider an emergent regime wherein multiple frameworks and references require articulation. I believe that Crispin’s adoption of Latour’s framework represents an effort to engage the challenge of systems thinking. That publics already confer agency on video games (as promoting violence) or flags (designating ownership as, for example, stated by the Principle of Effective Occupation expressed by the Berlin Conference in 1884-1885) already reveals the confusion engendered by the term. Copyright claims by artists against public generators (and those using them) highlight the breadth of authors conceived within this milieu and the challenge they present to individualism. In Crispin’s work, there is a sense that any image could be replaced by another. As a conceptual art project, this serves his aims, but audiences (and legal frameworks) continue to demand: who is the maker, aka the author? The problem stems from differing definitions of “making”: conceptual art legitimated the idea as the fundamental act of making rather than the object; appropriation art legitimated found objects and the placement of the object over its construction; some forms of performance and installation art legitimated audiences as creators or co-creators of the work, although those artists depend on conceptual art and legal frameworks that reinforce the idea as distinguishing the creative act; some have disputed the lack of acknowledgment given studio assistants and construction partners. With public generators, these conversations come to the fore again, and require reassessing the dependency on a singular author as agent.

The LM set exists within a network of relations to other actors (like the artist’s statement, institutional situation, platform of distribution, etc.) that inflect its agency. So, when an actor known as an artist takes an LM set and places it at Sotheby’s,

that LM set “does” something to art culture and the art market more than when another actor like my cousin texts an LM set to a friend with an LOL emoji. But that cousin’s text multiplied by many, many others doing the same thing likewise makes these LM sets in general “do” something to society more generally. That’s Debord’s point about the relations produced by images, which here applies more broadly to the complex of the LM set. This has nothing to do with the beingness of the image as art necessitating an author, which is the language and framework of prior regimes associated with modernity. The confusion around the authorship of an output strains to keep a discourse alive rather than considering how it must broaden. The LM set introduces systems thinking as a necessity for this aesthetic regime. Over the last fifty years, an interest in contextualizing an artist’s culture, background, and practice to explain the work has destabilized the residual authoritative judgements leftover from the academic hierarchies and practices of prior centuries’ making. Acknowledging that a work emerges amidst contexts reinforces the need for some form of systems thinking. Public generators make this explicit.

Most object to being in service to the machine, even though our activities provide data enabling the economies behind websites, social media, and apps. The abstraction of such data architectures and extractive economies challenged the imagination, but LM sets have broadened social recognition, perhaps because it impacts “white collar” work.³² Much of the hype around generative AI’s capabilities stems from techno-utopian claims of its closing the gap on “Artificial General Intelligence”, whereby the prompt-based machine or Siri/Alexa model evolves into an autonomous agent. So what do we mean by agent in these complex systems that include software designers, hardware limitations, coded productions that even designers can’t

32 Concerns are not dissimilar to previous denunciations of data scholars: Cathy O’Neil: *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*, New York 2016; Shoshana Zuboff: *The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power*, New York 2019; Bernard E. Harcourt: *Exposed. Desire and Disobedience in the Digital Age*, Cambridge 2015; Benjamin H. Bratton: *The Stack: On Software and Sovereignty*, Cambridge 2015; Lisa Nakamura, *Digitizing Race. Visual Cultures of the Internet*, Electronic Mediations 23, Minneapolis 2008; Safiya Umoja Noble: *Algorithms of Oppression. How Search Engines Reinforce Racism*, New York 2018.

explain, as well as already nebulous complex agents such as humans, society, and artwork? Amidst the pluralities already discussed surrounding these new systems, I wish to introduce the same for agency, using a concept proposed by philosopher Jennifer Corns of “distinct agentive forms” that helpfully broadens how we can conceive of ourselves as agents.

Corns introduces these forms in “Suffering as significantly disrupted agency”, where her argument orients to the forms of agency enjoyed by humans, but notes that her theory builds from one with a focus on artificial agents.³³ Agentive forms “are overlapping and dynamically related, but the systems, capacities, and norms which partly constitute these different agentive forms are nonetheless usefully distinguished,” with at least three extant by virtue of being bio-psycho-social beings.³⁴ We know that biological hunger can alter psychological states, which can impact social interactions. We then excuse such mindsets and cranky behavior by attributing it to hunger; this represents an attitude already open to Corns’ multiple agentive forms. Corns’ argument makes explicit a model of multiple and distinct agentive forms in which we are already participating every time we assess the complexity of agencies presented in our encounters with each other. As Corns indicates, “[o]ur agentive forms likely proliferate. So, for some instances, it seems to me that we have financial, aesthetic, familial, and creative agency.”³⁵ Corns elaborated in a webinar that “Whenever we can identify a range of activities, an environment in which the activities are done, and a standard of success for succeeding in that range of activities, we have identified a form of agency.”³⁶ For example, an artist who uses software and enjoys coding may experience it as enhan-

33 Email correspondence with author, 30-09-2024, making reference to: Xabier E. Barandiaran, Ezequiel Di Paolo, and Marieke Rohde: *Defining agency. Individuality, normativity, asymmetry, and spatio-temporality in action*. In: *Adaptive Behavior*, vol. 17, 2009, no. 5, pp. 367-386.

34 Jennifer Corns: *Suffering as Significantly Disrupted Agency*. In: *Philosophy and Phenomenological Research*, vol. 105, 2022, no. 3, p. 715.

35 Corns (see note 34), p. 716.

36 As quoted from webinar conducted by Jennifer Corns and Ben Colburn: *Understanding and Enriching Agency in Palliative and End of Life Care*, Glasgow, 13-09-2023, https://www.gla.ac.uk/colleges/arts/aboutus/news/headline_999985_en.html. They subsequently published: *How Can an Understanding of Agency Be Helpful in Alleviating Suffering?* In: *SPICE Spotlight|Solas Air SPICE*, 10-10-2024, <https://spice-spotlight.scot/2024/10/10/how-can-an-understanding-of-agency-be-helpful-in-alleviating-suffering/>.

cing their creative agency, with possibilities for broadening their social or financial agency. An illustrator who sees these programs as impacting their employment feels their creative agency devalued by employers preferring the speed of slick renderings in generative AI, with negative consequences for that illustrator's social and financial agency.

Returning to aesthetics, the LM set reveals a troubled agentive form: not so much that of creativity, which despite questionable efforts remains an ill-defined term, but that of aesthetic judgment. In 2022, when Jason Allen won first prize at the Colorado State Fair in the digital art category for a work produced using Midjourney, reactions were hostile, seemingly for "authorship" and "creative" reasons; but, I suggest the dissent came from a lack of knowledge that he had done so, that such submissions were possible, and confusion on the process (he applied at least 624 prompts to get the image he sought).³⁷ I propose this because far less distress occurs around projects that are explicit about their use of public generators, as for example Holly Herndon, who has been lauded in magazines ranging from *Art Review* to *The New Yorker* for her work. Though unarticulated, my suspicion is that the transition from assumptions around a single agent creating a work of art to the complex system of the LM set produce a sense of unease regarding who and what one is observing and therefore the agentive form harnessed for such assessments is destabilized. When one doesn't know that the "agent" is the complex of an LM set, the reaction is outrage akin to confronting a liar. The problem isn't just the machine, as such, but the uncertainty of how to judge such a joint effort. We do not have norms for such assessments.

Agentive norms indicate the range of conditions under which an agent constitutes its integrity as such, so with agentive

forms associated with human flourishing, as Corns' work initially investigates, that means a set of norms for biological, psychological, social, familial, etc. behaviors. Each form of agency involves norms that specify when one is succeeding as a friend, a daughter, a thinker, and so on. These standards of success will differ across people, cultures, and times, though there seem to be some loose generalizations possible. As she further explained: "These many forms are functionally integrated and regulated by a range of processes. Conscious decision making is one type of process through which one form is prioritized over another, but there are others."³⁸ The LM set instantiates several overt agents cooperating, despite themselves: the dataset as it establishes the statistical range of potential image production;³⁹ the software design that limits, or not, what the model can produce;⁴⁰ the human prompter.

Corns' argument aims to distinguish how an agent has multiple forms of agency, some of which may take precedence from one moment to the next. The bias critique, for example, emphasizes the agency of the training data set. Corns' argument also reveals how judgements around agency value one agentive form over another, from one context to another. When bills are due, one's financial agency is of greater value to the payee than any familial agency oriented towards caring for an aging parent. A corporation, like a human agent, has financial and social agency, but not biological agency. The latter does not undermine the corporation as being a potentially powerful agent within financial markets. Likewise, a corporation's social agency may be emphasized and valued over its financial agency by those pursuing environmental, social and governance (ESG) investment portfolios.

Crispin's *Spectacle* helpfully reveals this variety of agentive forms because its flat, quasi-neutral styling disrupts our expect-

³⁷ Sarah Kuta: Art Made With Artificial Intelligence Wins at State Fair. In: Smithsonian Magazine, 06-09-2022, <https://www.smithsonianmag.com/smart-news/artificial-intelligence-art-wins-colorado-state-fair-180980703/>.

³⁸ Email correspondence with author, 30-09-2024.

³⁹ See: Anthony Downey (ed.): Trevor Paglen. *Adversarially Evolved Hallucinations*. Berlin 2024; Prithvi Iyer: *Looking Beyond the Black Box. Transparency and Foundation Models*. In: Tech Policy Press, 24-10-2023, <https://techpolicy.press/looking-beyond-the-black-box-transparency-and-foundation-models/>.

⁴⁰ Efforts to mitigate bias re-introduce the questions surrounding content moderation. For more, see: Emilio Ferrara: *Should ChatGPT Be Biased? Challenges and Risks of Bias in Large Language Models*. In: First Monday, November 2023, <https://doi.org/10.5210/fm.v28i11.13346>.

tation of a distinct artistic voice, and so within the context of art invites the types of queries I have tried to unravel here. I cannot offer at this time an answer to the question of judgment within this new aesthetic system, but invite readers to consider the agentive forms they bring to the various LM sets they encounter, as well as those presented by the LM set. To blame the LM set as the product of a negative techno-capitalist power reiterates the necessity of the author, and ignores the contemporary complex of human, machine, and corporate entanglements. Authorial intentionality was crucial to one discourse around art, but that has been under increasing dispute for a century and is now evidently limiting the possibilities of our thought. The anger expressed by many about how these generators “undermine artistic creativity” aims to return to an earlier model of authorship and authority, one which can undermine efforts to engage generators’ impact where they are significantly different. Crispin’s reduction of Debord, output selection, and set production for *Spectacle* has a preposterous element that invites laughter, which can then operate as an oppositional affect—one that destabilizes the fixity of our position in relation to these generators and may even allow for new ways of encountering them. These aesthetic systems challenge established values and criteria of judgment, social models and concepts of agency. Amidst the interrelations of globalization, thinking in terms of systems and taking responsibility for the systems we select to acknowledge in considering a set has larger implications, which aesthetics has made sensible through public generators.

Conclusion

In an essay from 1997, Martin Jay identifies three scopic regimes of modernity, as part of an argument he later expanded upon in 2012.⁴¹ Here he presents the dominant regime as a Cartesian perspectivalism, though its assorted known tropes are more complicated than is claimed by those who implicate a fixed, uniform position of power in league with a domineering scientific worldview. Jay’s point, often forgotten, is that there have been and remain moments of unease in a dominant paradigm, contesting and complicating its status as *the* scopic regime; he also discusses regimes associated with the Baroque and the Netherlandish art of description, relying on Christine Buci-Glucksmann and Svetlana Alpers respectively for his understanding of both.⁴² Scopic regimes focus on the visual, however in the context of public generators, to speak of “aesthetic regimes” could allow us to address the complexity of the practice and production of these generators. By avoiding the language of the ocular, we can widen how we think about them and move beyond the argument of pictures and videos, of good and bad, of style and technique, of cores and margins.

Attempts to overcome dominant regimes often fall foul of the fantasy that this liberation will free us from politics as such. Different aesthetic regimes dominate in different contexts, representing distinct ideologies. Unauthored objects with destabilized meaning should not revert us back to the comfort of hierarchies enabled by Cartesian perspectivalism or even the Baroque. These public generators don’t frame the world as prior regimes determined, but instead plant us within a destabilized environment. Aesthetic defamiliarization cracks a legible politics. When an aesthetic presents a “moment of

⁴¹ Martin Jay: *Scopic Regimes of Modernity*. In: Hal Foster (ed.): *Vision and Visuality*, Bay Press 1988, pp. 3-23.

⁴² *Ibid.*, pp. 13-17.

unease” within the dominant scopic regime as Jacqueline Rose articulates it, or a “estrangement” as Viktor Shklovsky termed it, the aesthetic has fractured a dominant mode of reading—that is why it can seem illegible or illegitimate.⁴³ It disrupts the established systems for the organization of sensible information and in that moment of destabilization produces opportunities, danger, and anxiety. The legitimating structures of ethics or epistemology don’t easily slide onto a new aesthetic, but since aesthetic makes evident changes within the social and political which themselves necessitate revisiting by ethics and epistemology. An aesthetic mode of analysis offers an opportunity to consider what new tactics may be necessary to address the associated and insidious politics of an emergent technology. Public generators present us with a regime that evades authorship, veils lineages, lacks intentionality, collapses medium specificity, evacuates or confuses meaning, but also resists individualism, connects diverse actors, coheres plural modalities, encourages mobile frameworks, posits complex agentive forms, and emphasizes networked relations. Change is present. Aesthetics makes us sensible to an emergent social, political, and cultural system, and the criteria of judgement that we still lack.

Acknowledgment

The research to develop the conference presentation into this essay was supported by two grants: *Arts, Agency and Automation* from Google’s Artists + Machine Intelligence Research Awards; and *Arts, Agency and Automation: A Global Cultural Affair*, from National Endowment for the Humanities: Dangers and Opportunities in Technology, with recognition that any views, findings, conclusions, or recommendations expressed in this article do not necessarily represent those of the National Endowment for the Humanities.

Liquid Images. Holly Herndon, Mat Dryhurst, and Tyler Hobbs in conversation with Alex Estorick

In his book, *After Art*, David Joselit sought “to link the vast image population explosion that occurred in the twentieth century to the breakdown of the ‘era of art’”. The consequence, he argued, was a new kind of “image power [...] derived from networks rather than discrete objects.”¹ Given the growing number of artists developing transmedia practices, art is increasingly becoming a space for border thinking at the intersections of science and technology. For Joanna Zylińska, singular images such as photographs are “giving way to image and data flows” such that they are now “both objects to be looked at and vision-shaping technologies, for humans and machines.”² If generative artists are actively engaged in crafting an *output space* of multiple possibilities, those working with machine learning are curating an *input space* as the basis for new generations.

It was Tyler Hobbs who coined the term long-form generative art to refer to “a special class of artistic algorithm” that outputs hundreds of images, each transferred to a collector without any intervention or curation from the artist.³ This new development had been stimulated by the online platform Art Blocks, launched the previous year, which allowed for code to be minted as art on the Ethereum blockchain. Characterized by large populations of images, long-form generative art represents a departure from the historical tendency of artists working with code to curate small collections of prints or plots for exhibition in physical space.

1 David Joselit: *After Art*, Princeton and Oxford 2013, pp. 88 and 94.

2 Joanna Zylińska: *AI Art. Machine Visions and Warped Dreams*, London: Open Humanities Press, 2020, p. 106.

3 Tyler Hobbs: *The Rise of Long Form Generative Art*, 06.08.2021, <https://www.tylerxhobbs.com/words/the-rise-of-long-form-generative-art> [accessed 10/2024].

43 Jacqueline Rose: *Sexuality in the Field of Vision*, London and New York 1986, pp. 232-233, quoted in Jay (see note 41), pp. 3-4. Also, see Viktor Shklovsky: *Art as Device*. In: *Theory of Prose*, translated by Benjamin Sher, Funks Grove 1925, pp. 1-14.

In his original essay, Hobbs made the point that “[n]obody, including the collector, the platform, or the artist, knows precisely what will be generated when the script is run, so the full range of outputs is a surprise to everyone.” For this reason, the long form has often seemed to reveal the emergent possibilities of code, while offering collectors an experience personal to them. However, thus far, the discussion around human-coded generative art has tended not to overlap the conversation around generative AI, whereby advanced machine-learning (ML) algorithms trained on vast datasets produce new images, texts, and videos. Holly Herndon and Mat Dryhurst have been making ML models since 2017, contributing to the development of the AI image generator, DALL-E, while their independent and collaborative practices also focus on music and voice. Following the launch of their work, *The Call* (2024), at the Serpentine, Alex Estorick hosted them in a conversation with Tyler Hobbs about creativity at the edge of the human.

Alex Estorick: Tyler, how did your own practice shape your ideas about long-form generative art?

Tyler Hobbs: Shortly after I began making generative art ten years ago, I became very interested in the idea of seeing how far I could produce a varied and continuously fascinating stream of output from a single algorithm. With my own practice, I worked to understand how to grow the output space and the number of interesting outputs that I could get from a single algorithm. It took me a while to get even one interesting image from an algorithm, much less three, five, or ten. However, after a number of years, I released *Fidenza* (2021) with 999 uncurated outputs that ended up being quite popular (Fig. 1).

But I’ve also looked at versions of this style of work that involve curation. For example, I did a project called *QQL* the following year that was curated [by the collector] from a potentially unbounded stream of outputs from the algorithm (Fig. 2).

My essay pointed out the sudden rise in popularity of what I called ‘long-form’ generative art. This referred to the practice of crafting one complex algorithm from which you want to see hundreds or thousands of outputs versus what had typically been the practice of generative artists, which was to craft an algorithm where they might curate the single best output or else maybe a handful of select outputs. A simple way to put it is: “How many images are we looking at from the algorithm?” But a more interesting way to think about it is: “What is the complexity of the output space?”

AE: The idea of the long form has provoked a lot of discussion among artists. Jeff Davis has spoken of his interest in “narrow algorithmic spaces,” while Aleksandra Jovanić questions the premise of long-form on the basis that creating an algorithm already presupposes an infinite number of outputs.⁴ Julien Gachadoat queries whether ‘long-form’ is evocative because “what is ‘long’?”⁵

TH: That is a perfectly valid question, but just because it’s a spectrum rather than discrete classes of algorithms doesn’t mean that you can’t attempt to label the ends of the spectrum. I think that ‘short’ and ‘long’ are a decent place to start, if not perfect.

On the question of infinity, that is where the complexity of the output space matters. For example, if we take a pixel grid and randomize each pixel as black or white, we’re going to get

4 Alex Estorick: The Color of Code | Jeff Davis. In: Right Click Save, 08.05.2023, <https://www.rightclicksave.com/article/the-color-of-code-interview-jeff-davis> [accessed 10/2024].
5 Alex Estorick: The Power of the Plotter. In: Right Click Save, 14.11.2023, <https://www.rightclicksave.com/article/the-power-of-the-plotter-generative-art-aleksandra-jovanic-julien-gachadoat-feral-file-graph-interview> [accessed 10/2024].

a massive number of outputs but very little complexity in what we're seeing. Of course, it is difficult to define complexity—we might think in terms of *conceptual compressibility*, for example—but I think it's pretty easy for viewers to recognize complexity when they see it, and that is the substantial distinction that I'm talking about. Artists today are targeting substantially more complex output spaces than generative and algorithmic artists did in the past.

Mat Dryhurst: I think that the long-form frame is useful in distinguishing between producing a bunch of outputs and selecting a group as a collection. There's something useful in the essay establishing that this is actually a new kind of practice because I do think it is. Then there is the secondary consideration of the model, pioneered by Art Blocks, of creating scarce and collectible moments. Of course, that does have some relation to the market but it is also a legitimate proposal for how to value these things that I welcome. The analogy in our world is *modeling*. There's something powerful about an artist saying, "No, actually, my algorithm or my model was constructed over time with a bunch of different curated inputs with lots of trial and error in order to come up with a system that allows people to interact with it." That is where you really want optimal complexity or generalizability in the machine-learning space.

The striking thing to me is the analogy between this kind of long-form generative algorithmic process and what we've been thinking about in parallel, which is how to create a model as an artwork. That involves similar challenges. In developing *khairymutantx* (2024) for the Whitney Biennial, we ended up trying to produce a text-to-image model that would generate reliably general enough outputs that it was fun to play with

and you felt like your contribution was actually meaningful. I understand that it is more limited in the traditional generative sense than it would be in the text-to-image space but it seems like a complementary problem. To me, the question is: "When you are in a sea of abundant imagery or infinite possibilities, what is the unique or scarce element?"

Holly Herndon: We do both the open version and the curated version. One way of being open is being really playful with identity and allowing people to perform my identity (through *Holly+*). The scarcity is that there is only one me but the non-scarce thing is that I can let anyone perform me. There's an interesting push and pull there. Coming from a background defined by hypercuration where I wanted to control every second of sound that I put out into the world, it was definitely an aha moment when I heard people performing through my voice. It was an entirely new way of interacting both with other people and with my own voice. That really unlocked this duality that we have of both curated outputs, and openness with my IP by letting people perform through it. I think there's room for both.

TH: I think that core to these practices is shaping the model and, Holly, although I didn't get to actually experience your work at the Whitney in person, it felt like visitors were curating outputs from the entire potential output space of that model, cherry-picking little moments. Maybe they're not particularly skilled curators with that model because they're not so familiar with it, but they are doing informal curation. I think that the more generative a work is, the more important the moment of curation becomes in establishing artistic significance in a world where there are so many images.

MD: I think that's right, but I've also spent a lot of time pushing back against the idea that curation is the new creation.

HH: We're also using the term curation really liberally here. There's the simple act of choosing which bit gets out to the public. But there are also so many different layers to curation – you have the model and then different people perform the model, and what they choose to show is their own filter.

MD: One of the things that we've looked at a lot in recent years is embedding space, also known as latent space. For a reader, that would be how concepts are determined in large models. When you're talking about the embedded concept of a chair, generally speaking, there's a universal idea of a chair's form, and so if you take a million images of chairs and feed them to your model, when you type "chair", everyone will probably agree on the output. But when it comes to a person or a subjective term like beauty, these things get really hairy. Stable Diffusion's impression of beauty is going to be whatever was tagged beautiful the most times in its training set, which has this air of objectivity to it but ultimately is curated somewhere. As artists, that level of curation is further upstream than the image, so if you want a model to represent beauty in a way that encapsulates the voice or eye of an artist, you have to go further upstream in the development of tools, protocols, training sets, or an ensemble of algorithms in order to establish that concept of *beauty*.

AE: It seems to me that one of the exciting possibilities of long-form generative art is that it offers a pathway to new and emergent outcomes because of the role played by randomness in the creative process. Holly and Mat, your project *Infinite*

Images (2022) was informed by your work on the first iteration of DALL-E. From your experience, what is the emergent potential of machine learning?

MD: In machine learning there is a concept of *overfitting* where, for example, if you train a model poorly on a corpus of 20th-century art, then you might get locked into a groove where you can only produce things that look like Picasso. The ideal is that you would have a wide enough range of data to produce new outputs... I do think that these models are demonstrably capable of emergent outputs. I don't think that they're manifesting new stuff out of nowhere, rather the sheer complexity and volume of potential permutations in these large models is highly capable of novel combinations.

HH: With DALL-E, if that was still the model that everyone was using, it would be boring, but because it was this weird snapshot in time, it had this really particular grain and quality. We were almost playing it like an instrument, making aesthetic decisions based on what it would want to do. There was something really interesting about that.

The first album, *Movement* (2012), used the most up-to-date vocal processing available. Then, with *Platform* (2015), we were sampling the sounds of our browsing using an instrument called Net Concrete in order to capture our new life online. With *PROTO* (2019), we allowed the neural net outputs to be scratchy and dissatisfying in their timbre, which was a challenge because they were paired with really high-fidelity ensemble recordings. We had to degrade the other recordings so that they could live in the same universe together because we didn't want to release an album that was all about AI as shiny, clean, and perfect. Its weirdness is so much more beautiful.

Right now, so much sound work is retro and you can land in an aesthetic cul-de-sac where you're just circling the same 1970s synth over and over. There's something disheartening about that. Large models often capture a midness or average, which is on purpose because the goal is to have an accurate reflection of a cup and not necessarily a subjective personal relationship to your grandmother's cup. They're not really built for that. We make bespoke training data because we see that as artwork in itself, and if you view training data as beautiful mind-children that you can send to the future, the models can always be changing but you have this evergreen training data that can always be tapped into for new and interesting outputs.

MD: We use the term protocol art to avoid fetishizing a particular tool. The term algorithmic art is very laden with a particular toolset, but when you talk about protocols you could just as easily be talking about the Bible or the Quran. I would argue that the successful aspect of any artist's practice is the protocol, or practice, rather than whether they're using generative or machine-learning tools. It's also what differentiates them as more than a keyword associated with an image on a screen. Otherwise, it can be hard to individuate the artist and distinguish something we do, for example, from something a company does. If you regard an AI model as largely the abstract product of its training data, then you are overlooking how that data was coordinated, which might be a site of performance or even politics.

AE: A lot of machine-inflected images have a kind of beguiling polish, eliding the complexity of the training data into something seamless. Do you feel that?

TH: Yes, I do. There are so many traces to everything that's done by a machine as a simple consequence of the fact that, fundamentally, they're dealing with ones and zeros as well as grids, arrays, and loops that have such a big influence on everything that comes downstream from that. There are also other limitations like processing power or the amount of time or storage or memory that is available. Over time, that will shift but there will probably be some trace in the creations until far into the future. Just in terms of information density, the physical world has far more information than digital items tend to contain, so there is something of that artifact as well. But that limitation is also a source of novelty, including aesthetic novelty, which is part of the reason that I find images created by the computer to be highly interesting. There are also human contextual associations that we bring to our viewing of images that make them important to us that one can't discount.

MD: No matter how impressive the model, if you go further upstream you will likely find a human who is making really consequential creative decisions. That might be artists like us but it might also be someone at Meta making a decision about the UX of a piece of software or a protocol that we interact with every day. To me, this kind of upstream analysis of the site of the algorithm or model or UX is really useful. People are already developing a fluency with the images that they encounter as well as a resilience to deepfakes. All that is happening is that their lexicon of understanding of images and media is maturing very quickly. If you talk to an 18-year-old, they have so much more awareness about the provenance of media and its potential objectives than I did when I was that age. People don't need to nerd out about this stuff to accumulate that immunity. The human artist is going to be just fine.

Fig.1, Tyler Hobbs, Fidenza #438, 2021, Courtesy of the artist

Fig.2, Tyler Hobbs and Dandelion Wist, parametric artist Appleboy, @QL #154, 2022, Courtesy of the artists

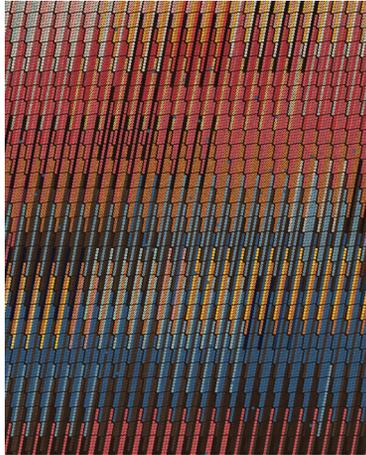


Fig.3, Holly Herndon and Mat Dryhurst, xhairymutantx, prompt response to "sam altman visiting an art gallery", 2024

Fig.4, Holly Herndon and Mat Dryhurst, xhairymutantx, prompt response to "jesus on the cover of artforum, koda-chrome", 2024



Holly Herndon and Mat Dryhurst are artists renowned for their pioneering work in machine learning, software, and music. They develop their own technology and protocols for living with the technology of others, often with a focus on the ownership and augmentation of digital identity and voice. These technical systems not only facilitate expansive artworks across media, but are proposed as artworks unto themselves. They were awarded the 2022 Ars Electronica STARTS prize for digital art. They have sat on ArtReview's Power 100 list since 2021. Holly holds a PhD in Computer Music from Stanford CCRMA. Mathew is largely self-taught. They have held faculty positions at NYU, the European Graduate School, Strelka Institute, and the Antikythera program at the Berggruen Institute. They publish their studio research openly through the Interdependence podcast, and co-founded Spawning, an organization building AI models on consenting data. Their critically acclaimed musical works are released through 4AD and RVNG Intl.

Tyler Hobbs is a visual artist from Austin, Texas who works primarily with algorithms, plotters, and paint. Hobbs' artwork focuses on computational aesthetics, how they are shaped by the biases of modern computer hardware and software, and how they relate to and interact with the natural world around us. Hobbs' Fidenza (2021) series profoundly impacted the generative art landscape, reshaping perceptions of "long-form" generative art. Hobbs' two most recent solo exhibitions were at Pace Gallery in New York and Unit in London. His work is in the collections of multiple prominent institutions including the Los Angeles County Museum of Art and the San Francisco Museum of Modern Art.

Alex Estorick is a media theorist who seeks to develop inclusive approaches to new technologies. As Editor-in-Chief at Right Click Save, he aims to drive critical conversation about digital art. He has developed some of the first academic courses on blockchain, NFTs, and Web3 and was lead author of the first aesthetics of crypto art. In 2023, artnet named his artistic collaboration with Ana María Caballero, artifacts, as one of the artworks of the year.

Herausgegeben von
Matthias Bruhn
Katharina Weinstock

DFG-Schwerpunktprogramm ‚Das digitale Bild‘



Erstveröffentlichung: 2025
Gestaltung: Lydia Kähny, Satz: Annerose Wahl, UB der LMU
Creative Commons Lizenz:
Namensnennung - Keine Bearbeitung (CC BY-ND)
Diese Publikation wurde finanziert durch die Deutsche
Forschungsgemeinschaft.
München, Open Publishing LMU

DFG Deutsche
Forschungsgemeinschaft

UB | Universitätsbibliothek
Ludwig-Maximilians-Universität München

Druck und Vertrieb im Auftrag der Autorin/des Autors:
Buchschniede von Dataform Media GmbH
Julius-Raab-Straße 8, 2203 GroÙbebersdorf, Österreic

Kontaktadresse nach EU-Produktsicherheitsverordnung:
info@buchschniede.at



DOI <https://doi.org/10.5282/ubm/epub.126472>
ISBN 978-3-99181-339-2

Reihe: Begriffe des digitalen Bildes
Reihenherausgeber
Hubertus Kohle
Hubert Locher



