

XVII. Die Kunst der Eingabe oder »A Shiba Inu dog wearing a beret and black turtleneck«

Der im Rahmen der Mikrokosmen des Merkens und Wohnens erhobene Befund von bestimmten Tendenzen zur Normalisierung und Habitualisierung in virtuellen Realitäten – ob auf der Ebene hyperbolischer Räume, unmöglicher Dinge oder unkonventioneller Erzählformen – ist auffallend und er ist alles andere als intuitiv eingängig. Zum einen steht er quer zu der phantasmatischen Überfrachtung mit Alteritätsversprechen, die gerade zu Beginn der Entwicklung mit einer gewissen Obstinate an das Virtuelle in Form des Cyberspace geknüpft wurde. Diese Ausrichtung auf einen Makrokosmos markiert einen gut überschaubaren Abschnitt in der noch jungen Historiographie einer technischen Virtualität. Zum anderen steht er aber auch quer zu einer anderen kulturellen Praxis, die seit geraumer Zeit dabei ist, die Ordnung sprachlicher wie visueller Bedeutungen auf den Kopf zu stellen. Die Rede ist von den künstlichen Intelligenzen und dem, was diesen als spezifische Kreativität attestiert wird. Während es in den Köpfen der mit einer virtuellen Mnemotechnik Befassten so aufgeräumt und katachrestisch gehegt zugeht wie in den stereotypen Wohnszenarien, in denen eine überaus übersichtliche und profane Ordnung des Imaginären herrscht, die den Dingen habitualisierte Orte zuweist, werden außerhalb derartiger Aufgeräumtheiten die Bilder und Bedeutungen mit aller Gewalt, im großen Stil und mit dem geballten Sachverstand dessen entfesselt, was künstliche Intelligenzen inzwischen zu leisten vermögen.¹ Hier werden Beziehungsstiftungen sichtbar, die sich in der Logik ihrer Bezüge scheinbar nicht mehr verantworten müssen. Die Gerichtshöfe der beziehungsstiftenden Vernunft, die sich im Rahmen der Historiographie der Gedächtniskunst immer wieder einberufen fühlten, sind gleichermaßen beredt wie stumm. Sie ergehen sich in der Versammlung von Beispielen, die scheinbar keiner eigenen Erklärung mehr bedürfen.

1 Zur historischen Herleitung des Deep Learning sowie zur Einteilung in drei Phasen (Kybernetik, Konnektionismus [Frank Rosenblatt und Perceptron], Deep Learning), vgl. das Einleitungskapitel in Goodfellow, Ian/Bengio, Yoshua/Courville, Aaron (Hg.) (2018), Deep Learning. Das umfassende Handbuch. Grundlagen, aktuelle Verfahren und Algorithmen, neue Forschungsansätze, mitp-Verlag: Frechen, v.a. S. 12ff. Vgl. zur historischen Einordnung auch Bajohr, Hannes (2022), »Gestalt und künstliche Intelligenz (1943 bis zur Gegenwart)«, in: Eva Geulen/Claude Haas (Hg.), Formen des Ganzen, Göttingen: Wallstein, S. 126-137.

Groß ist die Aufregung darüber, was ChatBots, was Programme wie DALL-E, Gemini und andere Verfahren zur *text-to-image production* mit scheinbar beträchtlicher Mühelosigkeit auf der Seite der Benutzer in die Welt zu setzen vermögen. In den Fokus der Aufmerksamkeit geraten damit DeepDream, DiffusionBee, Stable Diffusion, Mid-Journey, Imagen, CLIP-GEN und wie die Verfahren im Einzelnen alle heißen mögen.² Diese sind, blendet man einmal ihre gerne kulturkritisierte Funktion als Quatsch- und Fakegeneratoren aus, vor allem eines: Sie sind auf Dauer gestellte Veranstaltungen einer automatisierten Auffälligkeitserzeugung – nur mit dem Unterschied, dass sie nicht mehr länger dem Verwertungsmonopol einer auf Nützlichkeit angelegten Gedächtniskunst dienen, sondern dass sie Teil geworden sind einer sich spielerisch gebenden und einer exzessiven Flüchtigkeit prozessierenden Bildkultur des Teilens und Kommentierens, des Zirkulierens und einer gelegentlichen Verfestigung (etwa in Form von Memes, mit denen die algorithmisch erzeugten Bilder sistiert und in wieder in eine soziale Praxis zurückgebunden werden). Zu dieser Verfestigung vermerkt der DALL-E Eintrag bei Wikipedia: »Die surrealen Bilder des Modells wurden prompt von Internetnutzern zur Erstellung von Memes benutzt.«³ Was so freigesetzt wird, sind Umgangsweisen mit Bildern, sind Verfestigungen im Bilderhaushalt einer Kultur, die durch individuelle Habitualisierungen und soziale Praktiken gesteuert werden. Aber noch eine andere Bemerkung findet sich dort, die im weiteren Verlauf zunehmend Fahrt aufnehmen wird: der Hinweis auf schwerwiegende Fehler, die beim derzeitigen Stand der Technik keinem Menschen so unterlaufen würden. Die Bildproduktionen unterstehen einem Moment des Kompetitiven und einer Zeitlogik der Approximation. Noch hat der Mensch an bestimmten Stellen scheinbar die Nase vorn – so jedenfalls wird mit der Temporalisierung von Komplexität ein beruhigendes Moment des Aufschubs in Aussicht gestellt.

Die Beispiele, die im Zuge der Mnemotechnik etabliert wurden und die eine historisch wie konzeptuell kritisierfähige Kulturgeschichte der Beziehungsstiftung freisetzte, waren von keiner Ästhetik gedeckt, sondern ausschließlich auf Belange des Merkens bedacht und auf eine funktionale Vollkommenheit ihrer Bilder fixiert. Das war der Rahmen für das, was im bisherigen Verlauf dieses Buches immer wieder in Form von Auflistungen gefasst wurde: All die kleinen Szenarien und Episoden, mit denen aus unvollkommenen Bildern durch Dramatisierung vollkommene, weil uneingeschränkt merктаugliche Bilder wurden, all die Transformationsübungen, um mit deren Hilfe aus unvollkommenen Bildern vollkommene Bilder zu generieren, all die Techniken der Verlebendigung, um taugliche *imagines agentes* zu gewinnen, all die Etüden zur Größenanpassung, die aus Gründen der eigenen Sachdienlichkeit von der vorgestellten Verkleinerung großer und von der vorgestellten Vergrößerung kleiner Tiere handelte, und all die Beispiele, die bisher kaleidoskopartig in einem Panoptikum des Merkwürdigen versammelt wurden – sie alle wissen sich in der Sache und vom Resultat her mit den künstlich generierten Beispielen solidarisch. Ihre Beispiele haben diesen Text begleitet und ihm selbst eine kleine Wiederholungsstruktur verliehen: mit Trompetern, die von Wellen aus Berliner Kaffeehäuser gespült wurden, mit Löwen, die sich purzelbaumschlagend auf Beutezug ihren Weg nach Barcelona bahnten, mit

2 Zu Letzerem vgl. Wang, Zihao/Liu, Wei/He, Qian u.a. (2022), »CLIP-GEN: Language-Free Training of a Text-to-Image Generator with CLIP«, arXiv:2203.00386v1 [cs.CV] 1 Mar 2022.

3 Online unter <https://de.wikipedia.org/wiki/DALL-E> (letzter Eintrag: 30.09.23).

Schwänen, die unvermittelt in Hinterhöfen auftauchten, aber auch mit jener typographisch so auffällig gestalteten Reisegruppe und ihrem ungewöhnlichen Ziel (**Reisende mitten in Medien(l)and**).

Der kleinste gemeinsame Nenner all dessen liegt in der Formung von Geschichten: Das gilt für die Handlungsanmutungen im Umgang mit geometrischen Formen als Agenten eines Trickfilms, wie sie in der berühmten Studie von Fritz Heider und Marianne Simmel sichtbar wurden, oder für die von Gerhart Hauptmann beschriebene Herausforderung, unzusammenhängende Elemente mit einem narrativen Band zu verknüpfen und damit unter der Hand eine gleichermaßen unscheinbare wie pointierte Urszene von Literatur zu liefern – mit dem oben geschilderten Resultat einer grotesk anmutenden Episode um einen schlüsseltransportierenden Kalkhasen, der anlässlich eines Diplomateninars aufgetragen wurde und die Gäste dort zum Spucken brachte. Astronauten, die in weltfernen Habitaten auf Pferden reiten, gebärfreudige Pizzen und die Heerscharen von Tieren, die mit sonderbaren und ungewöhnlichen Accessoires ausgestattet sind (*A Shiba Inu dog wearing a beret and black turtleneck*), all die künstlich generierten Szenarien sind Teil dieser Logik, sie alle praktizieren, was die Psychologie der Aufmerksamkeitserzeugung in ihren Beispiellisten versammelte, was sie taxonomisch zu erfassen und was sie auf basale Konstruktionsprinzipien zurückzuführen suchte.⁴ Es ist auffallend, dass sich eine bestimmte Art der Verfügung bei den unterschiedlichen Verfahren zur Generierung durchhält und so inzwischen selbst eigene Topiken entstanden sind – wie im Fall der japanischen Hunderasse Shiba Inu ausgestattet mit Beret und Rollkragenpullover.⁵ Das Beispiel vom reitenden Astronauten fügt sich in ein Spektrum, bei dem Dinge und Kontexte, Anlässe und Accessoires auf markante Weise kombiniert oder durcheinander gewirbelt werden.⁶ Kängurus unter Einbezug der dazu sachdienlichen Werkzeuge in historisch akkurater Weise klöppeln und das in der Manier eines Vincent van Gogh dazustellen zu lassen, ist schon seit längerer Zeit keine erklecklich große Kunst mehr.

Die Semantik des Bizarren und Ungewöhnlichen, des in seiner Absurdität Steigerbaren und des gezielt Kontextenthoben erzeugte, wie man mit Blick auf die entsprechenden Diskurse sehen konnte, eigene Kleinstnarrative. In Form minimaler Sätze wird jenes Unfassbare fassbar, um das sich auch die Psychologie bemühte, wenn sie eigene Sätze aufschrieb und diese mit Alternativen versah, wenn sie Kategorien für den Kontextentzug angab und die Augmentierbarkeit des Auffälligen sichtbar hielt. Man muss aus der Liste der *dramatis personae* nur einfache Verwechslungen vornehmen und schon ist das Ungewöhnliche wirklich geworden (*The Horse drove the*

4 Das Beispiel dient als Belegbild auf Wikipedia. Online unter https://de.wikipedia.org/wiki/Datei:A_Shiba_Inu_dog_wearing_a_beret_and_black_turtleneck_DALLE2.jpg (letzter Zugriff: 08.08.23). Vgl. dazu Marcus, Gary (2022), »Horse rides astronaut«, in: Marcus on AI vom 28.05.2022, online unter <https://garymarcus.substack.com/p/horse-rides-astronaut> (letzter Zugriff: 21.10.22).

5 Vgl. dazu Ramesh, Aditya/Dhariwal, Prafulla/Nichol, Alex u.a. (2022), »Hierarchical Text-Conditional Image Generation with CLIP Latents«, arXiv:2204.06125v1 [cs.CV] 13 Apr 2022, sowie Bajohr, Hannes (2023), »Dumb Meaning: Machine Learning and Artificial Semantics«, in: Image. The Interdisciplinary Journal of Image Sciences 37/1, S. 58-70, und in deutscher Übersetzung Bajohr, Hannes (2022), »Dumme Bedeutung: Künstliche Intelligenz und artifizielle Semantik«, in: Merkur 76, S. 69-79.

6 Vgl. dazu etwa den Corgy, der eine rote Nikolausmütze trägt (corgi wearing a red santa hat). Nichol, Alex/Jun, Heewoo/Dhariwal, Prafulla u.a. (2022), »Point E: A System for Generating 3D Point Clouds from Complex Prompts«, arXiv:2212.08751v1 [cs.CV] 16 Dec 2022.

Car, The spokes of the wheel were kittens).⁷ Zum Behufe des Unsinnns scheinen Kontexte vermischt und Kohärenzen außer Kraft gesetzt, Erwartbares und Wahrscheinliches sind nicht mehr in Geltung – und das, obwohl allerorten Dispositive der Stabilisierung ihre Arbeit verrichten. So sind Lückentexte, die im Rahmen der Psychologie zur Anwendung gelangen, textuelle Maschinen zur Wahrscheinlichkeitsabsicherung, sind Schleppnetze des Plausiblen und des Gewohnten. Ihr Beifang und damit das, worauf die Mnemotechnik fußt, werden als Fehler markiert und zählen nicht.

Unter den Bedingungen künstlicher Intelligenzen wird das, was in der Kunst einmal ein in sich und für den Kunstbetrieb stabiles und stabilisierbares Exemplum für eine inzwischen hinreichend sanktionierte Kontextvermischung darstellte, zum auf Dauer gestellten Standardnormalfall: Marcel Duchamps Ready-mades – angeführt vom legendären Urinal aus dem Jahr 1917 – hätten somit ihr Alleinstellungsmerkmal eingebüßt: »Chirurgen operieren auf einer Kloschüssel. Jesus steigt in einen Ferrari. Batman wartet auf dem Fußballplatz auf seine Eltern.«⁸ Man könnte derartige Befunde mühelos ergänzen um solche, die, wie beim französischen Dichter Lautréamont und folgenreich für einen ihn bemühenden Surrealismus, absurde Metapher formen oder im Zuge einer *écriture automatique* formen lassen: Dabei kommt es bei der Beschreibung eines schönen Jünglings (und nicht wie bei Kant einer überblendungstechnisch gedachten Herstellung einer Normalidee) zu jener berühmten Formulierung, in der sich seine Schönheit mit der *Begegnung einer Nähmaschine und eines Regenschirms auf einem Sezientisch* beschrieben findet.⁹

Aber es sind nicht nur Versammlungen der nicht zusammengehörenden Dinge, also das, was der Mnemotechnikkritiker Reventlow anlässlich der mnemonischen Umtriebe um Machiavellis Todesjahr als hanebüchene »Conglomeration« beschrieben, an Bindfäden und Totenköpfen, Gabeln und Fledermäusen exemplifiziert und zur Karikierung die auffällige Rede vom Bohnenlied bemüht hat, es finden auch Formen von Brüchen statt, die zeitliche Ereignisse verwirbeln, Geschichtsklitterung betreiben und das Kontrafaktische in Bild und Szene setzen – kurz: die den Raum der bloßen und systemimmanent selbstzweckhaften Auffälligkeitserzeugung überschreiten in eine Episteme, die sich auf technischem Wege von der Wahrheit als epistemologisch verbindlichen Referenzsystem gelöst hat. Damit wird sichtbar, was hinter dem Potential des Auffälligen Fragen nach mutwilliger Täuschung und Lüge, nach Desinformation und Verschwörung, nach Manipulation und Fake News, nach contrafactuals und alternativen Wahrheiten aufwirft und was in der öffentlichen Auseinandersetzung als eine die Wahrhaftigkeit bedrohende alternative Ordnung der Dinge wahrgenommen wird. So groß ist die Sorge, dass entsprechende Gegenbewegungen und -maßnahmen nicht auf sich warten lassen – bis hin zu einer Initiative, die für ein Moratorium plädiert und eine zeitweise Außerkraftsetzung künstlicher Intelligenzen fordert.¹⁰

7 Vgl. dazu noch einmal die Skalierungen, etwa in den Abbildungen bei P. H. Marshall/K. Nau/C. K. Chandler, A functional analysis, und S. D. Cox/K. A. Wollen, Bizarreness and recall.

8 Breithut, Jörg (2022), »Künstliche Intelligenz. Die Text-zu-Quatsch-Generatoren«, in: Spiegel Netzwelt vom 18.06.2022, online unter <https://www.spiegel.de/netzwelt/gadgets/dall-e-2-und-google-imagen-die-text-zu-quatsch-generatoren-a-9caeeb86-7980-4064-b6be-a10020b56786> (letzter Zugriff: 05.01.24).

9 Lautréamont, Isidore L. (1992), *Die Gesänge des Maldoror*, 3. Aufl., Reinbek: Rowohlt.

10 Zur Akzeptanz und zum Erwerb einer AI Literacy vgl. Theophilou, Emily/Koyutürk, Cansu/Yavari, Mona u.a. (2023), »Learning to Prompt in the Classroom to Understand AI Limits: A Pilot Study«, in:

Nicht immer sind die Lektüreschlüssel und damit die Selbstschutzmöglichkeiten dabei so eindeutig wie bei Bildern, die Hitler zeigen, wie er sich mit den Minions einen Drink genehmigt, oder Gladiatorenkämpfe, die im antiken Rom mit moderner (und in ihrer Unzeitgemäßheit einfach erkennbaren) Technik aufgezeichnet werden – sie fügen sich im Modus historischer Bildbrüche in die Muster, in die Tradition und in die Kritik der Erzeugung von Auffälligkeiten.¹¹ Damit sind sie vergleichbar und aufeinander beziehbar wie die Beispiele aus Merkkunst und Psychopathologie, aus Literatur und den diversen Vorstellungs- und Wortverfügungsetüden, denen dieses Buch an so unterschiedlichen Schauplätzen und zu so unterschiedlichen Zeiten nachstellte. Täuschung und Wahnsinn gehen dabei Hand in Hand.¹² Die Beispiele »mit Zuschauern im alten Rom, die Gladiatorenkämpfe mit Smartphones fotografieren, Waschbären im Weltraum oder tanzenden Avocados mit Sonnenbrillen auf dem Kopf«, die oft mit dem Gestus einer skandalisierenden Bagatellisierung als Unsinn und Quatsch angeführt werden, drängen regelrecht danach, mit den Auffälligkeiten der Gedächtniskünste (stellvertretend für den Geltungsbereich einer gesamten Kulturgeschichte) gegeneinander gehalten und miteinander verglichen zu werden.¹³

Der Verweis auf die antiken Gladiatorenkämpfe und ihre Zeitgenossenschaft zu den technischen Medien der Gegenwart doppelt ein Muster, das auch in den Beispielen bei Eugen Bleuler, dem Psychiater und Nervenarzt zu finden war (»Epaminondas belagert die Akropolis mit Kanonenbooten und schlägt eine Panzerfregatte der Engländer«)¹⁴ und das, wie oben gezeigt, ob seiner Ähnlichkeit mit den Paradigmen der Mnemotechnik als Steilvorlage für entsprechende Pathologisierungen dienen konnte. Einmal mehr ist es neben dem Vorwurf ökonomischer Verschwendung der Vorhalt des Wahnsinns, der auf der Ebene der operativen Umsetzung gegen die Kulturtechnik des Merkens ins Feld geführt wird. In derartigen Konstellationen haben beiläufig wirkende Formulierungen wie die vom Schwachsinn mit barocker Phantasie ihren epistemologischen Ermöglichungsgrund: Was den Anschein einer beliebigen und vielleicht sogar idiosynkratischen Kasuistik hat, ist in eine Systematik eingebunden, die Kultur nicht nur nach Maßgabe zeitlicher Größen und mithin von Epochen beschreibbar macht, sondern die sich anschickt, heterogene Wissensbereiche auf einen für sie sachgemäßen Beschreibungsrahmen zu verpflichten.

So geläufig und so gehäuft sind die Beispiele, dass ihre Nennung selbst im Modus einer eigenen Performanz erfolgt: Sie bedient sich der Geste einer akkumulierenden

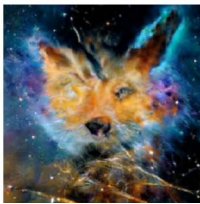
AIxIA 2023 – Advances in Artificial Intelligence: 22nd International Conference of the Italian Association for Artificial Intelligence, Proceedings, Cham: Springer, S. 481-496.

- 11 Aus der Fülle einschlägiger Bezüge vgl. stellvertretend Gess, Nicola (2021), *Halbwahrheiten. Zur Manipulation von Wirklichkeit*, Berlin: Matthes & Seitz.
- 12 Die Beschreibung hält sich an entsprechende Vorlagen und handelt von irren Bildern. Vgl. Marjanović, Petar (2922), »Blocher isst einen Cheeseburger« – Website kreiert irre AI-Bilder«, in: watson vom 16.06.2022, online unter <https://www.watson.ch/international/digital/678700729-dall-e-mini-ai-generiert-irre-fake-bilder> (letzter Zugriff: 23.06.22).
- 13 Vgl. zu einer kritischen Auseinandersetzung Tuschling, Anna/Sudmann, Andreas/Dotzler, Bernhard J. (Hg.) (2023), *ChatGPT und andere »Quatschmaschinen« Gespräche mit Künstlicher Intelligenz*, Bielefeld: transcript.
- 14 E. Bleuler, *Lehrbuch*, S. 44. Dieses Beispiel stellt für Bildherstellungssysteme keine sonderliche Herausforderung dar – ebenso wenig wie der Merksatz von dem unter den Linden von einer Welle ausgespülten Trompeter.

Listung. Fast nie hat es mit der Nennung von nur einem Beispiel dessen, was technisch möglich ist, sein Bewenden, in der Regel scheinen die Exempla in Abfolgen und Auflistungen angeordnet, so als ob ihr einzelnes Auftreten einem Selbstwiderspruch gleichkäme. Dieser Hang zum Plural und zur Pluralisierung wird mustergültig deutlich in einer Bildserie, erzeugt von GALIP: *Generative Adversarial CLIPs for Text-to-Image Synthesis*. Die Serie besteht aus vier nebeneinander angeordneten Beispielen, die den Kopf eines Hundes als Nebelexplosion, einen anzugtragenden Löwenlehrer vor einer Schultafel, eine niedliche Katze, die in einem Haus aus Sushi lebt, und einen unter blauen, aber bewölkten Himmel reitender Roboter zeigt.¹⁵

Derartige Konstellationen befördern in ihrem unvermittelten und in ihrem brachialen Nebeneinander zugleich auch eine bestimmte Form der Rezeption. In ihrer exzessiven Massierung begründen sie ein Reich der Gleichgültigkeit, eine durch die schiere Fülle erzwungene Haltung dessen, was man entgegen allen kulturkritischen Larmoyanzen vielleicht als eine Art positiver Indifferenz beschreiben könnte: Man ist weder willens noch in der Lage, den jeweiligen Produkten ob ihrer Massierung im Einzelnen zu folgen. Das Verhältnis zu den beliebigen Elaboraten kann daher notwendigerweise nur ein ausgesprochen flüchtiges und im Modus dieser Zerstreuung auch entsprechend kritisiertes sein.¹⁶ Die Haltung der Indifferenz als Modus von Rezeption ist dabei keine charakterologische Zuschreibung, ist also nicht auf ein Konto menschlicher Unarten und Unzulänglichkeiten zu buchen, sondern eine solche, die dem Stand der Datenverarbeitung, die also einem bestimmten Grad der Technisierung von Lebenswelt geschuldet ist. Wie in einem unausgesprochenen Konsens scheint sich die Kultur darauf verständigt zu haben, die Bildumtriebe als kollektive Praxis zu befördern und dabei ihre einzelnen Verfestigungen gar nicht mehr zu ignorieren.¹⁷

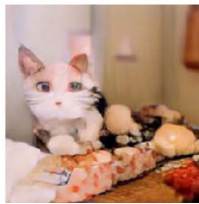
A corgi's head depicted as an explosion of a nebula



A lion teacher wearing a suit is in front of a blackboard.



A cute cat lives in a house made out of sushi.



A robot is riding a horse under the blue and cloudy sky.

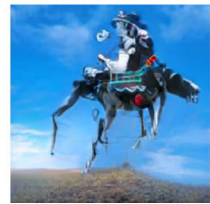


Abbildung 42: Massierung der Beispiele (nach Tao et al. 2023)

Doch es gibt noch eine andere Seite, die zu dieser unablässigen Proliferation scheinbar quer steht: Die Exzesse auf Seiten der unablässigen Verfertigung von Beispielen gehen zugleich mit einem Moment der Restriktion der sie ermöglichenden Algorithmen ein-

15 Dazu Tao, Ming/Bao, Bing-Kun/Tang, Hao u.a. (2023), »GALIP: Generative Adversarial CLIPs for Text-to-Image Synthesis«, arXiv:2301.12959v1 [cs.CV] 30 Jan 2023.

16 Zu dieser Kopplung vgl. P. Löffler, Verteilte Aufmerksamkeit.

17 Damit folgt sie bestimmten Graden der Sättigung von Aufmerksamkeit. Zu einer historischen Fallstudie vgl. Scholz, Joachim (2015), »Aufmerksamkeit im Schulmännerdiskurs der Sattelzeit«, in: Sabine Reh/Kathrin Berdelmann/Jörg Dinkelaker (Hg.), Aufmerksamkeit. Geschichte – Theorie – Empirie, Wiesbaden: Springer, S. 35–54.

her. Diese begründen in ihrer Nicht-Offenlegung ein eigenes Kuratel und schüren so einmal mehr das Moment einer grundlegenden Verdächtigung gegenüber den Verfahren ihrer Generierung – und ihrer ökonomisch, politisch, strategisch oder sonstwie getriebenen Sachwalter und Sachwaltungen: Die Proliferation des Skurrilen ist damit alles andere als harm- und belanglos. »Darth Vader mäht Rasen, eine Salatsoße sagt vor Gericht aus: Skurrile Motive aus den Text-zu-Bild-Generatoren DALL-E 2 und Googles hauseigener KI »Imagen« fluten das Netz. Der Zugang bleibt aus guten Gründen vorerst beschränkt.«¹⁸ Anlässlich der Preisverleihung an eine Gruppe von Wissenschaftlern, unter anderem an den britischen Informatiker und Kognitionspsychologen Geoffrey Hinton, ist gar die Rede von der *Deep Learning Mafia*. »Ehre für die »Deep Learning Mafia«. Für ihre Leistungen bei der Erforschung der künstlichen Intelligenz (KI) erhalten Yoshua Bengio, Geoffrey Hinton und Yann LeCun den Turing Award.«¹⁹

Die Art und Weise der Herstellung und die damit einhergehenden Praxeologien erlauben es, die Umtriebe der künstlichen Intelligenzen in eine größere Geschichte einzubinden, oder anders gesagt: Sie machen eine solche Rekonstruktion nachgerade notwendig – notwendiger vielleicht, als sich in unterschiedlichen Haltungen zu ihnen zu ergeben. So wird es möglich, ihre Produkte eben einmal gerade nicht vorrangig vor dem Hintergrund von arglistiger Täuschung und mutwilliger Manipulation, damit als Bedrohung gesellschaftlicher Zusammenhänge, als Verunsicherung politischer Verhältnisse und als Außerkraftsetzung einer Episteme der Wahrheit zu situieren und mit Faktencheck und Hintergrundrecherche, mit Debunking und Inokulation entsprechende Gegenstrategien zu entwickeln – und dabei doch weitgehend dem Schema der Kritik verhaftet zu bleiben.²⁰ Sichtbar wird in der Gegenüberstellung derartiger Beispiele vielmehr eine Kulturgeschichte des Auffälligen und eine Kulturgeschichte der Evozierung, der Hervorlockung von Auffälligem, die ihren eigenen Referenzrahmen aufweist und die auf diese Weise zumindest die von Luhmann ins Spiel gebrachten touristischen Einblicknahmen in die Verfahrensweisen und Sachstände vergangener Lebenswelten erlaubt: Die so jeweils erzeugten Produkte liegen maximal weit auseinander, nicht nur, dass sie im Zuge der Historiographie zeitlich weit voneinander entfernt sind und so ihr Potential einer diachronen Kritik entfalten können, darüber hinaus stellen sie doch den Übergang von einem menschengenerierten zu einem maschinell erzeugten (Un)Sinn dar. Ähnlich wie im Fall der Gedächtniskunst eröffnet das die Möglichkeit einer entsprechend kommentierenden Qualifizierung – etwa als Quatsch oder Pathologie, als Halluzination oder als Träumen, als Irrsinn oder als kreative Höchstleistung.

18 Breithut, Jörg (2022), »Künstliche Intelligenz. Die Text-zu-Quatsch-Generatoren«, in: Spiegel Netzwerk vom 18.06.2022, online unter <https://www.spiegel.de/netzwelt/gadgets/dall-e-2-and-google-imagen-die-text-zu-quatsch-generatoren-a-9caeb86-7980-4064-b6be-a10020b56786> (letzter Zugriff: 05.01.24).

19 Betschon, Stefan (2019), »Ehre für die »Deep Learning Mafia«, in: NZZ vom 04.04.2019, online unter: <https://www.nzz.ch/digital/ehre-fuer-die-deep-learning-mafia-ld.1472761> (letzter Zugriff: 27.05.2024). Einen anderen Weg schlägt die Forschungsgruppe CompVis der Ludwig-Maximilians-Universität München unter der Leitung von Björn Ommer ein. Deren System Stable Diffusion wird unter Bezeichnungen wie der von der Demokratisierung von KI gehandelt. Vgl. dazu Rombach, Robin/Blattmann, Andreas/Lorenz, Dominik u.a. (2022), »High-Resolution Image Synthesis with Latent Diffusion Models«, arXiv:2112.10752v2 [cs.CV] 13 Apr 2022.

20 Vgl. dazu etwa Kreps, Sarah/McCain, R. Miles/Brundage, Miles (2022), »All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation«, in: Journal of Experimental Political Science 9/1, S. 104-117.

Nicht, dass die kulturschaffende Kraft von Maschinen und Mechanik, von Daten und Algorithmen je in Frage gestanden hätte (sie waren immer schon in Geltung, aber über weite Strecken eben einfach nur übel beleumundet), jetzt setzen die Verfahren zur Erzeugung von Anomalien und Auffälligkeiten auf sorgfältig vorgelagerte Datensammlungen, auf Big Data und auf die Strategien der Stabilisierung durch große Zahlen.²¹ Kalküle des Wahrscheinlichen sind der Erzeugung von Unwahrscheinlichem vorgelagert. Von den unregelmäßigen Gedankengängen der Phantasie, denen auf die Spur zu kommen die Gedächtniskünste zeitübergreifend eine operative Handhabe boten, von den Imaginationsetüden, an denen sich der Aufklärungsphilosoph Christian Wolff beim Selbstversuch einer außergewöhnlichen Vorstellung noch die Zähne ausbiss, von Kants gedanklicher Vorwegnahme einer nachmaligen Überblendungsphotographie zum Behufe der Herstellung von Normalbildern und nicht zuletzt von den geregelten Gedankengängen einer in ihrer Vorstellungsgenerierung minutiös kalkulierten Literatur im Aufschreibesystem 1800, von all den damit verbundenen Fragen nach der Umsetzung, nach den quantitativen Aspekten einer solchen Transposition und mit Blick auf eine gerne unterstellte Konkurrenz von Medien ist hier keine Spur.

Das Moment einer historischen und kulturhistorischen Datenverarbeitung, jenes Moment, wie aus Wörtern auf dem Wege über (im Fall der Lektüre bei Mauchart gegen Ende des 18. Jahrhundert abzählbaren und im Fall von Benses Filmprojekt in der Mitte des 20. Jahrhundert kalkulierbaren) Vorstellungen ein Einbildungswissen erzeugt und wie es verhandelbar wird, bleibt weitgehend ausgeblendet. Zu diesem Geschäft der Verhandlung gehören auch die einschlägigen Versuche der Typenbildung, wie sie die Verwissenschaftlichung des Imaginierens gerade im Umfeld der Psychologie begleiten sollte (bis hin zur Feststellung eines Typus, der regelrecht als vorstellungsblind gilt). Die entsprechenden Befragungen halten sich durch und so werden mittels kurzen Listen von Adjektiven Imaginationstypen wie etwa derjenige des *high fantasizers* operationalisiert.²² Dabei wird die Frage nach dem Realitätsgehalt mitverhandelt – einmal mehr in Redeweisen, die sich typologisch der Realität anzunähern suchen und dabei zu Formulierungen und typographischen Markierungen greifen, die an Tautologie nur schwer zu überbieten sind: Gefragt wird nicht nur danach, ob die Einbildungsbilder von der Phantasie zugewandten Bilder tatsächlich realer sind (*Are the Mental Images of Fantasy-Prone Persons Really More »Real«*), und weil das Ganze mit Blick auf die Ambiguität von Kippfiguren verhandelt wird, darf auch eine ausschlaggebende Folgefrage nicht fehlen, nämlich die nach der Ambiguität von Einbildungswissen (*Can Mental Images Be Ambiguous?*).²³

21 Ihre Verfestigung finden derartigen Anliegen in einer *cultural robotics*. Vgl. dazu stellvertretend J.T.K.V. Koh/B. J. Dunstan/D. Silvera-Tawil u.a. (Hg.), *Cultural Robotics. Zur Vorgeschichte und namentlich zu den künstlerischen Avantgarden* vgl. Schramm, Helmar/Schwarte, Ludger/Lazardig, Jan (Hg.) (2008), *Spuren der Avantgarde: Theatrum machinarum. Frühe Neuzeit und Moderne im Kulturvergleich*, Berlin/New York: de Gruyter.

22 Vgl. dazu Marks, David F. (1973), »Visual Imagery Differences in the Recall of Pictures«, in: *British Journal of Psychology* 64/1, S. 17-24.

23 Vgl. dazu J. R. Council/T. A. Jundt/M. D. Good, *Mental Images*, und Chambers, Deborah/Reiserberg, Daniel (1985), »Can Mental Images Be Ambiguous?«, in: *Journal of Experimental Psychology: Human Perception and Performance* 11/3, S. 317-328.

Große Zahlen, die bei Kant und im gedanklichen Vorgriff auf die noch ausstehenden Kompositphotographie Galtons mit der für ihn gigantischen Einheit *Tausend* angegeben werden konnten, sind die Voraussetzung für zwei Operationen, die nur auf den ersten Blick im Widerspruch zueinander stehen: eine immer bessere Angleichung der Objekt- und Mustererkennung auf der einen Seite und das kalkulierte Unterlaufen der Bedeutungs- und Kontextinformationen auf der anderen. Algorithmen wie DALL-E sind sich selbst stabilisierende Generatoren zur unablässigen Erzeugung von Bildbrüchen und Kontextverlagerungen. Trainiert wurde DALL-E mit Bildern, die mit ihren horrenden Datensätzen zahlenmäßig die Beispiele bei Kant weit überflügelten. Es beruht auf einem anderen, ebenfalls von OpenAI entwickelten Verfahren, dem Generative Pre-trained Transformer 3 (GPT-3). In welchen Größenverhältnissen die Zahlen spielen, wird eindrücklich deutlich, wenn man die entsprechenden Trainingsumgebungen und damit die Kopplung unterschiedlicher Forschungsbestrebungen zu Rate zieht:

DALL-E wurde in Verbindung mit CLIP (Contrastive Language-Image Pre-training) entwickelt und der Öffentlichkeit vorgestellt. CLIP ist ein separates Modell, das auf Zero-Shot-Lernen basiert und mit 400 Millionen Bildpaaren mit Textbeschriftungen aus dem Internet trainiert wurde.²⁴

Ob Diffusionsmodelle oder GAN, ob *convolutional neural network* oder *generative adversarial network* und ungeachtet aller Binnendifferenzierungen (und Stabilisierungsanmutungen)²⁵ der einzelnen Verfahren – abgesichert durch das Prinzip der großen Zahl steht das generative Potential außer Frage.²⁶ Dieses wird unisono und unter Berücksichtigung der unterschiedlichen Verfahren bescheinigt.

In particular, Deep Dream (Mordvintsev et al. 2015) and Neural Style Transfer (Gatys et al. 2015) are algorithms, based on deep-learning convolutional neural networks (CNNs) (Krizhevsky et al. 2012), that blend visual qualities from multiple source images to create a new output image.²⁷

Und auch mit Blick auf die GAN bleibt das bilderzeugende Potential unstrittig: »Generative adversarial networks conditioned on textual image descriptions are capable of generating realistic-looking images.«²⁸ Die Wirkung solcher Elaborate wird allerorten

24 Wikipedia, Eintrag DALL-E (letzter Zugriff: 17.07.23).

25 Du, Chengbin/Li, Yanxi/Qiu, Zhongwei u.a. (2023), »Stable Diffusion is Unstable«, arXiv:2306.02583v2 [cs.CV] 6 Jun 2023.

26 Vgl. Borji, Ali (2022), »Generated Faces in the Wild: Quantitative Comparison of Stable Diffusion, Mid-journey and DALL-E 2«, arXiv:2210.00586v1 [cs.CV] 2 Oct 2022.

27 McCaig, Graeme/DiPaola, Steve/Gabora, Liane (2016), »Deep Convolutional Networks as Models of Generalization and Blending Within Visual Creativity«, in: François Pachet/Amílcar Cardoso/Vincent Corruble u.a. (Hg.), Proceedings of ICCV 2016, 7th International Conference on Computational Creativity, Paris: ACC (Association for Computational Creativity), S. 156-163, hier: S. 156.

28 Hinz, Tobias/Heinrich, Stefan/Wermter, Stefan (2022), »Semantic Object Accuracy for Generative Text-to-Image Synthesis«, in: IEEE Transactions on Pattern Analysis and Machine Intelligence 44/3, S. 1552-1565, hier: S. 1552.

mit anekdotischer Evidenz versehen und eröffnet ein semantisches Spektrum, das zwischen beeindruckend und unheimlich angesiedelt ist.²⁹ Nicht nur die Erzeugung von Bildern ist auf diese Weise möglich, auch die Gestaltung von komplexeren Szenarien ist umsetzbar wie in *SceneGenie: Scene Graph Guided Diffusion Models for Image Synthesis*.³⁰ Und selbst eigene Positionierungen und ironische Brechungen sind auf dem Feld künstlicher Kreativität längst an der Tagesordnung. Aufsehen erregte etwa ein algorithmisch erzeugtes Bild, das von einem Kollektiv mit dem Namen *Obvious* zum Verkauf angeboten wurde. Das *Portrait of Edmond de Belamy* konnte im Zuge dessen zum Anlass für ein Spiel mit Autorisierungsgesten, Wertschöpfungsketten und Vermarktungsstrategien werden.³¹

So sind die Porträts aus der Reihe *Obvious Belamy*, erstellt mit Hilfe von Generative Adversarial Networks (GANs) auf Grundlage von Algorithmen hergestellt, die Gesichter erkennen und modifizieren können. Zwar sind die Porträts mit einem Hinweis auf das Sujet gestaltet, signiert wird jedoch in ironischer Simulation künstlerischer Autorschaft mit dem entsprechenden Algorithmus.³²

Von derartigen Eingriffen in den Markt berichtet auch Matthew Kirschenbaum anlässlich einer neuen und unübersichtlichen Schwemme von Texten aus dem Bereich der Science-Fiction, die als Folge der Marktförmigkeit von Schreibgeneratoren den Buchmarkt fluteten. Eingebunden in das von Roland Barthes konstatierte Ende von Autorschaft wussten sich die Herausgeber entsprechender Organe nur noch dadurch zu helfen, dass sie die Schwemme neuer Texte weitgehend ignorierten und buchtechnisch ins Leere laufen ließen.

It was being flooded with stories self-evidently fabricated by artificial intelligence systems, especially ChatGPT launched ten weeks earlier by OpenAI. Publisher and editor Neil Clarke had been spending his days weeding the torrent of spammed submissions. But the effort was not sustainable.³³

29 Vgl. dazu Ramesh, Poornima (2022), GANs schön kompliziert: Applications of Generative Adversarial Networks, Diss. Tübingen, und zu DALL-E Meyer, Roland (2022), »Im Bildraum von Big Data. Unwahrscheinliche und unvorhergesehene Suchkommandos: Über Dall-E2«, in: Cargo 16/55, S. 50-53, sowie Meyer, Roland (2023), »The New Value of the Archive. AI Image Generation and the Visual Economy of ›Style‹«, in: IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft 37/19, S. 100-111.

30 Farshad, Azade/Yeganeh, Yousef/Chi, Yu u.a. (2023), »SceneGenie: Scene Graph Guided Diffusion Models for Image Synthesis«, arXiv:2304.14573v1 [cs.CV] 28 Apr 2023.

31 Vgl. dazu H. Bajohr, Schreibenlassen, S. 131ff. Zur Verschränkung von AI und Kunst vgl. Zylinska, Joanna (2020), *AI Art. Machine Visions and Warped Dreams*, London: Open Humanities Press.

32 Krewani, Angela (2021), »Technische Welten, Kunst und Künstliche Intelligenz«, in: Bernhard J. Dotzler/Berkan Karpat (Hg.), *Götzendämmerung – Kunst und Künstliche Intelligenz*, Bielefeld: transcript, S. 69-83, hier: S. 77. Zur Nähe von Diffusionsmodellen zur Kunst und zur Analyse von Kunstwerken vgl. etwa Wu, Yankun/Nakashima, Yuta/Garcia, Noa (2023), »Not Only Generative Art: Stable Diffusion for Content-Style Disentanglement in Art Analysis«, in: ICMR '23: International Conference on Multimedia Retrieval, New York: ACM, o. Pag. Auf diesem Feld ist auch Björn Ommer einschlägig.

33 Kirschenbaum, Matthew (2023), »Fictional Devices«, in: *Outland* vom 13.06.2023, online unter <https://outland.art/ai-writing-science-fiction/> (letzter Zugriff: 18.10.23).

Zwei Aspekte sind an dieser Episode aus dem Zentrum eines mit der Künstlichkeit potentieller Elaborate befassten Literaturbetriebes bemerkenswert: Zum einen die Wahl des Genres, das eine besondere Nähe zu künstlich erzeugten Gebilden jedweder Art nahezulegen und das in der Science-Fiction ob seiner Affinitäten besonders gut aufgehoben scheint. Was aber auch und gerade mit dieser Folie des Zukunftszugewandten ins Auge sticht, ist ein Blick in vergangene Systeme und die Art und Weise, wie dort Zugangsbedingungen auf den Buchmarkt organisiert waren, ein Vorgang, über den Walter Benjamin einen Text geschrieben hat, bei dem er gerade mit Blick auf die deliranten Bücher von Geisteskranken danach fragt, wie es diesen unter historisch anderen Veröffentlichungsbedingungen überhaupt gelang, »die Paßkontrolle dieses hunderttorigen Theben, der Stadt der Bücher« zu umgehen.³⁴

DALL-E, stellvertretend für das ganze Spektrum einschlägiger Verfahren, ist ein Agent der Rhetorik und ein Generator für die unablässige Herstellung von Bildbrüchen, der einfach und zu spielerischen Zwecken an die Hand gibt, wozu sich die mit der Mnemotechnik Befassten seit der Antike den Kopf zerbrochen haben: Antworten auf die Frage zu finden, wie man Auffälligkeiten und Unwahrscheinlichkeiten am laufenden Band (und mit dem Anspruch einer nicht versiegenden Eloquenz) produziert, wie man sie im Haushalt kultureller Bildproduktion vor Abnutzungseffekten und Verschleißerscheinungen schützt und wie man das Ganze nach Möglichkeit noch in einer so atemberaubenden Geschwindigkeit prozessiert, dass die Bildgenerierung dem Tempo des Schreiben nicht hinterherzuhinken braucht, so dass man, wie es bei Johann Ludwig Klüber zu Beginn des 19. Jahrhunderts anlässlich seiner Kommentierung einer Mnemotechnik aus dem 16. Jahrhundert heißt, »mit Bildern in Gedanken eben so fertig« soll schreiben können »als durch Buchstaben mit der Feder«.³⁵ Das ist kein geringer Anspruch, allerdings ist er in Algorithmen abgewandert und lässt den Menschen außen vor.

Was so nach außen tritt, sind Externalisierungen einer Phantasie, die von einer anderen Seite das Primat der Normalität sichtbar machen: Der Psychiater Bleuler wies darauf hin, was im Reich der Vorstellungen die Norm und was die Regel ist. Seine Beispiele, bei denen Menschenohren durch Blumen und Pferdehufe ersetzt und Hunde mit ungewöhnlichen Farben versehen werden, bleiben aus seiner Sicht die Ausnahme von der Regel, sie bleiben im Rahmen der Vorstellungsbildung und damit auf dem Feld des Einbildungswissens ungeläufig. Im Rahmen technischer Bilderzeugung stellen sie und damit das, was aus dem Rahmen fällt, allerdings eine neue Normalität vor Augen.³⁶ Diese hat sich in bestimmten Bereichen verselbständigt, ist sich selbst genug und muss sich nicht mehr vor den Karren einer bestimmten Funktion wie etwa derjenigen der Gedächtniskunst spannen lassen.

Aber noch etwas anderes fällt auf: Die Art und Weise, wie die Masken für die Bedienung zur Verfügung gestellt werden, wirkt selbst wie ein Verstärker, weil wie eine ausgesprochen niederschwellige Anreizung zu Diskursen, wie sie Foucault für andere

34 Vgl. dazu Benjamin, Walter (1972), »Bücher von Geisteskranken«, in: *Gesammelte Schriften*, unter Mitwirkung von Theodor W. Adorno und Gershom Scholem hg. von Rolf Tiedemann und Hermann Schweppenhäuser, Band 4/1, Frankfurt am Main: Suhrkamp, S. 615-619.

35 J. L. Klüber, *Compendium*, S. 30 (Anmerkung).

36 Vgl. dazu noch einmal die Bildreihe mit den ersetzten Ohren bei K. W. Ng/X. Zhu/Y-Z. Song u.a., *DreamCreature*.

Bereiche beschrieben hat. Die Verfügung von Wörtern zum Behufe der Produktion von Bildern und Bildfolgen wird ein eigenes und durchaus normal gewordenes Betätigungsfeld – ob aus spielerischen Gründen oder im Zuge einer Perfektionierung des Prompting als selbstständigem und sich professionell ausdifferenzierendem Berufsfeld.³⁷ Allerorten scheinen Menschen die Probe aufs Exempel zu machen, indem sie – mehr oder weniger reflektiert – die Eingabemasken mit Texten unterschiedlicher Provenienz füllen. Das Prompting wird für die Belange künstlicher Kreativität unhintergebar – und ist sogar im Modus der Negation belastbar. Unter dem Titel *How to use negative prompts?* wird im Fall von Stable Diffusion über die Gestaltung einer Welt gehandelt, in der bestimmte Dinge nicht der Fall sein sollen.³⁸

Kurzerhand und mit gehörigem Pomp wird das Prompting zur Urszene einer künstlichen Kreativität erklärt (*Am Anfang war der Prompt*).³⁹ Mit dieser Anmutung wird es möglich, den Bogen zu den Anfangsbedingungen dieses Buches zurückzuschlagen. Mit der Fertigkeit des Eingebens scheint auf eine gewisse Weise die Frage des Lesens gelöst, das als ideal konzipierte Konstruktion dem Aufschreibesystem von 1800 (eine innere Welt nach den Worten, wie es Novalis programmatisch für die Romantik verhieß) ebenso zugrunde lag wie der Frage nach der Herstellung statistisch unwahrscheinlicher Vorstellungen wie im Fall von Wolffs groteskem Mann. Die Frage, wie man aus Wörtern, wie man aus Konzepten und Bildern Bilder (oder ganze Bildfolgen) generiert, scheint – jedenfalls für eine bestimmte Bedürfnislage – technisch beantwortet und bewerkstelligt. Die Verfahren, die dabei zur Anwendung gelangen, sind dabei inzwischen selbst in Überblicksdarstellungen fass- und einteilbar, ein Indiz für den Stand der Sättigung auch auf Seite der Verfahren.⁴⁰ Das Word »«engineering«, das sich mit dem Prompting Bahn bricht, wird zu einer auf den Kopf gestellten Form der literarischen Produktion.⁴¹ Es trägt den Belangen des Literarischen auf eine Weise Rechnung, die bis hin zur Berücksichtigung von bestimmten Tonlagen, Herstellungs-

-
- 37 Zum Versagen Zamfirescu-Pereira, J. D./Wong, R. Y./Hartmann, B. u.a. (2023), »Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts«, in: CHI '23 Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, New York: ACM, o. Pag. Zum Gelingen White, Jules/Fu, Quchen/Hays, Sam u.a. (2023), »A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT«, arXiv:2302.11382v1 [cs.SE] 21 Feb 2023. Dort wird als eines der zu berücksichtigenden Muster der Kontext angeführt und mittels eines Context Managers behandelt. Zu Anwendungen im Kontext pädagogischer Wissensvermittlung vgl. Vogt, Andrea/Babel, Franziska/Hock, Philipp u.a. (2021), »Immersive virtual reality or auditory text first? Effects of adequate sequencing and prompting on learning outcome«, in: British Journal of Educational Technology 52/5, S. 2058-2076.
- 38 Der Beitrag ist online unter <https://stable-diffusion-art.com/how-to-use-negative-prompts/> (letzter Zugriff: 06.02.24).
- 39 Alexeev, Vladimir (2023), »Am Anfang war der Prompt. Die Wege zur Kreativität der Maschine«, in: Lothar B. Blum (Hg.), *Angewandte Data Science*, Wiesbaden: Springer, S. 51-80. Vladimir Alexeev publiziert auch unter dem Pseudonym Merzmensch. Vgl. dazu Merzmensch (2023), *KI-Kunst*, Berlin: Klaus Wagenbach.
- 40 Vgl. dazu noch einmal J. Agnese/J. Herrera/H. Tao u.a., *Survey and taxonomy. Für die Einbindung der text to image generation in Belange der Kunst* vgl. Berns, Sebastian/Broad, Terence/Cuckelsberger, Christian u.a. (2021), »Automating Generative Deep Learning for Artistic Purposes: Challenges and Opportunities«, arXiv: 2107.01858v1 [cs.LG] 5 Jul 2021.
- 41 Zu diesem Moment der Verkehrung vgl. Otto, Angela (2010), *IL-LUST-RATIO-N. Die Inversion des Illustrationsprozesses*, Bachelorarbeit Fachhochschule Vorarlberg, Studiengang Intermedia.

verfahren und Epochenvorgaben reicht (*painting, photo, sketch, cartoon, impressionist, abstract, renaissance*).⁴² Selbst die Kategorie *baroque* darf in den Listen möglicher Voreinstellungen neben *surreal* oder *anime* nicht fehlen. So ist es auch kein Wunder und eben nur konsequent, dass die Selbstbeschreibungen und Selbstperformanzen immer wieder die Nähe zur Kunst suchen – und sich wie in einigen oben angeführten Titeln einigermaßen selbstbewusst als *The Art of Prompting* zu Wort melden.⁴³

Aber mit dieser ausgestellten Nähe von *prompt engineering* und Kunst ist es nicht genug: Die technischen Verfahren lassen andere Modalitäten und Sinnlichkeiten zu – jenseits einer reinen Textzentrierung werden alternative Promptgestaltungen erwogen, die eine bessere Kontrolle über das Zielbild versprechen.⁴⁴ Darauf ausgerichtete Vorgehensweisen lassen es zu, das, was als nicht normal gilt, im Geschehen zu identifizieren – etwa in Form eines Verfahrens, das unter dem Akronym VAD (*video anomaly detection*) Bewegtbildsequenzen auf Unstimmigkeiten hin untersucht und es dabei mit all dem zu tun bekommt, was hier von Interesse war: mit Kontext und Ambiguität, mit Seltenheit und Normalitätsanmutungen, mit dem Erwartbaren und dem Gewöhnlichen.

Anomaly is commonly defined as a rare or unexpected or unusual entity, that diverges significantly from normality, which is defined as expected and common. Despite being sparse and diverse, the abnormal events are also very contextual, and often ambiguous, thus they challenge the performance of the VAD models.⁴⁵

Die neuen Anwendungsfelder betreffen das Träumen und das Erzählen von Märchen, die Produktion von Gesichtern und nicht zuletzt das Storytelling.⁴⁶ So gelangt MidJourney, ein AI-Kunstprogramm mit der Option, Phantasie ansichtig werden zu lassen, zur Anwendung – nicht abstrakt, nicht mit beliebig ausgewählten Prompts,

42 Liu, Vivian/Chilton, Lydia B. (2022), »Design Guidelines for Prompt Engineering Text-to-Image Generative Models«, in: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New York: ACM, Art. 384. Zu den Bildeigenheiten und den damit verbundenen Ästhetiken (*fluffy glamour glow*) vgl. Meyer, Roland (2023), »Es schimmert, es glüht, es funkelt – Zur Ästhetik der KI-Bilder«, in: 54 Books vom 20.03.2023, online unter <https://www.54books.de/es-schimmert-es-glueht-es-funkelt-zur-aesthetik-der-ki-bilder/> (letzter Zugriff: 11.07.23). Zu den Stilen vgl. Wang, Ning/Xie, Jiahao/Wu, Jihao u.a. (2023), »Controllable Image Captioning via Prompting«, in: Proceedings of the AAAI Conference on Artificial Intelligence 37(2), S. 2617-2625.

43 Vgl. Oppenlaender, Jonas/Linder, Rhema/Silvennoinen, Johanna M. (2023), »Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering«, CoRR abs/2303.13534.

44 Vgl. dazu Liu, Vivian (2023), »Beyond Text-to-Image: Multimodal Prompts to Explore Generative AI«, in: CHI '23: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, New York: ACM, Art. 482, und Qiao, Han/Liu, Vivian/Chilton, Lydia B. (2022), »Initial Images: Using Image Prompts to Improve Subject Representation in Multimodal AI Generated Art«, in: C&C '22, Proceedings of the ACM SIGCHI Conference on Creativity and Cognition, New York: ACM, S. 15-28.

45 Tur, Anil Osman/Dall'Asen, Nicola/Beyan, Cigdem u.a. (2023), »Exploring Diffusion Models for Unsupervised Video Anomaly Detection«, arXiv:2304.05841v2 [cs.CV] 2 Jul 2023. Zum Umgang von KI mit Kontexten und Situationen vgl. Marres, Noortje/Sormani, Philippe (2023), »KI Testen. ›Do we have a situation?‹«, in: zfm – Zeitschrift für Medienwissenschaft 29 (Test), S. 86-102.

46 Vgl. dazu Programme mit semantischen Anleihen an den Traum wie DeepDreamGenerator. Zur Generierung von Geschichten vgl. ferner X. Hong/A. Sayeed/K. Mehra u.a., Visual Writing Prompts.

sondern unter Verwendung ausgewählter Passagen aus ausgewählten Märchen der Brüder Grimm.⁴⁷ Wenn dabei, wie im Beispiel der Visualisierung von Rotkäppchen, Aschenputtel oder den drei Raben, die in die künstliche Intelligenz eingespeisten Textbausteine in ihrer kulturellen Verfestigung eigene (!) Bilder erzeugen, ist das Anliegen dieses Buches erreicht – wenngleich auf einem Umweg. Was vorliegt, ist nicht die Bebilderung einzelner Vorstellungen oder ganzer Vorstellungsfolgen, also das Einbildungswissen, das alle individuellen Lektüren in einer bestimmten als normal und ideal konzipierten Form soll begleiten können wie in der eingangs geschilderten Musterlektüre bei Immanuel David Mauchart, vielmehr sind es die Bilder der internen Lektüren einer Künstlichen Intelligenz.⁴⁸ Diese sind zwar nicht individuell, dafür aber sind ihre Resultate (wenn schon nicht die zugrundeliegenden Algorithmen) einsehbar. Vielleicht sollte man mit Blick auf die Komplexität besser von Bildprogrammen sprechen, gehen die entsprechenden Elaborate über einfache Wort-Bild-Relationen doch beträchtlich hinaus – wie im Fall eines Filmtrailers zu *Heidi (Cursed Heidi)*.⁴⁹ Was die Algorithmen umsetzen, ruft noch einmal den Ausgang dieses Unterfangens und damit das erste seiner vier strukturierenden Verben (Lesen, Merken, Wohnen, Träumen) in Erinnerung: Aus den Leseszenarien, die den Leser:innen die wortinduzierte Zustellung innerer Welten verheißen haben, ist ein Lesen ohne Leser:innen geworden. Die Systemstelle einer menschlichen Phantasie ist für diesen Akt der Rezeption nicht mehr vonnöten.⁵⁰

Damit tritt noch einmal jener Referenzrahmen in Erscheinung, der die historische Verhaltung zu dem ermöglichte, was an Bedeutung zulässig oder zu verwerfen ist, was als Sonde oder Brille für die Beobachtung des Besonderen diente, was als Fensteröffnung oder Fernrohr taugte, durch die man einen kritischen Blick werfen und mit diesem Blick die Selbstverständlichkeit der eigenen Operationen sichern und die Selbstverständigung über die eigene Logik der Bedeutungsproduktion betreiben konnte.⁵¹ Hier wurde verhandelt, was zu welcher Zeit und in der Differenz zu anderen Schauplätzen der Bedeutungsgenerierung als gelungen oder pathologisch, als an den Haaren herbeigezogen oder als Ausbund infantiler Albernheiten, was als hanebüchene Ungereimtheit oder was als Gipfel artistischer Beherrschung, was als ökonomisch ausgewogen oder was als Schwachsinn mit Anteilen barocker Phantasie galt. Hier wurden neben all den kleinmaschig wirkenden Bezügen grundlegende Ordnungen des Epistemologischen und des Ästhetischen sichtbar sowie großräumige Ambigui-

47 Vgl. dazu M. Ruskov, Grimm in Wonderland.

48 Zur Umsetzung von Bewegtbildfolgen und somit für das, was als DALL-E für Video von sich reden macht, vgl. Kim, Doyeon/Joo, Donggyu/Kim, Junmo (2020), »TIVGAN: Text to Image to Video Generation with Step-by-Step Evolutionary Generator«, in: IEEE Access 8, S. 153113-153122, sowie Vincent, James (2022), »Meta's new text-to-video AI generator is like DALL-E for video«, in: The Verge, online verfügbar unter: <https://www.theverge.com> (letzter Zugriff: 2. Juni 2024).

49 Online zugänglich unter <https://www.zueritoday.ch/videos/dieser-heidi-trailer-wurde-mit-kuenstlicher-intelligenz-erstellt-152478065> (letzter Zugriff: 25.07.23).

50 Und noch ein weiterer Aspekt ist festzuhalten: Es findet statt, was man als eine Inversion des Illustrationsprozesses beschreiben könnte. Vgl. dazu noch einmal A. Otto, IL-LUST-RATIO-N. Dort wird sehr stark mit Theoremen der Gestaltpsychologie argumentiert.

51 Zu dieser optischen Analogie vgl. auch Nowak, Stanislaw/Bartram, Lyn/Schiphorst, Thecla (2018), »A Micro-Phenomenological Lens for Evaluating Narrative Visualization«, in: IEEE Evaluation and Beyond – Methodological Approaches for Visualization (BELIV), Berlin: IEEE, S. 11-18.

tätstoleranzen und Kohärenzzumutungen, hier wurde aber auch wie bei Ernst Robert Curtius deutlich, dass eine klassische Kunst sich mit dem zahlenmäßig Einfachen zu bescheiden vermag, während eine nicht-klassische Kunst, wie der Romanist am Beispiel des Lipogramms als Prototypen aller Manierismen dekretiert, sich in einer horrenden Vielzahl beliebiger Weisen verliert, ihre Gegenstände auszusagen. Die Zahl der Möglichkeiten wird so zum Argument einer normativen Ästhetik.

Doch für die Umtriebe der Bedeutung sind nicht nur die Kunst im Allgemeinen und die Merkkünste im Besonderen zuständig, es gibt auch andere Schauplätze, an denen Grundlagenarbeit an der Produktion von Bedeutungen erfolgt. Einer dieser Schauplätze, der im bisherigen Verlauf immer schon alludiert wurde, ist der Traum. Auf der Suche nach der Einsichtnahme in die Welt innerer Bilder stand das Träumen als Sonderfall innerer Bildproduktion immer schon ganz oben auf der Liste. Den Traum zu observieren und zu analysieren, ihn zu typologisieren und zu unterscheiden, sich seiner Bildumtriebe und seiner Bedeutungsoperationen zu versichern, galten daher vielfältige Bestrebungen.⁵² Diese erfolgten ähnlich wie bei den Bemühungen um die Vorstellungen unter Einsatz aller nur verfügbaren Beobachtungs sonden, Medien und Theoriebildungen. Sie beließen es nicht bei der bloßen Aufzeichnung des Traumes als eines gegebenen Sachverhalts, den man dann nach allen Regeln der Kunst (und unter Beiziehung entsprechender Theorieansätze) analysieren konnte, sie versuchten ihn vielmehr auch durch alle nur vorstellbaren und denkbaren Experimente zu induzieren und zu manipulieren – und ihm unter erschwerten oder gar als paradox geltenden Bedingungen nachzustellen.⁵³

Mit dem Träumen als jenem letzten Verb, das neben dem Lesen, dem Merken und dem Wohnen diesem Buch sowohl seine Struktur als auch seine materiale Unterfütterung verleiht, steht ein kulturell sanktioniertes Residuum von Ungereimtheiten, von fehlender Kohärenz und Plausibilität, von Brüchen und Sanktionen, von Außerkraftsetzung von Gesetzmäßigkeiten und Vorgaben zur Disposition, das immer wieder anlässlich Benjamins Lesart des Traumes in *Erfahrung und Armut* durch unterschiedliche Bild- und Bedeutungsschichten verfolgt wurde. Die Fokussierung auf den Traum ist aber mehr als nur einem Interesse an bestimmten Formgebungs- und Formgebungsfreigabebemühungen geschuldet. Mit ihr und an allen Fragen der Bedeutungsstiftung vorbei wird der Mensch zur Organisationseinheit einer entsprechenden Epistemologie. Fortan steht er als lebendes, sprechendes und arbeitendes Wesen im Zentrum des Wissens.⁵⁴ Diese Fokussierung ist selbst zentraler Teil der Inauguration und war damit gerade auch Teil jener Aufmerksamkeitsgeschichte des Individuell-Anthropologischen, die seit dem 18. Jahrhundert die Ganzheit des Menschen in den Blick nahm und in Umsetzung dessen die Spezifika eines trauminhärenten Wissens allererst ans Licht

52 Dazu siehe die Einleitung. Für die entsprechenden Bemühungen vgl. Dieterle, Bernard/Engel, Manfred (Hg.) (2020), *Mediating the Dream. Les genres et médias du rêve*, Würzburg: Königshausen & Neumann, sowie Dieterle, Bernard/Engel, Manfred (Hg.) (2022), *Typologizing the Dream. Le rêve du point de vue typologique*, Würzburg: Königshausen & Neumann. Beide Bände sind Teil der inzwischen fünf Bände umfassenden Reihe *Cultural Dream Studies/Kulturwissenschaftliche Traum-Studien*.

53 Vgl. dazu etwa Heermann, Georg (1838), »Beobachtungen und Betrachtungen über die Träume der Blinden, ein Beitrag zur Physiologie und Psychologie der Sinne«, in: *Monatsschrift für Medicin* 3, S. 116-180.

54 Vgl. dazu M. Foucault, *Die Ordnung der Dinge*.

förderte – und damit untergründig das Spiel der Zahlen zum Ermöglichungsgrund, zum quantitativen Apriori menschlicher Individualität erhob.⁵⁵

Im Zuge einer Fokussierung auf den ganzen Menschen schien alles am Traum bemerkenswert – vor allem sein Formenreichtum und seine Verknüpfungsweisen, sein Überschuss und seine Eigenlogik, seine Vielsinnigkeit und Ambiguität, seine Bizarrerien und Phantastereien und nicht zuletzt seine Sanktion, Kausalität und Kohärenz, Plausibilität und Schlüssigkeit außer Kraft zu setzen – und sich dabei in Dimensionen eines Unbewussten verlieren zu können, das zum Fluchtpunkt diverser Ent- oder Transrationalisierungsbemühungen werden sollte. Für den Traum gilt, was auch dem Virtuellen gerne und mit einem gönnerhaften Gestus der Pauschalität bescheinigt wird – dass dort alles möglich und nichts unmöglich sein soll.⁵⁶ Der Traum wird zum Anlass, besondere Bilder und besondere Bildverknüpfungen zu beschreiben – changierend zwischen halluzinatorischem Charakter und Klartraum, zwischen der Zugänglichkeit durch das Bewusstsein und dem Entzug durch die unterschiedlich konzeptualisierten Arten eines Unbewussten und rational nicht Belangbaren, eines Untergründigen und Beiläufigen, eines Unterschwelligen und Latenten.⁵⁷ Der Traum wird zum Vorbild einer Welt, in der die Ordnung der Dinge nicht gilt, in der die Verfügung von physikalischen Gewichten und Formen keiner Sanktion mehr unterliegt: Weil das Traumgeschehen weder von Seiten physikalischer Gesetze noch von Seiten kultureller Konventionen belangbar ist, kann es im Moment einer solchen Kulturentblößung Möglichkeitsräume eröffnen, die Benjamin in *Erfahrung und Armut* als neues Barbarentum beschrieben hat – und die in jene anderen Formgebungen und Geschichten münden, die Benjamin entsprechend nachgezeichnet und damit Muster für andere Formen der Vergleiche und des Anschlusses eröffnet hat.

Im Zuge dessen wird zunehmend die Frage relevant, wie sich der Traum als vielfältig besetzbare Kulturthema zu den Möglichkeiten der technischen Immersion verhält, ob er diesen gegenüber abfällt oder ob er sich ihnen gegenüber als ebenbürtig erweist. Weil beide Phänomenbereiche mit einer ausgesprochen grundlegenden Möglichkeitsunterstellung versehen sind, stellt sich die scheinbar paradox anmutende Frage, ob sich dort, wo alles möglich sein soll, also in der Virtualität und im Traum, die jeweils attestierten Möglichkeiten die Waage halten oder eine der Möglichkeitsoptionen möglicher ist oder sein soll als eine andere. Diese Formulierung von der möglicheren Möglichkeit (und im Vorgriff auf eine historisch verbürgte Steigerung von Wahrscheinlichkeiten) ruft eine andere und ähnlich konstruierte Formulierung in Erinnerung, mit der Luhmann die eben nur scheinbar paradoxe »Steigerbarkeit der Individualität des Individuums« angeschrieben hat. »Oder um es im heutigen Jargon der Identität zu formulieren: wie sollte man es sich vorstellen, daß ein Individuum identischer ist als andere?«⁵⁸

55 Vgl. zur zugehörigen anthropologischen Formation immer noch H.-J. Schings (Hg.), *Der ganze Mensch*.

56 Vgl. dazu noch einmal H.-C. Jetter/R. Rädle/T. Feuchtner u.a., *Everything is possible!*.

57 Vgl. zur Skalierung des Bizarren vgl. Schredl, Michael/Erlacher, Daniel (2003), »The Problem of Dream Content Analysis Validity as Shown by a Bizarreness Scale«, in: *Sleep and Hypnosis* 5, S. 129-135, sowie zum Unbewussten noch einmal P. Fuchs, *Blindheit und Sicht*.

58 Luhmann, Niklas (1993), »Individuum, Individualität, Individualismus«, in: *Gesellschaftsstruktur und Semantik. Studien zur Wissenssoziologie der modernen Gesellschaft*, Band 3, Frankfurt am Main: Suhrkamp, S. 149-258, hier: S. 154.

Eben diese Frage nach der nicht ontologischen, sondern praxeologischen Ebenbürtigkeit des Virtuellen ist im bisherigen Verlauf immer wieder aufgetaucht und durchzieht die unterschiedlichen Diskussionen wie ein roter Faden. Sie hat ihre eigene Systematik und ihren eigenen Stellenwert dort, wo nach dem Verhältnis, nach der Unter- oder Überlegenheit von technischen Operationen und anthropologischen Residuen gefragt wird, wo davon gehandelt wird, ob die Immersion nicht nur der Imagination im Allgemeinen, sondern auch noch dem Träumen im Besonderen überlegen sein könnte.⁵⁹ Jenes Moment des Kompetitiven, das nach dem Verhältnis von normalen und virtuellen Gedächtniskünsten, nach dem Verhältnis von imaginären und nicht-imaginären Orten (*Loci Imaginary Versus Virtual Loci* – wie es oben anlässlich eines Virtual Memory Palace hieß) und nach der besonderen Rolle imaginärer Interfaces gefragt hat, macht sich einmal mehr auch hier breit.⁶⁰ An derartigen Fragen hat sich die Gediegenheit des Virtuellen zu bewähren.⁶¹

59 Und wie sich die einzelnen Positionen in einem um die Imagination erweiterten Spektrum anordnen lassen. Vgl. dazu noch einmal C. B. Stapleton/J. Davies, *Imagination*.

60 J.-P. Huttner/D. Pfeiffer/S. Robra-Bissantz, *Imaginary Versus Virtual Loci*.

61 Zur Metaphorik des Gediegenen als Ausweis einer Seinsform sowie zu ihrer Herleitung vgl. noch einmal S. Rieger, *Virtuelles Testen*.

