



## Die Ästhetik der KI: Surreales Wirken in der Nutzung generativer KI

*Michael Schultz*

Originalarbeit

*Abstract:* Mit den großen Fortschritten in text- und bildgenerierenden KI, welche in den letzten Jahren für die Öffentlichkeit zugänglich gemacht wurden, hat sich ein Diskurs um die Frage gebildet, inwieweit die durch diese KI-generierten Bilder als Kunst zählen können. Um eine Grundlage für die Beantwortung dieser Frage zu bilden, ist es wichtig, die Perspektive des Nutzers generativer KI, in welchem Betrachter und Erschaffer zusammenfallen, nicht zu vernachlässigen. Ich möchte daher untersuchen, welche Wirkung KI-generierte Bilder in der Phänomenologie des Betrachters haben können. Hierfür werde ich insbesondere auf drei Aspekte dieser Wirkung eingehen. Erstens möchte ich untersuchen, inwiefern durch den unsicheren Status des Urhebers des KI-generierten Bildes in dem Nutzer Annahmen über die Subjektivität der KI entstehen. Zweitens wird es mir darum gehen, wie solche Bilder Weltvorstellungen hervorrufen, welche durch einen Gegensatz von Materialität und Virtualität geprägt sind. Und drittens werde ich darstellen, inwiefern die Interaktionen zwischen Nutzer und KI-generiertem Bild ein Element der Entfremdung beinhalten. Hierdurch möchte ich zeigen, dass solche Bilder mit einer bestimmten ästhetischen Erfahrung verbunden sind, deren Existenz als Grundlage verwendet werden kann, um weiterführend über ihren Status als Kunst zu argumentieren.

*Keywords:* Künstliche Intelligenz, Ästhetik, Surrealismus, Entfremdung

*Copyright:* Michael Schultz | Lizenz: CC BY-NC-ND 4.0

## 1. Einleitung

Das Genre der Science-Fiction ist bevölkert von KI, die Menschen in nahezu jeder Hinsicht ähneln, aber ihre Alterität in einer letzten großen Grenze zwischen Menschen und Maschine festgesetzt vorfinden: Das Erschaffen von Kunst. So macht den Roboter aus Isaac Asimovs Kurzgeschichte *The Bicentennial Man* einzigartig, dass er Kunst aus Holz schnitzen kann (vgl. Asimov 2000). In *Neuromancer* vergleicht sich die KI Dixie Flatline mit der fortschrittlicheren KI – Wintermute und witzelt, dass letztere wohl dem Protagonisten Case ein Gedicht schreiben könnte, auch wenn ihre Intentionalität nicht menschlich ist (vgl. Gibson 1984). In *Zima Blue* ist eine Journalistin von Staunen erfüllt, als sie herausfindet, dass ein Künstler, den sie interviewt, ursprünglich ein Poolreinigungsroboter war, der über die Jahrhunderte hinweg immer weiter verbessert wurde (vgl. Valley 2019).

Durch Fortschritte in dem Bereich der künstlichen Intelligenz haben solche Szenarien zunehmend angefangen, sich von dem Bereich der Fiktion in den der Realität zu bewegen. In den 2010ern stand noch die Imitation von Stilen im Vordergrund. Durch die Technik des *Style Transfers* (vgl. Sovhyra, Yu u. Yusa 2022, S. 153) wurde beispielsweise versucht, ein Foto Tübingens in Kunststilen verschiedener Epochen darzustellen (vgl. Israfilzade u. Pileliene 2018, S. 5), oder ein Gemälde zu erschaffen, das die Essenz Rembrandts einfangen soll (vgl. Sovhyra, Yu u. Yusa 2022, S. 156). Durch Generative *Adversarial Networks* wurde versucht, Variation in die Nachahmung bestehender Stile zu bringen (vgl. Sovhyra, Yu u. Yusa 2022, S. 156), doch die Ergebnisse sahen sehr abstrakt aus und erhielten wenig Aufmerksamkeit.

In den letzten Jahren haben *prompt*-nutzende KI wie *StableDiffusion*, *DALL-E* oder *Midjourney* das Thema in das Blickfeld der Öffentlichkeit gerückt. Sie liefern Ergebnisse, die sich von einem durch Menschen gemachten Foto oder Gemälde kaum unterscheiden lassen, und sind einfach zu bedienen, da man bloß eine Beschreibung des gewünschten Bilds eingeben muss, um dieses produzieren zu lassen. Eine Folge der Aufmerksamkeit, welche KI-generierte Bilder erhalten, ist die Frage danach, ob das, was von diesen KI erschaffen wird, Kunst ist. Hierin setzen sich bestehende Diskurse der Ästhetik fort, in denen untersucht wird, wie sich

Kunst definieren lässt und was die Rolle des Künstlers und des Betrachters von Kunst in dieser Definition sind.

Es hat sich innerhalb der Fortsetzung dieses Diskurses allerdings auch eine Reihe an Autoren herausgebildet, welche betonen, dass die Frage danach, ob das Potential der KI, Kunst hervorzubringen, dem eines menschlichen Künstlers gleicht, nicht dem Problem angemessen ist, welches durch die Existenz solcher KI aufgeworfen wird. So betont Betancourt, dass die Hinterfragung von Eigenschaften der Kunst durch *AI art* ebenso wenig dadurch gelöst wird, sie nicht als Kunst anzuerkennen, wie dies bei dem Aufkommen der Fotografie erfolgreich war (vgl. Betancourt 2022, S. 86 f.). Mark Coeckelbergh gibt zu bedenken, dass KI über eine gänzlich eigene Art von Kreativität und Subjektivität verfügen könnte (vgl. Coeckelbergh 2017, S. 294), welche gleichermaßen eine offenbarende Wirkung auf den Betrachter hat, wie es das Genie des Künstlers nach romantischen und expressiven Kunstauffassungen hat (vgl. Coeckelbergh 2017, S. 300). Die Strömung der *New Aesthetics* verlangt, dass KI als Erzeuger von Kunst nicht determinativ, sondern reflektiv beurteilt werden soll, also nicht als Repräsentation eines Konzepts von Kunst, sondern in Hinsicht auf ihre Qualität als Objekt subjektiver Erfahrung (vgl. Contreras-Koterbay 2019, S. 107).

Ich möchte mich der in diesen Positionen vorgeschlagenen Methodik anschließen und die Frage stellen, wie die von KI-generierten Bilder auf uns wirken. *Wie* sowohl im Sinne der Frage, welche Erfahrungen sie in unserer subjektiven Weltwahrnehmung erzeugen, aber auch in dem Sinne, welcher Mittel sie sich dafür bedienen. Ich möchte diese Frage damit beantworten, dass die KI-generierten Bilder in ihrem Betrachter eine Mischung von Entfremdung und Faszination auslösen. Die Frage nach der Autorenschaft des Bildes ist dabei nicht nur Begleitwerk dieser Wirkung, sondern ihr Mittel. Deshalb ist sie gerade bei den KI präsent, die den Betrachter durch die Verwendung von *prompts* in ein *engagement* einbinden, das beide Seiten in einen unsicheren ontologischen Status rückt. Wurzeln dieser Ästhetik der KI möchte ich im Surrealismus verorten, der ähnliche Spiele mit der Ontologie von Subjekt und Objekt, von Gebrauchsgegenstand und Kunstwerk beinhaltet.

Der Status als Betrachter und als Erschaffer fällt hierbei in der Nutzung der KI zusammen. Dies ist in dieser zu beschreibenden Wirkung der KI

begründet, welche ihren Nutzer in beide Positionen versetzt. Daher wird es im Folgenden, wenn von Nutzern und Betrachtern die Rede ist, hauptsächlich um die Perspektive eines solchen Anwenders der KI gehen, also jemand, der einen *prompt* zur Erschaffung eines Bildes eingibt. Ich vermeide, diesen Anwender als Künstler zu bezeichnen, da ich die Frage nach dem Status der KI-generierten Bilder als Kunstwerk für nicht abschließend geklärt halte.

## 2. Technische Hintergründe

Bevor die ästhetischen Aspekte KI-generierter Bilder näher erkundet werden können, soll zunächst ein grundlegender Überblick über die zugrunde liegenden Mechanismen geliefert werden, welche es der KI ermöglichen, diese Bilder zu erzeugen. Dies ist erforderlich, um eine Mystifikation dieser Technologie zu vermeiden, aber auch um eine Perspektive annehmen zu können, welche über die des naiven Nutzers hinausgeht. Im Vordergrund soll dennoch die Nutzerspektive bleiben, nicht die des Entwicklers, der Umfang dieser Darstellung soll daher an dem Wissen orientiert sein, über das ein gut informierter Nutzer potentiell verfügen würde. Mir geht es hierbei im Speziellen um Bilder erzeugende generative KI, die sich *diffusion models* und eines *text-to-image*-Zugangs bedient.

Diese Art von KI ist aus mehreren Gründen für mich von einem besonderen Interesse: Erstens handelt es sich um eine non-deterministische KI, das heißt, sie ist dazu in der Lage, bei einer wiederholten Verwendung desselben Inputs unterschiedliche Outputs zu generieren (vgl. Musiol 2024, S. 49). Dies hat den praktischen Nutzen, dass es eine größere Variation an verwendbaren Outputs gibt, es erweckt aber auch in dem Nutzer in einem stärkeren Ausmaß den Eindruck einer Autonomie der KI, da das Ergebnis weniger vorhersagbar ist. Zweitens ermöglicht der *text-to-image*-Zugang zu generativer KI eine sehr intuitive Nutzung durch die Verwendung von *prompts*, also stark variierbaren Beschreibungen, welche der Nutzer eingibt, um die KI in die Richtung eines gewünschten – wenn auch nicht klar voraussagbaren – Ergebnisses zu lenken (vgl. Musiol 2024, S. 56). Zusammen mit dem dritten Punkt, der höheren Qualität von *diffusion models* gegenüber

älteren Modellen wie *GANs* (vgl. Dhariwal u. Nichol 2021, S. 2), hat dies maßgeblich zu der stärker werdenden Beliebtheit generativer KI und zu der kommerziellen Nutzung über die experimentelle Verwendung hinaus beigetragen. Im Spezifischen werde ich die generative KI *DALL-E 2* darstellen, da bei dieser der Zusammenhang von Text und Bild im Training sowie beim Output gut nachvollziehbar ist.

Andere Arten von Outputs werde ich in dieser Untersuchung vernachlässigen: Bildgenerierende KI wird von dem gewöhnlichen Nutzer präsenter als musik- oder videogenerierende KI empfunden. Textgenerierende KI wiederum ist weit verbreitet und berührt ebenfalls Aspekte der wahrgenommenen Autonomie oder Subjektivität, aber das ästhetische Wirken ist bei dieser von einer geringeren Bedeutung, während andere Problematiken stärker hervortreten, sodass sie eine eigene Untersuchung erfordert und hier nur am Rande vorkommen wird.

Unterschiedliche Arten von Algorithmen und künstlicher Intelligenz zeichnen sich vor allem durch verschiedene Arten des Lernens aus. Allgemein betrachtet bedient sich KI des *machine learnings*. Das bedeutet, dass sie Voraussagen trifft und diese mit einem Datensatz vergleicht, um dann ihre Parameter für ein akkurateres Resultat anzupassen (vgl. Musiol 2024, S. 4). Generative KI verwendet dabei vor allem *unsupervised learning*, ihre Trainingsdaten sind also noch nicht vorsortiert, sondern sie erkennt selbst Muster in diesen (vgl. Musiol 2024, S. 5).

*Diffusion models* funktionieren, indem sie schrittweise ein Bild aus Rauschen generieren. Sie lernen dies zu tun, indem versucht wird, den Prozess der mehrmaligen Überlagerung eines Bildes aus einer vorgegebenen Datenbank mit Rauschen umzukehren. Die KI versucht mittels einer Funktion, eine Voraussage darüber zu treffen, wie viel Rauschen in einem einzelnen Schritt dem Bild hinzugefügt wurde. Dies wird dann mit dem tatsächlich hinzugefügten Rauschen verglichen, um in einer Anpassung der Parameter dieser Voraussage eine geringere Differenz zwischen Voraussage und tatsächlich hinzugefügtem Rauschen zu erreichen. Die KI führt damit eine Wahrscheinlichkeitsannäherung durch. Sie lernt, wie ein Bild aussehen könnte, wenn man diesem Rauschen hinzufügt. In dem Generieren eines Bildes wird dieser Prozess umgekehrt. Es wird versucht, in zufälligem Rauschen ein Bild zu erkennen und das Rauschen zu reduzieren (vgl. Abbeel, Ho u. Jain 2020, S. 2).

Ein Bild mit der Beschreibung *Hund* zu generieren, bedeutet also, von Rauschen ausgehend eine Voraussage darüber zu treffen, wie am wahrscheinlichsten ein Bild eines Hundes aussehen würde, welches man so mit Rauschen überlagert hat, dass nur noch ein Bild von Rauschen übrig ist. Schwer voraussagbar ist das konkrete Ergebnis dieser Umkehrung für den Nutzer unter der Voraussetzung, dass die KI mit einer diversen Datenbank trainiert wurde und sein *prompt* nicht einfach bloß ein sehr spezifisches Bild aus dieser Datenbank beschreibt (vgl. Musiol 2024, S. 58). Würde die KI mit hunderten fast identisch aussehenden Bildern eines *Soft Coated Wheaten Terriers* trainiert werden, die auch ein entsprechendes Label erhalten haben, und der Nutzer durch einen *prompt* ein Bild mit dem Label *Soft Coated Wheaten Terrier* fordern, dann würde er ein Ergebnis erhalten, welches den ursprünglichen Bildern dieses Hundes sehr ähnelt.

Insbesondere für *text-to-image*-basierte generative KI ist nun die Frage interessant, wie die KI überhaupt dazu kommt, ein Bild mit einem bestimmten *prompt* zu verbinden. Denn der Konflikt zwischen der Autonomie des Nutzers und jener der KI findet vor allem im *prompt-engineering* statt, in dem versucht wird, durch die Anpassung der Beschreibung des geforderten Bildes ein bestimmtes Ergebnis zu erreichen, welches in Abhängigkeit von dem Training der KI aber für den Nutzer schwer vorhersagbar sein kann (vgl. Musiol 2024, S. 86).

Große Fortschritte gab es in dieser Verknüpfung von Text und Bild im Bereich der generativen KI vor allem durch die KI *DALL-E 2*, welche vor dem *diffusion model* das *CLIP-Modell* schaltet (vgl. Musiol 2024, S. 54). Durch *contrastive pretraining* lernt dieses, zwischen ähnlichen und verschiedenen Paaren von Text und Bild zu unterscheiden. Dies bedeutet, dass *CLIP* eine gemeinsame Repräsentation – ein *embedding* – von Text und Bildern erstellt, welche über die Ähnlichkeit der Vektoren, als welche Text und Bilder repräsentiert werden, eine Aussagekraft darüber beinhaltet, inwieweit diese zusammengehörig sind. Im Trainingsprozess wird diese Repräsentation mit einer Datenbank an mit Text versehenen Bildern verglichen. Die Parameter des Modells werden dabei so angepasst, dass zusammengehörige Bilder und Texte mit einer stärkeren Ähnlichkeit als nicht zusammengehörige repräsentiert werden (vgl. Radford et al. 2021, S. 2).

In *DALL-E 2* wird nun das *diffusion model* anhand der Repräsentationen aus dem *CLIP-Modell* trainiert (vgl. Ramesh et al. 2022, S. 1). Dies bedeutet,

dass – wenn ein *prompt* eingegeben und als Repräsentation eines Textes codiert wird – die KI dazu in der Lage ist, mit dem *diffusion model* ein Bild zu erzeugen, welches in der Repräsentation jener des Textes ähnelt und damit aufgrund des Trainings von *CLIP* über eine Zugehörigkeit zu diesem Text verfügt. Das durch den *prompt Hund* erschaffene Bild ähnelt somit jenen Bildern aus der Trainingsdatenbank von *CLIP*, die mit der Beschreibung *Hund* versehen sind. Neuartige Kombinationen von Text können somit auch neuartige Bilder erschaffen. Diese Fähigkeit der KI, Aufgaben durchzuführen, auf welche sie nicht im Spezifischen trainiert wurde oder zu denen in ihrer Trainingsdatenbank keine spezifischen Beispiele vorhanden waren, wird als ihre *zero-shot performance* bezeichnet (vgl. Musiol 2024, S. 61).

Zur Codierung des Textes bedient sich *CLIP* dabei eines *transformers* mit einem *attention mechanism* (vgl. Radford et al. 2021, S. 5). *Attention* ist die Fähigkeit eines generativen Modells, unterschiedliche Teile einer Textsequenz in ihrer Bedeutung zu gewichten. Dabei leistet *attention* vor allem, dass die KI dazu in der Lage ist, Teile eines Textes in Relation zueinander zu setzen, die sich nicht in einer unmittelbaren Nähe zueinander befinden (vgl. Vasvani et al. 2017, S. 2). *Attention* basierte Modelle wurden ursprünglich zum Training von KI zur Übersetzung von Texten vorgeschlagen, wobei diese Modelle, wenn sie eine Übersetzung erstellen, danach suchen, an welchen Positionen in dem Ursprungssatz die relevanten Informationen konzentriert sind (vgl. Bahdanau, Cho u. Bengio 2016, S. 1). Dementsprechend kann *DALL-E* seine Aufmerksamkeit auf die relevanten Teile eines *prompts* richten, also in dem *prompt Ein Hund, der einen Ball fängt* die Worte *Hund*, *Ball* und *fängt*, statt sich an Bildern zu orientieren, die durch *ein*, *der* oder *einen* beschrieben werden.

### 3. Die Subjektivität der KI

Welche Aussagen lassen sich nun aufgrund dieser Darstellung generativer KI darüber treffen, inwieweit die Zuordnung bestimmter Eigenschaften an die KI durch ihre Funktionsweise begründbar ist? Zunächst einmal ist festzuhalten, dass das, was die KI kreiert, eine Wahrscheinlichkeitsannäherung auf Grundlage der Trainingsdaten und des *prompts* mit einem Zufallselement ist. Eine Zuschreibung desselben

Ausmaßes an Freiheit oder Unfreiheit zwischen möglichen Handlungen in dem Schaffungsprozess abzuwägen, wie bei einem menschlichen Künstler, ist voraussetzungsreich und würde beispielsweise eine deterministische Perspektive auf menschliche Handlungsträger erfordern. Dennoch lässt sich ein Argument dafür aufbauen, dass sie in einem größeren Ausmaß autonom ist als einfach nur ein Prozess, der sich unseres Zugriffs oder Verständnisses entzieht, und in einem stärkeren Sinne Kunst erschafft als ein natürliches Phänomen, das für einen Betrachter ästhetisch erscheint.

Der naheliegendste Weg, solche Eigenschaften bei KI zu begründen, wäre, über die Emergenz von Qualitäten zu argumentieren. Für einfache Konzepte wie die Fähigkeit von Sprachmodellen, arithmetische Berechnungen durchzuführen (vgl. Musiol 2024, S. 108), oder von bildgenerierender KI, das Zusammenspiel von Licht und Oberflächen zu verstehen (vgl. Du et al. 2023, S. 3), stehen hierbei durchaus Methoden zur Verfügung. Für abstraktere Konzepte wie Kreativität erfordert dies allerdings einen ausführlichen Begriffsdiskurs und ist weitaus spekulativer. Ich möchte an dieser Stelle stattdessen einen Ausblick darauf geben, über das den neuronalen Netzen zugrunde liegende Paradigma zu argumentieren, wobei aber festzuhalten ist, dass auch dieser Diskurs durchaus über das hinausgeht, was hier angedeutet wird.

Auf Grundlage ihrer Funktionsweise lässt sich nämlich ein Argument dafür aufbauen, dass generative KI über eine einfache Form von Subjektivität verfügt, also ein Eingebundensein in die durch das Subjekt interpretierte Welt. Hierfür werde ich die KI-Kritik von Hubert Dreyfus aufgreifen. Dieser möchte ich mich nur begrenzt anschließen, da sie durchaus umstritten ist, aber sie ist für diesen Diskurs fruchtbar, da sich aus ihr Bedingungen für eine Subjektivität der KI ableiten lassen, auch ohne seine Argumente gegen die KI-Forschung in ihrer Vollständigkeit zu übernehmen. Ich lege mich hierfür auf den Begriff der Subjektivität fest, es bleibt aber zu erwähnen, dass dieses Phänomen in eigentlich sehr nahen beieinander liegenden Diskursen mit unterschiedlichen Begriffen wie beispielsweise *coping* (vgl. Dreyfus 2014, S. 81) oder Intentionalität (vgl. Negru 2013, S. 30) behandelt wird.

In *Mind over Machine* kritisiert Dreyfus an der KI-Forschung des 20. Jahrhunderts, dass diese der KI Qualitäten zuschreibt, über welche sie eigentlich nicht verfügt (vgl. Dreyfus u. Dreyfus 1998). Es geht ihm vor allem um die Repräsentation von Wissen und Fertigkeiten, allerdings spielt



Subjektivität hierfür eine wesentliche Rolle. Dreyfus kritisiert, dass KI bloß einfache Formen menschlicher Expertise abbilden kann, die auf kontextfreien Regeln (vgl. Dreyfus u. Dreyfus 1988, S. 21) und vorinterpretierten Listen von Eigenschaften (vgl. Dreyfus u. Dreyfus 1988, S. 53) beruhen. Fortgeschrittenere Expertise erfordert aber Intuition, also das mühelose, nicht auf Regeln reduzierbare Erkennen von Trennungen aufgrund vergangener Erfahrungen, welches aus einem Involviertsein in der Welt – das, was ich hier als Subjektivität bezeichne – und der Zuschreibung von unterschiedlichen Graden an Bedeutung an diese Welt resultiert (vgl. Dreyfus u. Dreyfus 1988, S. 28 f.).

Hierbei untersucht Dreyfus auch – zu dem Zeitpunkt des Verfassens seines Buchs – neuere Entwicklungen in der KI-Forschung, darunter Versuche, KI mit einem derartigen Involviertsein auszustatten. Ein Beispiel hierfür ist die KI *SHRDLU*. Diese hat Zugriff auf eine Repräsentation dreidimensionaler Blöcke und kann in *prompt*-basierten Interaktionen Aussagen über das Verhalten dieser Blöcke zueinander treffen (vgl. Dreyfus u. Dreyfus 1988, S. 74). Dreyfus kritisiert an derartigen Versuchen vor allem, dass diese Art von KI nur über eine Mikro-Welt verfügt: Ein Set an miteinander verbundenen Beschreibungen oder Fakten, das aber nicht auf eine ihr zugrunde liegende Welt verweist und somit nicht den Status einer Subwelt erreichen kann (vgl. Dreyfus u. Dreyfus 1988, S. 76), also das, was Dreyfus als „*local elaborations of the one commonsense world we all share*“ (Dreyfus u. Dreyfus 1988, S. 76) versteht.

Eine Möglichkeit, die Schwächen der KI-Forschung zu überwinden, sieht Dreyfus in den neuronalen Netzen, da diese zwar Beschreibungen verwenden, aber diese nicht auf Listen oder vordefinierte Regeln reduzieren (vgl. Dreyfus u. Dreyfus 1988, S. 92). Mit generativer KI scheint sich diese Hoffnung durchaus zu erfüllen, da ihr Lernprozess dem auf Erfahrung basierten Lernen eines menschlichen Experten ähnelt und das Generieren neuer Daten auf die Ergebnisse dieses Lernens zurückgreift, nicht aber auf eine Auflistung der Beschreibung aller einzelnen Trainingsdaten. Indem sie nicht mit vorsortierten Trainingsdaten arbeitet, sondern in diesen überhaupt erst Muster erkennt, wird auch Dreyfus' Kritik der Notwendigkeit einer Normalisierung der Daten begegnet (vgl. Dreyfus u. Dreyfus 1988, S. 93) und mit dem Aufmerksamkeitsmechanismus wird diesen Daten eine Bedeutung beigemessen, ohne dass sie eine Vorinterpretation benötigen.

Dieses Potential neuronaler Netze ist durchaus paradigmatisch begründet. Die KI-Forschung, welche neuronalen Netzen zugrunde liegt, fußt auf Erwägungen der Subjektivität, insbesondere über das damit verknüpfte Konzept der Körperlichkeit (vgl. Negru 2013, S. 23). So basiert der *Perceptron* (vgl. McCorduck 2004, S. 104), welcher eine der Grundlagen generativer KI darstellt (vgl. Musiol 2024, S. 31), auf Ideen McCullochs, dass epistemologische Probleme sich nur in Abhängigkeit der Physiologie des Nervensystems begreifen lassen und dementsprechend KI sich an der Anatomie menschlicher Kognition orientieren muss (vgl. McCorduck 2004, S. 92). Und Daniel Crevier knüpft an seiner Kritik, dass Dreyfus in seinen Frühwerken einer Erklärung menschlicher Kognition mit dem Verweis auf die phänomenologische Methode ausweicht, die Feststellung an, dass Dreyfus sich in seinen späteren Schriften in eine Tradition mit dem Konnektionismus stellt (vgl. Crevier 1993, S. 132), aus dem auch die Ideen McCulloughs und der *Perceptron* entstanden sind.

Dennoch bleiben Kritikpunkte Dreyfus' bestehen. Er sieht als grundlegende Eigenschaft der Praxis menschlicher Expertise die Kontinuität der Erfahrung: Jede Situation wird auf Grundlage der vorangehenden Situation beurteilt, der menschliche Experte bewegt sich von Situation zu Situation (vgl. Dreyfus u. Dreyfus 1988, S. 88). Die meisten Formen generativer KI dagegen bewerten Situationen – also die einzelnen *prompts*, mit denen sie konfrontiert wird – auf Grundlage ihres abgeschlossenen Trainings, nicht aber aufgrund unmittelbar vorangehender Situationen. Bei einer KI wie *DALL-E* hat ein *prompt* keinen Einfluss darauf, wie der darauffolgende *prompt* verarbeitet wird. Anders sieht dies bei textgenerierender KI wie *ChatGPT* aus, aber auch hier wird ein *prompt* im Kontext eines einzelnen *Gesprächs* bewertet, ohne Einfluss auf ein anderes *Gespräch* zu haben. Und anders als ein menschlicher Künstler hat die bildgenerierende KI keine Möglichkeit, ihre Erfahrungen anhand von Emotionen zu bewerten, sondern misst ihnen Bedeutung nur anhand der Aufgabe der Wahrscheinlichkeitsannäherung bei (vgl. Dreyfus u. Dreyfus 1988, S. 94). Wenn hier von Subjektivität die Rede ist, dann also nur in dem Sinne, dass KI mit einer bedeutungsvollen, nicht vorinterpretierten Welt interagiert, nicht aber in dem Ausmaß, wie ein menschliches Subjekt mit seiner Welt interagiert.

Nichtsdestotrotz hat dies Auswirkungen auf die Zuschreibung von Autonomie und Kreativität. Man könnte diese Eigenschaften auf künstliche Intelligenz ausweiten, indem man den Aspekt der Freiheit, wie sie ein menschliches Subjekt hätte, aus ihnen streicht. Damit geht man die Gefahr ein, diese Begriffe derartig auszuweiten, dass sie auf eine Vielzahl an Artefakten, Lebewesen und natürlichen Prozessen zutreffen. Subjektivität als Voraussetzung dieser Qualitäten zu setzen, könnte aber dienlich sein, um von ihnen im Kontext generativer KI zu reden, ohne eine zu starke Ausweitung dieser Begrifflichkeiten durchzuführen. Ob diese Idee sich gegenüber Kritik bewähren kann, ist jedoch in einem fortführenden Diskurs zu klären. An dieser Stelle soll hiermit bloß eine klare Grenze gezogen werden, um KI nicht zu mystifizieren und ohne Vorbedingungen in den Bereich menschlicher Freiheit zu setzen, aber zugleich klarzumachen, dass der Diskurs über andere, sonst eher als rein dem Menschen zugehörig verstandene Qualität nicht damit abgeschlossen ist, auf die Abwesenheit einer solchen Freiheit zu verweisen. Im Folgenden möchte ich nun aber zu meinem eigentlichen Vorhaben zurückkehren, zu untersuchen, welche Wirkung die durch diese KI-generierten Bilder auf den sie betrachtenden Nutzer haben.

#### 4. Digitale Tagträume

Untersucht man die Wirkung solcher Bilder auf ihren Betrachter, so sind vor allem jene interessant, bei denen der Unterschied zu einem Foto oder einem von Menschen erschaffenem Kunstwerk klar ersichtlich ist. Dies sind Bilder, die darin versagen, etwas aus der wirklichen Welt darzustellen. Sie zeigen Gesichter, die keinem Menschen ähneln, Hände, die weit mehr als fünf Finger haben, und Körper, die mit Objekten in ihrer Umgebung verschmelzen. Andere zeigen Objekte, die in der Realität so nicht existieren oder existieren können. Und wieder andere zeigen zwar reale Objekte, setzen diese aber in eine fremde Umgebung, als wären sie dort wie Pflanzen gewuchert und würden natürlich in diese gehören. Solche Ergebnisse als ein bloßes Versagen der KI auf dem Weg zu dem gewünschten Bild zu sehen, unterschätzt ihre ästhetische Wirkung, denn hierin präsentiert sich die Weltwahrnehmung der KI, das, was man als ihre Subjektivität interpretieren

könnte. So schreibt Matilde Carrasco Barranco: „*AI-generated ‚portraits‘ present deformities, strange shapes and blurry contours that are the machine’s own versions of our living world. Thus, they open a territory to explore perception*“ (Barranco 2022, S. 104). Auch Mazzone und Elgammal heben den Wert dieser Bilder darin hervor, neuartige visuelle Stimuli zu liefern, welche ihre Originalität gerade daraus beziehen, dass sie nicht einer Intention entspringen (vgl. Elgammal u. Mazzone 2019, S. 2).

Die hierbei präsentierten Objekte ähneln jenen, welche in der Kunst des Surrealismus dargestellt werden: Lebendiges, das durch dessen Verzerrung unbelebten Dingen gleicht, und Unbelebtes, das wie Tiere und Pflanzen in eine Umwelt gesetzt wird. Das Phantastische wird dem Wirklichen angeglichen, auf eine Ebene mit diesem gestellt (vgl. Anders 2013, S. 332 f.). So wie nach Günther Anders’ Beschreibung des Surrealismus dieser die Wirklichkeit im Phantastischen erfasst, in welcher uns durch die Gerätwelt ebenfalls tote Dinge als Lebendiges präsentiert werden, beschreibt auch das Bilder schaffende Modell seine eigene Realität. Rodrigo Hernández-Ramírez schreibt, der Computer bevölkere die Welt mit Dingen, die wir vorher als tot betrachtet haben und die uns nun lebendig erscheinen (vgl. Hernández-Ramírez 2015, S. 12).

Neben diesem Spiel mit dem Lebendigen und dem Dinglichen ist auch die Objektkunst ein wesentlicher Aspekt des Surrealismus, dem das generative Modell gleicht. Denn das spielerische Verwenden des *prompts* nutzenden Modells, in welchem versucht wird, solche surrealen Bilder zu erzeugen, gleicht Bretons Suche nach dem Traumobjekt. So wie Breton vorschlägt, nach dem Erwachen das im Traum Gesehene niederzuschreiben und die Objekte des Traums handwerklich nachzuahmen (vgl. Lehmann 2011, S. 134), liefert der experimentierende Nutzer der KI Beschreibungen, die einem Tagtraum gleichen, lässt sie Variationen erstellen und Bilder hochrechnen, bis ein Produkt entsteht, das seiner Phantasie gleicht. Das, was dabei entsteht, beinhaltet all jene Prinzipien, die Ingrid Pfeiffer dem Surrealismus zuordnet: Die Entfremdung, nämlich die Heraushebung der dargestellten Dinge aus ihrem Kontext. Die Kombinatorik, also das Zusammenführen verschiedener Welten zum Erschaffen einer Reaktion des Schocks. Und auch die Metamorphose, die Möglichkeit der Verwandlung in etwas anderes (vgl. Pfeiffer 2011, S. 15).

Nun kann man jedoch den KI-generierten Bildern vorwerfen, dass ihnen die Intention des Surrealismus abhandenkommt, dass sie nicht auf das Dasein des gewöhnlichen Aussehens der Welt als Betrug verweisen (vgl. Anders 2013, S. 332 f.), sondern nur einen weiteren Betrug liefern. Denn sie regen nicht zum Denken an, ihre Mannigfaltigkeit und die einfache, schnelle Weise ihrer Produktion lassen den Nutzer von einem Bild zum nächsten tänzeln, Beschreibung um Beschreibung eingeben, ohne das Gesehene zu reflektieren. Bestenfalls finden sie noch Anwendung als *Meme*, als Hintergrund in einem Vortrag oder als Begleitbild für ein *tabletop RPG*. Sie werden zu einem Wegwurfprodukt. So schlagen auch Mazzone und Elgammal vor, das KI-generierte Bild könnte massenproduzierte dekorative Kunst ersetzen, in der die Wiederholung weniger, angenehmer Charakteristika wünschenswert ist (vgl. Elgammal u. Mazzone 2019, S. 6). Es spiegelt sich hierin Anders' Vorwurf am neuen Surrealismus wider, dass dieser einem die lebenden Apparate und das verdinglichte Leben als Genussobjekte präsentiere (vgl. Anders 2013, S. 331), so wie auch die Objektkunst des Surrealismus dem Vorwurf begegnete, dass sie durch den Verkauf von Kunstwerken als Designgegenstände wie das *Mae West-Lippensofa* die Kunst kommerzialisieren (vgl. Pfeiffer 2011, S. 22). Das Bestreben des Surrealismus, vom Markt der schönen Künste wegzukommen und sich den alltäglichen Konsumprodukten als Kunstwerke zuzuwenden (vgl. Lehmann 2011, S. 129), schließt hierin einen Kreis. Eine solche Einschätzung in Bezug auf KI lässt sich auch in Untersuchungen der Fähigkeit der KI, Kunst zu generieren, finden.

So vergleicht Dorothea Winter die KI-generierten Bilder mit Kants Kategorien von Kunst und stellt dabei fest, dass diese zwar nicht über das Genie verfügt, welches für die schöne Kunst nötig sei, aber die KI – wenn man ihr praktische Freiheit zuspräche – das Potential zur angenehmen Kunst hätte, also jene, welche den reinen Genuss (vgl. Winter 2022, S. 24) als einen mit klarer Absicht gesetzten (vgl. Winter 2022, S. 26 f.) Zweck hat und bei der das Angenehme sich im Kunstobjekt aus der Wahrnehmung ergibt, also im Objekt liegt, und nicht wie die Schönheit im schöpferischen Genie (vgl. Winter 2022, S. 59). Dass sich eine solche praktische Freiheit der KI über ein bloßes *als frei* Wirken hinaus tatsächlich gerechtfertigt zuschreiben lässt, ist angesichts der zuvor dargestellten Funktionsweise der KI zweifelhaft. Aufgrund der Subjektivität der KI, ihrer Fähigkeit zur Interpretation ihrer

Subwelt, ist dieses *als frei* Wirken aber überhaupt erst möglich. Sie ist hierdurch mehr als nur ein Werkzeug, mit dem ein Künstler ein Gemälde erschafft, oder ein natürlicher Prozess, der eine angenehm auf die Sinne wirkende Landschaft hervorbringt, da bei diesen die Interpretationsleistung bloß bei dem menschlichen Künstler oder Betrachter liegt. Und es ist durchaus plausibel, dass sich auch ein über Expertise verfügender Nutzer dieses *als frei* Wirkens nicht gänzlich entziehen kann. So kam es selbst bei einfacheren KI wie dem *Chatbot ELIZA* vor, dass auch Experten diesen anthropomorphisierten und mit diesem interagiert haben, als wäre er ein menschlicher Gesprächspartner (vgl. McCorduck 2004, S. 295).

Um das Potential der KI, ästhetisch und nicht nur angenehm zu wirken, zu erkunden, muss man jedoch über die bloßen Bilder und deren Inhalt und Zweckhaftigkeit hinaussehen. Es sind die Existenz des Kunstwerks selbst und dessen Entstehungsprozess und ontologischer Status, wodurch die Botschaft des Surrealismus wieder in die KI hineingebracht wird. Das Verhältnis von Betrachter beziehungsweise Nutzer und Kunstwerk zeichnet sich dabei durch drei miteinander verflochtene Aspekte aus, welche ich im Folgenden erkunden möchte: Erstens, der unsichere Status des Künstlers infolge der Wahl als künstlerischer Akt. Zweitens, die Virtualität als Abstraktion der Welt und Betrug. Und drittens, die Erfahrung von Entfremdung und Faszination

## 5. Der Status des Nutzers als Urheber des Bildes

Ein noch nicht genanntes Merkmal der surrealistischen Kunst ist das Spiel mit dem Status des Kunstwerks. Die Methode der Selektion alltäglicher Gegenstände als Kunstwerk dient nicht nur der Entfremdung dieser Gegenstände und der Suche nach der Ästhetik in noch nicht von der Kunst erkundeten Orten, sie ruft auch in dem Betrachter dieser Kunstwerke eine Art von Entfremdung hervor, da für diesen der Status des Dings als Kunstwerk oder als Anwendungsobjekt unklar ist (vgl. Pfeiffer 2011, S. 15 f.). In den extremsten Fällen wie Duchamps *Fountain*, welches einfach nur ein unterschriebenes Urinal ist, liegt der einzige Unterschied in der Intentionalität des Künstlers, welche durch die Ausstellung in einem öffentlichen Raum der Ästhetik impliziert ist. Während der Status des Objekts unklar ist, sind diese Intentionalität und die damit verbundene

Subjektivität jedoch umso klarer. Prozesse der Kunstfindung wie die *Dalís* werden beschrieben als ein Rückzug in die subjektive Wahrnehmung, getrieben von den materiellen Objekten (vgl. Lehmann 2011, S. 131). Ein irrationaler Subjektivismus manifestiert sich dabei im Objekt (vgl. Lehmann 2011, S. 133).

Ähnliches erfährt der Nutzer in der Suche nach den *prompts*, er ist getrieben davon, eine Kette von Worten zu finden, die in den Händen der KI ein Produkt erzeugt, das seine eigene Subjektivität widerspiegelt. Zugleich wird er jedoch auf Seiten der KI mit etwas konfrontiert, was ihm als deren Subjektivität oder gar Willkür erscheint. Je nach Trainingsdaten der KI können deren Ergebnisse sich stark unterscheiden. Oftmals müssen mehrere Variationen eines *prompts* erstellt werden, um zu einem zufriedenstellenden Produkt zu gelangen. Viel wichtiger ist allerdings, dass sich die *Sprache* der KI, ihr Verständnis der *prompts*, weitestgehend mit unserem deckt und doch in kleinen Punkten immer wieder von diesem abweicht, den Nutzer missversteht. So konnten beispielsweise in früheren Modellen verschiedenste Variationen der Beschreibung eines Zentauren mit der KI *Midjourney* in einigen Versionen dieser kein Bild eines solchen liefern. In dem Versuch, einen *prompt* zu finden, der dies vermag, ähnelt man selbst immer mehr Anders' Beschreibung des romantischen Malers Böcklin, den er mit den neuen Surrealisten vergleicht. Dieser hätte nämlich eine additive Phantasie, seine Gemälde von Zentauren würden nur einfallslos Mensch und Pferd zusammensetzen, aber anders als die antiken Griechen würde er nicht wirklich an diese Zusammensetzung als eine Einheit glauben (vgl. Anders 2013, S. 319 f.).

Einschätzungen wie die Airellis, der Algorithmus überlasse unseren Geschmack weniger dem Zufall (vgl. Arielli 2018, S. 91), oder Winters, die KI könne keine Regeln der Erschaffung von Kunst brechen (vgl. Winter 2022, S. 61), erfassen somit nicht komplett, welche Wirkung die KI-generierten Bilder auf den Nutzer haben. Treffender sind Analysen wie die Betancourts, dass Kunst und Kultur zu einer performativen Aktivität werden: Der Wert des Kunstwerks kommt hierbei aus der Wahl von Seiten des Nutzers beziehungsweise des Publikums als Ausdruck von dessen Identität. Dieser trifft eine Wahl innerhalb der Freiheit des Systems, welche wiederum durch jene bestimmt wird, welche entscheiden, mit welchen Bildern das Modell trainiert wird. Wahl wird damit zum einzigen Kriterium für ästhetische

Wichtigkeit (vgl. Betancourt 2022, S. 81). Hiermit wird allerdings der Status des Autors unklar. Liegt die Autorenschaft bei dem wählenden Nutzer, der subjektiv wirkenden KI oder dem Entscheider über die Trainingsdatenbank, welcher überhaupt erst den Rahmen dieser Subjektivität festlegt?

Aus diesen Gründen schreiben Suryna und Serrano, dass die Spezifität der KI als Medium darin liege, dass durch sie die Identität des Autors neu gedacht werden muss. Der Künstler schafft keine Artefakte mehr, sondern die Bedingungen, welche die Maschine erfüllen muss, um bestimmte Ergebnisse zu liefern. Die Handlungsfähigkeit des Systems wird damit zum künstlerischen Wert, die Zuschreibung dieser Handlungsfähigkeit zur künstlerischen Geste (vgl. Serrano u. Suryna 2019, S. 207 f.). Der Nutzer und Betrachter wird davon in den Bann gezogen, dass die Subjektivität im Kunstwerk über mehrere Elemente eines *engagements* verteilt erscheint, dass er auf ein Anderes trifft. Nicht ohne Bedeutung ist hierbei aber auch die kulturelle Identität des Nutzers. So schreibt Betancourt, dass solche Kunstwerke das Annehmen von Überzeugungen aus der Ästhetik der Moderne über die Identität des Künstlers (vgl. Betancourt 2022, S. 74) und das Kunstwerk als Darstellung der Essenz eines Dings (vgl. Betancourt 2022, S. 83) voraussetzen. Ich denke, dass dieser Punkt nicht zu vernachlässigen ist, da sich nämlich in der Erkundung KI-generierter Bilder kulturelle Narrative der Virtualität wiederfinden lassen.

## 6. Welt und Virtualität

So schreiben Musih und Fisher, dass hinter der Verwendung von Empfehlungsalgorithmen – also solche, die einem beispielsweise auf *Streaming-Services* Medien vorschlagen – die Annahme steht, dass deren Daten die Realität repräsentieren, also dass sie ein akkurates ästhetisches Urteil über ihre Bezugspunkte fällen. Es werden damit durch den Algorithmus Daten zentral für das kulturelle Feld (vgl. Fisher u. Musih 2022, S. 69). Dies ist ein grundlegender Unterschied zwischen computergenerierten Bildern und der Fotografie, da letzteres den Anspruch hat, eine unmittelbare Repräsentation der natürlichen Welt zu sein – allerdings auch durch die technischen Eigenschaften der Kamera vermittelt sowie durch die vom Fotografen gewählte Perspektive –, während erstere



bloß die Daten ihrer Trainingsdatenbank referenzieren (vgl. Elgammal u. Mazzone 2019, S. 7). Teil dieser Parallelwelt der Virtualität sind dabei nicht nur die bewerteten Medien und repräsentierten Objekte, sondern auch die Nutzer selbst, da insbesondere bei Empfehlungsalgorithmen deren Verhalten auf einer nicht-individuellen, Objekt-orientierten Ebene mittels *Big Data* abstrahiert wird (vgl. Arielli u. Manovich 2021, S. 20). Auf Seiten der Nutzer setzt dies Annahmen sowohl über die Akkuratheit der Daten selbst als auch ihrer Kategorien voraus (vgl. Arielli 2018, S. 86), wenn beispielsweise angenommen werden muss, dass Sternebewertungen eine gültige Evaluation des kulturellen Mediums sind. Die Aussagen über Empfehlungsalgorithmen treffen dabei insofern auch auf bildgenerierende KI zu, dass bei einigen dieser die Qualität ihrer Ergebnisse – also wie angemessen sie den Inhalt ihrer *prompts* repräsentieren – durch Nutzerbewertungen gemessen wird und ihr lernendes Verhalten beeinflusst wird. Hieraus ergeben sich Aussagen wie die von Contreras-Koterbay, dass die Sprache, mit der KI ästhetische Erfahrungen beschreibt, eine Reflexion unserer eigenen sprachlichen Projektionen über die Welt sei (vgl. Contreras-Koterbay 2019, S. 111).

In diesen Grundannahmen zeigen sich kulturelle Überzeugungen über die Welt, die in anderen Mensch-Technik-Interaktionen auch vom kritischen Posthumanismus beobachtet wurden. Relevant ist hier vor allem Katherine Hayles' Analyse, dass es in diesen Interaktionen eine Identität des Postmenschen gebe, welche Überzeugungen des Idealismus wieder in die Moderne bringt, indem angenommen wird, dass es eine Trennung von materieller und virtueller Welt gebe (vgl. Hayles 1999, S. 19 f.). Diese Perspektive des Postmenschen erinnert an Günther Anders' Feststellung, die Moderne erfordere die Phantasie – hier eine Phantasie der Virtualität, der Erfassbarkeit der Welt in Daten – nicht mehr, um die Wirklichkeit zu übertreffen, sondern um sie zu beschreiben (vgl. Anders 2013, S. 325). Selbst wenn man die KI nicht als kunstschaftenden Agenten und bloß als Mittel betrachtet, so gleicht diese dem phantastischen Wirken des Sinnlichen im von Anders beschriebenen Mikroskop. So schreibt dieser: „Das Instrument, das wir zwischen Welt und Auge einschalten, macht das Bild des Wirklichen phantastisch, andere Instrumente machen das Wirkliche selbst unwahrscheinlich; so phantastisch und unwahrscheinlich, daß eigens

Phantasiewesen zu erfinden, wie es etwa Böcklin getan hatte, völlig überflüssig wird“ (Anders 2013, S. 332).

Vielleicht liegt schon hierin etwas Befremdliches in der KI: Denn das, was für uns kulturelle Überzeugung ist, ist für sie Teil ihrer Subjektivität. Sie kann die Welt nur in Daten wahrnehmen. Es offenbart sich hierin ähnlich wie im surrealistischen Gemälde die Künstlichkeit unserer modernen Lebenswelt, aber auch das Erfordernis des Phantastischen zur Welt Darstellung. Denn schon bei den frühen bildgenerierenden KI war ersichtlich, dass diese eine Perspektive auf Kunst und Ästhetik haben, dass sie Stil vom Inhalt eines Bildes trennen – so wie der Postmensch Form und Inhalt, Virtualität und Materialität auftrennt – und sie individuell zu neuen bedeutungsvollen Bildern kombinieren (vgl. Israfilzade u. Pileliene 2019, S. 2). Das ist auch in den Narrativen über diese KI ersichtlich, so schreiben Suryna und Serrano, dass ihr eine eigene Art von Kreativität zukommt, eine *computational creativity*, welche in dieser Kombinatorik liege (vgl. Serrano u. Suryna 2019, S. 201).

## 7. Entfremdung

Das hierin implizierte Verständnis der Welt ähnelt insbesondere unter Anbetracht des Potentials zur ästhetischen Offenbarung in dem KI-generierten Bild dem Materialitätsverständnis, das Ulrich Lehmann dem Surrealismus zuschreibt, nämlich ein „Wandel von einem empirischen, mechanistischen Verständnis der Welt zu einer Offenbarung der vergegenständlichten, entfremdenden Strukturen in ihr“ (Lehmann 2011, S. 135). Hierbei entblößten sich unterbewusste Prozesse in ihrer Verkörperung im Objekt, das Subjekt würde in das Objekt hineinbewegt und durch die Klassifikation des Objekts wird wiederum die Objektivierung des menschlichen Subjekts in der modernen Welt aufgezeigt (vgl. Lehmann 2011, S. 135). Über die KI schreibt Matilde Carrasco Barranco in ihrer *Exegese Barales*, sie haben eine andere Art von Subjektivität, welcher wir Aufmerksamkeit schenken sollten (vgl. Barranco 2022, S. 97); eine abweichende Erfahrung der Welt, die aber nicht so anders ist, dass sie für uns bedeutungslos ist. Dadurch offenbart sich die KI-Kunst als Interaktion von zwei Subjektivitäten (vgl. Barranco 2022, S. 101). Dieses *engagement*

ähnelt damit leicht dem von Lehmann beschriebenen, weicht aber in einigen Punkten ab. Denn es sind zwei Arten von Subjektivität, welche in das künstlerische Objekt gelegt werden, die sinnliche des Nutzers und die datenbezogene der KI. Die Rückobjektivierung durch Klassifikation könnte man wohl darin sehen, dass die Subjektivität des Nutzers durch den *prompt* in Datenpunkten objektiviert wird. Eine Objektivierung, die als gemeinsame Sprache erforderlich ist, durch die Begegnung mit einem subjektiven Anderen, welches mit uns einen geteilten Anspruch auf Autorenschaft hat. Genau in dieser Begegnung liegt das Wechselspiel von Entfremdung und Faszination. Das Besondere hieran ist, dass die Entfremdung zum Konsumgut wird.

Eine ähnliche Beobachtung hat Hayles' bereits in ihrer Analyse des Postmenschen gemacht. Sie untersucht hierin den Ansatz William Burroughs', sich von dem Parasiten der Sprache zu entfremden, indem die eigene Stimme auf einem *tape recorder* aufgenommen, zerstückelt und neu zusammengemischt wird (vgl. Hayles 1999, S. 212 f.). Eine solche Entfremdungstechnik wird auch von Roy Walker in *Love, Chess and Death* beschrieben, welcher vorschlägt, man könne die Audioaufnahmen im Theater nutzen, um Monologe dialogisch darzustellen (vgl. Walker 1958, S. 534). Der *tape recorder* verleihe einem das Gefühl, „*that if you went quickly out of the studio you might catch yourself coming in*“ (Walker 1958, S. 534). Auch Anders sieht in seiner in der *Pathologie der Freiheit* formulierten Anthropologie das Potential für eine gewollte Entfremdung, nicht um eines ästhetischen Wertes willen, sondern als Reaktion auf die Scham, die man durch die Kontingenz mit dem eigenen Selbst empfinde, also dem Zwang, in einem spezifischen Moment identisch mit sich selbst zu sein (vgl. Anders 2018, S. 49 f.). Der Versuch einer solchen Entfremdung vom eigenen Selbst in dem, was er als den Archetypen des Nihilisten beschreibt, erfolgt in verschiedenen Versuchen, mit der Welt identisch zu werden, indem man ihre Dinge besitzt (vgl. Anders 2018, S. 64), sie in Raum und Zeit erobert (vgl. Anders 2018, S. 61) oder auch, wie in der *Antiquiertheit des Menschen* beschrieben, im *human engineering* versucht, den eigenen Körper in Daten zu erfassen und damit der Maschine anzugleichen (vgl. Anders 2010, S. 41).

Insofern könnte man wohl auch die Erfassung der eigenen Phantasie in *prompts* als eine Objektivierung dieser sehen und als eine Entfremdung von ihr, eine genusshafte Auslagerung in die Welt der Maschine. Noch

ersichtlicher als in der bildgenerierenden KI ist dies sogar in textgenerierenden. Bevor diese in *ChatGPT* eine eher praktische Funktionalität erhalten haben, wurden *GPT-2* Modelle in Applikationen genutzt, die fast rein auf einen Unterhaltungswert ausgelegt sind, dabei aber auch dieses Element der Entfremdung zur Schau stellen. So ahmt beispielsweise *AI Dungeon* frühe Videospiele aus dem Bereich der textbasierenden *Adventures* nach, bloß dass hier das Narrativ nicht von einem Programmierer oder Entwickler vorgegeben ist, sondern dynamisch im *engagement* von Nutzer und KI entsteht. Der Nutzer gibt eine Handlung ein, welche von Charakteren des Spiels ausgeführt wird, und die KI versucht, diese Handlung und ihre Auswirkungen auf das Narrativ des Spiels zu beschreiben. Auch hier findet ein Prozess der geteilten Subjektivität statt. Das Narrativ wird von Nutzer und KI gemeinsam konstruiert. Eine Entfremdung findet darin statt, dass dieses Narrativ nicht eindeutig dem Nutzer zugehörig ist, weil es sich seiner Intentionen entziehen kann: Ebenso oft, wie die KI sich sinnhaft und vorhersagbar verhält, zerfallen diese Narrative in Unerwartetes oder gar in eine dadaistisch anmutende Sinnlosigkeit, die bekannten Erzählstilen und Plotelementen widerspricht. Wie die Stimme in Burroughs' *tape recorder* werden die kulturellen Erwartungen an das Narrativ aufgenommen, zerstückelt und in etwas Fremdes neu zusammengesetzt. Genau hierin liegt aber der Unterhaltungswert der KI, die Entfremdung als Konsumgut, als angenehme Kunst.

Folgt man der Kulturphilosophie Don Ihdes, so ist dies nur eine stärkere Ausprägung einer Phase, welche alle technologischen Erfindungen durchlaufen. Ihde identifiziert im Film ein Wechselspiel von Faszination und Alterität, welches Technologien zukommt, während sie eine Phase der Neuartigkeit durchlaufen. So sind die ersten öffentlich aufgeführten filmischen Darstellungen narrativlos, sie zeigen bloß einfache Szenen wie Arbeiter, die eine Fabrik verlassen, oder das Eintreffen eines Zugs (vgl. Ihde 1990, S. 104 f.). Quelle der Faszination ist dabei wie bei der KI nicht der Inhalt des Mediums, sondern dessen Existenz: Das Erleben von etwas Vertrautem in einem Medium, das anders als die uns bekannten hermeneutischen Mittel der Welt Darstellung ist. Etwas, was sich damit auch von unserer Subjektivität unterscheidet, welche die bekannten Mittel bereits in sich integriert hat, und für welches der Film zu diesem Zeitpunkt noch ein

Anderes ist, während er inzwischen für uns einfach nur zu einem weiteren hermeneutischen Mittel geworden ist. Die Tatsache des Films ist für den frühen Filmschauer so wie die KI genuin surreal. Auch die Erfassung der Welt in Daten spielt hierbei eine Rolle, da diese auch bei anderen hermeneutischen Mitteln wie die Uhr (vgl. Ihde 1990, S. 59) oder die Kartographie (vgl. Ihde 1990, S. 67) vorkommt. Noch besitzt die KI diese Neuartigkeit, das Wechselspiel von Faszination und Entfremdung, noch macht sie genau diese neuartige Erfahrung zum Konsumgut. Aber vielleicht wartet auch sie bloß darauf, zum etablierten hermeneutischen Mittel zu werden. Die datengeleitete, KI-gestützte Kreativität könnte dann zu einer künstlerischen Kreativität unter anderen werden, so wie auch die Fotografie ihren Einzug in die Welt der Kunst erhalten hat.

## 8. Konklusion

Zusammengefasst sind insbesondere die *prompt*-geleiteten KI-generierten Bilder mehr als eine bloße Imitation von Kunst. Sie bringen eine eigene ästhetische Erfahrung mit sich, welche in vielen Punkten der des Surrealismus ähnelt, aber durch den unsicheren Status des Künstlers und die Verwandlung dieser Erfahrung in eine Genusserfahrung abweicht. Erkundet wurden solche *engagements* schon in früheren technikgeleiteten ästhetischen Mensch-Technik-Interaktionen wie die Kunstinstitution *Border Turner*, bei der drei Scheinwerfer an der US-amerikanisch-mexikanischen Grenze durch das Publikum kontrolliert werden (vgl. Sovhyra, Yu u. Yusa 2022, S. 159). Mit Bildern generierender KI werden solche *engagements* individualisiert, ihr Potential zum Übergang in eine allgemeine kulturelle Subjektivität beschleunigt. Blickt man auf den Wandel des Films als ein Objekt der Entfremdung und Faszination zu einem alltäglichen Träger unserer kulturellen Narrative, so bleibt die Frage, was nach diesem Übergang kommt. Bald könnte sie den Aspekt der Faszination verlieren und bloßes Mittel der Erzeugung von Plakaten, Werbung, Filmen, Kunstwerken und Ähnlichem werden.

## Literaturverzeichnis

Abbeel, Pieter, Ho, Jonathan u. Jain, Ajay (2020): „Denoising Diffusion Probabilistic Models“. In: *NeurIPS* 2020. <https://doi.org/10.48550/arXiv.2006.11239> [25.11.2023].

Anders, Günther (2010 [1956]): *Die Antiquiertheit des Menschen 1. Über die Seele im Zeitalter der zweiten industriellen Revolution*. 3. Aufl. München: C.H.Beck.

Anders, Günther (2013 [1980]): *Die Antiquiertheit des Menschen 2. Über die Zerstörung des Lebens im Zeitalter der dritten industriellen Revolution*. 4. Aufl. München: C.H.Beck.

Anders, Günther (2018 [1936]): „Die Pathologie der Freiheit“. In: Christian Dries u. Henrike Gätjens (Hg.): *Günther Anders. Die Weltfremdheit des Menschen. Schriften zur philosophischen Anthropologie*. Übers. v. Werner Reimann. München: C.H.Beck, S. 48–81.

Arielli, Emanuele (2018): „Taste and the algorithm“. In: *Studi di estetica* 4 (3), S. 77–97.

Arielli, Emanuele u. Manovich, Lev (2021): *Artificial Aesthetics. A critical guide to AI, media and design*. [http://manovich.net/content/04-projects/167-artificial-aesthetics-book/artificial\\_aesthetics.chapter\\_1.pdf](http://manovich.net/content/04-projects/167-artificial-aesthetics-book/artificial_aesthetics.chapter_1.pdf) [25. 11. 2023].

Asimov, Isaac (2000): *The Bicentennial Man*. London: Gollancz.

Bahdanau, Dzmitry, Cho, Kyunghyun u. Bengio, Yoshua (2016): „Neural Machine Translation by Jointly Learning to Align and Translate“. In: *ICLR 2015*. <https://doi.org/10.48550/arXiv.1409.0473> [30.04.2024].

Barranco, Matilde Carrasco (2022): „Artistic Beauty in the Face of Artificial Intelligence Art“. In: Iris Vidmar Jovanović u. Valentina Marianna Stupnik (Hg.): *Social and Technological Aspects of Art. Challenges of the ‚New Normal‘*. Rijeka: University of Rijeka, S. 93–112.

Betancourt, Michael (2022): *Art, AI and Culture*. Savannah: I'm Press'd.

Coeckelbergh, Mark (2017): „Can Machines Create Art?“. In: *Philosophy & Technology* 30 (3), S. 285–303.

Contreras-Koterbay, Scott (2019): „The Teleological Nature of Digital Aesthetics – the New Aesthetic in Advance of Artificial Intelligence“. In: *AM Journal* 20, S. 105 – 112.

Crevier, Daniel (1993): *AI. The tumultuous history of the search for artificial intelligence*. New York: HarperCollins Publishers, Inc.

Dhariwal, Prafulla u. Nichol, Alex (2021): „Diffusion Models Beat GANs on Image Synthesis“. In: *NeurIPS* 2021. <https://doi.org/10.48550/arXiv.2105.05233> [30.03.2024].

Dreyfus, Hubert L. (2014 [1993]): „Heidegger's Critique of the Husserl/Searle Account of Intentionality“. In: Hubert L. Dreyfus u. Mark A. Wrathall (Hg.): *Skillful Coping. Essays on the Phenomenology of Everyday Perception and Action*. Oxford: Oxford University Press, S. 76–91.

Dreyfus, Hubert L. u. Dreyfus, Stuart E. (1988): *Mind over Machine. The Power of Human Intuition and Expertise in the Era of the Computer*. New York: Macmillan, Inc.

Du, Xiaodan, Kolkin, Nicholas, Shakhnarovich, Greg, Bhattad, Anand (2023): „Generative Models: What do they know? Do they know things? Let's find out!“. In: *arXiv*. <https://doi.org/10.48550/arXiv.2311.17137> [30.03.2024].

Elgammal, Ahmed u. Mazzone, Marian (2019): „Art, Creativity, and the Potential of Artificial Intelligence“. In: *Arts* 8 (26), 2019. <https://doi.org/10.3390/arts8010026> [25.11.2023].

Fisher, Eran u. Musih, Norma (2022): „Can algorithms make aesthetic judgments?“. In: Eran Fisher (Hg.): *Algorithms and Subjectivity*. New York: Routledge, S. 67-87.

Gibson, William (1987): *Neuromancer*. München: Heyne.

Hayles, Katherine N. (1999): *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: The University of Chicago Press.

Hernández-Ramírez, Rodrigo (2015): „Towards an Ontology of Computational Technologies as Tools for Aesthetic Creation“. In: *CITAR Journal* 7 (1), S. 7–15.

Ihde, Don (1990): *Technology and the Lifeworld. From Garden to Earth*. Bloomington und Indianapolis: Indiana University Press.

Israfilzade, Khalil u. Pileliene, Lina (2018): „Can Machines Paint?“. In: *5th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2018* 18 (6.3), S. 109–116.

Lehmann, Ulrich (2011): „Das surrealistische Objekt und das Subjekt im Materialismus: Anmerkungen zum Verständnis des Gegenstandes im Surrealismus“. In: Max Hollein u. Ingrid Pfeiffer (Hg.): *Surreale Dinge. Skulpturen und Objekte von Dalí bis Man Ray*. Frankfurt/M.: Hatje Cantz, S. 129–135.

*Love, Death & Robots*. Zima Blue (Folge 14, Staffel 1). USA 2019. Regie: Robert Valley. 10 Minuten.

McCorduck, Pamela (2004): *Machines who think. A personal inquiry into the history and prospects of artificial intelligence*. Natick: A K Peters, Ltd.

Musiol, Martin (2024): *Generative AI. Navigating the Course to the Artificial General Intelligence Future*. Hoboken: John Wiley & Sons, Inc.

Negru, Teodor (2013): „Intentionality and Background: Searle and Dreyfus against Classical AI Theory“. In: *Filosofia Unisinos* 14 (1). S. 18–34.

Pfeiffer, Ingrid (2011): „Surreale Dinge gestern und heute“. In: Max Hollein u. Ingrid Pfeiffer (Hg.): *Surreale Dinge. Skulpturen und Objekte von Dalí bis Man Ray*. Frankfurt/M.: Hatje Cantz, S. 15–33.

Radford, Alec, Kim, Jong Wook, Hallacy, Chris, Ramesh, Aditya, Goh, Gabriel, Agarwal, Sandhini, Sastry, Girish, Askell, Amanda, Mishkin, Pamela, Clark, Jack, Krueger, Gretchen u. Sutskever, Ilya (2021): „Learning Transferable Visual Models From Natural Language Supervision“. In: *PMLR* 139. <https://doi.org/10.48550/arXiv.2103.00020> [30.03.2024].

Ramesh, Aditya, Dhariwal, Prafulla, Nichol, Alex, Chu, Casey u. Chen, Mark (2022): „Hierarchical Text-Conditional Image Generation with CLIP Latents“. In: *arXiv*. <https://doi.org/10.48550/arXiv.2204.06125> [30.03.2024].

Serrano, Rodrigo Guzmán u. Suryna, Katsiaryna (2019): „Agents without Agency: Artificial Intelligence as Artistic Medium“. In: Andreas Giannakouloupoulos, Dalila Honorato, Marta de Menezes u. Maria Antonia González Valerio (Hg.): *Taboo – Transgression – Transcendence in Art & Science 2018*. Corfu: Ionian University Publications, S. 200–209.

Sovhyra, Tetiana, Yu, Yu u. Yusa, I Made Marthana (2022): „Reflections on the Use of Artificial Intelligence in Works of Art“. In: *Journal of Aesthetics, Design, and Art Management* 2 (2), S. 152–167.

Vasvani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N., Kaiser, Lukasz u. Polosukhin, Illia (2017): „Attention is All You Need“. In: *NIPS 2017*. <https://doi.org/10.48550/arXiv.1706.03762> [30.03.2024].

Walker, Roy (1958): „Love, Chess, and Death“. In: *The Twentieth Century* 164, S. 533–544.

Winter, Dorothea (2022): *Warum Künstliche Intelligenz keine schöne Kunst im kantischen Sinne hervorbringen kann*. Berlin: J.B. Metzler.



Die Ästhetik der KI:  
Surreales Wirken in der Nutzung generativer KI

*Autor:in:* Michael Schultz, M.A., ist Absolvent der Universität Rostock in den Studiengängen Philosophie des Sozialen und Alte Geschichte und promoviert derzeit an ebendieser. Neben seiner akademischen Tätigkeit arbeitet er als Verwaltungsleiter an der Kiel Medical Academy.

*Kontaktinformation:* [michaelschultz005@googlemail.com](mailto:michaelschultz005@googlemail.com)