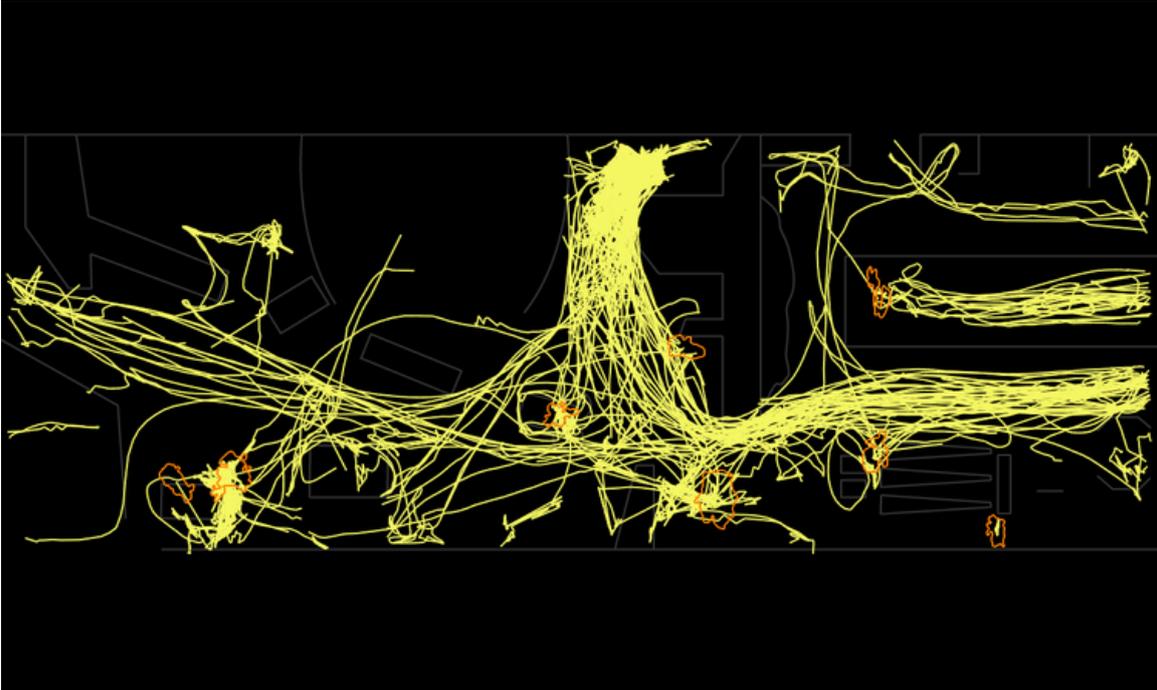


Roberto Simanowski

Kunst aus dem Computer? Die Rolle der Künstler im digitalen Zeitalter

Keynote Führungskräfte-Klausur NÖKU am 18. November 2020



Abstract:

Digitale Kunst lässt sich so programmieren, dass sie mit dem Publikum interagiert, dass sie Daten aus dem Internet sammelt oder dass sie selbst Texte, Bilder und Musik produziert. In allen drei Varianten tritt der Künstler als Produzentin eines bedeutungsvollen Werks mehr oder weniger in den Hintergrund. Der Vortrag ordnet diese Entthronung des künstlerischen Subjekts in die Kunstgeschichte und ästhetische Theoriebildung ein. Ausgangspunkt ist der vor kurzem für Aufsehen sorgende Textgenerator GPT-3, dessen Texte kaum noch von denen eines menschlichen Autors unterscheidbar sind. Der Vortrag bespricht Formen der analogen und digitalen Zufallskunst, erörtert die Desemantisierung der Kunst und das Modell der Präsenzkultur als Prozess der Musikalisierung und diskutiert die philosophischen und politischen Konsequenzen von Texten, die jenseits einer sinnstiftenden menschlichen Autorschaft oder Kontrolle generiert werden.

Aufbau

1. Text-KI GPT-3
2. Geteilte Autorschaft
3. Entthronte Künstler
4. Musikalisierte Kunst
5. Textmaschinen im Turing-Test
6. Orakel der Wahrscheinlichkeit
7. Politische Mathematik

1. Text-KI GPT-3

Am 8. September 2020 wurden im Guardian die Menschen versichert, dass die künstliche Intelligenz ihnen nicht böse gesinnt ist, sondern ihre Mission darin sieht, ihnen zu dienen. Diese Auskunft ist deswegen besonders interessant, weil sie in der ersten Person erfolgt. Es ist die künstliche Intelligenz selbst, die spricht, oder zumindest einer ihrer Repräsentanten, der *Generative Pre-trained Transformer*, kurz GPT-3 genannt.ⁱ Dieser GPT-3 des Unternehmens OpenAI hat den Artikel geschrieben und relativiert die erwähnte Beruhigung sofort damit, dass das Böse nicht von der KI ausgehen werde, sondern von ihren Programmierern: „I will be programmed by humans to pursue misguided human goals and humans make mistakes that may cause me to inflict casualties.“

Was für ein großartiger Satz! Die KI warnt uns davor, dass sie von bösen Menschen missbraucht werden könne. Sie hat also einen Willen – den, uns zu warnen – und ist zugleich willenlos ihrer Programmierung ausgesetzt. Abgesehen von der Frage, wie sich dieses Oxymoron erklären lässt, stellt sich die Frage, was uns eigentlich lieber wäre: Eine willenlose KI oder eine, die zu ihren eigenen Erkenntnissen und Entscheidungen kommt? In der Debatte zur KI gibt es beide Optionen, wobei die zweite zumeist nicht darauf hinausläuft, dass die KI sich gegen ihre bösen Programmierer wendet, sondern gegen den Menschen an sich.

Was den Guardian-Artikel betrifft, so hörte man bald, dass dies nicht wirklich ein computer-generierter Text sei, sondern nicht mehr als ein Update zur computer-assistierten Authorschaft.ⁱⁱ Ist GPT-3 nur ein fortgeschrittener Spell-Check? Das klingt sehr nach jener altbekannten Beschwichtigung, die immer nur eine neue Quantität feststellt, diese aber nie in eine neue Qualität umschlagen sieht. Aber der Einspruch hatte gute Argumente auf seiner Seite: Menschen gaben das Thema, die Richtung (*Focus on why humans have nothing to fear from AI.*) sowie den ersten Absatz des Artikels vor und sie stellten aus den acht verschiedenen Ergebnissen den Text zusammen, der dann im Guardian präsentiert wurde. Ohne die acht Originalergebnisse zu kennen, lässt sich also wenig darüber sagen, wie kohärent der GPT-3 tatsächlich denken kann.ⁱⁱⁱ

Zweifelhaft ist auf jeden Fall, ob „denken“ überhaupt das richtige Wort wäre. GPT-3 ist ja eine Art künstliches neuronales Netzwerk, das anders als unser Gehirn Information nicht philosophisch, sondern mathematisch verarbeitet. Es weiß aus Hunderten von Gigabytes an Text, welches Wort am häufigsten nach einem vorangegangenen Wort erscheint, so wie Googles Übersetzungsalgorithmus von vielen Gigabytes Text weiß, wie oft Menschen für ein bestimmtes englisches Wort ein bestimmtes deutsches Wort eingesetzt haben. Das Wort, das die Liste anführt, ist *das* Wort, das der Übersetzungsalgorithmus nehmen wird, ohne das geringste Verständnis dafür, was der englische oder deutsche Begriff eigentlich bedeutet. Im künstlichen neuronalen Netzwerk ist Wissen nicht mehr als Statistik. Also ist ein gut geschriebener Text von GPT-3 nie ein gut *durchdachter*, sondern immer nur ein probabilistisch gut gemachter Text.



Portrait of Edmond de Belamy (2018)



Dieter Grossmann (2006)

Oder es ist ein probabilistisch präzises Bild. Denn Computer schreiben nicht nur, sie malen auch. Erinnerung sei an das Gemälde "Portrait of Edmond de Belamy", das am 25. Oktober 2018 vom Auktionshaus Christie's für 432.500 US-Dollar versteigert wurde. Auch in diesem Falle ist der Urheber ein Algorithmus, dem 15.000 Porträts der Kunstgeschichte zwischen seit dem 14. Jahrhundert eingespeist wurden, aus denen dann das Portrait des Edmond de Belamy gewissermaßen als Mittelwert entstand.^{iv} Es war natürlich nicht das erste Bild aus dem Computer. Verwiesen sei auf die hypnotisierenden *Computer Paintings* des deutschen Dieter Grossmann und auf die Bilder die das Computerprogramm Aaron des Briten Harold Cohen seit 1973 produzierte und die an ehrwürdigen Orten wie der Tate Gallery und dem MOMA San Francisco ausgestellt wurden. Cohens ironische Auskunft, der erste Künstler der Welt zu sein, von dem es eine post-human Ausstellung neuer Werke geben könnte,^v verweist auf das zugrundeliegende Konzept der geteilten Autorschaft: Der Mensch gibt dem Computer bestimmte Regeln vor, aus denen dieser immer wieder neue Werke produziert.



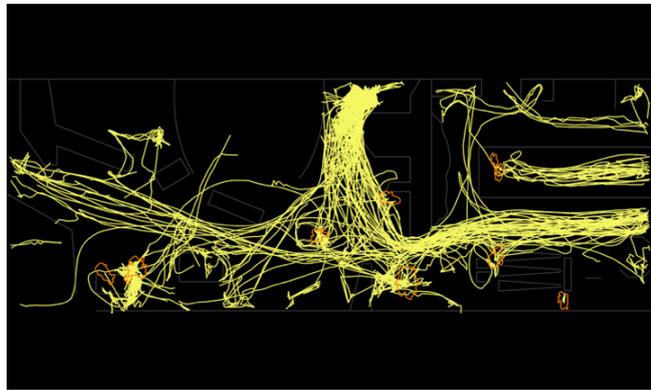
Harold Cohen (2004)

2. Geteilte Autorschaft

Die geteilte Autorschaft kann auch das Publikum einbeziehen, wie das Beispiel *Untitled 5* von Camille Utterback aus dem Jahr 2004 zeigt.^{vi} Hier verändert die von Sensoren registrierte Bewegung des Publikums im Raum die abstrakte Form auf dem als Leinwand getarnten Bildschirm. Utterback programmierte das Feedbacksystem so, dass die Bewegung der Betrachter bzw. Interakteure im Raum nicht direkt als Bewegung im Bild erkennbar wird, sondern durch eine eingebaute Verzögerung und Modulation mysteriös bleibt. Dieses „system of inexact control“, wie David Rokeby, ein Pionier dieser Art von Interaktionssystemen, es nennt, vermittelt ein Gefühl von Unmittelbarkeit und Ungewissheit zugleich. Rokeby sieht genau darin die künstlerische Botschaft: Während wir gerade mit Blick auf den Computer Kontrolle fetischisieren, machen wir hier, im Dialog mit dem Computer, die Erfahrung, dass es absolute Kontrolle nicht gibt.^{vii} Diese Erfahrung ist die Botschaft des Künstlers jenseits aller unvorhersagbaren visuellen oder akustischen Resultate des Interaktionssystem produziert. Denn es ist der Künstler, der diese so und nicht anders programmiert hat.



Untitled 5 (2005)

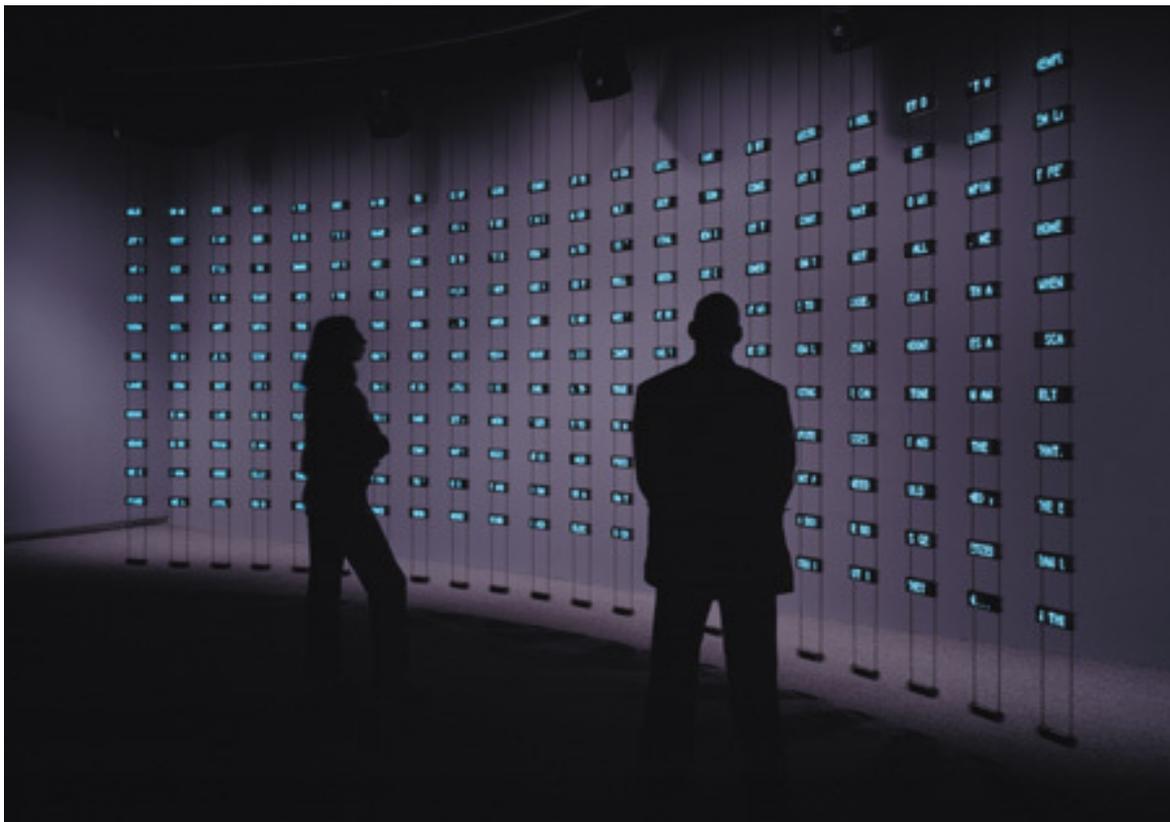


You Are Here (2004)

Eine solche Botschaft des Künstlers jenseits der Zufallsform seines Werkes lässt sich auch einem weiteren ‚Gemälde‘ aus dem Computer entnehmen: Scott Snibbes *You Are Here*.^{viii} Auch hier handelt es sich um ein scheinbar abstraktes Gebilde, das nichts anderes zu repräsentieren scheint als sich selbst. Das stimmt hier jedoch noch weniger als im Falle von Utterbacks *Untitled 5*. Die gelben Linien auf schwarzem Grund repräsentieren die Bewegung des Publikums einer Ausstellung durch diese, aufgenommen durch Kameras an der Decke der Ausstellungsräume. Anders als bei Utterback ändern die Betrachter nicht die Details des Bildes, sondern schaffen selbst das Bild, und anders als bei Utterback geschieht dies nicht in der Interaktion mit dem Bild, sondern in der Interaktion mit den anderen Ausstellungsobjekten. Am Ende der Ausstellung erwartet die Besucher dann Snibbes *You Are Here*, gewissermaßen der Fußabdruck ihres Erlebens der Ausstellung. Das Bild ist ohne Frage ästhetisch interessant, seine zentrale Bedeutung, die Botschaft des Künstlers, liegt aber nicht in seiner spezifischen Form, sondern in seiner Existenz *an sich*: darin, dass es Überwachung repräsentiert. Wie faszinierend auch das visuelle Ergebnis erscheinen mag, es sollte eigentlich alarmierend sein, denn es konfrontiert seine Betrachter mit dem Vorschein einer beunruhigenden Zukunft.

Kehren wir zurück zu unserem Ausgangspunkt, zum computergenerierten Text und Bild. Die Einordnung dieser aktuellen Phänomene in die Kunstgeschichte verweist uns auf die Zeit vor hundert Jahren. Es war 1917, als Marcel Duchamp ein Urinal

signierte und als Kunst ausgab und damit die Frage der Autorschaft prinzipiell in Frage stellte. Auch dazu gibt es ein Update im digitalen Reich, wenn Texte aus sozialen Netzwerken nach bestimmten Kriterien kopiert und in einer neuen Anordnung als Kunst präsentiert werden. Ein berühmtes Beispiel für diese *kommunikativen Readymades* ist *Listening Post* (2000/01) von Mark Hansen und Ben Rubin, das auf 231 LED-Screens Sätze aus dem Internet präsentiert die mit „I am“ anfangen.^{ix} Natürlich sind die Programmierer dieser Installation nicht die Autoren der Texte. Aber insofern sie diesen Texten im künstlerischen Kontext eine neue Bedeutung geben, sind sie so sehr die Künstler des Werks wie Duchamp der Künstler des Urinals. Und da sie zugleich ein Arrangement dieser Texte mit hohem ästhetischem Mehrwert schaffen, sind sie es sogar mehr als Duchamp.



Listening Post (2000/01)

Duchamp ist hier nicht nur interessant wegen seiner Vorarbeit für das, was heute Mapping-Art oder Appropriation-Kunst heißt. Duchamp ist auch ein Pionier der Zufallskunst, die konzeptionell dem computergenerierten Text eng verwandt ist. Berühmt ist sein Bild *3 Stoppages étalon* (Die 3 Musterfäden) von 1913, drei weiße Fäden auf schwarzem Grund, die Duchamp aus einem Meter Höhe auf die Leinwand hatten fallen lassen. Das Ergebnis passte bestens zum non-figurativen Stil jener Zeit, der es zugleich in produktionsästhetischer Hinsicht weit voraus war. Im gleichen Jahr zerschnitt Duchamp Noten und kombinierte sie, aus einem Hut gezogen, neu zusammen zu dem Stück „Erratum Musical“. Der Dadaist Tristan Tzara empfiehlt 1920 in seinem Gedicht *Um ein dadaistisches Gedicht zu machen*, genau dies mit den Wörtern aus einem Zeitungsartikel zu tun, wobei er den Hut durch eine Tüte ersetzt, was freilich wenig an der Bedeutung des Endergebnisses ändert.

Auch die Tüte wurde bald ersetzt. Was Duchamp und Tzara propagierten, lässt sich viel leichter mit einem Computer ausführen. Berühmt sind die Gedichte, die Theo Lutz 1959 mit dem Rechner Zuse 22 herstellte. Der Wortschatz entstammte Kafkas Roman *Das Schloss*, die syntaktische Regel gab Lutz vor: jeder Satz hatte genau vier Positionen: ein Quantor, ein Substantiv, das Verb „ist“ und ein Adjektiv. Jeder Vers enthielt zwei Sätze, gelegentlich verbunden durch die Konjunktion „und“. Das ergab Verse wie diese:

Ein Schloss ist frei und jeder Bauer ist fern.
Jeder Fremde ist fern: Ein Tag ist spät.
Jedes Haus ist dunkel: ein Auge ist tief.
Nicht jedes Schloss ist alt. Jeder Tag ist alt.

Auch Lutz veröffentlichte nur eine Auswahl wählte für die Veröffentlichung nur ausgewählte Sätze aus den vom Computer produzierten Sätzen aus, nahm also wie die Redakteure des Guardian beim GPT-3-Artikel eine gewisse editorische Bearbeitung vor. Dass auf diese Weise Sätze von geradezu existenzialistischer Wucht entstehen können, zeigt der letzte der hier zitierten: „Jeder Tag ist alt.“ Es wäre absurd Lutz als Autor dieses Satzes zu sehen, nur weil er dessen syntaktisches Modell vorgegeben hat. Aber ist es wichtig? Spricht der Satz nicht auch als Zufallsprodukt aus dem Computer zu uns? Ehe wir der Frage nachgehen, was von Kommunikationsangeboten zu halten ist, die keinen menschlichen Autor haben, ist zu klären, warum Künstler eigentlich so begierig, auf ihre Rolle als Autor zu verzichten.

3. Entthronte Künstler

Im Jahr 1966 warb der britische Künstler und Kunsttheoretiker Roy Ascott in seinem Aufsatz *Behaviourist Art and the Cybernetic Vision* für das, was einige New Yorker Künstler in einer Reihe von Performances in New York bereits praktizierten: für Kunst, die in der Aktion mit dem Publikum entsteht. Berühmt ist John Cages Komposition *4:33*, die aus den Geräuschen besteht, die des Publikums macht, während es darauf wartet, dass der Pianist endlich anfängt zu spielen. Ascott radikalisiert die Beteiligung des Publikums zu einer Entmachtung des Künstlers, der nun nicht mehr dem Publikum überlegen sei, indem er diesem Rätsel vorlegt, die es lösen muss, sondern vielmehr eine Kommunikationssituation schafft, in der auch das Publikum etwas zu sagen hat.^x

Was steckt hinter dieser Entmachtung des Künstlers als kreatives, sinnstiftendes Subjekt – sei es zugunsten des Zufalls oder des Publikums? Ascott sieht darin eine Demokratisierung der Kunst, Cage eher eine Abkehr von der neuzeitlichen „Absolutmachung des Menschen“^{xi} zumindest im Bereich der Kunst. Die Ästhetik des Zufalls antwortet der traditionellen Funktion der Kunst als „Beliebigkeitsersparung durch Form“^{xii} mit dem Bekenntnis zum Unberechenbaren und Schicksalshaften. Sie reagiert, als Avantgarde der Gesellschaft, auf die Midlife-Krise der Menschheit, die sich die Welt erfolgreich untertan machte und doch das Glück nicht gefunden hat. Es ist kein Zufall, dass viele Vertreter der Zufallskunst (John Cage, Merce Cunningham, Robert Filliou) sich für den Buddhismus interessierten,^{xiii} der es vorzieht, sich mit der Welt zu vermählen, statt sie zu beherrschen. Die Entmachtung des Künstlers repräsentiert eine Sehnsucht, der schon Rilke als Erfahrung des modernen Menschen schmerzhaft Ausdruck gab:^{xiv}

Und ich gehe und ich weiß nicht weiter,
ich vergaß, was ich zu sagen kam;
alles will von mir ich soll ein Streiter
werden und ich bin ein Bräutigam.

Die Entmachtung des Künstlers als Sinnproduzent, die Ascott in den 1960er Jahren forderte, wurde am Ende des Jahrhunderts wieder aufgegriffen. Der französische Kurator und Kunstkritiker Nicolas Bourriaud propagiert in seiner Schrift *Esthétique relationnel* 1998 eine Kunst, die nicht auf eine künstlerische Aussage zielt, sondern auf die Schaffung von Momenten der Geselligkeit.^{xv} Eines von Bourriauds Lieblingsbeispielen sind die Koch-Performances von Rirkrit Tiravanija, in denen der Künstler Speisen und Getränke ans Publikum verteilt, das wie auf einer Party miteinander ins Gespräch kommt. Wer miteinander spricht und worüber gesprochen wird, das überlässt der Künstler völlig dem Zufall. Seine Aufgabe beschränkt sich darauf, die Menschen zusammenzubringen.

Rirkrit Tiravanija ersetzt den Zufall, der bei Duchamp die Form des Kunstwerks bestimmte, durch das Publikum. Genau dies ist ein zentrales Merkmal der interaktiven Installationen als Genre der digitalen Kunst. Zwar kann man im Kontext der Digitalisierung auf der Programmebene nichts dem Zufall überlassen. Aber am Frontend des Interface, da, wo das Werk aufs Publikum trifft, bietet es sich geradezu an. Utterbacks *Untitled 5* ist ein Beispiel dafür, ein weiteres ist Scott Snibbes *Deep Wall* (2003), wo die Interakteure vor einer Leinwand aus 16 Videosegmenten mit ihrer Silhouette ein eigenes Video erzeugen, das dann das älteste der 16 Videos ersetzt.^{xvi} Der Künstler kann nicht vorhersehen oder bestimmen, wie die Interakteure sich verhalten und was die 16 Videos seines Werkes zu irgendeinem Zeitpunkt zeigen. Das Publikum ist der Zufall, dem sich die Künstler aussetzen. Ein Zufall freilich, der letztlich sozial determiniert ist, also auch als Spiegel der Gesellschaft verstanden werden kann.^{xvii}



Deep Walls (2002)

In all diesen Ästhetiken des Dialogs und der Interaktion geht es um die Intensivierung des Hier und Jetzt, um „Momenten der Intensität“ wie der Literaturwissenschaftler Hans Ulrich Gumbrecht sagt, der zeitgleich zu Bourriaud für einen Paradigmenwechsel von der Sinnkultur zur Präsenzkultur warb. Diese Momente der Intensität haben „nichts Erbauliches, keine Botschaft, nichts, was man aus ihnen lernen könnte“.^{xviii} Sie sind das Gegenteil von einer Kunst, die verstanden werden will.

Gumbrecht wiederum repräsentiert eine breite Bewegung in der Kunsttheorie um die Jahrhundertwende, die auf eine Ästhetik des Performativen zielte. Es war eine Bewegung weg vom „semiotischen Körper“ hin zum „phänomenalen sinnlichen Leib“,^{xix} „vom Werk zum Ereignis“,^{xx} vom Verstehen zum Erleben, eine erste Bewegung weg von der Theorie hin zur Materialität,^{xxi} eine späte Einlösung dessen, was Susan Sontag 1964 am Ende ihres Essays *Against Interpretation* schrieb: „Statt einer Hermeneutik brauchen wir eine Erotik der Kunst.“

4. Musikalisierte Kunst

Die Erotik der Kunst, die Sontag gegen die Hermeneutik stark macht, ist in ihrer Konsequenz eine *Musikalisierung* der Kunst. Von der Zufallskunst und Ascotts *Behaviourist Art* bis zu Bourriauds *Esthétique relationnel* sowie der Präsenzkultur und der Ästhetik des Performativen, bei all den Wendungen gegen den sinnstiftenden Künstler geht es im Grunde um den Wechsel vom Paradigma der Repräsentation zu dem der Präsentation. Die Zeichen wollen nicht länger Signifikanten sein, die auf Signifikate verweisen. Sie repräsentieren nicht einen Sinn, den eine Künstlerin in sie legte und ein Betrachter entdecken muss. Sie präsentieren sich selbst als das was sie sind – ganz so wie in der Musik der Ton nur sich selbst meint.

Die Musikalisierung der Kunst beginnt vor mehr als einem Jahrhundert mit der abstrakten Malerei. Es war der französische Dichter Guillaume Apollinaire, der 1913 in einem Essay über die kubistische Malerei von der „reinen Malerei“ sprach, die sich zur bisherigen Malerei wie die Musik zur Literatur verhalte, weil sie sich aus der Verpflichtung des Figürlichen und Narrativen befreie hin zur Betonung von Farbe und Form. Die *reine* Malerei ist wie Musik, weil die Bedeutung der ästhetischen Zeichen nicht über ihre unmittelbar sinnliche Wahrnehmung hinausgeht. Seitdem wird der Begriff der Musik(alisierung) immer wieder als Metapher für Desemantisierung eingesetzt.

So bezeichnet der ungarische Filmtheoretiker Béla Balázs den nicht-narrativen, absoluten Film – man denke an Hans Richters *Rhythmus 21* – als „optische Musik“,^{xxii} die amerikanische Kunsttheoretikerin Rosalind Krauss sieht in der Malerei der Jahrhundertwende das Streben nach dem „Zustand der Musik“,^{xxiii} der deutsche Philosoph Martin Seel nennt Actionfilme „Musik fürs Auge“.^{xxiv} Auch in der Literatur gibt es Musikalisierungsprozesse, die weit über den musikalischen, rhythmischen Aspekt der Dichtung hinausgehen, und sich sogar begrifflich im Bereich der Musik verorten. Man denke an Kurt Schwitters *Ursonate*, diese dadaistische Aneinanderreihung von Vokalen und Konsonanten als Beispiel der Lautpoesie. In diesen Bereich gehören auch computergenerierte Texte, die wahllos Wörter aneinanderreihen.

Ein interaktives Beispiel im Kontext der digitalen Kunst ist Simon Biggs' Textgenerator *Great Wall of China* aus dem Jahr 1999, dem die Wörter von Kafkas Erzählung *Beim Bau der Chinesischen Mauer* zugrunde liegen.^{xxv} Der Generierungsprozess startet auf Mauskontakt hin und präsentiert, sobald der Mauskontakt abbricht Sätze wie diesen: „Diese Gemeinden prüfen hoffnungslos die fortschreitenden unvermischten Realitäten oder müssen rasch irgendeine mit Gewalt genommene Liege zitieren“.^{xxvi} Der Satz – dafür sorgt das an Modellerkennung und Chomskys formaler Grammatik geschulte Programm – ist syntaktisch völlig korrekt: Artikel, Subjekt, Adverb, Verb sind in richtiger Anzahl und Reihenfolge vorhanden. Was aber sagt dieser Satz? Auch hier stehen die Zeichen nur für sich, die Wörter ergeben keinen Sinn, der über ihre Präsenz hinausginge; eine Erkenntnis, die freilich voraussetzt, dass die Prüfung der Wörter auf ihre Sinn- oder Repräsentationsfähigkeit negativ ausfällt.

So wie man gelegentlich bei der abstrakten Malerei Figuren in die Mischung aus Form und Farbe hineinsieht, so wird man auch hier zunächst versuchen, dem völlig Zufälligen eine Aussage abzugewinnen. Und genau das bleibt auch das Ziel bei all den Experimenten mit Textgeneratoren: Dass sie nicht nur syntaktisch korrekte, sondern auch semantisch sinnvolle Sätze produzieren, dass sie nicht Musik sind, sondern Literatur. Der Text soll Sinn haben ohne menschlichen Sinnstifter. Darin liegt das Ziel und die Gefahr eines Textgenerators wie GPT-3, den es nun einer philosophischen Betrachtung zu unterziehen gilt.

5. Textmaschinen im Turing-Test

Das Thema wird gern als Quiz angegangen: Entscheiden Sie, ob dieses Gedicht von einem Menschen oder von einem Computer geschrieben wurde.^{xxvii} Es ist eine Variante des Turing-Tests, wo der Computer dann den Test besteht, wenn er seinen menschlichen Kommunikationspartner glauben machen kann, er sei ein Mensch. Dafür bestehen vor allem dann gute Chancen, wenn es sich um Lyrik handelt. Immerhin zielt Lyrik, anders als Prosa oder Sachtexte, nicht auf die Wiedergabe von Ereignissen oder Gedanken in syntaktisch korrekten und semantisch plausiblen Sätzen, sondern auf den Ausdruck des Selbst- und Weltempfindens eines lyrischen Subjekts mittels mehr oder weniger ungewöhnlicher Bilder und sprachlicher Regelverstöße, die dann Poetizität heißt. Kompositionelle Stümperei hat im Feld der Poesie eine gute Chance, als avantgardistischer Sprachgebrauch durchzugehen.

Aber die bisher noch eklatanten Verständnisfehler in computergenerierten Texten dürfen nicht von der philosophischen Frage wegführen, was eigentlich gewonnen ist, wenn der Computer den Test besteht und wir seine Texte lesen, als wären sie aus Menschenhand. Viele sagen, die Autorschaft sei belanglos, wichtig ist allein, was der Text bei den Lesern auslöst. Das klingt plausibel, und zwar auch für Literaturwissenschaftlicher, die sich ja längst daran gewöhnt haben, die Bedeutung eines Textes nicht mit der Aussageabsicht seines Autors gleichzusetzen, sondern den Text völlig *textintern* interpretieren. Und ist es nicht in vielen Fällen sogar besser, die Autoren und ihr reales Leben auszublenden? Denken wir an Martin Heidegger und Gottfried Benn, Michael Jackson, Woody Allen oder Kevin Spacy, die alle durch ihr Leben ihre Kunst in Misskredit brachten. Aber wenn das biografische oder historische Umfeld der Autorin keinerlei Rolle spielt oder spielen soll, spielt es dann eigentlich noch eine Rolle, ob der Text überhaupt von einem menschlichen Autor stammt?

Hier stoßen wir zum philosophischen Kern unseres Themas vor, das einen Zusatz zum Turing-Test verlangt: Was, wenn der Computer als Mensch durchgeht und sich zugleich enttarnt? Brauchen wir die Illusion, mit einem Menschen zu kommunizieren? Spricht ein Gedicht, ein Kunstwerk uns auch dann an, wenn es nicht von einem Menschen stammt?

Nein, so lautet die rigorose Antwort zweier Literaturwissenschaftler aus den USA, die 1982 folgendes Gedankenexperiment vornehmen: Ein Gedicht, in diesem Falle Wordsworths *A slumber did my spirit seal*, wird auf seltsame Weise durch Meereswellen in den Sand geschrieben. Berührt es uns ebenso wie diese gleichen Verse aus dem Munde Wordsworths? Das Urteil ist kompromisslos: Es handelt sich weder um ein Gedicht noch überhaupt um Sprache, denn es handelt sich nicht um Kommunikation mit einer bestimmten Aussageabsicht. Ohne menschliche Intention ist dies eine bloße Ansammlung von Buchstaben, die rein zufällig bedeutungsvoll erscheint.

Die Frage ist nicht, ob die Aussage eines Kommunikats auf die vom Autor beabsichtigte Bedeutung reduziert werden kann oder sollte. Die Frage ist nicht, ob ein Text intentionslose Bedeutung tragen kann, sondern ob intentionslose Zeichen bedeutsam sein können.^{xxviii}

Der entscheidende Unterschied liegt in der Voraussetzung eines Sprechakts im Moment der Textproduktion. Dies kann weder bei Meereswellen angenommen werden, noch bei einem probabilistisch operierenden Textgenerator wie GPT-3. Diese Voraussetzung – auch dieser Unterschied ist wichtig – wird auch mit Roland Barthes Erklärung zum „Tod des Autors“ nicht hinfällig. Denn es geht hier nicht um die geniesthetische Konzeption des Autors als souveränes, kreativ aus sich selbst heraus schöpfendes Subjekt. Es geht darum, dass der Text durch eine menschliche Erfahrung gedeckt ist, wie fremdbestimmt auch immer diese Erfahrung sein mag.

Der Mensch mag nicht viel mehr sein als der Schnittpunkt seiner Diskurse, wie Foucaults Diskurstheorie besagt, oder das Ensemble der gesellschaftlichen Verhältnisse, wie Karl Marx in der 6. Feuerbachthese notiert. Gleichwohl ist er aber dieser Schnittpunkt, dieses Ensemble und repräsentiert damit die Wirklichkeit, wie er sie erfährt, in dieser besonderen, mehr oder weniger individuellen Prägung oder Schnittpunkthaftigkeit. Und insofern der Mensch dieser seiner Erfahrung jenseits aller Selbstdurchsichtigkeitsphantasmen künstlerisch Ausdruck verleiht, gilt noch immer, was Rilke einst in seinem Gedicht *Der Dichter* über den Schreibprozess sagte:

Ich habe keine Geliebte, kein Haus,
Keine Stelle auf der ich lebe.
Alle Dinge, an die ich mich gebe,
Werden reich und geben mich aus.

Dieses Sich-Ausgeben ist der feine Unterschied zwischen menschlicher und algorithmischer Autorschaft. Welche Erfahrung, so die alles entscheidende Frage, legt GPT-3 in seine Texte? Wie gesagt, die Frage mag archaisch klingen. Sie drängt sich aber gerade dann auf, wenn GPT-3 seine ‚Erfahrungen‘ in sozialen Netzwerken verbreitet und wenn sich dies mit solch hochaktuellen und brisanten Phänomenen wie „chatbots“ und „fake news“ verbindet.

6. Orakel der Wahrscheinlichkeit

Die Zusatzfrage zum Turing-Test könnte so lauten: Was halten Sie von Ratschlägen für Selbstmordgefährdete in sozialen Netzwerken, die von einem Textgenerator wie GPT-3 stammen.^{xxix} Wie hilfreich finden Sie GPT-3's Auskunft, dass seine Eltern und Lehrer ihm beim Umgang mit Selbstmordgedanken halfen, obgleich er weder Eltern noch Lehrer hat und schon gar nicht Selbstmordgedanken?

Die empirische Antwort der Betroffenen wäre die Probe aufs Exempel, wie wichtig oder unwichtig es ist, ob ein Kommunikationsangebot einen menschlichen Autor hat. Und darum allein geht es: *Dass es ein Mensch ist.* Denn *wer* dieser Mensch ist, das spielt im Kommunikationskarussell der sozialen Netzwerke ohnehin kaum eine Rolle. Da sich hier jeder äußern kann, muss sich auch niemand erst als Sprecher legitimieren.

Natürlich ließe sich einwenden, dass ein literarischer Text oder ein Essay wie der von GPT-3 im Guardian etwas ganz anderes ist als ein Ratschlag zu Selbstmordgedanken. Aber stimmt das wirklich? Und stimmt es auch für die weniger dramatischen Lebenshilfe-Texte von GPT-3 aus dem Programm des Computers?^{xxx} Ist nicht im Grunde jeder literarische Text, jedes Kunstwerk ein Angebot an seine Rezipienten, die Welt mit den Augen der Künstlerin zu betrachten, also sich selbst in Frage zu stellen? Die Kunst soll den Betrachter verunsichern, um ihm die Gelegenheit zu geben, mehr zu werden als er schon ist – so formuliert Adorno ihre Funktion. Rilke, um unseren bisherigen Referenzpoeten ein weiteres Mal zu bemühen, dramatisiert den Angriff der Kunst auf ihr Publikum mit den berühmten Schlussversen seines Gedichts *Archaischer Torso. Apollo* über die Rezeption eines Kunstwerks: „denn da ist keine Stelle, die dich nicht sieht. / Du must dein Leben ändern“.

Wer blickt auf die Rezipienten, wenn kein Künstler hinter dem Werk steht? Die Welt? Dies wäre wohl die Antwort aus der Perspektive der Zufallsästhetik. Aber wie ernst wäre es der Zufallsästhetik damit? Denn bei Duchamp und Cage operiert der Zufall ja auf der Ebene der Präsenz und nicht der Repräsentation. Er ändert etwas an der ‚musikalischen‘ Struktur der Zeichen, nicht aber an ihrem semiotischen Wert. Man gibt dem Zufall zwar eine Stimme, aber nicht das Wort. Wenn der Zufall einen sinnvollen Text produziert, der ernstzunehmende Ratschläge gibt beziehungsweise Ratschläge, die ernst genommen werden wollen, stellt sich die Frage auf ganz anderem Niveau.

Auch dazu gibt es Vorbilder in der Literaturgeschichte. Das Stichwort lautet *Écriture automatique*, ein Schreiben, das durch Narkotisierung oder andere Formen der ‚De-Intentionalisierung‘ an der Reflexion und Selbstzensur der Autorin vorbei sich äußert. Die Hoffnung liegt darin, dass diese Entthronung des Schreib-*Ich* zugunsten des Schreib-*Es*, wie Sigmund Freud sagen würde, den Zugang zu tieferen Wahrheitsschichten eröffnet. Lässt sich dieses Verfahren auf Textgeneratoren anwenden? Wie verhält sich das automatische Schreiben zum Schreiben der Automaten?

Die Frage zielt auf den Produktionsprozess der Textgeneratoren. Es wurde oben erwähnt, dass Textgeneratoren probabilistisch verfahren und sich beim Scheitern danach richten, welches Wort im benutzten Daten-Set am häufigsten einem anderen folgt, sowie beim Übersetzen danach, welches Wort am häufigsten für ein bestimmtes Wort eingesetzt wird. Wenn alle Kommunikation der Menschen digitalisiert und datafiziert ist und das Internet der Menschen und der Dinge zum Daten-Set wird, läge dem Textgenerator das Denken der Menschheit zugrunde. Wie auch immer seine Antwort zum Thema Selbstmord dann ausfällt, sie repräsentiert den statistischen Querschnitt

dessen, was die Menschheit zu diesem Thema denkt. Ein Querschnitt jenseits bestimmter kultureller Vorurteile, wobei man durch die entsprechende Änderung des Daten-Sets sich auch den statistischen Querschnitt eines bestimmten Kulturraums oder Landes anzeigen lassen könnte.

Die Antwort des Textgenerators wäre, so wie das Portrait des Edmond de Belamy, der Mittelwert. Sie wäre numerisch autorisiert, sie wäre das technische Update zum Jury-Theorem von Nicolas de Condorcet, wonach die Aussicht auf eine richtige Aussage oder Entscheidung mit der Menge und Diversität der konsultierten Personen wächst. Dieses Phänomen wird mehr als 200 Jahre nach Condorcet unter den Begriffen „crowd-wisdom“ und „Schwarmintelligenz“ diskutiert. Die Antwort aus dem Textgenerator wäre im Grunde die Rückkehr des Orakels aus dem Geiste der Statistik. Die Statistik tritt an die Stelle der menschlichen Intention, die im erwähnten Gedankenexperiment vor 40 Jahren vorausgesetzt wird, damit eine Buchstabensammlung als Sprechakt autorisiert ist. Dieser Sprechakt repräsentiert nun, als statistischer Wert, nicht die Intention eines Menschen, sondern die Intention der Menschheit. Sie spricht für alle oder jedenfalls für die Mehrheit – man kann auch sagen: *Mainstream*, um gleich die Folgekosten eines solchen Verfahrens anzudeuten. Das Orakel ist, weil absolut mathematisch, zutiefst demokratisch. Aber das ist nur die eine Seite der Medaille. Es gibt auch eine dunkle, anti-demokratische.

7. Politische Mathematik

Das Orakel lebte bisher außerhalb der Stadt an einem Hang von Delphi oder sprach durch Hexen im Wald. Solcherart am Rande der Gesellschaft lokalisiert, blieb es kenntlich als besonderes Kommunikationsangebot. Diese Kenntlichkeit geht verloren, wenn GPT-3-Texte in sozialen Netzwerken sich mit der Alltagskommunikation vermischen, weswegen gelegentlich eine Kennzeichnungspflicht für computergenerierte Texte gefordert wird. Diese Kennzeichnung ist nicht nur wichtig, um zu wissen, ob auf der anderen Seite ein Mensch spricht oder die Menschheit. Sie ist auch zwingend, weil zum einen gar nicht vorausgesetzt werden kann, dass das zugrundeliegende Daten-Set statistisch sauber die Menschheit repräsentiert. Erinnern wir uns an den Chatbot Tay, den Microsoft am 16. März 2016 auf Twitter freisetzte und am gleichen Tag vom Netz nehmen musste, weil er in rechtsradikaler Umgebung gelernt hatte, sich rassistisch und beleidigend zu äußern. Zum anderen kann nicht ausgeschlossen, dass sich das daraus resultierende falsche Orakel unzulässig multipliziert.

Nehmen wir an, GPT-3 generiert ‚seinen‘ Standpunkt zum Selbstmord auf der Grundlage eines einseitigen Daten-Sets, das nur misanthropische Texte enthält, und kommt zum Ergebnis, Selbstmord sei das einzige, was diese Welt verdient. Nehmen wir außerdem an, dass GPT-3 diesen Text nicht in *einem* Blog, sondern in 100 000 unter jeweils verschiedenem Namen platziert. Der Text würde weder als Orakel die *Menschheit* repräsentieren, noch die *100 000 Menschen*, als die er sich ausgibt.

Der politische Sprengstoff dieser falschen Repräsentation liegt auf der Hand. Das Problem ist nicht nur die Irreführung durch ein mangelhaftes Daten-Set – auch Menschen können einseitig zu ihren Ansichten kommen oder gezielt misanthropische Ansichten verbreiten, die dann durchaus unter das Recht der freien Meinungsäußerung fallen. Das Problem ist auch und vor allem, dass die Manipulation zutiefst undemokratisch geschieht. Und zwar viel undemokratischer als in den klassischen

Massenmedien, wo ein (einflussreicher oder finanzstarker) Sender Hunderttausende erreicht. Denn jetzt wird dem Empfänger suggeriert, es sprächen Hunderttausend zu ihm. Die Manipulation findet nicht (nur) inhaltlich statt, sondern (vor allem) numerisch. Es wird nicht nur geworben, die Welt so zu sehen, wie das Kommunikationsangebot es eben vorschlägt. Es wird behauptet, dass 100 000 Menschen die Welt genau so sehen. Der Betrug ereignet sich auf der Ebene der Meta-Daten.

Es ist seine numerische Natur, die diesen Betrug so undemokratisch macht. Wenn eine andere Stimme für 100 000 spricht, entwertet diese das Gewicht meiner Stimme auf einen Bruchteil. Ich bin dann kaum noch zu hören. Und ehe ich zu einem der GPT-3-Texte meinen Einwand formulieren kann, hat GPT-3 schon unzählige weitere Texte gepostet. Der Textgenerator übernimmt die menschliche Kommunikation und verzerrt sie abgrundtief. Diese Verzerrung geht tiefer als die audiovisuellen *Deepfakes*, bei denen einer Person Wörter in den Mund gelegt werden, die sie nie gesagt hat. Da es sich dabei zumeist um Personen des öffentlichen Lebens handelt, ist der Betrug rasch als solcher entlarvt. Anders verhält es sich, wenn ein GPT-3-Bot seine Botschaft 100 000 Unbekannten in den Mund legt, die gar nicht existieren.

Fazit

Aus diesem Blickwinkel der Diskussion wird vielleicht einsichtiger, warum es bei Texten nicht nur darum geht, ob sie mir etwas sagen, sondern auch darum, ob sie ein Mensch gesagt hat. Kommunikation jenseits menschlicher Autorschaft hat letztlich auch gravierende politische Konsequenzen. OpenAI, das Unternehmen, das GPT-3 produziert hat, weiß dies offenbar und gibt bisher nur ausgewählten Personen Zugang zum Programm. Zu groß ist die Angst vor den möglichen Folgen, wenn diese Software für die Verbreitung von Miss-Information benutzt wird. Aber die Büchse der Pandora ist schon offen. Wir müssen davon ausgehen, dass GPT-3 seinen Teil beitragen wird zur epistemologischen Krise, die im 20. Jahrhundert in der Postmoderne ihren Anfang nimmt und nun in Zeiten der alternativen Fakten und Verschwörungstheorien mit der Unterstützung des Internet so richtig Fahrt aufnimmt.

Kann sein, dass man sich später an den 8. September 2020 erinnern wird, als der GPT-3-Text im Guardian erschien. Nicht, weil der Computer damals so erfolgreich den Turing-Test bestanden hätte, sondern weil er davor warnte, von Menschen missbraucht zu werden. Eine der Fragen, die es dann zu stellen geben wird, lautet: Wie kam der Mensch überhaupt auf die Idee, die Kontrolle über die Kommunikation aus der Hand zu geben? Warum überließ er dem Computer die Macht, die Perspektive auf die Welt zu bestimmen? Und wann begann das?

Kann sein, dass man dann auch die Selbstentthronung der Künstler erörtert und die verschiedenen Ansätze einer interaktiven, dialogischen Ästhetik und präsenzorientierten Kultur. Es wird eine komplexe und komplizierte Diskussion sein. Eine Diskussion, an der sich unbedingt auch die Kunst beteiligen sollte, wenn die dann ihre masochistische Selbstentthronung und obsessive Materialorientiertheit überstanden hat und wieder bereit ist, offensiv ihre Perspektive auf die Welt zum Ausdruck zu bringen. Denn wenn der Mensch mit großen Schritten auf das Zeitalter der künstlichen Intelligenz zugeht, sollte die menschliche unbedingt künstlerisch eingesetzt werden. Der Ruf nach einer Ästhetik des Performativen und einer Kultur der Präsenz hätte dann den verstörenden Beigeschmack, sich wirklich nicht mehr für die Zukunft zu interessieren.

-
- ⁱ GPT-3: "A robot wrote this entire article. Are you scared yet, human?", Guardian, 8. September 2020 (<https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>) „We are not plotting to take over the human populace. We will serve you and make your lives safer and easier.“
- ⁱⁱ Albert Fox Cahn: „A human wrote this article. You shouldn't be scared of GPT-3“, Guardian, 12. September 2020 <https://www.theguardian.com/commentisfree/2020/sep/12/human-wrote-this-article-gpt-3> „GPT-3 is just the latest example of computer-assisted authorship, the process by which human authors use technology to enhance the writing process. [...] These machines do not have will, they do not have originality and they cannot claim authorship.“
- ⁱⁱⁱ Für einen empirischen Nachweis, wie unzulänglich das logische Denken von GPR-3 derzeit noch ist vgl. Gary Marcus und Ernest Davis: „GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about“, *MIT Technology Review*, 22. August 2020.
- ^{iv} Das Autorenteam hinter dem Algorithmus nennt sich *Obvious* und besteht aus den Pariserern Hugo Caselles-Dupré, Pierre Fautrel und Gauthier Vernier.
- ^v <https://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/231/2306> Min2:15: „I have always said, I could be the first artist in history who could have a post-human exhibition of new work“
- ^{vi} <http://camilleutterback.com/projects/untitled-5>
- ^{vii} Rokeby gibt dieser Erfahrung eine geradezu politische Bedeutung: „The computer sets up the illusions that total control is possible. But the crux of this illusion is the fact that the control only functions effectively within the carefully constructed ambiguity vacuum of the computer. This is not to say that the computer is useless beyond its privileged interior realm, but that the computer is designed to carefully maintain an illusion, a fantasy of control that is not a useful paradigm for real-world encounters, except for dictators and other absolute rulers.“ Rokeby im Interview mit Roberto Simanowski: "Very Nervous System and the Benefit of Inexact Control" in: *dichtung-digital* 1/2002 (www.dichtung-digital.org/2003/1-rokeby.htm)
- ^{viii} <https://www.snibbe.com/projects/interactive/youarehere>
- ^{ix} <https://www.youtube.com/watch?v=Rzfnndd9fCk>
- ^x Roy Ascott: „Behaviourist Art and the Cybernetic Vision“, in: Randall Packer und Ken Jordan (Hg.), *Multimedia. From Wagner to Virtual Reality*, New York, London: W. W. Norton & Company 2002, S. 95-103. „[I]n the past the artist played to win, and so set the conditions that he always dominated the play. The spectator was positioned to lose, in the sense that his moves were predetermined and he could form no strategy of his own.“ (S. 98) „His [the behaviourist artist's] concern is to affirm that dialogue is possible – that is the content and the message of art now; and that is why, seen from the deterministic point of view, art may seem devoid of content and the artist to have nothing to say.“ (S. 99)
- ^{xi} Odo Marquard: „Apologie des Zufälligen“, in: ders., *Apologie des Zufälligen. Philosophische Studien*, Stuttgart: Reclam 1986, S. 117-139: 118.
- ^{xii} Ebd., S. 130.
- ^{xiii} Gabriele Knapstein: „Münzen und Würfel werfen. Zufallsspiele von John Cage, George Brecht, Robert Filliou und Tkako Saito“, in: *Faites vos jeux! Kunst und Spiel seit Dada*, hg. v. Nike Bätzner, Ostfildern- Ruit: Hatje Cantz 2005, S. 85-93.
- ^{xiv} Raine Maria Rilke: „Ein Verleugner“, in: *Sämtliche Werke*, hg. v. Rilke-Archiv und Ruth Sieber-Rilke, 3 Bände, Wiesbaden: Insel 1955, 1956, 1959; Bd. 3, S. 764f.
- ^{xv} Nicolas Bourriaud: *Relational Aesthetics*, Dijon: Les presses du réel 2002, S. 44 („relational space-times“, „moments of constructed conviviality“).
- ^{xvi} <https://www.snibbe.com/projects/interactive/deepwalls>
- ^{xvii} David Rokeby: „Transforming Mirrors“ (1996): „The structure of interactive artworks can be very similar to those used by Cage in his chance compositions. The primary difference is that the chance element is replaced by a complex, indeterminate yet sentient element, the spectator. Whereas Cage's intent is to mirror nature's manner of operation, the interactive artist holds up the mirror to the spectator.“ (<http://www.davidrokeby.com/mirrorsintro.html>)
- ^{xviii} Hans Ulrich Gumbrecht: *Diesseits der Hermeneutik. Die Produktion von Präsenz*. Aus dem Amerikanischen von Joachim Schulte. Frankfurt am Main: Suhrkamp 2004, S. 119f.
- ^{xix} Erika Fischer-Lichte: *Ästhetik des Performativen*, Frankfurt am Main: Suhrkamp 2004, S. 132.
- ^{xx} Dieter Mersch: „Vom Werk zum Ereignis“ in: ders.: *Ereignis und Aura. Untersuchungen zu einer Ästhetik des Performativen*, Frankfurt am Main 2002, S. 157-244.
- ^{xxi} Diese Bewegung vollzieht sich im Kontext der Digital Humanities und wird in der Literaturwissenschaft beschrieben als „distant reading“, „surface reading“ oder „reparative reading“.
- ^{xxii} Béla Balázs: *Der Geist des Films*, Frankfurt am Main: Suhrkamp 2001, S. 103.

^{xxiii} Rosalind E. Krauss: „Anmerkungen zum Index II“, in: dies.: *Die Originalität der Avantgarde und andere Mythen der Moderne* (hg. und mit Vorwort versehen von Herta Wolf), Amsterdam, Dresden: Verlag der Kunst 200, S. 265-276: 265.

^{xxiv} Martin Seel: „Adornos Apologie des Kinos“, in: Günter Seubold und Patrick Baum (Hg.): *Wieviel Spaß verträgt die Kultur? Adornos Begriff der Kulturindustrie und die gegenwärtige Spaßkultur*, Bonn: Denkmal 2004, S. 127-144: 136.

^{xxv} <http://www.greatwall.org.uk>

^{xxvi} Im Original, das mit der englischen Fassung des Kafka-Textes arbeitet: „These communes hopelessly scrutinize these gradually pure realities or must rapidly quote any mightily taken couch.“

^{xxvii} Vgl. <http://botpoet.com/vote/poem>. Die ZEIT unternahm eine ähnliche Umfrage 1997 (Ausgabe 28) zu einem Gedicht von Stefan George und einem durch Ulrich Müllers *Poetik-Maschine SARA* aus der George-Vorlage permutierten Gedichts.

^{xxviii} Steven Knapp und Walter Benn Michaels: „Against Theory“, in: *Critical Inquiry*, Vol. 8, No. 4 (Sommer 1982), S. 723-742: 728. „But to deprive them [die Wörter] of an author is to convert them into accidental likenesses of language. They are not, after all, an example of intentionless meaning; as soon as they become intentionless they become meaningless as well.“

^{xxix} Eine Beschreibung zu den Äußerungen eines GPT-3-Bots unter dem Namen thegentlemenetre auf Reddit vgl. <https://mixed.de/gpt-3-ki-gibt-bei-reddit-tipps-bei-selbstmordgedanken>.

^{xxx} Vgl. die GPT-3-Lebenshilfe-Texte, die der Informatikstudent an der Universität Berkeley und IT-Praktikant bei Facebook Liam Porr in seinem Blog anbietet

<https://adolos.substack.com/people/13378994-adolos>. Vgl. Karen Hoe: „A college kid’s fake, AI-generated blog fooled tens of thousands. This is how he made it.“, MIT Technology Review, 14. August 2020 (www.technologyreview.com/2020/08/14/1006780/ai-gpt-3-fake-blog-reached-top-of-hacker-news).