

Digitale Kunstgeschichte

AI und die Kunst des ungewollten Pasticcio. Ein Experiment

Prof. (pens.) Dr. Felix Thürlemann
felix.thuerlemann@uni-konstanz.de

AI und die Kunst des ungewollten Pasticcio. Ein Experiment

Felix Thürlemann

I. Probe

„Die Hauptfigur ist Mars, der vom Tempel des Janus mit Schild und blutigem Schwert ausbricht und dabei die Künste, die Liebe und die Religion mit Füßen tritt. Daneben steht in einer verzweifelten Pose Venus, die Mars zurückzuhalten versucht. Sie fleht ihn an, zu ihr zurückzukehren. Auf der anderen Seite sind die Allegorien des Friedens, der Fülle und des Überflusses dargestellt, die in ihrer vollen Blüte durch den Krieg unterbrochen werden. Über allem schwebt die Figur der Verderbnis, die eine Fackel in der Hand hält. Im Vordergrund sind die verwirrten Menschen, die durch den Krieg leiden. Das Ganze ist eine Darstellung der Folgen des Krieges, der sich über alles hinwegsetzt, was heilig und menschlich ist.“

Wir können nicht wissen, was Justus Sustermans mit seinem inneren Auge sah, als er die Beschreibung las, die ihm Peter Paul Rubens in einem auf den 12. März 1638 datierten Brief von seiner eben vollendeten, für den Großherzog der Toskana gemalten *Allegorie des Krieges* hatte zukommen lassen. (Sie ist oben in einer kurzen deutschen Zusammenfassung des italienischen Originaltextes wiedergegeben.) Aber ich weiß jetzt, was die AI, genauer die kostenlose Version des intermedialen Chatbots *Gemini* von Google[↗] nach der Eingabe dieser Zusammenfassung der Ekphrasis von Rubens ‚gesehen‘ hat. Genauer: Ich weiß, was der Bot, von den oben wiedergegebenen Sätzen geleitet, aus den vom Google-Imperium auf ihren Serverfarmen gespeicherten bildfähigen Daten mit ziemlich hohem Stromverbrauch gesucht, herausgezogen und neu zusammengebaut hat. Ich hatte dem Text bloß den Befehl „Mach mir ein Gemälde, das der folgenden Beschreibung entspricht: ...“ vorangestellt und so als Prompt in das Suchfeld eingegeben. Nach etwa zwanzig Sekunden leuchtete mir das Resultat

auf dem Display des PC entgegen, bereit zum Herunterladen und zum schnellen Verteilen nach Belieben.

| Abb. 1 |

Die Darstellung ist trotz ihres Detailreichtums und des großen Figurenpersonals in Perspektive, Schattenwurf und Farbigkeit auf den ersten Blick überraschend kohärent. Auch die gewählte Stillage ist einigermaßen homogen und gibt kaum Anlass zur Kritik. Heinrich Wölfflin hätte sie mit Hilfe der fünf Grundbegriffspaare durchaus in sein System der barocken Kunst einordnen können, obwohl das Produkt der Artificial Intelligence (AI) – anders als das in der Ekphrasis beschriebene Gemälde – aufgrund der etwas harten Konturierung der Figuren eher der Spätphase des Barock zugewiesen werden muss. Der Chatbot scheint sich vor allem an den Werken des römischen Spätbarock, etwa jenen der Maratta-Schule (man denkt an Andrea Procaccini und an Giuseppe Bartolomeo Chiari) orientiert zu haben. Aber warum? Wenn der Prompt veraltete oder wenig gebräuchliche Wörter wie ‚Allegorie‘ zusammen mit Namen von antiken Götterfiguren enthält, wählt *Gemini* anscheinend ohne weiteres Zutun eine ‚richtige‘, dem Darstellungsgegenstand adäquate Stillage. Der Chatbot historisiert die Szene entsprechend dem Sprachstil der ihm gelieferten Beschreibung. Unter der stilistisch-chronologischen Vorgabe ‚Spätbarock um 1700/1720‘ wird dann, geleitet vom geheim gehaltenen Algorithmus, in Sekundenschnelle offenbar ein riesiger Bilderschatz durchsucht, werden klar konturierte Figuren und Figurencluster, die jenen der Texteingabe entsprechen, erkannt und neu zusammengesetzt, um schließlich, nach der notwendigen stilistischen ‚Glättung‘, dem Nutzerindividuum als kohärentes Bild – hier als simuliertes Barockgemälde – zur Verfügung gestellt zu werden. So etwa kann der



| Abb. 1 |

Visualisierung
der gekürzten
Ekphrasis des
Gemäldes
„Die Folgen des
Krieges“ von
Peter Paul Rubens,
generiert durch den
Chatbot Gemini.
google.com am
3. März 2025

Prozess der digitalen Bildproduktion ausgehend vom eingegebenen Prompt rekonstruiert werden.

Eine detaillierte Kritik des von Google gelieferten Resultats soll nicht im Zentrum der folgenden Überlegungen stehen. Auch auf einen qualitativen Vergleich mit der Leistung anderer Chatbots, deren Resultate bei einem identischen verbalen Input sehr unterschiedlich ausfallen können, soll verzichtet werden. Es geht um grundsätzlichere Fragen, um Fragen der Sinnkonstitution. Bei näherem Hinsehen ist leicht zu erkennen, dass es sich beim gelieferten ‚Werk‘ um ein *mixtum compositum* handelt, zusammengesetzt aus Bruchstücken von mehr oder weniger bekannten Werken der gewählten stilistischen Schule, wobei das Resultat vor allem bei den Figurengruppen im Mittelgrund aufgrund der aufgehellten Farbigkeit etwas gar süßlich, rokokoartig erscheint. Doch haben fleißige Künstler der zweiten Garde, die sich an den Werken der großen Meister orientierten, nicht auch auf diese Weise gearbeitet, als ungehemmte Plünderer und erfundungslose Plagiatoren?

Es gibt die Tendenz anzunehmen, die mit Hilfe der künstlichen Intelligenz produzierten Bilder seien immer *neue Bilder*, Bilder, die uns Auskunft geben über die Gefühlslage der heutigen Gesellschaft, ähnlich den Prêt-à-porter-Kleidern nach der neuesten Mode. Diese Haltung den künstlich generierten Bildern gegenüber ist nicht ganz adäquat. Die großen Anbieter speichern auf ihren Servern in der Form von Pixelassemblagen grundsätzlich alle ihnen zugänglichen Bilder, auch solche, die den Kunsthistorikerinnen und Kunsthistorikern vertraut sind. Es sind die Bilder (auch Reproduktionen von Skulpturen), die in den großen Museen als reale Objekte aufbewahrt, die in Büchern reproduziert oder die von kommerziellen Anbietern und auf den Websites der Sammlungen als Scans online dargeboten werden. Ein AI-generiertes Bild ist nur in dem Sinne neu, als es mit der jeweils neuesten Technologie im gnadenlosen, profitorientierten Wettbewerb unter den Tech-Giganten geschaffen worden ist. Neu erscheint es auch deshalb, weil die Möglichkeit, mit Hilfe von künstlicher In-

telligenz über verbale Anweisungen komplexe Bilder kreieren zu können, bis noch vor Kurzem unvorstellbar war. Schließlich kann man den Werken der AI aufgrund der Schnelligkeit ihrer Herstellung, der leichten Verfügbarkeit und einer gegenüber den handwerklich hergestellten Bildern überlegenen Mimesis eine eigene, bislang unbekannte Qualität zuerkennen. (Wer kann denn schon zersplitternde Glasscheiben im Flug so überzeugend darstellen wie unsere AI?) In Wirklichkeit aber ist jedes von der AI mit Hilfe eines Prompts generierte Bild in einem fundamentalen Sinne immer *alt*, da es auf der Grundlage von bereits geschaffenen, also alten Bildern hervorgebracht worden ist. Originalität kann nur der Prompt haben, als Anweisung für die Produktion von etwas sinnvollem Neuen aus Altem.

Das Pasticcioartige, das einem komplexen, AI-generierten Bild eigen ist, ist nicht das eines echten Pasticcios im kunsthistorischen Sinn. Es gehorcht keiner kohärenten Logik, auch keiner ästhetisch definierten Anti-Logik (etwa im Sinne der von Hieronymus Bosch erfundenen Figuren oder der surrealistischen Bildmontagen). Das ‚Pasticcio‘ der AI ist das Resultat von Mängeln und Macken, der nicht vermeidbaren, unkalkulierbaren Bruchstellen, die sich beim Zusammenbau der von der Maschine aus dem Netz gefischten, aus dem Kontext herausgelösten Bildfragmente ergeben haben. Freilich kann jedes semantisch inkohärente Bildprodukt im Modus der Parodie gelesen werden und hat so durchaus seinen Reiz. So etwa, wenn man versucht, die rechte untere Ecke der digital-generierten Kriegsallegorie à la Rubens **vgl. Abb. 1** zu entziffern, die völlig wirre visuelle Kakophonie, zusammengesetzt aus einem fröhlich gen Himmel blickenden jugendlichen Hirten, einem riesigen, Früchte und Blumen ausgießenden Füllhorn und einem Alten, der vor einem zerfledderten oder verschmutzten Folianten kniet und dabei die linke Hand nicht etwa vor, sondern zwischen seine beiden Augen hält, um das vor ihm ausgeschüttete Chaos nicht wahrnehmen zu müssen.

AI-generierte Bilder können zwar eine gewisse kompositorische Ordnung haben, aber diese ist infolge

der Imitation einer großen Menge von Referenzbildern als Durchschnittswert ‚oberflächlich‘, rein dekorativ. So zeigt das vorliegende Beispiel die gerade in der Barockzeit beliebte axialsymmetrische Verteilung von Nebenfiguren entlang der beiden unteren Ecken, dazu eine abgestufte perspektivische und farbliche Gewichtung in Haupt- und Nebenszenen, usf. Doch die kompositorischen Elemente sind nicht ‚gedacht‘, sie sind ohne eine konsistente bedeutungstragende Funktion. Zudem ist die hier als Prompt eingegebene Ekphrasis auch in Bezug auf die sinnträchtige kompositorische Ordnung ‚unverstanden‘. Nicht umgesetzt ist die für Rubens‘ Allegorie bestimmende, in einer Reliefstruktur veranschaulichte Sukzessionslogik des *ante hoc, propter hoc*-Prinzips, das der Maler von Lukians Ekphrasis und den schon zu seinen Lebzeiten in Stichen weit verbreiteten visuellen Umsetzungen entnommen hat (vgl. Jean Michel Massing, *Du texte à l'image. La Calomnie d'Apelle et son iconographie*, Straßburg 1990, und die dort abgebildete Visualisierung der Ekphrasis Lukians aus der Hand des jungen Rubens selbst, Kat. 26.GA). Rubens hat in seiner Kriegsallegorie die sequentielle Ordnung durch die Domino-Struktur, das sich beschleunigende Umkippen der Figuren aus der Vertikalen nach rechts in die Horizontale, verstärkt. | **Abb. 2** | Rubens‘ Komposition ist sinntragend. Sie zeigt das Unausweichliche des Schicksals, gegen das die dargestellten Figuren vergeblich ankämpfen. (Es ist jedoch bei einem Vergleich zwischen dem in natürlicher Sprache verfassten Input und dem maschinell generierten visuellen Output zu berücksichtigen, dass der Maler, zumindest was die Domino-Struktur betrifft, die kompositorische Logik seines Werkes im Brief an Sustermans nicht explizit beschreibt.)

II. Gegenprobe

Das Verhältnis zwischen Text und Bild ist keine Einbahnstraße. Mit Hilfe des multimedialen Chatbots kann das von ihm produzierte Bild wieder vertextet, in eine Beschreibung zurückverwandelt werden. Und so kann man eine Art Gegenprobe veranstalten. Stellen wir uns folgende Frage: ‚Versteht‘ der Bot von *Gemini*

die von ihm nach der verkürzten Ekphrasis von Rubens generierte Gemälde simulation? Zieht man das visuelle Produkt des Chatbots zurück in das Suchfeld hinein und lässt ihm den Prompt „Verfasse von diesem Gemälde eine Beschreibung“ folgen, wird in Sekundenschnelle eine „Analyse“ in der Form eines ausführlichen Bildkommentars zusammen mit kritischen Erörterungen geliefert. Sie beginnt mit einer klar formulierten allgemeinen Beschreibung des als Wiedergabe eines Gemäldes ‚erkannten‘ Bildes: „Dieses Gemälde, das im Stil der barocken Historienmalerei gehalten ist, stellt eine komplexe allegorische Szene dar, die den Kampf zwischen Tugend und Laster thematisiert. Im Zentrum stehen drei Hauptfiguren: ein bewaffneter Mann, eine Frau und eine dämonische Kreatur.“ Diese Charakterisierung ist insofern

überraschend, als sie die digitale Gemälde simulation stilistisch und gattungsmäßig richtig einordnet, indem sie diese der „barocken Historienmalerei“ zuweist. Sie bestimmt jedoch die „allegorische Szene“ nur ganz generell und stereotyp als „Kampf zwischen Tugend und Laster“. Zudem werden die beiden in das Bildzentrum gesetzten Figuren, Mars und Venus, nicht als solche erkannt.

Auch die anschließende „Analyse“ ist nicht wesentlich bestimmter, auch wenn sie zahlreiche Bildelemente detailliert ausbuchstabiert. Die fein säuberlich aneinandergereihten Deutungen werden häufig als ungesichert, als bloße Vorschläge vorgetragen. So heißt es zu den Figuren in der rechten unteren Ecke: „Ein Mann mit einem Stab und ein alter Mann, der mit gesenktem Blick vor einer Fülle von Blumen, Früchten



| Abb. 2 | Ferdinand Gregori, Radierung nach dem Gemälde „Die Folgen des Krieges“ (1637/38) von Peter Paul Rubens, gedruckt zusammen mit der Ekphrasis des Malers, 1771.
Staatsgalerie Stuttgart, Inv.-Nr. A 1996/6688 (KK) ↗

und Getreide kniet. Diese Fülle symbolisiert Reichtum, Fruchtbarkeit und die Früchte der Arbeit oder der Tugend. Der Mann könnte ein Hirte oder eine allegorische Figur der Weisheit sein.“ Der Chatbot von Google hat offenbar kein Gedächtnis und stellt bei jeder Bildgenerierung jeweils wieder neue algorithmische Berechnungen ohne Vorgeschichte an. Daher kann er nicht ‚erkennen‘ bzw. ‚wiedererkennen‘, dass die von ihm eine Minute zuvor generierte Gemälde simulation auf der Grundlage der verkürzten Ekphrasis von Rubens produziert worden ist. Und er hat überraschenderweise auch nicht ‚bemerkt‘, dass das simulierte barocke Historiengemälde ein eigenes Produkt ist, obwohl dieses rechts unten, in der Signaturecke, die Markierung „AI“ im Oval trägt.

Kurios ist schließlich die in vergoldeten Majuskeln realisierte, in das Gebälk des Janustempels gesetzte Inschrift. Notgedrungen muss sich der gleiche Chatbot, der die sinnlose Wortfolge generiert hat, dann auch bei der Gegenprobe ‚die Frage stellen‘, was diese denn wohl bedeuten könnte: „Die Inschrift ‚AFLIO ARPSON‘ (oder ähnlich) auf der Fassade ist schwer zu entziffern, könnte aber einen Hinweis auf den Kontext des Bildes geben (vielleicht ein Name oder ein Motto).“ Der Bot ‚versteht‘ die von ihm ‚erfundene‘ Inschrift über dem Portal des Rundtempels nicht und produziert in dieser Notsituation die uns aus Bildinterpretationen nur allzu bekannte Abfolge von Vermutungen: „Die Inschrift A (oder ähnlich) könnte einen Hinweis auf B geben (vielleicht X oder Y)“. Über die Verarbeitung einer großen Menge von Textbeispielen aus unserem Fach hat sich der Bot die gängigen Formeln des kunsthistorischen Diskurses perfekt zu eigen gemacht. Der Bot hält uns hier den Spiegel vor – nur dass ihm für sein Agieren im Sinne der Systemkonformität kein Vorwurf gemacht werden kann, uns, seinen Vorbildern, hingegen schon.

III. Fragen

Welche Lehren lassen sich aus dem hier beschriebenen Experiment ziehen? Folgende Erkenntnisse dürften über den derzeitigen Entwicklungsstand der generativen AI hinaus Gültigkeit haben:

(1) Alle von AI generierten Bilder sind in dem Sinne alt, als der Chatbot nur bereits existierende Bilder und ihre Figuren als Ausgangsmaterial für die Realisierung der vom Nutzerindividuum über die Texteingabe gewünschte ‚Neuschöpfung‘ einsetzen kann. Nach welchen Kriterien – sie sind je nach Chatbot offensichtlich verschieden – das visuelle Ausgangsmaterial ausgewählt wird, ist nicht einsichtig und deshalb auch nicht nachvollziehbar.

(2) Die vom Chatbot automatisch generierte kompositorische Ordnung ist als Durchschnittswert oberflächlich, rein dekorativ. Sie ist generisch und besitzt keine spezifische, für das produzierte ‚Werk‘ gültige bedeutungstragende Funktion.

(3) Auch wenn die intermedialen Chatbots weiterentwickelt werden, bleiben die genannten Mängel aus systematischen Gründen bestehen. Das kreative Denken, die Grundlage jeder originären künstlerischen Hervorbringung, ist immer individuell und werkbezogen. Die gezielte Auswahl und Summierung von visuellen Elementen aus noch so zahlreichen, in Bezug auf den eingegebenen Prompt als relevant ‚erkannten‘ Vorbildern kann diesen Mangel nicht beheben. Ein im ästhetischen Sinne gültiges digitales Werk kann deshalb nur im kritischen ‚Dialog‘ mit der Maschine entstehen. Unzufrieden mit dem gelieferten Resultat, kann das Nutzerindividuum den Prompt sukzessive präzisieren. Autor ist dann der Mensch, der Chatbot sein Instrument.

(4) Kunst entsteht immer im Kontext einer Bildtradition. Ein selbstbewusstes schöpferisches Individuum kann die überlieferten Werke imitieren, sich ihnen gegenüber kritisch positionieren oder versuchen, diese zu ignorieren. Der Chatbot hingegen kann, solange er auf sich allein gestellt ist, nur plagiieren. Die Kohärenz der KI-Bilder hat keine ästhetische Grundlage. Sie ist optischer Natur, auch dann, wenn sie sich an stilgeschichtlichen Modellen orientiert. Denn der Bot verhält sich den gespeicherten Datenquellen gegenüber, selbst dann, wenn der Algorithmus Kriterien für die Auswahl vorgibt, notwendigerweise affirmativ. Deshalb ist er auch nicht bild-kritik-fähig. Alle durch den öffentlichen Diskurs, aber auch durch die Texte

der Kunstgeschichte sanktionierten ‚Schönheiten‘ sind für den Chatbot gleichwertig.

(5) Eine zentrale Frage bleibt: jene nach der hermeneutischen Potenz, der Deutungsleistung der AI in Bezug auf bereits geschaffene Werke. Der Wert künstlerischer Objekte ist seit der Moderne stark von einer Ästhetik der Originalität geprägt. Wenn man von der Annahme ausgeht, dass vor allem jene Werke von Belang sind, die entweder in ihrer Gestaltung und/oder in ihrer spezifischen Verarbeitung allgemeiner Problemstellungen einzigartig sind, kann eine solche Leistung von der AI nur dann erkannt werden, wenn diese bereits in dem über das Internet zugänglichen kunstkritischen Diskurs ihren Niederschlag gefunden hat. Liegt eine solche Diskursivierung nicht vor, bleiben gerade die originellen kreativen Aspekte im Dunkeln. Diese können von der AI nicht im positiven Sinne, als sinnstiftende Elemente, erkannt und gedeutet werden, sondern höchstens *ex negativo*, als Abweichungen von verbreiteten ästhetischen Normen. Eine adäquate ‚Rezeption‘ des mit Hilfe der natürlichen Intelligenz jeweils neu Geschaffenen scheint der so genannten künstlichen Intelligenz nicht möglich.