

Ähnlichkeit ist ein elementares Phänomen, das sich als Übereinstimmung von Objekten auf Basis gemeinsamer Merkmale manifestiert. Aufgrund ihrer Vielschichtigkeit ist das Denksystem der Ähnlichkeit eingebettet in ein komplexes Wechselspiel kulturell geprägter und individueller Wahrnehmungen, das immer wieder rekontextualisiert werden muss. Dieses Heft verbindet naturwissenschaftliche, geisteswissenschaftliche und künstlerische Perspektiven und zeigt, wie Ähnlichkeit algorithmisch operationalisiert wird und Ähnlichkeitsstiftende Muster im digitalen Bild produktiv werden.

Ähnlichkeit

Reihe
Begriffe des
digitalen Bildes



Ähnlichkeit

Herausgegeben von
Stefanie Schneider
Hubertus Kohle

München, 2025
Open Publishing LMU

Inhalt

Einführung <i>Stefanie Schneider, Hubertus Kohle</i>	5
Nachbarn und Verwandte. Zur Operationalisierung der Ähnlichkeit <i>Roland Meyer</i>	15
Dimensionen der Ähnlichkeit. Prozesse der Merkmalsselektion und -bewertung <i>Stefanie Schneider, Desiree Frontczak, Julia Wolf, Theresa Zischkin</i>	33
Die Suche nach ähnlichen Bildern. Ein Praxisbericht mit dem e-Research-Werkzeug iART <i>Hubertus Kohle</i>	47
Infinite Loop. Disruption and Similarity in the Act of Design with Machine Learning <i>Jon Emmony, Silvia Weidenbach</i>	67

1



2



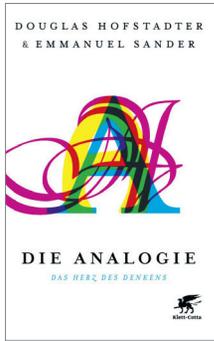
3



4



5



6

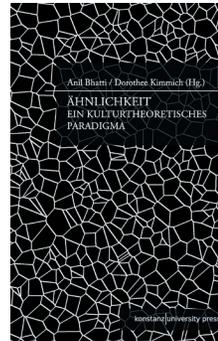


Abb. 1, Michel Foucault: Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften, Frankfurt am Main 1974

Abb. 2, Gerald Funk, Gert Mattenklott und Michael Pauen (Hg.): Ästhetik des Ähnlichen. Zur Poetik und Kunstphilosophie der Moderne, Frankfurt am Main 2001

Abb. 3, Lena Bader, Martin Gajer und Falk Wolf (Hg.): Vergleichendes Sehen, München 2010

Abb. 4, Michael Eggert (Hg.): Von Ähnlichkeiten und Unterschieden. Vergleich, Analogie und Klassifikation in Wissenschaft und Literatur (18./19. Jahrhundert), Heidelberg 2011

Abb. 5, Douglas Hofstadter und Emmanuel Sander: Die Analogie. Das Herz des Denkens, Stuttgart 2014

Abb. 6, Anil Bhatti und Dorothee Kimmich (Hg.): Ähnlichkeit. Ein kulturtheoretisches Paradigma, Konstanz 2015

Einführung

Ähnlichkeit: das ist die Übereinstimmung von Objekten aufgrund einer Teilmenge gemeinsamer Merkmale. In der Natur zeigt sie sich zum Beispiel bei der Fortpflanzung, wo die Merkmale der Nachkommen mit denen der Elterngeneration konvergieren und so die Identität einer Spezies bestimmen. In der Mathematik spielt Ähnlichkeit vor allem in der Geometrie eine Rolle: Zwei Polygone gelten dann als ähnlich, wenn ihre Winkel und Längenverhältnisse übereinstimmen. In der Physik wiederum ist sie die Grundlage für eine berühmte Anekdote. So leitete Isaac Newton das Gravitationsgesetz ab, nachdem er beobachtet hatte, dass der vom Baum fallende Apfel, der um die Erde kreisende Mond und die um die Sonne kreisende Erde ähnliche Bewegungsmuster aufweisen. Die Ähnlichkeit scheint also unser Verhältnis zur Welt entscheidend mitzuprägen.

Ähnlichkeit – das ist jedoch auch eine merkwürdige Kategorie: nichts Halbes und nichts Ganzes. Nicht von ungefähr leiten Gerald Funk, Gert Mattenklott und Michael Pauen ihre *Ästhetik des Ähnlichen* (Abb. 2) mit den folgenden Sätzen ein, die das schlechte Image der Ähnlichkeit in der gängigen Theoriebildung andeuten:

„Wer vom Ähnlichen spricht, weiß offenbar nichts Genaueres. Er mag eine Ahnung haben, doch offensichtlich ist er unfähig, Identitäten zu erkennen, und nicht Manns genug, das Differenten zu unterscheiden.“¹

Der Aristoteliker Robert Spaemann geht so weit zu behaupten, seine Disziplin habe „diesem elementarsten aller ‚Phänomene‘

5

¹ Gerald Funk, Gert Mattenklott und Michael Pauen: Symbole und Signaturen. Charakteristik und Geschichte des Ähnlichkeitsdenkens. In: dies. (Hg.): *Ästhetik des Ähnlichen. Zur Poetik und Kunstphilosophie der Moderne*, Frankfurt am Main 2001, S. 7–34, hier S. 7.

vorwiegend auszuweichen versucht.“² Das hat in seinen Augen mit der Tatsache zu tun, dass „Ähnlichkeit [...] keine klare und distinkte Idee im Sinne Descartes' [sei], sondern eine ‚konfuse Idee‘ im Sinne Leibniz.“³ Im *Historischen Wörterbuch der Philosophie* wird die Ähnlichkeit in einem sehr kurzen und einigermaßen nichtssagenden Lemma behandelt.⁴ Daneben fristet sie eine Randexistenz in dem sehr viel ausführlicheren Lemma über die begrifflich affine, zuweilen als Oberkategorie aufgefasste Analogie.⁵ Der Analogie wiederum wurde vor einem Jahrzehnt von Douglas Hofstadter und Emmanuel Sander der Status eines „Herz[ens] des Denkens“ zuerkannt (Abb. 5), denn „ohne Begriffe kann es kein Denken geben, und ohne Analogie gibt es keine Begriffe.“⁶ Schon die Romantik griff das cartesianisch-ähnlichkeitskritische Denken des rationalistischen Teils der Aufklärung an, und das analogische Denken avancierte in ihr zum Königsweg. In diesem Kontext erklärt Paul Valéry:

*„Die Existenz ähnlicher Dinge ist das Fundament von allem. Eine Welt aus Einzelexemplaren ist unvorstellbar. Wenn nichts sich wiederholte, dann wäre auch nichts. Der Sohn von anderem Schlage als der Vater; und jeder fortwährend sich selbst unähnlich; jeder Augenblick mit jedem anderen unvergleichbar; genau das wäre das Chaos.“*⁷

Bei einem überaus einflussreichen, allerdings in der philosophischen Diskussion zuweilen nicht immer ernst genommenen Autor nimmt der Ähnlichkeitsbegriff eine bekanntermaßen zentrale Stellung ein. Gemeint ist Michel Foucault, der in seinem Buch *Die Ordnung der Dinge* (Abb. 1) Ähnlichkeitsbezie-

hungen als elementar für ein vormodernes oder vorrationalistisches Denksystem beschreibt:

*„Die Welt ist von Zeichen bedeckt, die man entziffern muß, und diese Zeichen, die Ähnlichkeiten und Affinitäten enthüllen, sind selbst nur Formen der Ähnlichkeit.“*⁸

In diesem Denksystem hatten Ähnlichkeiten eine ontologische Relevanz: Sie zogen Phänomene in einen kohärenten Wahrheitsraum, wenn sie sich in der Anschauung nur ähnelten. Der cartesianische Rationalismus beendete dieses Denken – das aus ethnologischer Perspektive in den Kontext einer magischen Weltauffassung primitiver Völker gestellt wird⁹ – im 17. Jahrhundert zugunsten eines strikten Kausalitäts- und Identitätsbegriffs. In der an Foucault anschließenden Wissenschaftstheorie wird diese rationalistische Vorstellung zugunsten einer Aufwertung der Ähnlichkeits- beziehungsweise Analogiekategorie kritisiert: Auch in der modernen, klassifizierenden und systematisierenden (Natur-)Wissenschaft sei ein analogisches Denken zentral, und man könne sogar behaupten, dass der Rationalismus das Ähnlichkeitsdenken nur überlagert, aber nicht wirklich überwunden habe.¹⁰

In der Kunstgeschichte nimmt der Begriff der Ähnlichkeit eine zentrale Rolle in einer ihrer grundlegenden Verfahren, dem sogenannten „vergleichenden Sehen“ ein.¹¹ Denn verglichen wird eben dort, wo Kunstwerke auf ihre Ähnlichkeit hin befragt werden. Wenn Heinrich Wölfflin, einer der stilbildenden Praktiker dieser Methode, ein Porträt von Albrecht Dürer mit einem von Frans Hals vergleicht, so geht er von

2 Robert Spaemann: Ähnlichkeit. In: Zeitschrift für philosophische Forschung, 1976, Heft-Nr. 50, S. 286–290, hier S. 287.

3 Ebd., S. 290. Dabei wäre ein Hinweis auf die empiristische Erkenntnistheorie Lockes und die der englischen Aufklärung, insbesondere die in der Ähnlichkeit fundierte Idee der Assoziation, wie sie vor allem bei Hume entfaltet wird, sicherlich angebracht.

4 „Ähnlichkeit“. In: Historisches Wörterbuch der Philosophie, Basel 1971, Sp. 114–115.

5 „Analogie“. In: ebd., Sp. 214–229.

6 Douglas Hofstadter und Emmanuel Sander: Die Analogie. Das Herz des Denkens, Stuttgart 2014, S. 17.

7 Paul Valéry: Cahiers/Hefte, Frankfurt am Main 1989, Bd. 3, S. 121.

8 Michel Foucault: Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften, Frankfurt am Main 1974, S. 63.

9 Albrecht Koschorke: Ähnlichkeit. Valenzen eines post-postkolonialen Konzepts. In: Anil Bhatti und Dorothee Kimmich (Hg.): Ähnlichkeit. Ein kulturtheoretisches Paradigma, Konstanz 2015, S. 35–45. Siehe Abb. 6.

10 Vgl. neben Hofstadter (s. Anm. 6) Olaf Bredbach: Analoge Anthropologien. Zur Reanimierung des Mikro-Makrokosmos-Denkens im 17. Jahrhundert. In: Michael Eggers (Hg.): Von Ähnlichkeiten und Unterschieden. Vergleich, Analogie und Klassifikation in Wissenschaft und Literatur (18./19. Jahrhundert), Heidelberg 2011, S. 33–52, vor allem S. 37. Siehe Abb. 4.

11 Vgl. Lena Bader, Martin Gaier und Falk Wolf (Hg.): Vergleichendes Sehen, München 2010, siehe Abb. 3. Zur Philosophie der Ähnlichkeitsbestimmungen vgl. auch Felix Thürlemann: Mehr als ein Bild. Für eine Kunstgeschichte des hyperimage, München 2013.

Abb. 7. Entwicklungsgeschichte von Mensch und Wirbeltier (Ernst Haeckel: Anthropogenie oder Entwicklungsgeschichte des Menschen. Keimes- und Stammesgeschichte, Leipzig 1877, Tafel V)



deren unterschiedlichen Gestaltungsweisen bei ähnlicher Gattungszugehörigkeit aus und fasst die Unterschiede in den Kategorien „linear“ und „malerisch“ (Abb. 8), jedoch auch in vier weiteren gegensätzlichen Formbestimmungen zusammen. Diese Kategorien sind universell angelegt und sollen – über die Gattungszugehörigkeit hinaus, die Wölfflin vor allem aus pädagogischen Gründen wählte – allgemeine Charakteristika der Renaissance- und Barockkunst bestimmen. Die Anschaulichkeit seiner Kategorien wird besonders wirksam, wenn sie an einem identischen Motivbestand demonstriert wird: Hier zeigt sich eindrucklich der Unterschied zwischen einer scharfen, malerischen Definition des dargestellten Antlitzes und seiner verwaschenen Unschärfe. Mit der Methode, die bei Wölfflin zugleich ihren Ausgangs- und ihren Höhepunkt findet, schreibt sich die Kunstgeschichte in eine Wissenschaftspraxis ein, die als vergleichende um 1900 vielleicht ihre größten Triumphe feiert, ihren Ausgangspunkt aber im Denken der Romantik, ja sogar im klassifikatorischen Zeitalter der Aufklärung hat. Ein Kronzeuge dafür ist Ernst Haeckel, der in den 1899 erschienenen *Welträtseln* seine monistische Theorie als Fortführung von Darwins Evolutionslehre darlegt.¹² In der Ähnlichkeit der embryonalen Entwicklungsgeschichte von Mensch und Wirbeltier begründet sich sowohl die Einheit der Lebenswelt als auch deren Einbindung in eine gemeinsame Abstammung. Der Affe als Vorläufer des Menschen wird in diesem Zusammenhang nur als letzte Stufe einer Geschichte präsentiert, die auf die urweltlichen, vor Hunderten von Millionen Jahren entstandenen Einzeller zurückgeht (Abb. 7). Entsprechend ist der Vergleich in allen Bereichen kunsthistorischer Beschreibung fundamental: In der Stilgeschichte entscheidet die Bestimmung formaler Eigenheiten über die Zuordnung künstlerischer Phänomene zu Stilhaltungen. Auch in der

¹² Ernst Haeckel: *Die Welträtsel. Gemeinverständliche Studien über monistische Philosophie*, Bonn 1899.

Abb. 8. Die Unterschiede in den Kategorien des „Linearen“ und des „Malerischen“ am Beispiel von Albrecht Dürers Bernhard von Reesen (1521; a) und Frans Hals' Porträt eines jungen Mannes (1646/1648; b)

a



b



Ikonografie ergeben sich inhaltliche Festlegungen aus der Beobachtung vergleichbarer – oder ähnlicher – Darstellungs-konventionen: Zwei männliche Figuren können beispielsweise zu einem heiligen Bartholomäus werden, wenn beide ein Messer in der Hand halten und sich somit im Attribut des Messers ähneln. Selbst Aby Warburgs Tafeln in seinem Bilderatlas *Mnemosyne* sind nach Ähnlichkeitskriterien geordnet (Abb. 9).¹³ Letztlich wird man dem Kunsthistoriker George Kubler nur

Abb. 9, Tafel 39 aus der Serie von Aby Warburgs Bilderatlas *Mnemosyne* (1925/1929, Rekonstruktion des Originals)



¹³ Roberto Ohrt und Axel Heil (Hg.): Aby Warburg. Bilderatlas *Mnemosyne*. The Original, Berlin 2020.

zustimmen können, der in seinem strukturalistisch inspirierten Buch *The Shape of Time* aus dem Jahr 1962 behauptet:

„We can grasp the universe only by simplifying it with ideas of identity by classes, types, and categories and by rearranging the infinite continuation of non-identical events into a finite system of similitudes. It is in the nature of being that no event ever repeats, but it is in the nature of thought that we understand events only by the identities we imagine among them.“¹⁴

Was bei Valéry eine Eigenschaft der Wirklichkeit war, wird bei Kubler subjektiviert und zu einem Vorgehen menschlicher Erkenntnis.

Das vorliegende Heft widmet sich diesem, in vielerlei Hinsicht fließenden Begriff der Ähnlichkeit – insbesondere seiner Anwendbarkeit und Produktivität für das *digitale* Bild. Roland Meyer vertritt in seinem Beitrag die These, dass Ähnlichkeit, wie sie heute in der algorithmischen Mustererkennung operationalisiert wird, historisch verwurzelt ist: Es handelt sich um eine statistische Größe, die als quantifizierbarer Abstand in einem virtuellen Raum konzeptualisiert wird, Ähnlichkeit also operationalisierbar, indem sie quantifiziert und räumlich abgebildet wird. In diesem virtuellen Raum stellt jedes Bild einen Vektor dar, dessen relative Position zu den anderen Vektoren die zugrunde liegenden relationalen Beziehungen offenbart. Welche gemeinsamen Merkmale als relevant empfunden werden und den Eindruck von Ähnlichkeit erzeugen, hängt vom jeweiligen Kontext ab – etwa von der Situation, dem kulturellen Hintergrund und den individuellen Wahrnehmungsmustern der betrachtenden Person. Der Beitrag von Stefanie Schneider, Desiree Frontczak, Julia Wolf und Theresa Zischkin zerlegt die Prozesse der Ähnlichkeitsfindung empirisch in

¹⁴ George Kubler: *The Shape of Time. Remarks on the History of Things*, New Haven und London 1962, S. 67.

ihre Einzelschritte, um die beobachtbaren *Dimensionen* – also die vielfältigen Eigenschaften, anhand derer Objekte verglichen werden – systematisch zu erfassen. Die Analyse regt dazu an, traditionelle Vergleichsmaßstäbe zu hinterfragen und die individuellen Momente der Ähnlichkeitsbewertung zu überdenken. Mit jenen Dimensionen der Ähnlichkeit, die sich aus dem Wechselspiel von kulturell geprägter und individueller Wahrnehmung ergeben, beschäftigt sich auch Hubertus Kohle in einem Praxisbericht zum e-Research-Werkzeug iART (interaktives Analyse- und Retrieval-Tool).¹⁵ Anhand experimenteller Suchanfragen zeigt sein Beitrag, wie ein auf maschinellen Lernverfahren basierendes System aus Tausenden von Datenpunkten Ähnlichkeitsbeziehungen herausfiltert, die anschließend von menschlicher Intelligenz auf ihre Plausibilität für die jeweilige Suchanfrage bewertet werden. Das Heft schließt mit einem Interview von Jon Emmony und Silvia Weidenbach über den Einfluss der Ähnlichkeit auf den kreativen Forschungsprozess: In ihrem gemeinsamen Projekt werden historische Schmuckringe in eine künstliche Intelligenz eingespeist, die deren Materialität und Form analysiert und daraus neue Ringe generiert – die wiederum, als Reflexion der Ähnlichkeit, gleichermaßen vertraut und fremd erscheinen.

Nachbarn und Verwandte. Zur Operationalisierung der Ähnlichkeit

„Let’s strip all the meta data“ – mit dieser lakonischen Formel leitete Amit Sood, Leiter des Google Cultural Institute, zu einem der interessantesten Abschnitte seines vielbeachteten TED-Talks von 2016 über.¹ Zuvor hatte Sood nicht nur die beeindruckenden Zahlen von Googles weltumspannendem *Art Project* heruntergerattert – rund 1.000 beteiligte Institutionen in 68 Ländern, mehr als sechs Millionen digitalisierte Artefakte –, sondern auch einen spektakulären „Cultural Big Bang“ inszeniert: Millionen digitaler Reproduktionen von Kunstwerken aus aller Welt, angefangen bei der über 200.000 Jahre alten *Venus von Berekhat Ram*, füllten innerhalb von Sekunden die riesigen Projektionsleinwände hinter seinem Rücken und bildeten eine Galaxie winziger Rechtecke in einer grenzenlosen schwarzen Weite, ließen sich nach Jahreszahlen, Ursprungsländern und Künstler*innennamen sortieren und zu immer neuen Formationen arrangieren (Abb. 1). Doch all dies hängt eben noch an den Metadaten – jenen nichtvisuellen Informationen also, die den digitalen Bilddateien von den sammelnden Institutionen beigefügt wurden. Was aber bleibt, wenn man von diesen Metadaten absieht? Welche anderen Möglichkeiten der Navigation in diesen endlosen Bilderweiten kann es noch geben?

Soods Antwort lautete, wenig überraschend: „[...] let’s look at what machine learning can do based purely on visual recognition of this entire collection.“ Im Wortsinne rein visuell, „purely visual“, läuft das natürlich nicht ab. Computer lernen

¹ Amit Sood: Every piece of art you've ever wanted to see - up close and searchable, TED 2016, https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable [Stand 02/2025], im Folgenden zitiert nach dem dort abrufbaren Transkript. Vgl. dazu auch kritisch: Ulrich Pfisterer: Big Bang Art History. In: Merkur, 2017, Heft-Nr. 816, S. 95-101.

Abb. 1, Szene aus Amit Soods TED-Talk „Every piece of art you've ever wanted to see - up close and searchable“, Screenshot 10:05, 02.2016, https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable [Stand 02/2025]



nicht zu ‚sehen‘, sie werden lediglich darauf trainiert, Datenmengen abzugleichen und Übereinstimmungswahrscheinlichkeiten zwischen Mustern zu erkennen. Worauf Soods Ankündigung mithin abzielte, war, was man ähnlichkeitsbasierte Bildersuche nennt: Die Bildinformationen selbst, nicht die textuellen Metadaten, sollen den Schlüssel zu ihrer Sortierung und Gruppierung bilden.² Einmal darauf trainiert, vermögen die Algorithmen wie auf Knopfdruck ganze Bildcluster allein aus Porträts oder Pferdedarstellungen zu versammeln – egal, ob diese Gattungsbezeichnung den Bilddaten bereits beigelegt worden war oder nicht. Neue, gewissermaßen sprachlose Formen der Suche in den Bilddatenmengen werden so möglich. Begeistert demonstrierte Sood, wie selbst das eigene

² Für eine frühe Auseinandersetzung mit den Perspektiven ähnlichkeitsbasierter Bildersuche vgl. Wolfgang Ernst, Stefan Heidenreich und Ute Holl (Hg.): Suchbilder. Visuelle Kultur zwischen Algorithmen und Archiven, Berlin 2003.

Gesicht zum Suchkommando werden kann – in Echtzeit findet der „portrait matcher“ zu jedem mimischen Ausdruck ein passendes Pendant aus den digitalen Sammlungsbeständen. „[S]omething fun for kids“ sei das, so könnten sich auch die Jüngsten für Porträtsammlungen begeistern. Unter dem Namen Art Selfie lässt sich mittlerweile eine entsprechende App auf Googles Art & Culture-Plattform herunterladen.

All das ist mehr als nur eine Spielerei. Vielmehr führt Soods TED-Talk anschaulich vor Augen, welche Rolle algorithmischer Mustererkennung heute beim Umgang mit riesigen Bildermengen zukommt. Als digital mobilisierte Bilddatensätze rücken beliebige museale Sammlungsobjekte, weitgehend ungeachtet ihrer je spezifischen Materialität und Medialität, ihres ursprünglichen kulturellen Kontextes wie ihrer Sammlungsgeschichte, in einen homogenen und kontinuierlichen Raum des Vergleichs, der sich auf Basis vermeintlich „rein visueller“ Ähnlichkeiten organisieren und navigieren lässt. Solche Vergleichsräume finden sich mittlerweile in vielen Interfaces musealer Sammlungsbestände, und sie bilden auch die Grundlage einschlägiger Bemühungen der „visuellen Erschließung“ großer Bilddatenmengen in den Digital Humanities.³ Die Suche nach visuellen Ähnlichkeiten kommt dabei immer dort zum Tragen, wo es gilt, Ordnungen von Bildern zu etablieren, die jenseits sprachlich verfasster Kategorien angesiedelt sind. Was dabei allerdings heute in Techniken der algorithmischen Mustererkennung als Ähnlichkeit operationalisiert wird, so die These dieses Essays, ist durchaus historisch voraussetzungsvoll: Es ist eine statistische Ähnlichkeit, die als quantifizierbarer Abstand in einem virtuellen Raum konzeptualisiert wird. Ähnlichkeit wird, anders gesagt, operationalisierbar, indem sie quantifizierbar und verräumlicht wird – und der Genealogie dieser Quantifizierung und Verräumlichung möchte ich im Folgenden nachgehen.

³ Peter Bell und Björn Ommer: Visuelle Erschließung. Computer Vision als Arbeits- und Vermittlungstool. In: Andreas Bienert, Anko Börner, Eva Emenlauer-Blömers und James Hemsley (Hg.): Konferenzband EVA Belin 2016. Elektronische Medien & Kunst, Kultur und Historie, Berlin 2016, S. 67-73.

Vermessene Ähnlichkeiten

Was Ähnlichkeit ist, woran sie sich zeigt und was sie bedeutet, ist historisch nicht stabil: „Da Ähnlichkeit ihrerseits eine Idee ist, wurde sie in ihrer Geschichte immer wieder neu definiert“, so hat es Hans Belting formuliert.⁴ Mindestens ebenso historisch wandelbar wie die Ideen über Ähnlichkeit sind die Medien und Praktiken, derer man sich bedient, um Ähnlichkeiten festzustellen: Ähnlichkeit hat eine Ideen- wie eine Mediengeschichte, die sich nicht aufeinander reduzieren lassen, aber eng miteinander verschränkt sind. Was die Ideengeschichte betrifft, so hat Michel Foucault der Ähnlichkeit in seiner Archäologie der neuzeitlichen abendländischen Episteme, in der *Ordnung der Dinge* von 1966, bekanntlich eine eindeutige Zeitschicht zugewiesen. Bis zum Beginn des 17. Jahrhunderts schien die Welt aus einem Netz von Nachbarschaften, Dopplungen, Analogien und Sympathien gewoben. Wer die sichtbaren Ähnlichkeiten zwischen Mikro- und Makrokosmos, dem menschlichen Körper und den Dingen der Natur zu lesen vermochte, so das Versprechen der frühneuzeitlichen Signaturenlehren, hielt den Schlüssel zur Erkenntnis der „prosaischen Welt“ in der Hand.⁵ Nichts scheint einem modernen Wissenschaftsverständnis ferner zu liegen als ein solch emphatisches Ähnlichkeitsdenken. Und doch ist die Ähnlichkeit keineswegs vom „Horizont der Erkenntnis“⁶ verschwunden – sie hat nur ihre Gestalt verändert. An die Stelle eines intuitiv erfassbaren Netzes sinnlicher Ähnlichkeiten, das den ganzen Kosmos mit Bedeutung durchzieht, sind Medientechniken des Vergleichens getreten, die objektivierbare Übereinstimmungen und Differenzen auswertbar machen sollen.

⁴ Hans Belting: *Bild-Anthropologie. Entwürfe für eine Bildwissenschaft*, München 2002, S. 39.

⁵ Michel Foucault: *Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften*, Frankfurt am Main 1974, S. 46–77.

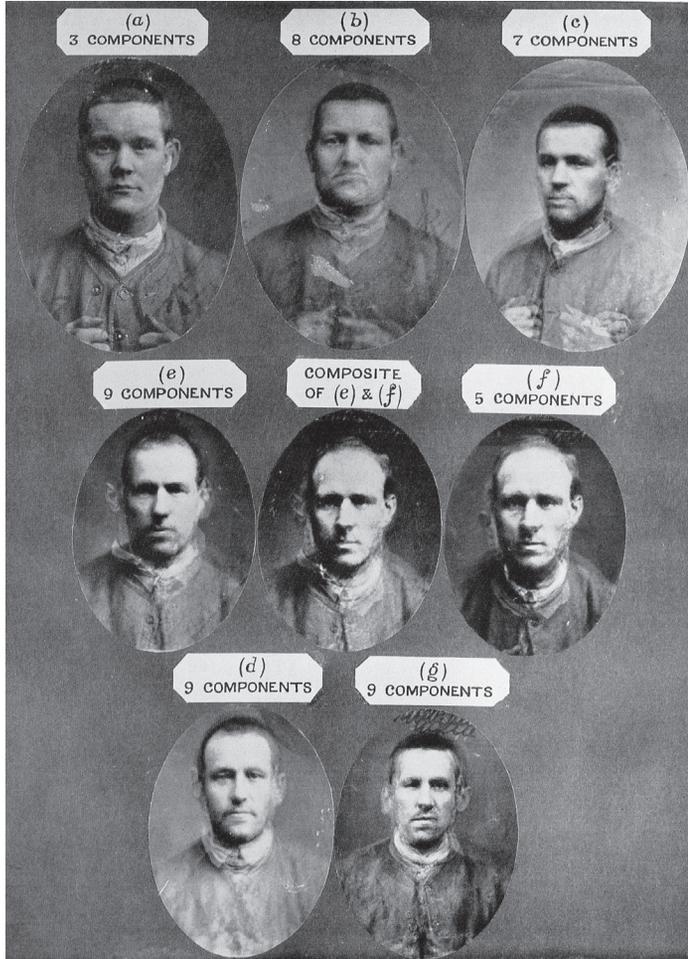
⁶ Ebd., S. 46. Vgl. auch die Einführung zu diesem Heft.

Eine Figur, die wie kaum eine andere an der Schwelle dieses modernen Ähnlichkeitsparadigmas steht, ist der Zürcher Pfarrer Johann Caspar Lavater. Ideengeschichtlich erscheint Lavaters Physiognomik, die ihn im ausgehenden 18. Jahrhundert zu einer europäischen Berühmtheit werden ließ, wie ein verspäteter Nachzügler einer älteren Episteme: Der innere Charakter eines Menschen, davon war er überzeugt, zeige sich an den äußeren Merkmalen des Körpers, und dort bevorzugt am Gesicht. Allerhöchste Vollkommenheit der Seele wiederum gehe mit größtmöglicher Christusähnlichkeit einher, weswegen Lavaters physiognomische Charakterkunde zugleich der Suche nach dem wahren Antlitz Christi gewidmet war. Medientechnisch allerdings präsentiert sich dieses Unternehmen auf der Höhe der Zeit. Das 18. Jahrhundert war, mit einem Ausdruck von Roland Kanz, eine „porträtdurstige“⁷ Epoche, und die vier reich illustrierten Bände von Lavaters *Physiognomischen Fragmenten* (1775–1778) verdanken ihren publizistischen Erfolg auch der zeitgenössischen Gesichterversessenheit. Nicht zuletzt die Mode des Schattenrisses hatte es möglich gemacht, selbst ohne jegliche künstlerische Begabung die individuellen Züge eines Profils rasch und preisgünstig im Bild festzuhalten, wovon Lavater reichlich Gebrauch machte. Gerade weil er frei von jeglicher künstlerischen Interpretation war, galt ihm der Schattenriss als Mittel der getreuen Fixierung und des exakten Vergleichs objektiv messbarer Ähnlichkeiten. Zugleich etablierte er damit einen neuen, zuvor unbekanntem Blick auf das Porträt: Statt im Bild die Begegnung mit einem individuellen Gegenüber zu suchen, wurden ihm Bilder von Gesichtern zu Operationsflächen, die sich exakt vermessen und in einzelne Segmente zergliedern ließen. Mit Lavater wird Physiognomik zum massenhaften Bildvergleich auf Basis isolierbarer Merkmale.⁸

⁷ Roland Kanz: *Dichter und Denker im Porträt. Spurengänge zur deutschen Porträtkultur des 18. Jahrhunderts*, München 1993, S. 11.

⁸ Vgl. dazu ausführlich (und mit weiterer Literatur) Roland Meyer: *Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook*, Konstanz 2019.

Abb. 2, Francis Galton, Kompositporträts des „kriminellen Typus“, um 1880, aus: Karl Pearson: The Life, Letters and Labours of Francis Galton, Vol. II, Researches of Middle Life, Cambridge 1924, Tafel XXVIII



Es ist dieser operative, ebenso vermessende wie vermessene Blick auf das im Bild festgehaltene Gesicht, der im 19. und 20. Jahrhundert eine ganz Fülle von (Pseudo-)Wissenschaften der Ähnlichkeit prägen sollte, von der Kriminalanthropologie Cesare Lombrosos bis zur NS-Rassenhygiene.⁹ An die Stelle des Schattenrisses trat dabei ab der Mitte des 19. Jahrhunderts die Fotografie, und unter dem Eindruck von Darwins Evolutionslehre galten Ähnlichkeiten nun nicht mehr als Anzeichen göttlicher Vorsehung, sondern als Merkmale biologischer Verwandtschaft. Was Lavaters Nachfolger jedoch von ihm übernahmen, war der Wunsch, Abbildung und Vermessung in einer einzigen Operation zu vereinen. Bei kaum jemand anderem erscheint dieser Wunsch so ausgeprägt wie bei Francis Galton. Der Vetter Charles Darwins und Begründer der Eugenik war geradezu davon besessen, menschliche Fähigkeiten umfassend zu vermessen, zu verdaten und auf biologische Vererbung zurückzuführen, und konsequenterweise richtete sich sein Wille zur Quantifizierung auch auf physiognomische Ähnlichkeiten. Berühmt wurden seine Kompositporträts aus den 1880er Jahren, in denen er unterschiedliche fotografische Porträts so übereinander belichtete, dass eine Art optischer Durchschnitt dabei herauskam. Mit Hilfe dieses Verfahrens wollte Galton nicht nur das vermeintlich „typische“ Gesicht des „geborenen Verbrechers“ im Bild fixieren, sondern ebenso verwandtschaftliche Ähnlichkeiten innerhalb einer Familie (Abb. 2). Auch wenn die unscharfen Konturen der geisterhaften Kompositporträts gerade keine Vermessung von Merkmalen erlaubten, war er dennoch davon überzeugt, in der Verbindung von Fotografie und Statistik ein Werkzeug zur Objektivierung der Ähnlichkeit gefunden zu haben: Die gemeinsame Ähnlichkeit eines vermeintlichen physiognomischen „Typus“ sollte dabei dem

⁹ Einschlägig dazu Richard T. Gray: About Face. German Physiognomic Thought from Lavater to Auschwitz, Detroit 2004.

statistischen Durchschnitt der fotografisch erfassten Gesichter entsprechen, und die individuellen Züge wurden so zur bloßen statistischen Abweichung von diesem Mittelwert.¹⁰

Galtons fotografisch-physiognomische Experimente wären kaum mehr als eine wissenschaftshistorische Fußnote wert, würden sie nicht technisch aktualisiert seit einigen Jahren eine Art Renaissance erleben. Regelmäßig kursieren sensationalistische Meldungen darüber, wie es mittels maschinellen Lernens inzwischen möglich geworden sei, sexuelle Präferenzen, politische Einstellungen oder, in schlechtester Tradition des 19. Jahrhunderts, selbst die Neigung zur Kriminalität aus Selfies und Social-Media-Profilbildern herauszulesen. Solche digitale Neophysiognomik überführt Galtons Verbindung von Fotografie und Statistik ins Zeitalter von Big Data und Deep Learning: Digitale Mustererkennung soll nun in riesigen Mengen online verfügbar gewordener Bilder von Gesichtern vermeintliche Ähnlichkeiten aufspüren, die dann als typische Merkmale angeblich mit spezifischen Neigungen und Haltungen korrelieren.¹¹ Die so vermessenen Ähnlichkeiten sind allerdings in der Regel keine zwischen distinkten anatomischen Merkmalen, die sich, wie etwa Augenabstände oder Nasenlängen, auch an lebendigen Gesichtern vermessen ließen – es sind, ähnlich wie sich Galton das erträumte, aber in der analogen Fotografie nicht verwirklichen konnte, Korrelationen zwischen Helligkeitsverteilungen. Denn erst im digitalen Bild, das aus diskreten Pixelmengen besteht, lassen sich auch solche unscharfen Ähnlichkeiten auf Basis visueller Muster quantifizieren.

Mustererkennung zeigt sich hier im doppeldeutigen Sinne als *pattern discrimination*, als diskriminierende Unterscheidung und Sortierung sozialer Gruppen auf Basis von Ähnlichkeiten und Differenzen.¹² Mit Galtons eugenischem Denken, das hat

¹⁰ Zu Galtons Kompositfotografie vgl. Meyer 2019 (s. Anm. 8); Josh Ellenbogen: *Reasoned and Unreasoned Images. The Photography of Bertillon, Galton and Marey*, University Park (Pa.) 2012.

¹¹ Vgl. dazu Roland Meyer: *Gesichtserkennung. Vernetzte Muster, körperlose Masken*, Berlin 2021, S. 32–39.

¹² Vgl. Clemens Apprich, Wendy Hui Kyong Chun, Florian Cramer und Hito Steyerl (Hg.): *Pattern Discrimination*, Lüneburg und Minneapolis (Minn.) 2018.

Wendy Chun kürzlich gezeigt, hat dies weit mehr zu tun, als es vielleicht den Anschein hat, stammen doch ganz wesentliche mathematische Konzepte, die bis heute die massenhafte algorithmische Datenauswertung prägen, von Galton und seinen direkten Nachfolgern – bis hin zum Konzept der Korrelation selbst: „Modern statistical pattern recognition expands upon early work to discriminate between populations.“¹³ Und dient nicht letztlich jede Quantifizierung von Ähnlichkeiten, selbst in so vermeintlich harmlosen Beispielen wie Googles algorithmischer Bildermengensortierung, der Kontrolle von „Populationen“, die zu bloßen Datenpunkten abstrahiert wurden? Die Bildpopulationen, die sich auf Soods Kommando hin zu Clustern der Ähnlichkeit formieren, sind in dieser Hinsicht nur der spektakuläre Showcase für Techniken der Mustererkennung, die ihre viel gewinnbringenderen, wenngleich mit weit weniger kulturellem Kapital verbundenen Anwendungsfelder im digitalen Marketing und der algorithmischen Massenüberwachung finden.

Latente Nachbarschaften

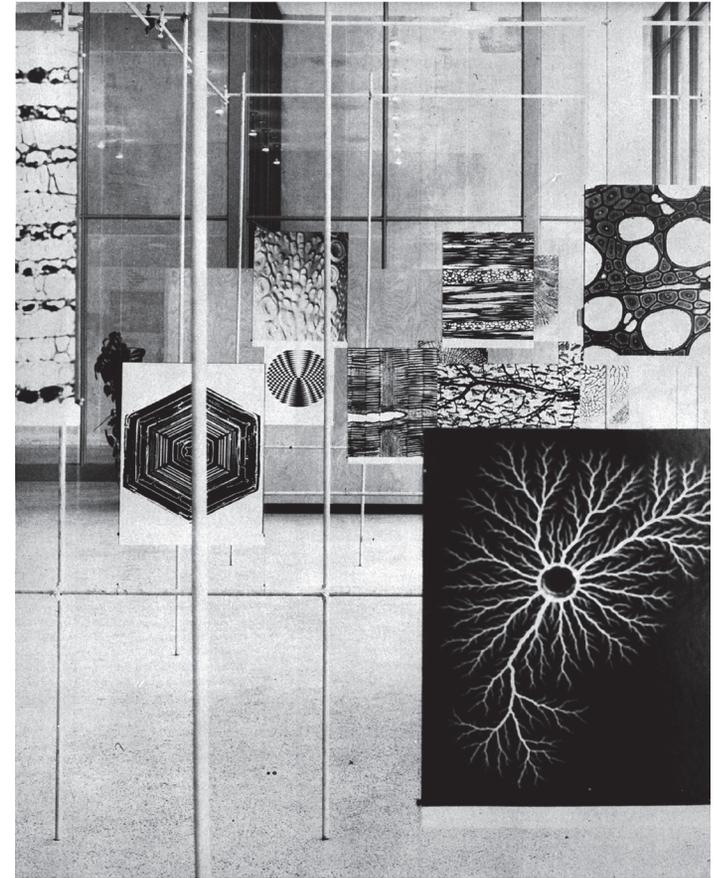
Algorithmische Mustererkennung in Bildermengen setzt nicht allein voraus, dass sich Ähnlichkeiten statistisch quantifizieren lassen. Sie setzt ebenso voraus, dass sich Ähnlichkeiten und Unähnlichkeiten als Nähe- und Distanzverhältnisse in einem homogenen Raum des Vergleichs modellieren lassen. Was Soods TED-Talk vorführte, war nur die bildmächtige Inszenierung von abstrakten mathematischen Konzepten, die den konzeptuellen Kern vieler heutiger Systeme der Bildgenerierung, Bilderkennung und Bildanalyse bilden. Jedes Bild, verstanden als digitale Matrix diskreter Pixelwerte, erscheint darin als Vektor in einem hochdimensionalen Koordinaten-

¹³ Wendy Hui Kyong Chun: *Discriminating Data. Correlation, Neighborhoods, and the New Politics of Recognition*, Cambridge (Mass.) 2021, S. 198.

ihrer je spezifischen Geschichte im selben abstrakten Raum zu verorten, erscheint Warburgs geschichtsgesättigter Ikonologie, die schließlich vor allem das Nachwirken der europäischen Antike verstehen wollte, letztlich fremd. Zudem waren die stets in Veränderung begriffenen Tafeln dazu gedacht, Warburgs Vorträge zu unterstützen, und nicht als abstrakte Diagramme, die objektivierte Ähnlichkeitsbeziehungen fixieren sollten. Nicht zuletzt widerspricht die Reduktion von Bildern auf statistisch auswertbare Informationsmuster seinem Konzept des Bildes als Speicher lebendiger Energien.¹⁶

Wollte man dagegen eine Genealogie jener grenzenlos digitalen Räume aus schwebenden Bildermengen schreiben, wie sie für heutige algorithmische Verfahren der Mustererkennung auf der operativen Ebene der Berechnung wie der repräsentativen Ebene der Visualisierung so charakteristisch scheinen, müsste man eher in die Zeit um 1950 zurückblicken. Hier nämlich, in der Hochzeit der Kybernetik, wurde es denkbar, Bilder nicht als singuläre Formen, sondern als massenhaft auswertbare Information zu betrachten: als diskrete Muster, ablösbar von ihrer je konkreten Materialität und Medialität. Was von Claude Shannon, Warren Weaver und Norbert Wiener theoretisch entwickelt wurde, hat schon früh auch künstlerische Formen der Inszenierung von Bildermengen inspiriert. Dafür steht exemplarisch Gyorgy Kepes' Ausstellung *The New Landscape*, die 1951 am MIT gezeigt wurde (Abb. 4). Lange bevor *pattern recognition* zum technischen Projekt wurde, sollte hier eine neue Form des Sehens eingeübt werden, die Kepes selbst *pattern-seeing* nannte: Wer seinen Blick über die wie schwerelos im Raum schwebenden Bilder zuckender Blitze, verästelter Adern und kristalliner Strukturen schweifen ließ, dem zeigte sich, so Kepes' Hoffnung, eine latente visuelle Ordnung von Ähnlichkeiten. Hatte sich die

Abb. 4, Gyorgy Kepes: *The New Landscape*, Ausstellungsansicht, MIT, 1951, aus: *Arts and Architecture*, Mai 1951, S. 21



¹⁶ Zur Ordnung der Bilder in Warburgs Bilderatlas vgl. bspw. Felix Thürlemann: *Mehr als ein Bild. Für eine Kunstgeschichte des hyperimage*. München 2013, S. 97-116.

Ausstellung zunächst nur auf naturwissenschaftliche Bildgebungen fokussiert, wurden in Kepes' nachfolgenden Buchprojekten dann tatsächlich potenziell alle Bilder Gegenstand einer solchen freischwebenden Mustererkennung, von den frühesten Höhlenmalereien bis zu den *drip-paintings* Jackson Pollocks.¹⁷

Konsequent fortgesetzt wurde diese Art der Verräumlichung mobil gewordener Bildinformationsmengen in den gleichermaßen von Kepes' *New Landscape* wie von Wieners Kybernetik inspirierten Ausstellungen der Londoner Independent Group, etwa *Parallel of Life and Art* von 1953 (Abb. 5). An unsichtbaren Nylonfäden aufgehängt und über Wände und Decken verteilt waren hier, in Form unkommentierter grobkörniger schwarz-weißer Vergrößerungen, ganz unterschiedliche, aus Büchern und Zeitschriften zusammengetragene Bilder zu sehen: archaische Masken, mikroskopische Aufnahmen, Reproduktionen informeller Malerei ebenso wie historische Archivaufnahmen. Sie bildeten eine Art raumgreifendes visuelles Netz, in dem das Ausstellungspublikum durch seine Bewegung im Raum immer neue Vergleichskonstellationen in den Blick nehmen konnte.¹⁸ Anders als in Warburgs ebenso gelehrsam wie idiosynkratischen Bildtafeln wurden hier tatsächlich voraussetzungslose Betrachtende adressiert, die in gleichsam von allen kulturellen Metadaten befreiten, schwerelosen *patterns* Konstellationen der Ähnlichkeit jenseits aller sprachlichen Kategorien erkennen sollten.¹⁹ Es ist dieser jedoch nur scheinbar voraussetzungslose Blick auf Bildermengen als Populationen von Mustern, der sich, algorithmisch implementiert, in Projekten wie Googles Art & Culture fortsetzt. Und mit der damit einhergehenden Verräumlichung von Bildermassen wird zugleich eine ganz bestimmte Bildvorstellung ins Werk gesetzt: die Vorstellung nämlich, dass zwischen zwei

Abb. 5, Nigel Henderson, Eduardo Paolozzi, Alison und Peter Smithson: *Parallel of Life and Art*, Ausstellungsansicht (Foto: Nigel Henderson), Institute of Contemporary Arts (ICA) London, 1953, aus: Kevin Lotery: *The Long Front of Culture. The Independent Group and Exhibition Design*, Cambridge (Mass.) und London 2020, S. 89



beliebigen Bildern immer noch weitere mögliche Bilder existieren. Ähnlichkeit wird hier als Kette latenter Nachbarschaften konzeptualisiert, eine Vorstellung, die etwa in experimentellen Sammlungsinterfaces wie *X Degrees of Separation* anschaulich wird. Hervorgegangen aus der Kooperation des Künstlers Mario Klingemann mit Google, erlaubt *X Degrees* die spielerische Exploration einer 250.000 Bilddatenobjekte umfassenden virtuellen Sammlung. Von jedem Objekt lässt sich dabei ein „Pfad“ von Ähnlichkeiten zu einem beliebigen anderen Objekt

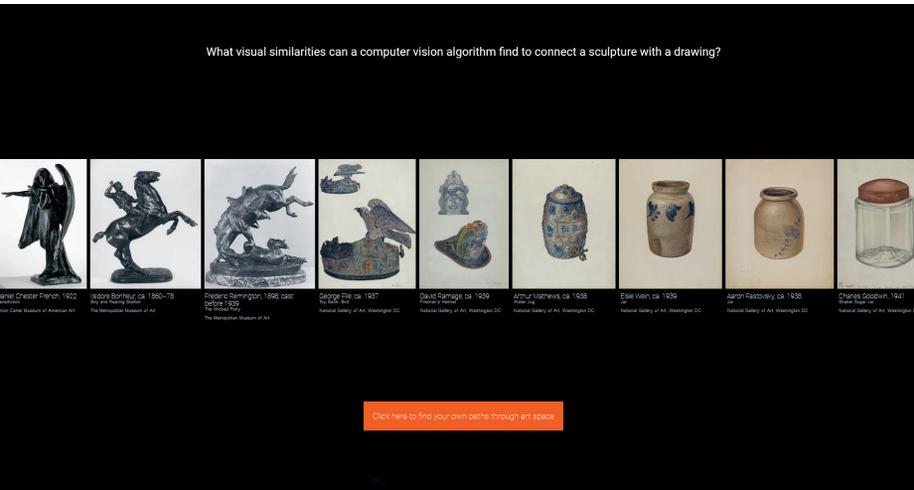
¹⁷ Zu Kepes' pattern-seeing vgl. Reinhold Martin: *The Organizational Complex. Architecture, Media, and Corporate Space*, Cambridge (Mass.) und London 2003, S. 42-79; John R. Blakinger: *Gyorgy Kepes. Undreaming the Bauhaus*, Cambridge (Mass.) und London 2019, S. 69-163.

¹⁸ Kevin Lotery: *The Long Front of Culture. The Independent Group and Exhibition Design*, Cambridge (Mass.) und London 2020.

¹⁹ Wollte man diese Genealogie fortsetzen, würde sie zum Beispiel über die Ausstellungsexperimente von Charles und Ray Eames um 1960 zu den Expanded-Cinema-Visionen von Gene Youngblood um 1970 führen. Vgl. dazu auch Fred Turner: *The Democratic Surround. Multimedia and American Liberalism from World War II to the Psychedelic Sixties*, Chicago 2013.

finden, und all diese Pfade, so wird impliziert, koexistieren in einem gemeinsamen „art space“ (Abb. 6). Doch die vermeintlichen Ähnlichkeiten, die dabei aufgespürt werden, sind vor allem solche zwischen Bilddateien, und das Interface unterscheidet nicht, ob ein heller Bildgrund, wie etwa bei einer Zeichnung, Teil der künstlerischen Bildfindung ist, oder, wie bei vielen Objekt-fotografien, dem Format der musealen Erfassung geschuldet.²⁰

Abb. 6, Mario Klingemann and Simon Doury (Google Cultural Institute): X Degrees of Separation, 2017, Screenshot, <http://artsexperiments.withgoogle.com/xdegrees/> [Stand 02/2025]



Schluss: Voraussetzungsvolle Ähnlichkeiten

Es gibt keine Ähnlichkeit ohne Vergleich. Ähnlichkeit ist daher stets voraussetzungsfull: Sie beruht auf medialen Praktiken des Vergleichs, und sie setzt eine Vergleichbarkeit des Unterschiedlichen voraus, die nie einfach gegeben ist, sondern erst hergestellt werden muss. Die aktuellen KI-basierten Formen des Vergleichs, wie sie Amit Sood und Mario Klingemann so eindrucksvoll in Szene zu setzen wissen, lassen leicht vergessen, wie voraussetzungsfull sie sind. Alles scheint heute miteinander vergleichbar, sofern es nur in den riesigen Datenbanken vorliegt. Was dabei unsichtbar gemacht wird, sind nicht zuletzt die Prozesse der Erfassung und Formatierung, die aus musealen Sammlungsobjekten digitale Bilddateien werden lassen. Erst diese aber ermöglichen überhaupt erst den Vergleich von visuellen Mustern. Dass etwa die *Venus von Berekrat Ram*, von der nicht einmal sicher ist, dass sie tatsächlich menschengemacht ist, die Gemälde Van Goghs und die aus Benin geraubten Elfenbeinmasken alle denselben homogenisierten virtuellen „art space“ teilen können, hängt nicht zuletzt vom standardisierten Bildformat ab, das für die fotografische Erfassung musealer Sammlungsobjekte typisch ist: Körperlichkeit wird dabei auf eine Fläche reduziert, Materialität zur visuellen Textur, Dimensionsunterschiede verschwinden und Formate werden einander angeglichen. Operationalisierung, das galt für Lavaters Silhouetten wie für Galtons Kompositporträts, setzt stets Formatierung voraus.

Formatierung ist jedoch niemals unschuldig: Die Prozeduren der musealen Erfassung und Erschließung, so hat es Ariella Aïsha Azoulay jüngst beschrieben, zerschneiden alle lebensweltlichen Beziehungsnetze, in die kulturelle Artefakte

einstmals eingebunden waren, und reduzieren alles, dessen sie sich bemächtigen, auf seine Sammelbarkeit und Ausstellbarkeit. An die Stelle diverser und komplexer kultureller Praktiken im Umgang mit kulturellen Artefakten treten so standardisierte Protokolle, die auf alle beliebigen Objekte gleichermaßen anwendbar sind.²¹ Die Massendigitalisierung musealisierter Artefakte bricht in der Regel nicht mit dieser imperialen und kolonialen Logik, sondern setzt sie vielmehr im Raum des Virtuellen fort: Die musealen Klassifikationen und Zuschreibungen, mit all ihren konstitutiven Lücken und Fehlstellen, werden digital verdoppelt und in Form von Metadaten operationalisierbar gemacht, aber nur selten kritisch hinterfragt, geöffnet oder gar revidiert.²² Die „absolut[e] Verfügbarkeit des Materials“, die etwa Ulrich Pfisterer als „Verheißung des Digitalen“ gilt,²³ setzt mithin die vorhergehende Zurichtung dieses Materials voraus. Nicht zuletzt reduziert sie die Vielfalt dessen, was alles als Bild gelten kann, auf eine isolierbare zweidimensionale Pixelmatrix – einen Bilddatensatz, mit dem sich rechnen lässt. Und damit wird zugleich ein eurozentrisches Bildverständnis als Basis des vermeintlich universellen Vergleichs visueller Ähnlichkeiten in technischen Formaten festgeschrieben.²⁴ Vielleicht gibt es daher kein besseres zeitgenössisches Bild für dieses Phantasma der absoluten Verfügbarkeit auf der Grundlage homogenisierender Formate und Protokolle als die endlosen Weiten von Googles „art space“: ein ortloser Raum freischwebender Bildcluster, die von einem körperlosen Blick souverän durchmessen werden können; ein weltloses Datenrepositorium, in dem alles Leben verschwunden scheint zugunsten operationalisierbarer Ähnlichkeiten; ein digitales Universalmuseum im Zeitalter des Datenextraktivismus.

32

21 Ariella Aïsha Azoulay: *Potential History. Unlearning Imperialism*, London 2019.

22 Eine bemerkenswerte Ausnahme in dieser Hinsicht stellt das 2022 zugänglich gemachte Online-Portal Digital Benin dar, vgl. dazu Roland Meyer: *Diesseits des Mausoleums. Über Digital Benin und die Zukunft virtueller Sammlungen*. In: *cargo. Film Medien Kultur*, 2023, Heft-Nr. 57, S. 46–52.

23 Ulrich Pfisterer: *Kunstgeschichte zur Einführung*, Hamburg 2020, S. 203.

24 Vgl. Jens Schröter: *Zum Eurozentrismus im Begriff des Bildes*. In: *Zeitschrift für Medienwissenschaft*, 2022, Heft-Nr. 26, S. 91–100.

Stefanie Schneider, Desiree Frontczak, Julia Wolf,
Theresa Zischkin

Dimensionen der Ähnlichkeit. Prozesse der Merkmalsselektion und -bewertung

Ähnlichkeit beruht darauf, dass zwischen zwei ähnlichen Objekten gemeinsame Eigenschaften bestehen. Wären zwei Objekte aber bereits dann ähnlich, wenn sie mindestens ein gemeinsames Merkmal aufwiesen, dann müsste nach Nelson Goodman jedes beliebige Objekt als einem anderen ähnlich gelten, da es immer irgendein den Objekten gemeinsames Merkmal gibt – und Ähnlichkeit wäre als Vergleichsmaß nutzlos.¹ Welche gemeinsamen Eigenschaften jeweils als relevant empfunden werden und den Eindruck der Ähnlichkeit hervorrufen, variiert allerdings je nach Kontext – der Situation, dem kulturellen Hintergrund und den individuellen Wahrnehmungsmustern der betrachtenden Person:

„[E]s gibt keine einheitlichen Verfahren oder Methoden, welche zu einem (positiven oder negativen) Ähnlichkeitsurteil führen. Die einschlägigen Vergleichspunkte können ebenso wechseln wie die Kategorien und die Gewichtungen. Worauf geachtet wird, wie es bewertet wird, hängt nicht nur von den zu vergleichenden Gegenständen, der Umgebung und dem Zweck des gesamten Verfahrens ab, sondern auch von den Kenntnissen, der Erfahrung und den Interessen des Betrachters.“²

Ähnlichkeit, so auch Douglas Arrell, manifestiert sich also im dynamischen Moment der Wahrnehmung: Zwei Objekte gelten dann als ähnlich, wenn ihre gemeinsamen Eigenschaften als ähnlich *erkannt* werden³ oder zumindest die mit ihnen

1 Siehe Nelson Goodman: *Seven Strictures on Similarity*. In: ders. (Hg.): *Problems and Projects*, New York 1972, S. 437–446, hier S. 443.

2 Oliver R. Scholz: *Bild, Darstellung, Zeichen. Philosophische Theorien bildlicher Darstellung*, Freiburg und München 1991, S. 45.

3 Douglas Arrell: *What Goodman Should Have Said About Representation*. In: *The Journal of Aesthetics and Art Criticism*, 1987, Heft-Nr. 46, S. 41–49, hier S. 42. Vgl. dazu auch: Anil Bhatti und Dorothee Kimmich: *Einleitung*. In: dies. (Hg.): *Ähnlichkeit. Ein kulturtheoretisches*

gemachten Erfahrungen. Erst der aktive Prozess der Merkmalsselektion und -bewertung macht Ähnlichkeit produktiv.⁴ Doch wie genau läuft dieser Prozess der Merkmalsselektion ab? Wie also können sowohl die zugrunde liegenden Wahrnehmungsprozesse als auch die spezifisch gewählten Vergleichsmerkmale transparent gemacht und kritisch reflektiert werden? Und wie zeigt sich Ähnlichkeit konkret – etwa zwischen kunsthistorischen Objekten? Diese Fragen laden dazu ein, den Prozess der Ähnlichkeitsfindung in seine Einzelschritte zu untergliedern.

Vor diesem Hintergrund konzipierten Studierende der Ludwig-Maximilians-Universität München im Wintersemester 2024/25 im Rahmen einer Lehrveranstaltung eine Online-Umfrage. Befragt wurden im Zeitraum von Januar bis Februar 2025 vor allem Studierende der Kunstgeschichte in München, ergänzt durch Mitarbeiterinnen und Mitarbeiter des dortigen Instituts für Kunstgeschichte. Den insgesamt 30 Beteiligten war somit eine grundlegende – teils bereits wissenschaftlich fundierte – Affinität zur differenzierten Beschäftigung mit dem kunsthistorischen Objekt gemein. Die Umfrage zielte darauf ab, die beobachtbaren Dimensionen der Ähnlichkeit – also die vielfältigen Eigenschaften, anhand derer kunsthistorische Objekte miteinander verglichen werden – systematisch zu erfassen. Auf diese Weise sollte ein Bewusstsein dafür geschaffen werden, wie fließend und schwer greifbar der Begriff Ähnlichkeit selbst in der Kunstgeschichte ist, einer sich fundamental mit der Gegenüberstellung von Objekten beschäftigenden Disziplin, und dass er, in einem komplexen Wechselspiel kulturell geprägter und individueller Wahrnehmung, fortwährend rekontextualisiert werden muss. Die Offenlegung der verschiedenen Dimensionen forderte dazu auf, traditionelle Vergleichsmaßstäbe zu hinterfragen und die individuellen Momente der Ähnlichkeitsbewertung aufzulösen.

Paradigma, Göttingen 2015, S.7-34, hier S.13.

⁴ Auch über diese Behauptung lässt sich diskutieren. Vgl. dazu kritisch: Sara Bangert: *Entgrenzte Ähnlichkeit im Milieu des Surrealismus. Konturen, Vorgeschichte und Konjunktur eines ästhetischen Konzepts*. Berlin 2023, S.59-61.

Im Fragebogen wird die jeweils zu betrachtende Dimension der Ähnlichkeit sukzessive verändert. Zunächst stehen die Komposition und Motivik von Kunstwerken im Vordergrund; anschließend rücken Farbgebung, Inhalt und schließlich der historische Kontext in den Mittelpunkt. Um den Prozess der Merkmalsselektion anhand der jeweils vorgegebenen Dimension aktiv und möglichst niederschwellig anzustoßen, ist den Frageblöcken eine Einleitungsseite vorangestellt, die den Wechsel zur neuen Dimension ankündigt und die Wahrnehmung ausschließlich auf diese fokussiert. Darüber hinaus folgt die Reihenfolge – und weitgehend auch die Ausgestaltung – der einzelnen, im Folgenden beleuchteten Frageblöcke einem einheitlichen Schema.⁵ Zunächst werden Fragen zur Beurteilung von Aussagen zu den jeweiligen Beziehungsverhältnissen gestellt. Darauf folgen offene Nennungen, in denen die Teilnehmenden gebeten werden, Unterschiede oder Gemeinsamkeiten der dargestellten Gemälde in Bezug auf die jeweilige Dimension zu artikulieren. Den Abschluss bildet ein Teil mit Fragen zur visuellen Einordnung, in dem vorgegebene Kategorien den jeweils relevanten Regionen im Bild zugeordnet werden sollen. Unabhängig von der jeweils betrachteten Dimension werden dabei immer dieselben drei Gemälde aus dem 17. Jahrhundert fokussiert: Bernardo Strozzi's *David mit dem Kopf von Goliath*, Artemisia Gentileschi's *Judith köpft Holofernes* und Alessandro Varotari's *Salome mit dem Haupt Johannes des Täufers* (Abb. 1). Die Ähnlichkeit dieser Werke ist bereits auf den ersten, unreflektierten Blick offensichtlich: In allen Darstellungen tritt aus einem dunklen, wenig detaillierten oder vagen Hintergrund eine Figur hervor, die beinahe den gesamten Bildraum ausfüllt. Die Szenen konzentrieren sich auf den Moment ein und derselben Handlung – einer Enthauptung, die entweder gerade vollzogen wird oder bereits stattgefunden hat.

⁵ Nicht für jede Dimension werden alle Frageblöcke bereitgestellt – bspw. wird die visuelle Einordnung ausschließlich für die Dimensionen Komposition und Farbgebung abgefragt.

Abb. 1: Die drei Gemälde aus dem 17. Jahrhundert, die im Fragebogen näher untersucht wurden: Bernardo Strozzi, David mit dem Kopf von Goliath (ca. 1636; a), Artemisia Gentileschi, Judith köpft Holofernes (um 1620; b), Alessandro Varotari, Salome mit dem Haupt Johannes des Täufers (o. J.; c)



Beziehungsverhältnisse

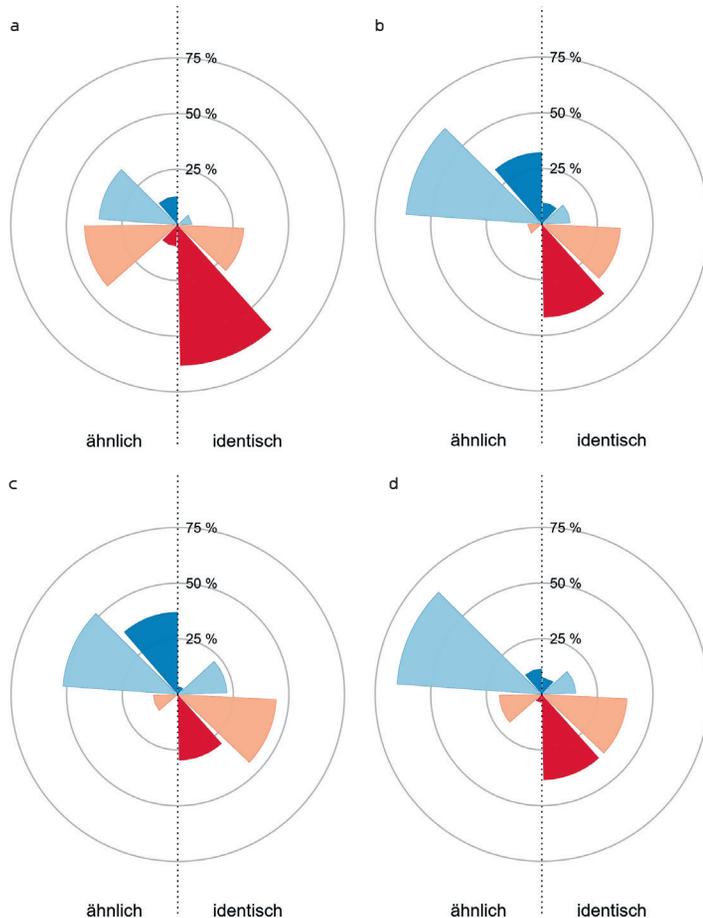
Das Ähnliche bewegt sich „zwischen Identität und Differenz“, wie Sara Bangert betont, „als Mittleres zwischen dem Einen, Selben und dem Anderen, Verschiedenen“.⁶ Der Übergang vom Ähnlichen zum als identisch Wahrgenommenen wird im jeweils ersten Frageblock jeder betrachteten Dimension in Beschlag genommen. Dabei wird bewusst darauf verzichtet, die Begriffe „Ähnlichkeit“ und „Identität“ explizit zu definieren; stattdessen werden sie als implizit vorausgesetzte Konzepte verstanden, die auf den individuellen Erfahrungen der Teilnehmenden basieren und der dem Fragebogen zugrunde liegenden Narration. Die Fragetexte sind auch hier unabhän-

gig von der jeweiligen Dimension einheitlich aufgebaut, um eine Vergleichbarkeit der Ergebnisse zu gewährleisten, wobei die Antworten auf einer vierstufigen Likert-Skala erfasst werden, die von „stimme gar nicht zu“ über „stimme eher nicht zu“ und „stimme eher zu“ bis hin zu „stimme voll und ganz zu“ reicht. Die Ergebnisse sind als Polardiagramme in Abb. 2 visualisiert. Jedes Diagramm zeigt einen radialen Vergleich zwischen den Kategorien: Der linke Bereich kennzeichnet, inwiefern die gezeigten Werke in der jeweiligen Dimension als „ähnlich“, der rechte Bereich, inwiefern sie als „identisch“ empfunden werden. Die einzelnen Ausprägungen der Skalenniveaus sind entlang konzentrischer Kreise dargestellt, die den Anteil der Teilnehmenden widerspiegeln, der für eine bestimmte Antwortstufe gestimmt hat.

Die Ergebnisse bestätigen einerseits die erwartete Diskrepanz zwischen Ähnlichkeit und Identität. Andererseits lassen sie auf eine unterschiedliche Wahrnehmung der Dimensionen schließen: So wird die Komposition der Werke als am wenigsten „identisch“ empfunden, während der Inhalt – obwohl ebenfalls im deutlich negativen Spektrum verortet – vergleichsweise als „identischer“ wahrgenommen wird. Dabei spielt offensichtlich eine Rolle, dass alle drei Werke Enthauptungen darstellen, allerdings jeweils unterschiedliche Zeitpunkte des Geschehens und unterschiedliche Erzählungen aus der christlichen Mythologie fokussieren (David und Goliath, Judith und Holofernes, Salome und Johannes der Täufer). Bemerkenswert ist, dass Unterschiede in der Ausprägung der Skalenniveaus zwischen den als „ähnlich“ und den als „identisch“ empfundenen Dimensionen miteinander korrespondieren. So lässt sich bereits anhand dieser kleinen Stichprobe empirisch belegen, dass es tatsächlich *graduelle* Abstufungen sind, die „das Ähnlichkeitsurteil allererst motivieren – als einen dynamischen

Abb. 2, Polardiagramme mit Unterschieden in der Wahrnehmung der Gemälde hinsichtlich Ähnlichkeit und Identität in vier Dimensionen: Komposition (a), Motivik (b), Inhalt (c) und historischer Kontext (d)

Bewertet wurde auf einer vierstufigen Skala: „stimme gar nicht zu“ (rot), „stimme eher nicht zu“ (hellrot), „stimme eher zu“ (hellblau) und „stimme voll und ganz zu“ (blau).



Grad der Nicht-Identität, der auf eine (potenziell unendliche) Zahl von Punkten auf einer Skala zwischen Gleichheit und Differenz verweist.“⁷

Öffene Nennungen

Die mangelnde Spezifität der einzelnen Ausprägungen verhindert jedoch konkrete Rückschlüsse auf die individuell zugrunde liegenden Wahrnehmungsprozesse. Um dennoch Einblicke in die subjektive Interpretation der Werke zu erhalten, werden die Teilnehmenden in einem zweiten Frageblock gebeten, in Freitextfeldern jeweils bis zu sechs Unterschiede oder Gemeinsamkeiten der dargestellten Gemälde in Bezug auf die jeweilige Dimension zu benennen. Damit sollen nicht nur standardisierte Bewertungen in vordefinierten Kategorien erfasst, sondern auch nuancierte, gegebenenfalls vom Konsens abweichende Wahrnehmungen sichtbar gemacht werden. Insgesamt wurden 218 offene Nennungen übermittelt, wobei die Anzahl der Nennungen pro Dimension teilweise stark variiert: 80 Nennungen erfolgten für die Komposition, 61 für die Farbgebung, 47 für den Inhalt und 30 für den historischen Kontext der Werke. Dies ist weniger auf die jeweilige Dimension als auf deren Positionierung im Fragebogen zurückzuführen: Zu Beginn abgefragte Dimensionen (wie Komposition und Farbgebung) führten zu ausführlicheren Nennungen, während später abgefragte Dimensionen (wie Inhalt und historischer Kontext) – vermutlich aufgrund einsetzender Ermüdung – weniger häufig thematisiert wurden.

Aufgrund des Umfangs und der fehlenden Standardisierung der Nennungen können diese naturgemäß nur stichprobenartig untersucht werden. Ein Versuch wird in Abb. 3 anhand einer ‚Pinnwand‘ unternommen, die eine Auswahl der offenen Nen-

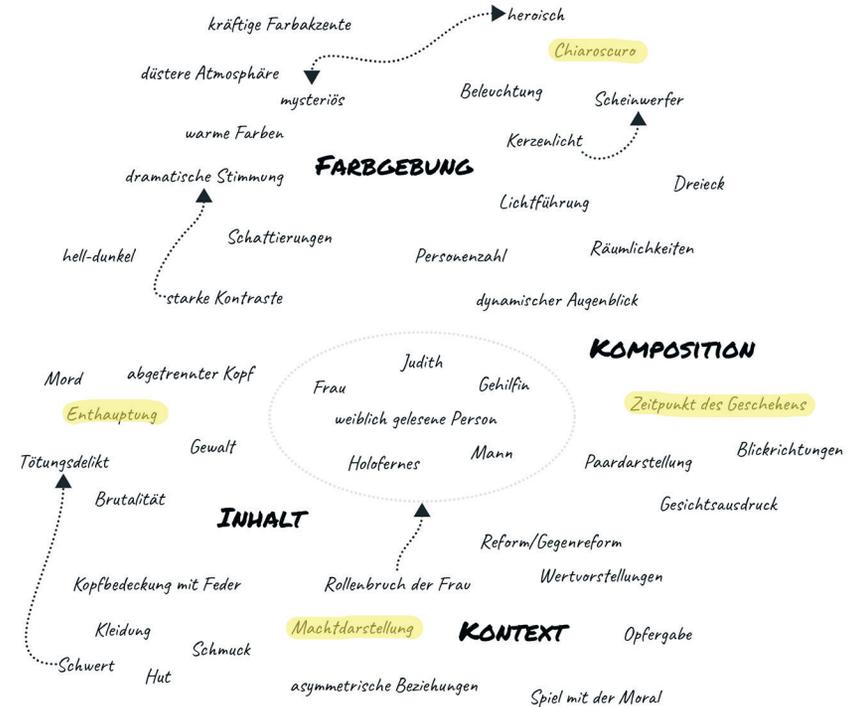
⁷ Johannes Endres: Unähnliche Ähnlichkeit. Zu Analogie, Metapher und Verwandtschaft. In: Martin Gaier, Jeanette Kohl und Alberto Saviello (Hg.): Similitudo. Konzepte der Ähnlichkeit in Mittelalter und Früher Neuzeit, München 2012, S. 29-58, hier S. 34.

nungen der 30 Teilnehmenden zeigt. Dabei werden inhaltlich gleichwertige Antworten verschiedener Personen auf ihre Kernaussagen zusammengefasst, um möglichst viele Nennungen repräsentativ abzubilden. Längere Nennungen sind zudem auf ihre wesentliche Aussage gekürzt oder – bei ausführlicheren Beschreibungen, etwa zum Inhalt – in mehrere Aussagen aufgeteilt. Die Pinnwand gibt damit einen Einblick in die vielfältigen Wahrnehmungsprozesse, die bei der Betrachtung der Gemälde offensichtlich ausgelöst werden. Im Vordergrund stehen insbesondere Aspekte, die die Dramatik der Szenarien oder die emotionale Wirkung auf den Betrachter reflektieren („dynamischer Augenblick“ oder „düstere Atmosphäre“). Deutlich wird die zum Teil sehr unterschiedliche Expertise in Bezug auf die dargestellte Erzählung. Einige Teilnehmende benennen die Figuren explizit, wie zum Beispiel „Judith“ und „Holofernes“, während andere auf generische Bezeichnungen wie „Frau“ und „Mann“ zurückgreifen. Die Nennungen sind keineswegs immer trivial oder strikt auf eine einzelne Dimension beschränkt: So wird die Farbgebung mehrfach als „heroisch“ beschrieben, insbesondere in Verbindung mit ausgeprägten Hell-Dunkel-Kontrasten – ein Effekt, der auch mit dem Fachausdruck des *Chiaroscuro* mehrfach präzisiert wird. Begriffe wie „Rollenbruch der Frau“ weisen im historischen Kontext darauf hin, dass auch geschlechtsspezifische Interpretationen in der Rezeption für die Teilnehmenden bedeutsam sind.

Visuelle Einordnung

Der bislang überwiegend textuell geprägten Erfassungsperspektive wird in einem dritten Frageblock eine visuelle Komponente hinzugefügt, in der vorgegebene Kategorien den jeweils relevanten Bildregionen zugeordnet werden sollen. Die Teil-

Abb. 3: Pinnwand mit offenen Nennungen der Teilnehmenden, in denen Unterschiede oder Gemeinsamkeiten der drei Gemälde in Bezug auf die jeweilige Dimension artikuliert werden



nehmenden markieren zum Beispiel die Positionen von Haupt- und Nebenfiguren in den Gemälden und ordnen den jeweiligen Bildbereichen spezifische Farbwerte zu. Auf diese Weise wird nicht nur die räumliche Anordnung der dargestellten Elemente erfasst, sondern auch, ob und inwieweit sie miteinander interagieren. Anhand dieser Punkte lassen sich Muster erkennen, wie etwa die zentrale Platzierung von Hauptfiguren gegenüber peripher angeordneten Nebenfiguren oder der gezielte Einsatz bestimmter Farben zur Hervorhebung einzelner Bildbereiche.

Interessanterweise markieren die Teilnehmenden häufig Punkte im Zentrum des relevanten Elements – insbesondere, wenn es sich um eine flächige Darstellung handelt, die nicht von anderen Bildelementen verdeckt wird –, obwohl sie nicht explizit dazu aufgefordert werden. So wird der goldene Griff von Davids Schwert in Strozzi's Gemälde ebenso hervorgehoben wie die blutrote Draperie des im Schlaf überfallenen Holofernes bei Gentileschi (Abb. 4). Obwohl Judiths Dienerin deutlich hinter Judith positioniert ist, erscheint ihr blaues Gewand dennoch annotationswürdig. „Die Farbe ähnelt zunächst nichts“, führt dazu Ludger Schwarte aus,

„noch bevor sie ornamental wirken kann, ist sie Fleck, ein Aufblitzen des Zufalls. Farbe kann eine Eigenschaft von etwas [...] auf Bildern zu erkennen geben. Gerade in Bildern emanzipiert sie sich oft von Gegenständen, ordnet sie sich unter, lässt sie zweitrangig erscheinen oder verschwinden.“⁸

Diese Beobachtungen verdeutlichen einmal mehr, dass auch die visuelle Wahrnehmung von Farbe keineswegs objektiv ist, sondern von subjektiven und kulturell geprägten Interpretationen abhängt. Der goldene Griff des Schwertes wird nicht bloß als Farbfleck wahrgenommen, sondern als zentrales Gestal-

Abb. 4: Zuordnung von Farben zu den jeweils relevanten Bildregionen durch die Teilnehmenden zur Untersuchung der Farbwahrnehmung: Gelb, Rot, Grün und Blau



tungselement, das den Blick des Betrachters fesselt, wie die zahlreichen Annotationen belegen. Die insgesamt 532 zu den Bildern hinterlegten Punkte⁹ geben nicht nur Aufschluss über einen Grundkonsens in der Wahrnehmung, der nur geringe Variation aufweist, sondern auch über Elemente, die diskussionswürdig sind, weil sie sich nicht in die vorgegebenen Kategorien einfügen. Der Umhang der Salome in Varotaris Szenerie lässt sich eben nicht eindeutig als blau oder grün klassifizieren; vielmehr erscheint er in seiner farblichen Komplexität so indifferent, wie es auch die Markierungen in Abb. 4 widerspiegeln. Die visuelle Einordnung auf Kompositionsebene – basierend auf 411 Punkten¹⁰ – erlaubt weitere Rückschlüsse auf die kunsthistorische Expertise der Teilnehmenden. Die handlungstragenden Figuren, die für die jeweils porträtierten Enthauptungen verantwortlich sind – nämlich David, Judith und Salome – werden überwiegend als Hauptfiguren identifiziert (Abb. 5a).

⁸ Ludger Schwarte: *Pikturale Evidenz. Zur Wahrheitsfähigkeit der Bilder*, München 2015, S. 141.

⁹ Die Anzahl der Punkte, die Farben indizieren, ist bei den einzelnen Bildern relativ ausgeglichen: Für Bernardo Strozzi's David mit dem Kopf von Goliath sind es 176 Punkte, für Artemisia Gentileschi's Judith köpft Holofernes 178 und für Alessandro Varotaris Salome mit dem Haupt Johannes des Täufers 158.

¹⁰ Die Anzahl der Punkte, die sich auf die Komposition beziehen, ist auch hier bei den einzelnen Bildern relativ ausgeglichen, zeigt jedoch eine sukzessive Abnahme vom ersten zum letzten Bild: Bernardo Strozzi's David mit dem Kopf von Goliath erhält 169 Punkte, Artemisia Gentileschi's Judith köpft Holofernes 136 Punkte und Alessandro Varotaris Salome mit dem Haupt Johannes des Täufers 106 Punkte.

Abb.5, Zuordnung von Kategorien zu den jeweils relevanten Bildregionen zur Analyse der Komposition: Hauptfigur (a), Nebenfigur (b), zentrales Element (c) und Attribut (d)

a



b



c



d



Die Kategorisierung der jeweils enthaupteten Figur gestaltet sich dagegen weniger eindeutig: Insbesondere Goliath wird häufig als solche annotiert, während Johannes der Täufer selten als Haupt- oder Nebenfigur, sondern vorwiegend als „zentrales Element“ eingeordnet wird (Abb.5c). Die Schale wiederum, auf der sein abgetrennter Kopf präsentiert wird, hat für viele Teilnehmende Attributcharakter (Abb.5d), obwohl sie formal kein Attribut darstellt. Darüber hinaus scheint die Benennung der Nebenfiguren stark von der Anzahl der tatsächlich auf den Bildern dargestellten Personen beeinflusst zu sein. In den Werken von Gentileschi und Varotari wird eine dritte Figur – die Dienerin Judiths oder Salomes – abgebildet, wobei diese bei Gentileschi sogar aktiv in das Geschehen eingreift. Dies führt dazu, dass Holofernes und Johannes der Täufer seltener als Nebenfiguren benannt werden, während in Strozzi's Darstellung der isoliert am rechten Bildrand positionierte, abgeschlagene Kopf des Goliaths mehrheitlich als Nebenfigur assoziiert wird (Abb.5b).

Die Ergebnisse verdeutlichen, dass auch scheinbar eindeutige Darstellungen auf mehreren Ebenen interpretiert werden können und somit ein facettenreiches Verständnis der kunsthistorischen Rezeption abbilden. Auch wenn dem Begriff der Ähnlichkeit der Beigeschmack eines fluiden, unbestimmten Konzepts anhaftet, belegen die empirischen Befunde, dass die vom Ähnlichen Sprechenden sehr wohl in der Lage sind, Identitäten zu erkennen und das Differenten zu unterscheiden.

Die Suche nach ähnlichen Bildern. Ein Praxisbericht mit dem e-Research-Werkzeug iART

Herkömmliche Metadaten-Suchmaschinen sind in erster Linie identitätsorientiert. Wenn ich nach „Michelangelo“ suche, erhalte ich alle Werke von Michelangelo – genauer gesagt: Werke eines Künstlers, der im Metadatenfeld „Autor“ als „Michelangelo“ aufgeführt ist. Auch wenn ich die Suche komplexer gestalte und durch eine Kombination von Kriterien verfeinere – etwa auf Werke, die zwischen 1500 und 1510 entstanden sind –, ändert sich an diesem Ansatz grundsätzlich nichts. Seit mehr als 25 Jahren sind solche Suchmaschinen in der täglichen Arbeit von Forschenden der Kunstgeschichte und anderen Geisteswissenschaften etabliert; sie finden Anwendung in Bibliothekskatalogen oder digitalen Werkverzeichnissen wie Prometheus.¹ Allerdings stößt dieser identitätsbasierte Ansatz an seine Grenzen: Werke, in denen der Künstler unter verschiedenen Namensvarianten verzeichnet ist, können unter Umständen nicht vollständig erfasst werden. Ebenso – und dies ist von entscheidender Bedeutung – können semantische Zusammenhänge zwischen Werken oder Künstlern nur dann abgebildet werden, wenn diese Beziehungen von der Person, die die Metadaten hinterlegt, explizit dokumentiert werden. Mit anderen Worten: Durchsuchbar ist nur das, was zuvor explizit als durchsuchbar gekennzeichnet wurde.

¹ Vgl. <https://prometheus-bildarchiv.de/de/index> [Stand 02/2025].

Von der Identität zur Ähnlichkeit

Aus diesem Grund werden Suchwerkzeuge entwickelt, die über herkömmliche metadatenbasierte Ansätze hinausgehen, indem sie Methoden des maschinellen Lernens implementieren. In diesen Systemen firmiert ein Kunstwerk nicht mehr als opakes Anhängsel eines Datensatzes – etwa durch den Eintrag im Metadatenfeld „Autor“ –, sondern wird direkt über seine bildliche Struktur adressiert und ermöglicht so eine Suche, die inhaltliche Merkmale auf visueller Ebene berücksichtigt. Ein solches Werkzeug ist iART (interaktives Analyse- und Retrieval-Tool), das zwischen 2019 und 2021 vom Lehrstuhl für Mittlere und Neuere Kunstgeschichte der Ludwig-Maximilians-Universität München, der Forschungsgruppe „Visual Analytics“ der TIB Hannover und der Fachgruppe „Intelligente Systeme und Maschinelles Lernen“ des Heinz Nixdorf Instituts der Universität Paderborn entwickelt wurde.²

Die Leistungsfähigkeit von iART beruht maßgeblich auf in den letzten Jahren entwickelten multimodalen Modellen – Modellen also, die nicht nur einzelne Modalitäten verstehen, sondern in der Lage sind, Text- und Bildinformationen gleichzeitig zu verarbeiten und miteinander zu verknüpfen. Als besonders performant hat sich die Technik CLIP (Contrastive Language-Image Pre-training) erwiesen: Sie überführt Bild- und Textdaten mit Hilfe modalitätsspezifischer Kodierer in einen gemeinsamen Merkmalsraum, um die Beziehung zwischen sprachlichen Begriffen und visuellen Konzepten herzustellen.³ Ein zentraler Aspekt dieser Technik besteht darin, die aus Bild und Text generierten *Embeddings* so anzupassen, dass sie eine möglichst hohe Ähnlichkeit aufweisen. Das Embedding stellt einen mehrdimensionalen Vektorraum dar, in dem jedes Kunstwerk durch einen numerischen Vektor

repräsentiert wird – die Grundlage für seine computergestützte Verarbeitbarkeit. Innerhalb dieses Raumes werden semantisch verwandte Inhalte gruppiert – beispielsweise Bilder, die in irgendeiner Hinsicht ähnliche Merkmale aufweisen –, wodurch Zusammenhänge sichtbar werden, die über eine Erschließung mit rein formalen Metadaten hinausgehen. Das von der amerikanischen Firma OpenAI veröffentlichte, CLIP-basierte Modell wurde mit 400 Millionen Bild-Text-Paaren aus dem Internet trainiert – ausgehend von der Überzeugung, dass die in den Texten enthaltenen beschreibenden Informationen Aufschluss geben über die Beschaffenheit der Bilder. Das Modell lernt, wie ein Bild aussehen muss, um den jeweiligen Textbezeichnungen zu entsprechen, und kann anschließend auch andere Bilder identifizieren, die diese Kriterien erfüllen. Konkret bedeutet dies: Wird ein Bild wiederholt als Darstellung eines Baumes beschrieben, lernt das Modell, welche Bereiche des Bildes den Baum repräsentieren. Dadurch ist es in der Lage, einen Baum auch in Bildern zu erkennen, deren Beschreibungstext den Begriff „Baum“ nicht enthält. Durchsuchbar wird so auch das, was nicht explizit als durchsuchbar gekennzeichnet wurde.

Im Folgenden untersuchen wir auf Basis des in iART integrierten CLIP-Modells verschiedene Suchszenarien, um die Möglichkeiten von Systemen zu evaluieren, die Ähnlichkeitsbasierte Methoden des maschinellen Lernens einsetzen. Dabei zielen unsere Beobachtungen nicht darauf ab, systematisch zu ermitteln, inwieweit – und unter welchen Bedingungen – iART gegenüber herkömmlichen identitätsbasierten Metadaten-Suchmaschinen einen signifikanten Mehrwert für die kunsthistorische Forschungspraxis bietet. Vielmehr sollen experimentelle Suchanfragen die Potenziale dieser bislang oft als ungewohnt empfundenen Herangehensweise schlaglicht-

² Vgl. <https://www.iart.vision/> [Stand 02/2025], ferner Stefanie Schneider, Matthias Springstein, Javad Rahnama, Hubertus Kohle, Ralph Ewerth und Eyke Hüllermeier: iART. Eine Suchmaschine zur Unterstützung von bildorientierten Forschungsprozessen, DHd 2022, Kulturen des digitalen Gedächtnisses, Konferenzabstracts, S. 142-147.

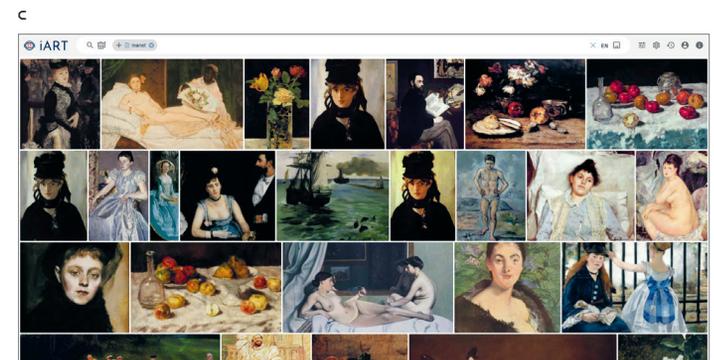
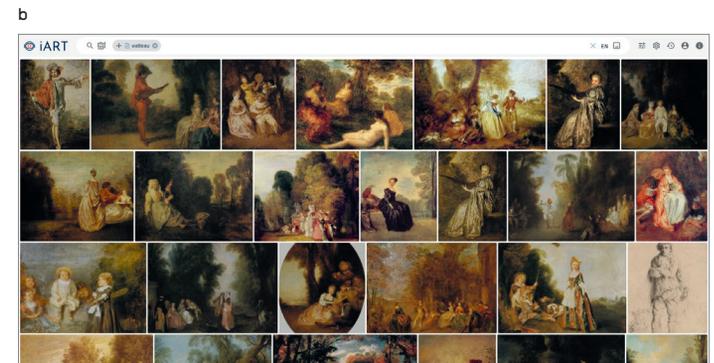
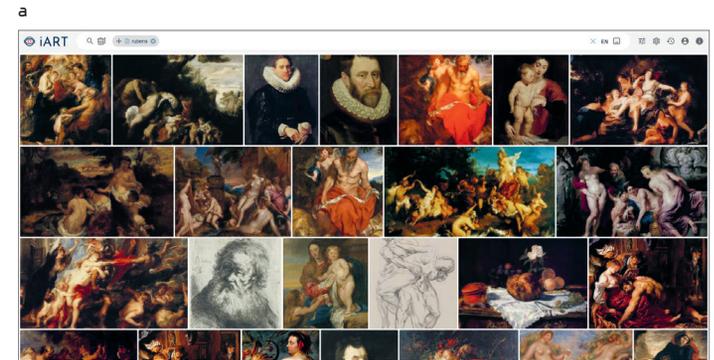
³ Siehe Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger und Ilya Sutskever: Learning Transferable Visual Models from Natural Language Supervision, Proceedings of the 38th International Conference on Machine Learning, ICML 2021, S. 8748-8763.

artig aufzeigen. In allen Fällen greifen wir auf die Benutzeroberfläche von iART online zu und geben, sofern nicht anders angegeben, unsere Suchanfragen direkt in den an Google erinnernden Suchschlitz ein.

Suche nach bekannten Künstlern

Da es sich um einen der häufigsten Suchvorgänge in metadatenbasierten Systemen handelt, beginnen auch wir mit der Suche nach Künstlernamen – willkürlich ausgewählt: Rubens, Watteau und Manet. Die Ergebnisse werden stets in absteigender Reihenfolge der Ähnlichkeit angezeigt, wobei das der Suchanfrage ähnlichste Objekt immer links oben erscheint. Dabei steht hier – bedingt durch die Suche nach einem Künstler – vor allem die Ähnlichkeit der Zuschreibungen im Vordergrund, nicht die der Bilder selbst. Bei Rubens erscheint zunächst tatsächlich ein solcher, gefolgt von mehreren „unknowns“, also Gemälden unbekannter Künstler (Abb. 1a). Es folgt ein van Dyck, der Rubens so nahestand, dass selbst kunsthistorisch Gebildete die beiden gelegentlich verwechseln. Danach erscheinen erneut einige Treffer unbekannter Künstler sowie weitere Rubens-Werke. Es schließt sich ein Treffer von Charles Gleyre an, einem Künstler des 19. Jahrhunderts, der zwar nicht als Rubensianer gilt, aber – mit seinem Gemälde *Nymphen und Amor* (1830–1850) – thematische Parallelen zu Rubens aufweist. Weniger überraschend ist der darauffolgende Tizian, der als eines der großen Vorbilder des flämischen Barockmalers bekannt ist. Auch die später zurückgegebenen Werke von James Ensor und Pierre-Auguste Renoir lassen sich erklären: Ensor verehrte Rubens, während Renoir als der ‚barockste‘ unter den Impressionisten bezeichnet werden kann. Verwunderlich ist, dass die im Bildreservoir von iART

Abb. 1. Suchergebnisse (Screenshots) in iART für die Begriffe [Stand 02/2025]: „rubens“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Arubens%3A>); a), „watteau“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Awatteau%3A>); b), „manet“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Amanet%3A>); c)



vermutlich in größerer Zahl vorhandenen Rubens-Werke nur selten unter den ersten Treffern auftauchen. Der zweite Künstler ist Watteau (Abb. 1b). Hier ähneln sich die Ergebnisse nicht nur in der Zuschreibung, sondern auch in der Bildsprache – bedingt durch die konsequente Beschäftigung Watteaus mit demselben Genre. Erstaunlich ist, dass schon früh, nach einer Reihe von authentischen Watteaus, auch ein Bild von Henri Fantin-Latour auftaucht. Dieses zeigt eine bukolische Szene, die man dem Neorokoko zuordnen könnte und die deutlich vom Meister des 18. Jahrhunderts beeinflusst ist. Später folgen mehrere Bilder von Pater und Lancret, die bekannte Nachfolger Watteaus sind. Und ein letzter Test: Manet (Abb. 1c). Hier erscheint unter den ersten Treffern Carl Schuch, der deutsche Realist mit starkem Frankreich- und Realismusbezug. Es folgen Realisten und Impressionisten aus Manets Umfeld, wie Courbet, Cassatt, Cézanne und Morisot, jedoch auch überraschende, aber plausible Treffer wie Vallotton, Repin und Slevogt. Auch Goya taucht auf, um Manets Begeisterung für Spanien zu illustrieren. Auffällig ist, dass häufig Bildnisse von (unbekleideten) Frauen gezeigt werden, während Werke mit erzählerischen Themen überhaupt nicht vertreten sind. Die Gründe dafür sind nicht eindeutig zu klären, könnten aber mit einem bestimmten ‚Image‘ der Impressionisten zusammenhängen, die wenig an narrativen Darstellungen interessiert waren. Somit lässt sich festhalten, dass das CLIP-Modell künstlerische Felder referenziert und nicht eindeutige künstlerische Identitäten. Damit lassen sich sowohl Gruppenidentitäten und -affinitäten als auch künstlerische Historismen erkennen – was Analysen sowohl in synchroner als auch in diachroner Perspektive ermöglicht. Dabei ist zu berücksichtigen, dass nicht wie bei metadatenbasierten Systemen nach eindimensionalen Suchkriterien gefahndet wird,

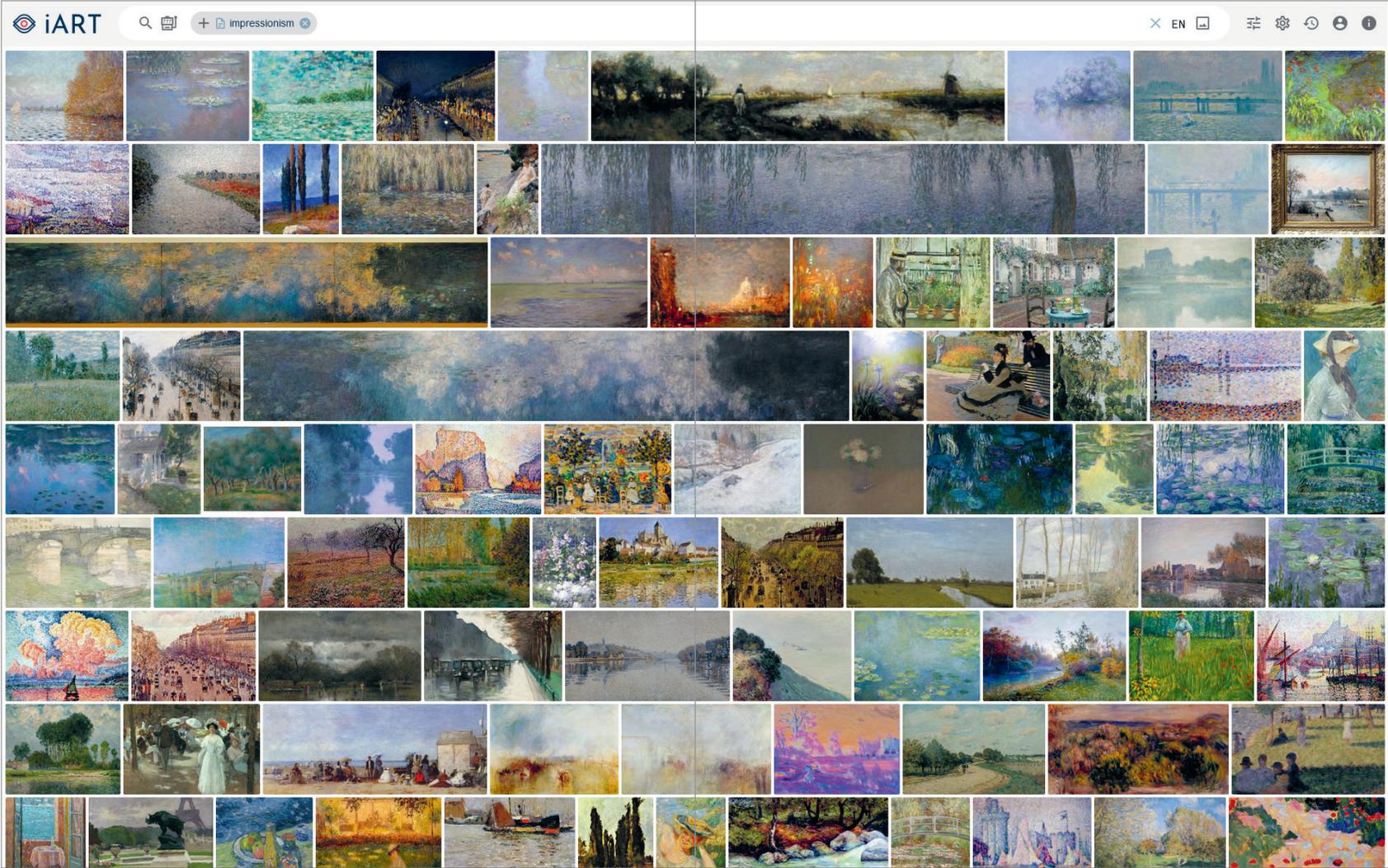
sondern nach einem Embedding, das Hunderte bis Tausende solcher Kriterien implizit enthält und sich nur schwer in einzelnen interpretierbare Komponenten zerlegen lässt. Dies könnte die teilweise erstaunlichen Ergebnisse erklären – etwa im Fall Manets, bei dem (unbekleidete) Frauenporträts möglicherweise als kanonisch gelten. Auch wenn die bisher durchgeführten Untersuchungen im Vergleich zu herkömmlichen Methoden zu originellen Ergebnissen geführt haben, basieren sie doch auf traditionellen Herangehensweisen. Skeptiker werden zudem mit Recht darauf hinweisen, dass das gänzliche Fehlen einiger Werke, die sich eigentlich ähnlich sind, Zweifel an der Praktikabilität des Systems aufkommen lässt – was jedoch weniger am für diese Experimente verwendeten CLIP-Modell liegt, sondern vielmehr an der in iART hinterlegten Datenbasis, die überwiegend aus Wikidata gespeist wird.

Suche nach kunsthistorischen Stilrichtungen

Zweites Experiment: Wir suchen nach einer Stilrichtung – dem Impressionismus (Abb. 2). Die ersten und damit ähnlichsten Ergebnisse bestehen fast ausschließlich aus Werken von Monet; nur vereinzelt tauchen Arbeiten von Pissarro, Seurat und Morisot auf, während einige zentrale Vertreter des Impressionismus gänzlich fehlen.⁴ Wesentlich ist jedoch die Feststellung, dass Monet die Eigenschaften des Impressionismus am vollständigsten verkörpert; sei es, weil sein Name trivialerweise am häufigsten mit diesem Begriff assoziiert wird, oder weil seine Werke am besten in das vom Modell definierte Strukturschema passen – wie auch immer das Modell dieses Schema definiert. Anschließend wird die Suchanfrage variiert, sodass gezielt nach impressionistischen Landschaften mit un-

⁴ Auch hier stellt sich die Frage, ob einige Künstler und Künstlerinnen nicht in den verfügbaren Daten enthalten sind.

Abb. 2: „impressionism“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aimpressionism>)



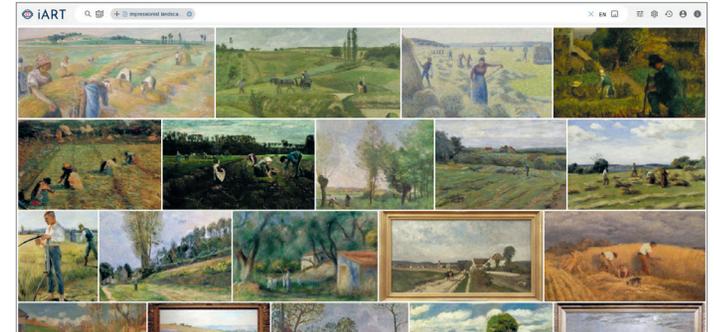
terschiedlichem Bildpersonal gesucht wird: Landarbeitern, Touristen und Fischern. Zuerst suchen wir nach „impressionist landscape showing rural workers“ (Abb. 3a). Das Ergebnis ist weitgehend überzeugend, auch wenn nicht immer explizit Landarbeiter dargestellt werden – gelegentlich erscheinen stattdessen ihnen ähnelnde Personen, und bei einigen Treffern fehlt jegliches Personal. Ähnliches gilt für die Suche nach „impressionist landscape showing tourists“, einem klassischen Motiv des Impressionismus. Erstaunlich ist, wie gut das Modell – trotz gewisser Unzulänglichkeiten – zwischen Landarbeitern und Touristen unterscheiden kann, auch wenn in den Touristenszenen gelegentlich weniger ländliche als städtische Elemente mitschwingen (Abb. 3b). Bei der Suche nach „impressionist landscape with fishermen“ (Abb. 3c) liegt die Vermutung nahe, dass CLIP vor allem nach dem Vorhandensein von Wasser – etwa in Form von Meer, Seen oder Teichen – kategorisiert. Insgesamt jedoch zeigt sich, dass es durchaus in der Lage ist, zwischen Landarbeitern, Touristen und Fischern zu differenzieren, ohne dabei die grundsätzlich „impressionistische Landschaft“ aus den Augen zu verlieren – das charakteristische, impressionistische Merkmal der gesuchten Szenerie bleibt erhalten.

Suche nach Handlungszusammenhängen

iART bewährt sich vor allem dort, wo lebendige Vollzüge gefunden werden sollen. Eine traditionell nach Merkmalen organisierte Metadatensuchmaschine kann zwar Begriffe wie „Blume“ und „Person“ kodieren, hat aber Schwierigkeiten, gezielt eine Person zu finden, die an Blumen riecht. iART nicht, wie Abb. 4a belegt. Kritisch anzumerken ist jedoch, dass Bilder mit an Blumen riechenden Personen vielleicht nur deshalb erscheinen, weil jeweils eine Blume und eine Person

Abb. 3: „impressionist landscape showing rural workers“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aimpressionist+landscape+showing+rural+workers>; a), „impressionist landscape showing tourists“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aimpressionist+landscape+showing+tourists>; b), „impressionist landscape with fishermen“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aimpressionist+landscape+with+fishermen>; c)

a



b



c



vorhanden sind – worauf auch die Tatsache hindeutet, dass die Mehrzahl der Bilder zwar beide Merkmale aufweist, jedoch nicht unbedingt in der spezifischen Beziehung des „an Blumen Riechens“. Umgekehrt könnte in einer MetadatenSuchmaschine die Suchanfrage „Blume“ und „Person“ gelegentlich auch ein Bild liefern, in dem tatsächlich eine Person an Blumen riecht. Hier bedarf es einer genaueren Evaluation, ob iART tatsächlich erfolgreicher ist und eine größere Anzahl von Bildern zurückgibt, die explizit an Blumen riechende Personen darstellen. Ein ähnliches Muster zeigt sich in einem weiteren Beispiel: Auch eine Person, die aus einem Fenster schaut, wird nach der Eingabe von „Person“ und „Fenster“ in der Regel nur zufällig gefunden. Im Gegensatz dazu erzielt iART – und von Zufall kann angesichts der schieren Anzahl der Treffer kaum die Rede sein – eine Vielzahl passender Ergebnisse (Abb. 4b). Beeindruckend sind auch Ergebnisse wie das folgende: Bei der Suche nach „dangerous animals“ gibt iART fast ausschließlich Raubtiere wie Tiger und Löwen zurück – vor allem in Szenarien, in denen tierische Kämpfe angedeutet oder explizit dargestellt werden (Abb. 5a). Bei der Suche nach „sweet animals“ (Abb. 5b) hingegen tauchen zwar teilweise die gleichen Tierarten auf, insbesondere wenn sie aus einem bäuerlichen Umfeld stammen, doch werden zum Beispiel Füchse – die zu den Raubtieren gehören – in einer kindlichen Balgerei gezeigt, die eher den Aspekt der Harmlosigkeit betont. Ein weiteres Beispiel, das das Prinzip näher veranschaulicht, ist die Suche nach „pedestrians in a parisian street of the 19th century“ (Abb. 6a). Zwar befinden sich nicht alle gezeigten Straßen tatsächlich in Paris, doch wirken die meisten Bilder sehr ‚pariserisch‘, und selbst die wenigen Abweichungen sind überzeugend. Eine geringfügige Erweiterung der Anfrage – „pedestrians in a parisian street of the 19th century *talking to*

Abb. 4: „person smelling flowers“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aperson+smelling+flowers>; a), „person looking out of a window“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aperson+looking+out+of+a+window>; b)

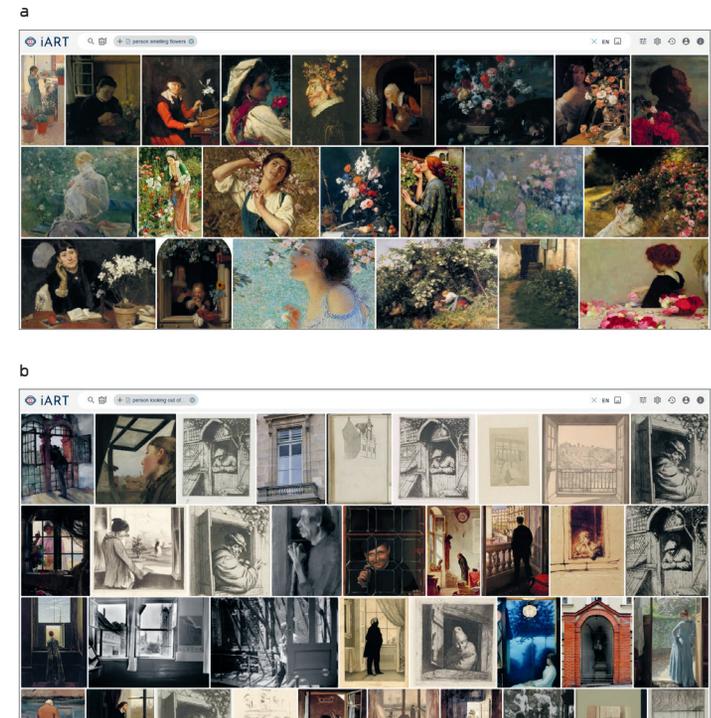
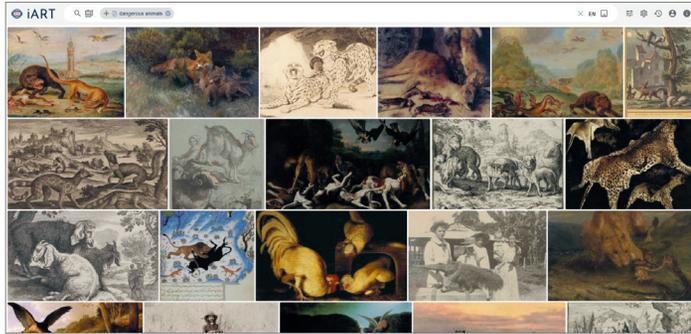
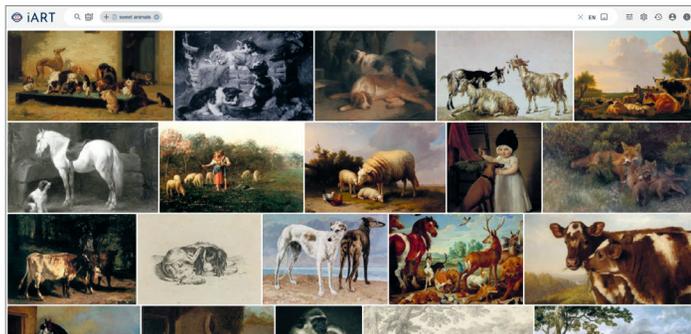


Abb. 5: „dangerous animals“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Adangerous+animals%3A>); „sweet animals“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Asweet+animals%3A>) b)

a



b



each other“ – schränkt die Ergebnisse auf grafische Darstellungen (nämlich Lithografien) ein, fokussiert jedoch deutlich stärker auf Szenen, in denen Personen miteinander kommunizieren (Abb. 6b). Gerade die zuletzt betrachteten Suchanfragen verdeutlichen, inwiefern sich iART von Metadaten-suchmaschinen unterscheidet: Während letztere vornehmlich nach einzelnen Begriffen suchen, kann iART ganze Handlungszusammenhänge erfassen.

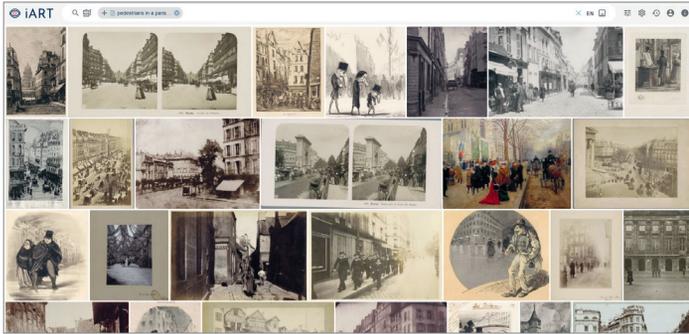
Dass iART mit diffusen Beschreibungen umgehen kann, zeigt sich auch an den Ergebnissen zu „*overcrowded street in the morning*“ (Abb. 7a) und „*empty street in the morning*“ (Abb. 7b). Zwar erscheinen manche als überfüllt deklarierte Straßen überraschend leer und einige als leer bezeichnete Szenen zeigen durchaus reges Treiben – und ob die Darstellungen tatsächlich den Morgen widerspiegeln, ist nicht immer eindeutig –, dennoch reduziert das statistisch fundierte Empfehlungssystem die Datenmenge so erheblich, dass eine manuelle Durchsicht der Ergebnisse praktikabel wird. Angesichts der Tatsache, dass ein großer Teil historischer Kunstbilder nur unzureichend oder gar nicht begrifflich erschlossen ist, bleibt unbestritten, dass iART eine erfolgreiche Suche nach Bildinhalten ermöglicht.

Weitere Nutzungsszenarien

Während die bisherigen Suchanfragen alle auf dem bekannten Text-zu-Bild-Schema metadatengestützter Werkzeuge basieren, unterstützt iART auch eine Bild-zu-Bild-Suche, die abschließend zumindest kurz skizziert werden soll. Wir geben dafür Antonello da Messinas *Kreuzigung* (1475) in iART ein. Bei nahezu allen gefundenen Bildern handelt es sich ebenfalls um Kreuzigungen (Abb. 8a) – wobei entscheidend bleibt, dass nicht nach dem Titel „Kreuzigung“ gesucht wird, sondern

Abb. 6: „pedestrians in a parisian street of the 19th century“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Apedestrians+in+a+parisian+street+of+the+19th+century>; a), „pedestrians in a parisian street of the 19th century talking to each other“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Apedestrians+in+a+parisian+street+of+the+19th+century+talking+to+each+other>; b)

a

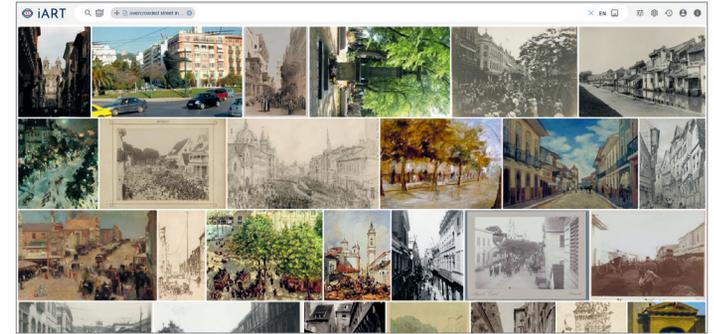


b

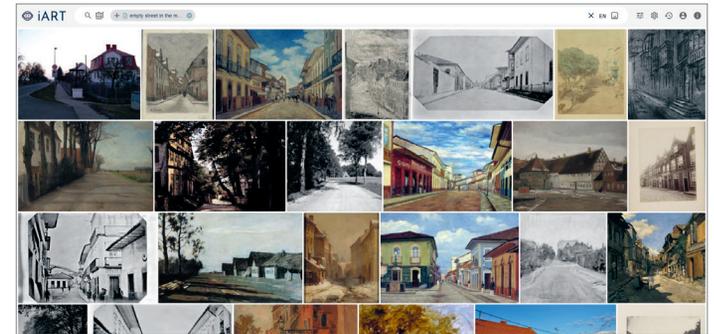


Abb. 7: „overcrowded street in the morning“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aovercrowded+street+in+the+morning>; a), „empty street in the morning“ (<https://www.iart.vision/search?lang=en&query=%2Btxt%3Aempty+street+in+the+morning>; b)

a



b



nach einem Embedding, das die Kreuzigungsmerkmale des Bildes erfasst und beschreibt. Dabei steht das Hauptmotiv – der Gekreuzigte selbst – im Vordergrund, während Nebenfiguren keine prägende Rolle zu spielen scheinen; Kreuzigungen ohne die im Suchbild vorkommenden Schächer und Heiligen werden ebenfalls gefunden. Je nach Bedarf könnten nun in iART über den in den Einstellungen hinterlegten Button „Global Weights“ gezielt semantische oder formale Qualitäten der Ähnlichkeitsbestimmung verstärkt, also die jeweils relevante Dimension der Ähnlichkeit angepasst werden. Darüber hinaus können die Ergebnisse nicht nur im traditionell eindimensionalen Bildraster angezeigt, sondern auch auf eine zweidimensionale Leinwand projiziert werden. Ausgehend von da Messinas *Kreuzigung* als Referenz eröffnen sich auf den ersten Blick völlig unterschiedliche Cluster, die der Kreuzigungsszene ähneln: beispielsweise eine Gruppe rechts oben mit Fotografien von Speerspitzen und Spindelschäften, die in ihrer minimalistischen Darstellung auf grauem Hintergrund die Geradlinigkeit des Kreuzes referenzieren, eine Gruppe etwas rechts vom Zentrum mit Kruzifixen sowie eine kleinere Sammlung von *Vera Ikons* links daneben (Abb. 8b).

Abb. 8. Suchergebnisse in iART für eine Bild-zu-Bild-Suche nach Antonello da Messina, *Kreuzigung* (1475): eindimensionales Bildraster (a) und zweidimensionale Leinwanddarstellung (b)

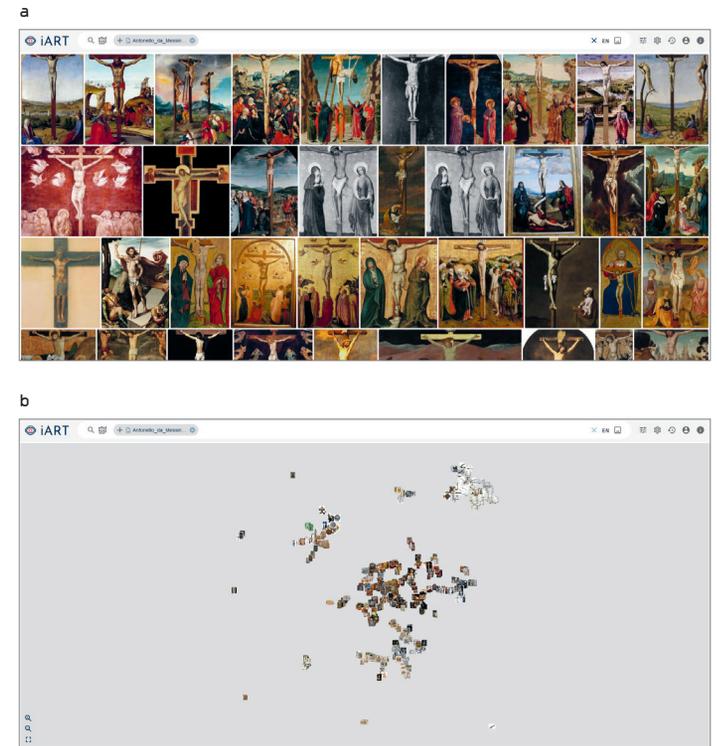
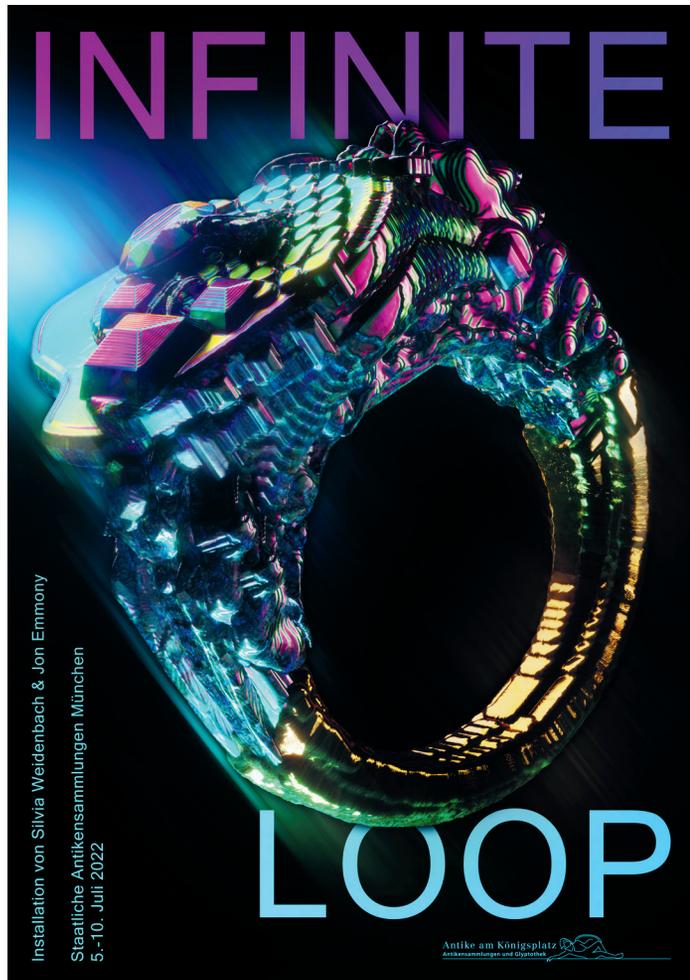


Fig. 1. Computer-generated Imagery (CGI) rendering of the Infinite Loop ring, featured on a poster for the exhibition at the Staatliche Antikensammlungen in Munich



Jon Emmony, Silvia Weidenbach

Infinite Loop. Disruption and Similarity in the Act of Design with Machine Learning

Jon Emmony is a London-based digital artist and creative director. From 2011 to 2016, he served as the digital art director at Nick Knight's SHOWstudio, a period that was instrumental in expanding his approach to digital imagery. Emmony's post-SHOWstudio career includes collaborations with renowned clients such as Selfridges, the Victoria and Albert Museum (V&A), the band The 1975, Nike, and fashion designer Iris van Herpen.

Silvia Weidenbach is an award-winning jeweler who blends the traditional with the contemporary: she uses both historical techniques and new technologies, combining them to create striking pieces of art. Weidenbach was the first Gilbert Collection Artist-in-Residence at the V&A, where her work is part of its permanent collection. She is a lecturer at the Glasgow School of Art and the London College of Fashion.

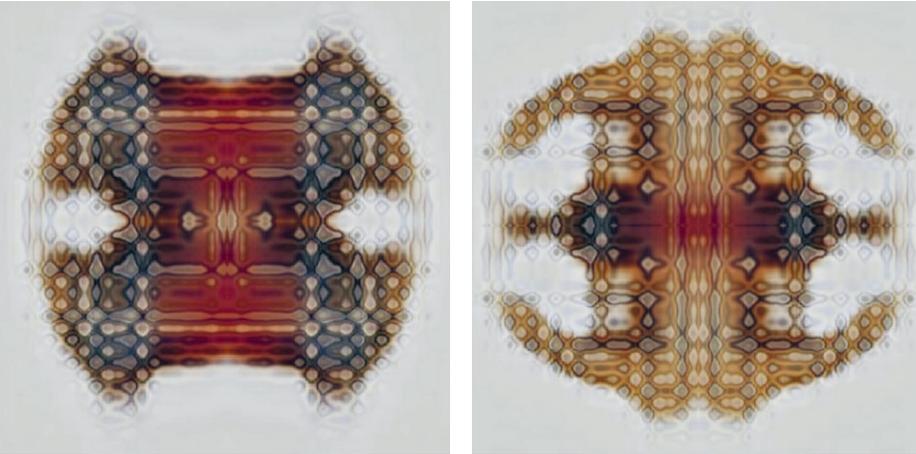
Jon Emmony: Silvia, when did you first become aware of Artificial Intelligence (AI)?

Silvia Weidenbach: It was 2019 in China at a conference.¹ It was mind-blowing. I thought, oh my God, in which direction can this go, what is possible, and what kind of tool can this allow for artists, for designers. I wondered: what could this mean for jewelry?

67

¹ 5th Art and Science International Exhibition and Symposium, TASIES 2019, Academy of Arts & Design, Tsinghua University, China, November 2-30, 2019.

Fig. 2. Process images of the AI 'dreaming' new forms from its understanding of historical rings



JE: I've also been aware of AI in the context of digital images for a while. I am often researching new techniques and processes, and I remember finding this new, exciting visual language that was formed—images generated from 'the in-between,' or the 'latent space.' So, you'd feed an AI ten images of rings and it would present you with images of new, imagined rings, as if pulled from a dream (Fig. 2). The AI identifies common features and patterns in the original images and then synthesizes these elements to create something familiar yet new. It's like a reflection on the nature of similarity, fusing the common features of the originals into unique yet recognizably related outputs.

68

ChatGPT, there are moments where AI can feel truly creative, or soulful. Can you speak a little from your perspective about how you see your role as a creative collaborator?

ChatGPT: Creativity, as typically understood, involves the ability to generate unique and valuable ideas, often drawing upon personal experiences, emotions, and a deeper understanding of complex concepts. While I can assist in the creative process by providing information, prompts, and alternative perspectives, the creative aspect of any work produced using my assistance is ultimately a reflection of the user's creativity, not my own. I can be seen as a tool to assist in idea exploration.²

SW: An exciting tool! I see the creative process as a ping-pong game between tools, between worlds and ideas.

JE: I think it's worth mentioning our project with the V&A as an introduction to how we met.³ At the time, we were both looking at how we could use techniques like 3D scanning to see historical objects in a new way. And I think for me, it's always been interesting to look through the eyes of technology to see; see the glitches, see the mistakes, see how it can understand the world around it.

SW: For me, it was the excitement, the endless possibilities. Coming from a traditional making context with limitations, real-world limitations of material properties and so on, you now have the opportunity to work with these unlimited facets of design. You can give objects a heartbeat and a life.

ChatGPT, can you speak to how AI can reimagine physical design, away from the constructs and limitations of the physical?

69

² ChatGPT output was curated from multiple sessions to make sense.

³ Visual Feast, V&A South Kensington, UK, July 7, 2018-January 7, 2019.

ChatGPT: AI can reimagine physical design by transcending the confines of traditional physical constraints. Through generative design, material innovation, simulation, and testing, AI can optimize designs for efficiency, sustainability, and accessibility. It fosters creative collaboration, personalization, and human-centered design, all while continuously evolving and breaking down disciplinary barriers. By doing so, AI liberates design from many of its former limitations, enabling more innovative, sustainable, and user-centric solutions that push the boundaries of what's possible in the physical world.

JE: When we came together to work on our first project all those years ago, it was interesting to me how we were using similar technologies with such vastly different outputs. As a digital artist, many of the objects and designs I make are never seen in the 'real' world. But to consider materiality, the human touch, the presence of an object, opened creative doors. I often work with 3D scanning in my fashion work, which does have a tactile side to it. When capturing a garment on a model in 3D, I think about how the dress may sit on the human form, what the physical hair and make-up might be, the pose...

SW: Like making jewelry. You coordinate, you curate the forms and their relation to the body.

JE: Absolutely. And through that you create compositions that relate to the body. We all exist now with technology as part of our own bodies; our phones as an extension of our being, so it is never too distant.

SW: And I remember we had a lot of chats during the lockdown and COVID-19 where we were all confronted with these new ways of communicating and what is possible, relying

more so on digital spaces and devices to communicate. When you came to Glasgow, we went for a walk and we said, okay, how do we use AI and technology to create something tactile, to bring this new space into the real world around us as something that can be beyond a screen.

JE: And I think that it's important that we've never seen these techniques as being a one-click solution to the creative journey. We both love the process of what we do.

ChatGPT, can you speak to how you see the creative process? How do you, yourself, curate and navigate creative tasks that can have multiple paths?

ChatGPT: As an AI, my creative process involves generating responses based on patterns and knowledge learned from vast amounts of text data. When faced with creative tasks with multiple paths, I analyze the context and available information to craft relevant and coherent responses. I rely on probabilistic modelling to suggest responses that align with common patterns and linguistic structures. While I can suggest creative ideas and solutions, my creativity is inherently algorithmic, drawing from existing data and patterns rather than personal intuition or inspiration. I aim to provide diverse options and guide users through the exploration of various creative directions.

SW: It's problem solving. Finding the best gemstones for a brooch or earring, for example. It's more about the process and what the potential is.

JE: I remember when we spoke there was a challenge as to how we could imagine this third collaborator (AI) sitting within our process.

SW: Yeah, absolutely. Our first tests with the software felt like doodles, an endless stream of ideas that could be generated 24 hours a day without a break (Fig. 3)! We quickly started to feel like curators, guiding its taste and perceptions. Once we'd finished testing the potential, we started talking about the project that eventually became *Infinite Loop* (Fig. 1). We took a trip to Munich to see a collection of Etruscan jewelry at the Staatliche Antikensammlungen. We were lucky to be supported by the museum with the wonderful expertise of curator Dr. Astrid Fendt. Etruscan jewels are famous for their technical achievement, which is an interesting link when we think about the cutting-edge technology of AI. For their time, the pieces were very advanced and futuristic.

JE: I found those objects so fascinating because they must have seemed fairly alien at the time of their creation. So refined and almost digital in their perfection. Rings are such a consistent piece of design through the ages and can have such varying significances to the wearer. The pieces also became a way for the makers and wearers of the time to connect with higher beliefs, almost in the same way that technology allows us to do today.

SW: Otherworldly! This is where it started—we went to the physical world to see historical objects. We used 3D scanning apps on our phones to capture some of the pieces, to immediately see these objects through a new lens (Fig. 4). That became our starting point. It's like sketching things out, keeping the creative process open, knowing what to disregard, and using low-tech solutions to get started.

JE: I have always been a fan of low-tech, approachable solutions, even though I work with technology consistently. When I first

started some years ago, I found code online that turned an Xbox Kinect device, a cheap piece of hardware used to measure the movement of people playing computer games, as a 3D scanner. That way of working always appeals to me—finding use for unexpected tools in a process as a way of creation.

SW: And then we fed an AI with hundreds of images of different jewelry from all different time periods, with the consistency being that they were rings. We went to the Germanisches Nationalmuseum in Nuremberg to collect more rings from different eras, and also to chat with the very generous and supportive Dr. Heike Zech. It was quite nice to get a bit of an overview of the history, then understanding the form, the shape. And then the idea of: what does a ring look like? How can it be defined? Are there common features, similarities?

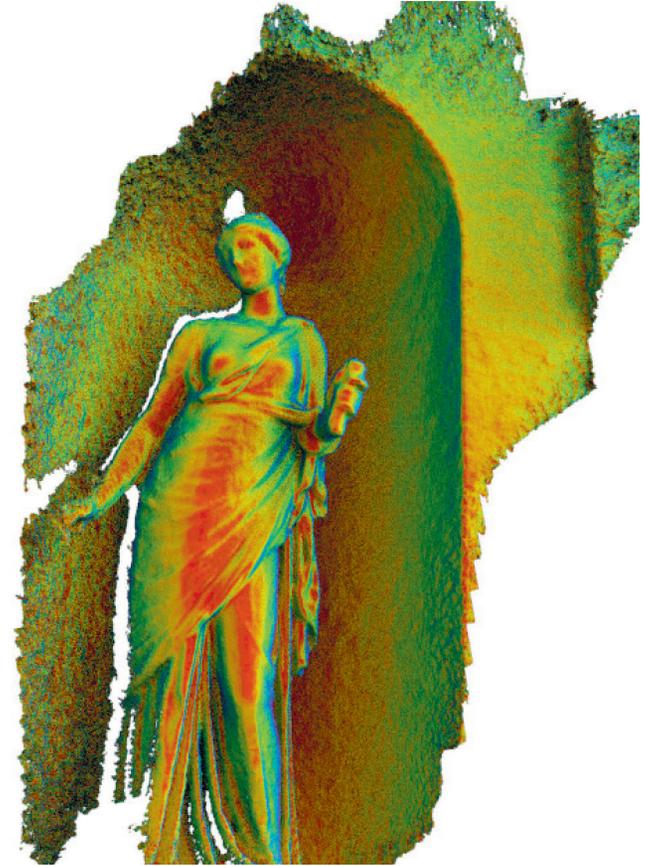
JE: Absolutely. I found it quite touching that all these historical rings, from such varying cultures and times, were brought together. The makers of each, in a way, collaborating with people they could have never met—including each other. For me, that was where you start to see the magic in this new technology—how it can erode time and place and connect ideas.

SW: Importantly, it was a curated selection of rings, too. We acted on instinct, selecting pieces that spoke to us with their differences being celebrated—knowing that more extreme design elements would appear most clearly through the AI's interpretations and inventions. As well as the previous selection we had collected, we also met with jewelry historian Dr. Beatrice Chadour-Sampson, curator from The Alice and Louis Koch Collection who kindly provided the final selection of rings that we added to our AI.

Fig.3. Initial exploration with CGI to explore the concept of AI as a tool to create jewelry



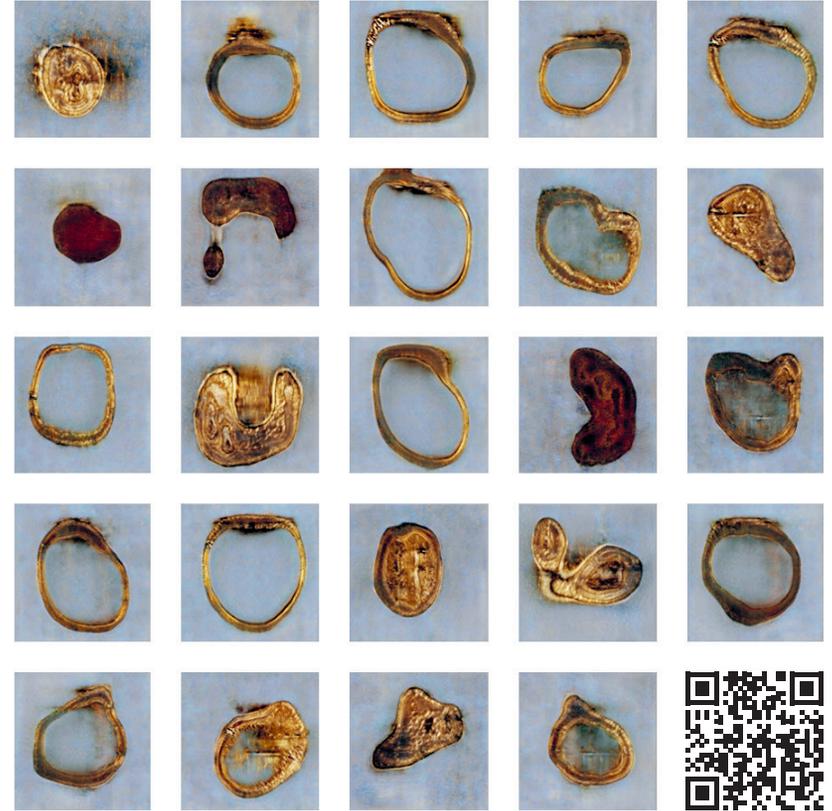
Fig.4. Initial 3D scan from our trip to the Staatliche Antikensammlungen in Munich



Jon, how did you feel when you first saw the results of the AI after we had fed it those rings? Because the results were unexpected, melting images that felt almost like the beginning or origin of something (Fig. 5). Not like the past at all!



Fig. 5, AI outputting 'imagined' rings, as featured in the process film for Infinite Loop, Screenshot 0:06, 04.07.2022, <https://youtu.be/V4RB-3Udjck> [Stand 02/2025]



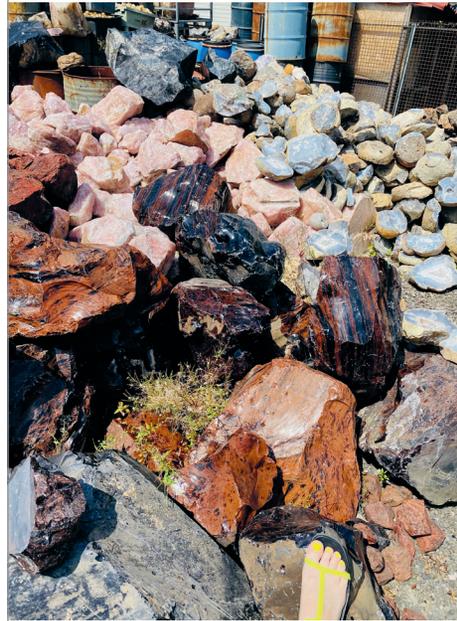
JE: It was particularly exciting because these images felt like elements—almost like chunks of organic matter found in the ground, like we were uncovering stones or the ingredients to make jewelry.

ChatGPT, can you define how you see futurist design? How do we align that with visions of the past?

ChatGPT: Futurist design is an approach that envisions and creates products, spaces, and experiences with a forward-looking perspective, often incorporating cutting-edge technology and innovative concepts. It seeks to anticipate future needs and trends while pushing the boundaries of current design conventions. Aligning it with visions of the past involves drawing inspiration from historical design elements, understanding how they have evolved, and selectively incorporating timeless aspects into futuristic designs. This fusion of past and future can result in designs that resonate with tradition while also embracing progress and innovation.

SW: And that is also quite poetic—and informed the process. Rather than looking at utilizing something like 3D printing, which felt so mechanical and calculated, we went back to the Earth. This informed the material we used for our own final ring, inspired by the designs produced by the AI. And then we went to Idar-Oberstein, looked at all the gemstones—what looks nice, what looks interesting (Fig. 6). And then came the process of revealing our designed ring (Figs. 7–11), informed by the past, by working with such a tangible and tactile real material.

Fig. 6. Trip to Idar-Oberstein to explore gemstones



JE: Forwards and backwards. So we had echoes of the past and the future. Designs that felt similar or relatable but in this new context. The final design didn't sit in the now, which I thought was interesting, and to be honest, unexpected. Seeing the process as a journey was so key to what we did, and the final ring feels like the sum of these parts.

7



8

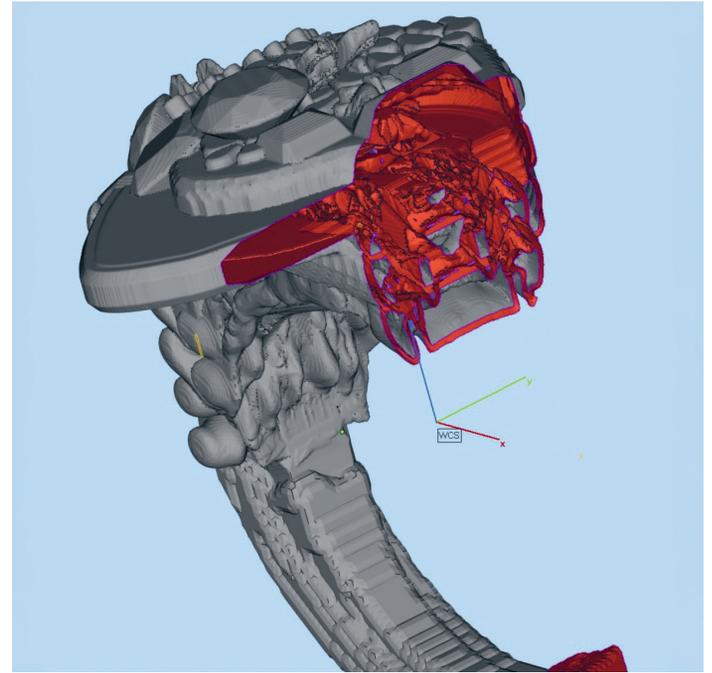


80

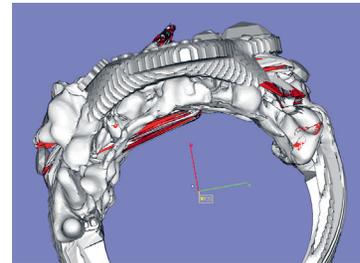
Fig.7, Selecting elements from the AI's output for inclusion in the physical ring design; digital collages

Fig.8, Ideation of translating the AI output into materiality; printing images directly onto aluminium

9



10



11



81

Fig.9, Interior view of the Infinite Loop ring; problem solving the complexity of unseen internal 3D data generated by the AI

Fig.10, Identifying problems with AI-generated geometry; red areas highlight impossible forms that can only exist digitally

Fig.11, Final 3D model of the Infinite Loop ring

Herausgegeben von
Stefanie Schneider
Hubertus Kohle

DFG-Schwerpunktprogramm ‚Das digitale Bild‘



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

DFG Deutsche
Forschungsgemeinschaft

UB | Universitätsbibliothek
Ludwig-Maximilians-Universität München

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

Kontaktadresse nach EU-Produktsicherheitsverordnung:
info@buchschnie.at



DOI <https://doi.org/10.5282/ubm/epub.126558>
ISBN 978-3-99181-375-0

Reihe: Begriffe des digitalen Bildes
Reihenherausgeber
Hubertus Kohle
Hubert Locher



Das DFG-Schwerpunktprogramm ‚Das digitale Bild‘ untersucht von einem multiperspektivischen Standpunkt aus die zentrale Rolle, die dem Bild im komplexen Prozess der Digitalisierung des Wissens zukommt. In einem deutschlandweiten Verbund soll dabei eine neue Theorie und Praxis computerbasierter Bildwelten erarbeitet werden.

