

5.6 CULTURAL ANALYTICS: DIE QUANTIFIZIERUNG KULTURELLER PHÄNOMENE

Im Jahre 2007 beantwortete Lev Manovich die Geburt von *Google Analytics* mit einem gleichermaßen ambitionierten wie streitbaren Text, den er programmatisch mit *Cultural Analytics* betitelte und in dem er nicht weniger einfordert als eine fundamentale Reform im Umgang der Kulturvermittlung und Kulturwissenschaften mit dem Wissen über die Welt, in der wir leben. Es gelte, so seine These, sich hier Technologien zu öffnen, die in anderen Bereichen – wie den ›harten‹ Wissenschaften, der Wirtschaft und diversen Regierungsagenturen – längst essentiell geworden seien. Er meint damit nicht etwa nur die ›neuen Medien‹ als Präsentationsplattform, sondern mit ihnen einhergehend vor allem jene häufig unter dem Begriff *big data* gehandelten Methoden des *data mining*, der statistischen Auswertung und der simulatorischen Visualisierung, mit denen Muster, Regelmäßigkeiten und Anomalien in extrem großen Datensätzen ausfindig gemacht werden sollen. Der Kulturbetrieb, ob er nun profitorientiert sei oder nicht, komme nach Manovichs Dafürhalten nicht länger umhin, sich diese Werkzeuge ebenfalls zu eigen zu machen – schließlich fände bereits ein Großteil (wenn nicht gar der größte Teil) unserer Auseinandersetzung mit kulturellen Inhalten in digitaler und vernetzter Form statt. Zum Zeitpunkt der Entstehung seines Textes hatten allein *Google* und *Amazon* bereits zehntausende von Büchern eingescannt und im Netz abrufbar gemacht, und zur Illustration der schiereren Datenmasse verweist Manovich auf die Fotoplatform *Flickr*, auf der im Juni 2007 nicht weniger als 560 Millionen Bilder samt Metadaten abrufbar waren. Das Ziel, das Manovich mit den *Cultural Analytics* verfolgt, erscheint dabei als eine extrem konsequente technische Umsetzung dessen, was Friedrich Kittler 1980 folgenreich die »Austreibung des Geistes aus den Geisteswissenschaften« (vgl. Kittler 1980: 8ff.) nannte: nämlich die Ablösung der Kulturarbeit von dem, was man klassischerweise den ›Inhalt‹ nennen würde, und eine Hinwendung zu den situativ-medialen Umständen seiner Rezeption. Gegenstand von *Cultural Analytics* sind für Manovich nicht Inhalte, sondern die Daten, die im Zuge ihrer Verwendung generiert werden. Das System soll dreierlei in den Fokus rücken: Erstens die existierenden Metadaten, mit denen kulturelle Inhalte ins Netz hochgeladen werden, zweitens die neuen Metadaten, die aus dem Umgang mit im Netz verfügbaren Inhalten entstehen (ein einfaches Beispiel wären hier z.B. Bildanalysen, die sich auf digital abrufbare Bilder beziehen und ihrerseits im Web publiziert sind), und drittens schließlich die virtuellen Spuren von Netz-Nutzern, die kulturelle Inhalte online produzieren, veröffentlichen, anschauen, verändern, diskutieren, konsumieren usw. Die so gewonnene Datenflut gilt es dann laut Manovich nicht etwa unters metaphorische Mikroskop zu legen und im hermeneutischen Detail zu untersuchen – dies liefe den Implikationen des Mediums völlig zuwider –, sondern

vielmehr je nach Interessenlage der jeweiligen Untersuchung auf zweckdienliche Art zu visualisieren (vgl. Manovich 2007: 1f.).

Es geht Manovich also um eine neue Form der Geisteswissenschaften und der kulturellen Vermittlung, die sich nicht länger am vermeintlich ›Wesentlichen‹ kultureller Gegenstände abarbeitet, sondern vielmehr nach ihrer Verbundenheit – oder eben *connectedness* – in Massenbewegungen des Wissens fragt, die womöglich rein gar nichts Wesentliches mehr an sich haben. Was dahinter steht ist eben die McLuhansche Erkenntnis, dass das Medium die Botschaft ist, oder in diesem Falle präzisiert: Dass Such- und Abrufsysteme für Information Determinanten sind für deren Wirken in unserer sozialen Umwelt. Die Suchmaschine ist für Manovich ausdrücklich kein indexikalisches Auffindewerkzeug, dessen Wirkung und Anwendbarkeit sich auf das spezifische technische System beschränkt, in dem es als Software ausgeführt wird. Weil der Computer nach Manovich ja ein Filter für die Gesamtheit aller Kultur geworden ist, wird die Suchmaschine unweigerlich zu einem ihrer zentralen Interfaces. Und auch wenn Empfehlungssysteme überwiegend noch die Domäne von kommerziellen Dienstleistern wie eben Internet-Kaufhäusern sind, so werden doch auch sie zunehmend zu maßgeblichen Navigatoren, die uns im Dschungel der Virtualitäten zu dem führen, was uns interessiert und gefällt (oder die – das wäre die kulturpessimistischere Lesart – womöglich aus der assoziativen Logik des Webs heraus unsere Interessen nicht etwa abbilden, sondern überhaupt erst produzieren). Kultur, davon ist Manovich überzeugt, ist letztlich keine Ansammlung dezidiert gemachter Fakten, sondern vielmehr *data* – das gefundene (und somit *gegebene*) Ergebnis des Umgangs mit Gegenständen und ihren Bedeutungen, die in der Rezeption erst aktualisiert werden. Das revolutionäre Potential digitaler Medien besteht nun darin, dass diese Aktualisierungsprozesse nicht mehr nur in den uneinsehbaren Köpfen von Rezipienten ablaufen, sondern messbare Spuren hinterlassen, weil jeder Abruf von Information dem Netzwerk und seinen Abrufsystemen gegenüber eine Aussage darstellt und somit selbst *data* wird, die in großem Stil ›geminet‹ werden kann. Naturwissenschaftliches, kommerzielles und gouvernementales Wissen wird längst auf diese Art produziert – und wenn sich die Kultur- und Geisteswissenschaften solche Methoden nicht ebenfalls aneignen, dann laufen sie nach Manovich Gefahr, über kurz oder lang hoffnungslos hinter ihre eigenen Untersuchungsgegenstände zurückzufallen (vgl. ebd.: 2f.).

Damit will Manovich zugleich eine Kulturpädagogik und -wissenschaft anstoßen, die ihren Blick anstatt auf die Geschichte auf die Gegenwart richtet. Jede wissenschaftliche Auseinandersetzung mit der kulturellen Welt hat ja unweigerlich mit ihrer eigenen Nachträglichkeit zu kämpfen. Wer kulturelle Trends an Inhalten ablesen will, der kann dies immer nur im Nachhinein tun, wenn die entsprechenden Diskurse und die mit ihnen einhergehenden Bedeutungsverschiebungen längst stattgefunden und mediale Zeugnisse in genügendem Ausmaß hinterlassen haben. Manovich sieht die Attraktivität von *Cultural Analytics* nicht zuletzt darin begründet, dass sich mit ihnen

kulturelle Prozesse in Echtzeit beobachten und beschreiben lassen sollen – und mit diesem Umstand verbinden sich für ihn auch ganz neue Bilder von den Dispositiven der Geisteswissenschaften und der Kulturvermittlung. Die Zukunft der *humanities* ist seines Erachtens nicht der individuelle Forscher, der in einer Bibliothek über einzelnen Büchern oder in einem Museum über einzelnen Kulturgegenständen brütet. Stattdessen soll unser Verständnis davon, was unsere Kultur ist und wie sie funktioniert, bald in riesigen Datenzentren produziert werden, die laufend ein unvorstellbar großes Kommunikationsvolumen erfassen und auswerten. Der Geisteswissenschaftler der Zukunft sitzt nach Manovichs Vorstellung in einer Kommandozentrale vor riesigen Bildschirmen, auf welchen ständig sekundenaktuelle Visualisierungen von Entwicklungen innerhalb großer Datenmengen ablaufen, und er illustriert diese Vision mit dem Foto einer Verkehrsüberwachungszentrale in Florida (vgl. ebd.: 3).

Dabei sind diese Ideen nicht vollkommen neu. Bereits recht früh in der zweiten Hälfte des zwanzigsten Jahrhunderts entstand innerhalb der Geisteswissenschaften ein Forschungszweig, der zunächst den Namen *Humanities Computing* trug und in den 1990er Jahren zunehmend unter der Bezeichnung *Digital Humanities* firmieren sollte. Als Initialzündung wird hier meist das Mitte der 1940er Jahre (und damit abermals zeitgenössisch zu Vannevar Bush und André Malraux) vom italienischen Jesuitenpater Roberto Busa angestoßene Projekt genannt, einen vollständigen Index sämtlicher Worte in den mittellateinischen Originaltexten des Thomas von Aquin anzulegen. Busa war offenbar nur allzu bewusst, dass dieses Vorhaben in rein menschlicher Kopf- und Handarbeit kaum zu bewältigen sein würde, was ihn dazu veranlasste, sich 1949 um die Unterstützung des damaligen IBM-Geschäftsführers Thomas J. Watson zu bemühen. Mit der Schützenhilfe des bereits seit 1911 in der Herstellung von Rechenmaschinen tätigen Konzerns wurde in den folgenden Jahren und Jahrzehnten das gesamte zu erschließende Textvolumen auf Lochkarten übertragen, die sich anschließend maschinell auswerten ließen. Der *Index Thomisticus* erschien ab 1974 zunächst in gedruckter Form, bevor er 1992 erstmals digital auf CD-ROM veröffentlicht wurde (vgl. Hockey 2004; vgl. Niewerth 2013: 4).

Die Institutionalisierung dieser Form statistischer Texterschließung mit maschinellen Hilfsmitteln als legitimer disziplinärer Zugang der Geisteswissenschaften geschah 1987 mit der Gründung der *Text Encoding Initiative*. Diese Interessengruppe von Wissenschaftlern war geographisch wie fachlich gleichermaßen weit verstreut, kommunizierte vorwiegend über Mailinglisten miteinander und hatte sich zunächst ein wenig utopisch klingendes Ziel gesetzt: Es ging ihr darum, ein Standardformat für die digitale Codierung von geistes-, sozial- und kulturwissenschaftlichen Texten auszuarbeiten, das elektronische Forschungsressourcen für Wissenschaftler in den entsprechenden Disziplinen leichter verfügbar werden lassen sollte. Die 1994 erstmals vorgestellten *TEI Guidelines* besitzen bis heute Gültigkeit (vgl. Text Encoding Initiative 2014; vgl. Niewerth 2013: 4). Der Begriffswechsel vom *Humanities Computing* zu den *Digital Humanities*, der zum Ende der 1990er Jahre einsetzte, war vor

allem in einer ausgeweiteten Perspektive und einem gewachsenen Selbstbewusstsein dieses Forschungsansatzes begründet: Während *Humanities Computing* noch weitgehend als ein weiteres Werkzeug unter vielen im bestehenden Repertoire der Geisteswissenschaften konzipiert war und sich im Ansatz sehr bestimmt auf die Nutzung des Computers zur Beherrschbarmachung von Texten konzentrierte, wollten die *Digital Humanities* das Potential des Computereinsatzes in der geisteswissenschaftlichen Forschung sehr viel umfassender in Augenschein nehmen – gegebenenfalls bis hin zum völligen Umdenken über Funktion, Rolle und Auftrag der eigenen Disziplin in der digitalisierten Welt (vgl. Schreibman, Siemens u. Unsworth 2004; vgl. Nierwerth 2013: 4).

Dementsprechend sind die *Digital Humanities* getragen von einer doppelten Einsicht: Nämlich erstens jener, dass der wissenschaftliche Umgang mit einer digitalisierten Kultur auch digitaler Methoden bedarf, und zweitens jener, dass die Methoden einer Wissenschaft niemals nur Werkzeuge sind, die nicht auf ihren Nutzer zurückwirken, sondern vielmehr die Struktur jedweder wissenschaftlichen Erkenntnis vorab mitbestimmen. Diese beiden Einsichten erschöpfen sich indes nicht in ihrer Bedeutung für die geisteswissenschaftliche Forschung, weil sie darüber hinaus auf einen Tatbestand von noch umfassenderer Bedeutung verweisen, welcher in Malrauxs Konzeption des Imaginären Museum bereits anklingt: Die Reproduktion kultureller Inhalte ist niemals nur ein Prozess der Vervielfältigung und Kopistik, sondern notwendigerweise immer auch einer der Transformation. Der transformative Charakter der Reproduktionen liegt darin begründet, dass anhand ihrer andere Formen der Rezeption möglich werden, als die Originale sie erlauben. Der Buchdruck befreite die Literatur aus den Bibliotheken von Adel und Klerus und ermöglichte ihre Rezeption in bürgerlichen Settings, die mechanische Fotografie nach Malraux befreite die Kunst aus den Museen und Galerien, durchspülte die Alltagswelt mit ihr und hob virtuell die Voneinander-Geschiedenheit der Einzelobjekte auf. In beiden Fällen ›verändern‹ sich die kulturellen Ursprungsobjekte unwiederbringlich. Die von neuen Medientechniken ermöglichten Verkehrsformen bereits bestehender kultureller Inhalte operieren nicht folgenlos innerhalb bestehender Dispositive, sondern verändern das Gefüge kulturellen Verstehens insgesamt: Nie wieder wird ein Mensch das Alte Testament so lesen können, wie es ein katholischer Geistlicher vor Luther und Gutenberg vermochte, und nie wieder werden wir die Mona Lisa mit den Augen eines Menschen sehen können, für den sie nur im *Louvre* existierte. Jedoch hat die analoge Reproduktion eines niemals angetastet: Die völlige Verortung kultureller Objekte innerhalb der Kulturwelt. Ein Foto der *Erschaffung Adams* aus dem Deckenfresko der Sixtinischen Kapelle ist natürlich nicht identisch mit dem Deckenfresko selbst, aber auch auf der glänzenden Oberfläche eines Polaroids bleibt es ein Bild, das nur in kulturellen Zusammenhängen gelesen werden kann – nämlich zum einen als das Abbild eines berühmten Kunstwerkes, und zum anderen als eine Darstellung der biblischen Schöpfungsgeschichte. Unser erlerntes Sehen und Deuten erlaubt es uns überhaupt nicht,

die Darstellungsfunktion des Objektes und seiner Reproduktionen infrage zu stellen. Auch die Wilde Semiose nach Aleida Assmann bricht nicht aus der Kulturwelt aus – sie setzt ja gerade voraus, dass in den Dingen eine Lesbarkeit vermutet wird, eine Verweisstruktur, die das Gesehene mit dem Gewussten verbindet (vgl. Assmann 1988: 240). In der digitalen Reproduktion hingegen verlassen die Dinge den *culture layer* und werden nicht nur in etwas transformiert, was sie eben *eigentlich* nicht sind (das gilt im strengen Sinne ja auch für die Fotografie), sondern in etwas, das für uns ohne den strategischen Einsatz von Interfacetechnologien wie eben dem Computerbildschirm überhaupt nicht mehr als ein ›Etwas‹ zu erkennen ist.

Was sich in den *Digital Humanities* also andeutet und was Manovich mit den *Cultural Analytics* aufgreifen und im großen Maßstab weiterdenken will, ist nicht nur eine Erschließung neuer Analysemethoden für veränderte Arbeits- und Rezeptions-situationen. Vielmehr wird hier ein fundamentaler Wandel im Wesen kultureller Objekte und Äußerungen thematisiert, welcher sich eben im Prozess der Digitalisierung vollzieht und dessen Analyse in einen völlig neuen Kontext stellt: Kulturelle Phänomene sind im Digitalen nicht länger nur interpretier-, sondern nunmehr auch quantifizierbar, bzw. vielmehr noch: In ihrer Aufbereitung als digitaler Code sind sie bereits quantifiziert, denn nur so können sie auf dem *computer layer* überhaupt gespeichert, prozessiert und transcodiert werden. Ein Text ist nicht länger ›nur‹ eine sprachliche Äußerung, die gelesen und verstanden werden kann, sondern repräsentiert eine genau vermessbare Menge an Information, die sich wiederum statistisch beschreiben lässt. Ein Gemälde ist nicht länger einfach Farbe auf Leinwand, die kulturell als eine Anhäufung von bedeutsamen Zeichen verstanden werden kann, sondern eine mathematisch aufschlüsselbare Verteilung von Farbwerten über ein Raster. Kulturelle Objekte, die sich zuvor scheinbar nur qualitativ und sprachlich verhandeln ließen, können nunmehr auch quantitativ und somit in Zahlen beschrieben und verarbeitet werden – und der Modus dieser Beschreibung ist die von Manovich geforderte zweckorientierte Visualisierung, die jeden beliebigen Aspekt der zugrundeliegenden Datenmenge hervorheben und verbildlichen kann.

Dies eben ist der Verlust des ›Eigentlichen‹ in der Digitalisierung: Unter den zahlreichen Arten, auf welche sich digitale Information visualisieren lässt, ist keine besonders privilegiert. Eine Bilddatei der Mona Lisa *als* Mona Lisa in Form bunter Pixel auf einen Bildschirm zu projizieren ist grundsätzlich keine ›richtigere‹ Form ihrer Darstellung als jede andere. Insofern entstehen die *trails of interest* nicht nur im Hypertext zwischen mehr oder weniger diskreten digitalen Objekten. Vielmehr sind die digitalen Objekte Huis und die Attributobjekte Novaks selber Produkte strategischer Visualisierung von Daten, die man sich auch ganz anders erschließen könnte. Für virtuelle Museen bedeutet dies ganz konkret, dass sie ihre epistemische Architektur auf zwei Etagen entfalten müssen: auf dem *culture* und dem *computer layer*. Soll so etwas wie museales Erleben im Web überhaupt möglich sein, dann müssen erstens abstrakte digitale Daten auf eine Art und Weise auf dem Monitor visualisiert

werden, die sie als kulturelle Objekte erkennbar werden lässt (und in vielen Fällen wird das heißen: als Abbildungen von physisch existierenden Dingen), und zweitens diese digitalen Objekte auf eine Art zueinander in Beziehung gesetzt und mit einem Abrufinterface versehen werden, das einen ähnlichen Kompromiss von didaktischer Führung und navigatorischer Assoziierbarkeit ermöglicht, wie es der physische Museumsraum tut.

Lev Manovichs Hauptmotivation für die Konzeption der *Cultural Analytics* als Alternative zu den *Digital Humanities* ist dabei seine Beobachtung, dass letztere immer noch allzu befangen seien in klassischen, analogen Vorstellungen von Kultur, und damit auch in praxis- und lebensfernen Expertendiskursen. Das offensichtlichste Symptom dafür sieht er darin, dass existierende Analysesoftware für digitale Kulturgüter diese vornehmlich über Meta-Daten erschließbar zu machen trachtet: Sie blickt, so stellt Manovich fest, nicht *in* die eigentlichen Daten und auch nicht auf die Nutzungsmuster, in welche sie eingebunden sind, sondern auf ein die Daten umgebendes, von Menschen katalogisiertes Verwaltungswissen. Insofern füge sie den bestehenden Zugängen nur ein etwas schnelleres Abrufsystem hinzu (vgl. Manovich 2007: 6). Darüber hinaus erschöpften sich abseits aller visionären Vorstellungen von einer neuen Form der Kulturwissenschaft die tatsächlich Anwendungen der *Digital Humanities* in der statistische Analyse von Texten – und zwar vor allem von hochkulturellen Texten, deren Auswahl immer noch nach den bestehenden Vorurteilen und Relevanzvorstellungen einer Wissenschaft stattfindet, die dem gedruckten Wort verhaftet sei. Die *Digital Humanities* tun sich, so moniert Manovich, schwer mit all jenen digitalen Datenformen, die keine Schriftsprache abbilden – und mit der gesamten Fülle kultureller Artikulationen, die überhaupt erst aus digital-vernetzter Kommunikation heraus entstehen können (vgl. ebd.: 10). In anderen Worten: Die *Digital Humanities* wollen klassische geisteswissenschaftliche Zugänge um digitale Erschließungsmethoden erweitern, deren Nutzen aber unweigerlich in die etablierte Domäne der Geistes- und Kulturwissenschaften zurück münden muss, nämlich den *culture layer*. Paradoxerweise entstehen neue kulturelle Artikulationsmöglichkeiten aber gerade dort, wo kulturelle Kommunikation die Sphäre der Kultur verlässt und als digitaler Code eben in *alles* transformierbar wird. In der Digitalität können kulturelle Praktiken entstehen, welche den etablierten und didaktisch legitimierten Umgang mit Inhalten unterlaufen, gerade weil diese Inhalte als etwas anderes visualisiert, rezipiert und genutzt werden können, als das, was sie »eigentlich« zu sein scheinen. Das mag nun tatsächlich das Bild eines »Museums von Babel« evozieren, und mit ihm die Dystopie von der Beliebigkeit allen medial codierbaren Wissens. Laut Manovich ist dies jedoch gerade nicht der Fall.

Denn während die Bibliothek von Babel das furchteinflößende Produkt einer Aktualisierung all dessen ist, was ihr Codesystem virtuell-kombinatorisch hergibt, ist das Web eben kein Produkt, sondern Prozess. Zwar sind kombinatorisches Compu-

ter-Kauderwelsch und von menschlichen Absichten bedingte und geleitete Programmiertätigkeit ontisch nicht voneinander zu unterscheiden, funktional lässt sich aber sehr wohl differenzieren zwischen dem, was ein Computer aus seinem formallogischen Aufbau heraus gewissermaßen ›für sich selbst‹ tut, und dem, was er für die kulturelle Rezeption durch Menschen über Interface-Systeme visualisiert. Manovich möchte seinen analytischen Blick nicht auf die Masse der Virtualitäten des Webs und der vernetzten Computer richten, sondern auf die Aktualitäten dessen, was Menschen mit Computern machen. Der Zugang dieser Analyse ist dabei kurioserweise gerade *doch* das virtuelle Element digitaler Systeme, aus dem heraus Aktualisierung erst möglich wird: *Cultural Analytics* reißen, indem sie sich der Visualisierung digitaler Daten als Werkzeug zur Erforschung digital mediierten Kulturphänomene bedienen, die Grenze zwischen Methode und Untersuchungsgegenstand völlig ein.

Im und um das World Wide Web hat sich längst eine ganze Industrie entlang der Notwendigkeit konstituiert, die digital-formalen Abbilder kulturell motivierter Entscheidungsprozesse vermessbar und aufeinander beziehbar zu machen, um sie dann wieder in kulturelle Zusammenhänge zurück zu überführen. Die *Cultural Analytics* greifen die unleugbare Tatsache auf, dass dynamische ›Navigatoren‹ wie Suchmaschinen und Empfehlungssysteme längst zu Architekten von Sinngefügen im Netz geworden sind, und dass es damit keine feststehende ›Ontologie‹ mehr gibt, innerhalb derer sich digitale Objekte konstituieren und deren einmalige Entschlüsselung die Gesamtheit aller Modalitäten der Sinnproduktion im World Wide Web offenlegen könnte. Stattdessen ändern sich mit den Objekten des Wissens auch laufend die Umstände und Richtlinien ihrer Hervorbringung.

