

„... A DISPLAY (NOT A REPRESENTATION)...“

Zur Sichtbarmachung von Daten¹

VON ARNO SCHUBBACH

Der Begriff des Displays ist in mehrerlei Hinsicht von Interesse. Zum einen ist seine Geschichte eng verknüpft mit der Entwicklung von Darstellungstechniken, die meist in Konkurrenz zueinander stehen und stets auch das Verhältnis zwischen Mensch und Maschine verändert haben.² Zum anderen scheint der Begriff geeignet, um die neuen computerisierten von älteren Bildtechniken abzuheben. Denn auf Displays erscheinen trotz ihrer Ursprünge in der Radartechnik längst Bilder, die auf Algorithmen und Daten verschiedenster Art beruhen. Daraus wurde die These abgeleitet, dass diese Bilder sich von einer Wirklichkeit, die sich in optisch-chemische Photographien vielleicht noch hätte einschreiben können, gelöst oder sie durch Simulationen gar ersetzt hätten.³ Obwohl sich diese These also darauf stützt, dass Bilder berechnet werden, wurden die konkreten Verfahren und Algorithmen in den verschiedensten Feldern der Visualisierung kaum auf differenziertere Weise diskutiert. Eine solche Diskussion ist jedoch notwendig, um Displays zu charakterisieren, die auf der Basis von Daten und Algorithmen etwas sichtbar machen, damit der Nutzer möglichst leichthändig mit dem Computer interagieren kann. Displays lassen sich weder ausschließlich auf die zugrunde liegenden Algorithmen und Daten zurückführen noch auf ihre jeweils zu sehenden Ausgaben. Sie sind Scharniere zwischen vielfältigen algorithmischen Sichtbarmachungen durch den Computer und der phänomenalen Sichtbarkeit für den Nutzer.⁴

-
- 1 Dieser Text geht auf einen Vortrag zurück, den ich an der Tagung „DISPLAY. Die kulturelle Entfaltung digitaler Bilder“ (14./15. Juli 2006) gehalten habe. Ich danke Hans-Christian von Herrmann und allen TeilnehmerInnen für zahlreiche Anregungen.
 - 2 Vgl. dazu Thielmann, Tristan: „Statt einer Einleitung: Eine Mediengeschichte des Displays“, in: *Navigationen. Zeitschrift für Medien- und Kulturwissenschaften*, „Display I: Analog“, hrsg. v. Jens Schröter/Tristan Thielmann, Jg. 6, H. 2, 2006, S. 13-30.
 - 3 Vgl. z.B. Kittler, Friedrich: „Fiktion und Simulation“, in: Barck, Karlheinz u.a. (Hrsg.): *Aisthesis. Wahrnehmung heute oder Perspektiven einer anderen Ästhetik*, Leipzig 1998, S. 196-213, hier S. 200ff. Vgl. auch Baudrillard, Jean: *Simulacres et Simulation*, Paris 1981, S. 12ff., allerdings ohne expliziten Bezug auf technische Medien.
 - 4 Vgl. dazu auch Schröter, Jens: „Statt einer Einleitung: Versuch zur Differenz zwischen dem Medialen und dem Display“, in: *Navigationen. Zeitschrift für Medien- und Kulturwissenschaften*, „Display I: Analog“, hrsg. v. Jens Schröter/Tristan Thielmann, Jg. 6, H. 2, 2006, S. 7-12, hier S. 7f. Den Begriff der Sichtbarmachung entlehne ich den Arbeiten von Hans-Jörg Rheinberger. Vgl. z.B. Rheinberger, Hans-Jörg: „Objekt und Repräsentation“, in: Heintz, Bettina/Huber, Jörg (Hrsg.): *Mit dem Auge denken. Strategien der Sichtbarmachung in wissenschaftlichen und virtuellen Welten*, Zürich u.a. 2001, S. 55-61.

I DAS DISPLAY: DATEN, ALGORITHMEN, BILDER

Die wesentlichen Funktionen des Displays zeigen sich von Beginn an. Es wurde erstmals im Luftverteidigungssystem SAGE (*Semi-Automatic Ground Environment*) der USA im großen Maßstab eingeführt, um Daten zu visualisieren, die via Telefonleitung von Radarstationen übermittelt und vom Computer verarbeitet wurden. An zahlreichen Arbeitsplätzen wurde auf diese Weise ein Verteidigungssektor mitsamt Flugzeugen und ihren Bewegungen übersichtlich dargestellt, um eine schnelle Reaktion des militärischen Bedienungspersonals zu ermöglichen und diese zeitnah in neue Ausgaben umzusetzen. Eine solche Interaktion war zuvor unmöglich gewesen, weil die Eingabe mittels Lochkarten und die Ausgabe von kaum verständlichen Papierausdrucken zeitlich weitgehend voneinander entkoppelt waren.⁵ Die Displays von SAGE haben somit ein neuartiges Zusammenspiel von algorithmischer Datenverarbeitung und menschlicher Entscheidung eröffnet und damit das Bedienungspersonal auf neue Weise in das Luftüberwachungssystem eingebunden.

Das Display hat aber auch die Form der Ein- und Ausgabe selbst verändert: Es löste die bis dahin vorherrschende alphanumerische Ausgabe ab und präsentierte in erster Linie Graphiken. Mit dem Display hält das Bild Einzug in die Welt von Computer und Nutzer. Dies zeigt eindrücklich der zwölfminütige Film *ON GUARD!*, der ungefähr 1956 von der *Military Products Division* der *IBM Corporation* gedreht wurde.⁶ In propagandistischer Absicht führt er neben anderen neuen Kriegstechnologien auch SAGE vor und setzt dessen bahnbrechende Displays ins Bild. Nachdem die Datenverarbeitung des begehren Computers vor Augen geführt wurde, bereitet die männliche Off-Stimme den Schnitt auf einen einzelnen Arbeitsplatz vor: „Beyond the fantastic capacity for calculation and memory, SAGE possesses the newest and most revolutionary advance in data processing: the display scope, a computer generated visual display, on call as needed.“⁷ Um diesen revolutionären Fortschritt ins Bild zu setzen, folgt ein Rückblick auf die um 1956 noch gegenwärtige Vergangenheit: Der Film zeigt in zwei Einstellungen Frauen, die um einen Drucker herumgehen und sich um dessen Stapelausdrucke

5 Vgl. zum so genannten *batch processing* Ceruzzi, Paul E.: *A History of Modern Computing*, Cambridge, Mass. u.a. 2003, S. 71-78. Die Technik der Interaktion geht dem Begriff jedoch voraus: „Construction of a realistic direction center depended heavily on the development of a versatile display console that provided the operators the information they needed to make decisions and also provided them the means to send commands to the computer. (The modern term *interactive* applies but was not in common use in the early 1950s.)“, so Wieser, C. Robert: „The Cape Cod System“, in: *Annals of the History of Computing*, Jg. 5, Nr. 4, 1983, S. 362-369, hier S. 367.

6 Der Film entstand in Zusammenarbeit mit dem *Department of Defense* der USA, der *U.S. Air Force* und der *Boeing Airplane Company* und ist in verschiedenen Formaten verfügbar unter <http://www.archive.org/details/OnGuard1956>, 09.07.2007. Ich entnehme auch die angeführten Angaben dieser Website.

7 *ON GUARD!* (USA ca. 1956, IBM Corporation), ab 3'45”.

zu kümmern scheinen. Die Off-Stimme unterstreicht, wie ungenügend diese Technik ist: „Until SAGE the miracle of the computer was its ability to calculate in split seconds and then provide printed information. But SAGE needed more than this.“ Einige Sätze später wird dagegen die schon gegenwärtige Zukunft des Displays hervorgehoben:

Air defense required split second presentation as well as split second calculation. Given this objective, IBM filed the latest extension of data processing: the display scope, a giant picture tool, on which computer results are instantaneously and continuously translated into graphic images. In SAGE, airmen have the battle visualized for them on the computer generated display.⁸

Wir sehen unterdessen zunächst den riesigen Rechenraum und sodann zwei IBM-Ingenieure, die an einer Arbeitsstation von SAGE lehnen. Die Kamera gleitet schließlich an ihnen hinab, bis das kreisrunde, leere und reflektierende Display in frontaler Aufnahme nahezu das ganze Bild füllt.⁹ Nach einem Blick in einen abgedunkelten Raum voller Arbeitsplätze, an denen das Bedienungspersonal mit Lightguns hantiert, wird das Display, das ca. alle 2 ½ Sekunden aufgefrischt wird, nochmals in Funktion gezeigt.

ON GUARD! führt so vor Augen, wie die graphische Ausgabe zu einem besseren Überblick über die verfügbaren Daten verhilft. Der Film spielt die Stärke seiner bewegten Bilder aus, um die Vorteile der interaktiven Visualisierung von SAGE zu zeigen. Er bestimmt aber trotz aller propagandistischen Absichten recht präzise, dass diese Visualisierung darstellt, was der Computer berechnet hat. Denn während wir erneut militärisches Personal an den Arbeitsplätzen sehen, charakterisiert die Off-Stimme die Funktion der Displays durchaus zutreffend: „Although they look like the offspring of a marriage between a television tube and a radar screen, display scopes do not show physical images transmitted from elsewhere. They display the results of the computer’s findings.“ Displays sind Orte der Sichtbarmachung von Berechnungen und Daten.

Bereits bei SAGE haben Displays somit zwei Seiten: Zum einen präsentieren sie Bilder, damit der Nutzer die Lage möglichst schnell erfassen, eine Entscheidung treffen und sie dem Computer zur Weiterverarbeitung leicht mitteilen kann. Zum anderen beruhen sie oft auf sehr umfangreichen Datenmengen, die der Computer erfasst, algorithmisch verarbeitet und darstellt.¹⁰ Der praktische Sinn

8 ON GUARD! (wie Anm. 7), 4’14”.

9 Auf die Rollen, die Frauen und Männern in dieser Schnittfolge zugewiesen wird, wie auch auf andere politisch-ideologische Implikationen des Films kann hier nicht eingegangen werden.

10 Vgl. zu dieser Doppelseitigkeit von Bildern Nake, Frieder: „Das doppelte Bild“, in: *Bildwelten des Wissens. Kunsthistorisches Jahrbuch für Bildkritik*, Jg. 3, Nr. 2, 2005, S. 39-50, hier insbesondere S. 45ff. und ders.: „The Display as a Looking-glass: Zu Ivan E. Sutherlands früher Vision der grafischen Datenverarbeitung“, in: Hellige, Hans-Dieter (Hrsg.):

solcher Displays ist die Ermöglichung eines produktiven Zusammenspiels von Mensch und Maschine. Denn der Nutzer wird entlastet von der mühseligen und manuell oft unmöglichen Erfassung und Verarbeitung von Daten und kann umso besser darauf reagieren, was ihm der Computer übersichtlich darstellt. Die Visualisierungen zielen somit auf die Entscheidungen des menschlichen Nutzers ab und sind deshalb nur notwendig, wenn der Computer aufgrund der Komplexität oder des Umfangs der Daten nicht auch die notwendigen Maßnahmen errechnen kann. In diesem Sinn ‚sieht‘ der Computer nicht das dargestellte Bild, anhand dessen der Nutzer Entscheidungen trifft.¹¹ Das Display muss deshalb als komplexe und irreduzible Kopplung zwischen Daten, ihrer algorithmischen Sichtbarmachung und den sichtbaren Bildern gefasst werden.

Im Hinblick auf diese Bestimmung des Displays versprechen gegenwärtige Forschungen genauere Einblicke. Alle bildgebenden Verfahren setzen mit Hilfe komplexer Algorithmen Daten in Bilder um. Der hochgradig berechnete, also vermittelte Charakter der Bilder tritt jedoch oft in den Hintergrund zugunsten einer Darstellung von anscheinend realen Gegenständen, weshalb das Display, wenn nicht wie ein Fernsehapparat oder Radarschirm, so doch wie ein Fernrohr, Mikroskop oder Röntgenapparat zu funktionieren scheint. Eine präzisere Bestimmung des Verhältnisses von Daten und Bildern eröffnet dagegen die Informationsvisualisierung. Sie verfolgt seit Ende der 1980er Jahre in verschiedensten Anwendungsbereichen den Ansatz, Daten zu visualisieren, um Einsichten in ihre inneren Zusammenhänge zu gewinnen.¹² Die Informationsvisualisierung bezeichnet also „a computer-aided process that aims to reveal insights into an abstract phenome-

Geschichten der Informatik. Visionen, Paradigmen, Leitmotive, Berlin 2004, S. 339-365, hier insbesondere S. 345ff. Nake beschreibt den Bezug der algorithmischen und phänomenalen Seite des Bildes jedoch mitunter als Entsprechung und identifiziert dann meist mit Rekurs auf den Begriff des Zeichens, was der Mensch sieht und der Computer in Daten erfasst. Vgl. ebd., S. 360f. Anhand der Visualisierung von Informationen möchte ich im Folgenden dagegen zeigen, dass hier der Mensch *idealiter* gerade sehen soll, was der Computer nicht formal zu erfassen vermag.

- 11 Displays haben somit dort einen praktischen Sinn, wo der Computer berechnet, was er weder in symbolischen Gesetzen erfassen noch in einem *pattern recognition* synthetisieren kann – und wo deshalb nicht auf den Menschen zu verzichten ist. Vgl. dagegen Kittler, Friedrich: „Die Welt des Symbolischen – eine Welt der Maschine“, in ders.: *Draculas Vermächtnis: Technische Schriften*, Leipzig 1993, S. 58-80, hier insbesondere S. 58-62 und 72ff. Der jüngere Aufsatz „Schrift und Zahl. Die Geschichte des errechneten Bildes“, in: Maar, Christa/Burda, Hubert (Hrsg.): *Iconic Turn. Die neue Macht der Bilder*, Köln 2004, S. 186-203 trägt der Rolle von Visualisierungen stärker Rechnung, löst sie aber vom menschlichen Blick, um sie zumindest *idealiter* in die Reflexivität und Subjektivität von Computern einzufassen, die ihre Bilder selbst sehen und analysieren würden. Vgl. ebd., S. 199ff.
- 12 Folgende Einführungen geben mit ihren unterschiedlichen Schwerpunkten eine umfängliche und aktuelle Übersicht über die Informationsvisualisierung: Chen, Chaomei: *Information Visualization. Beyond the Horizon*, London ²2004; Spence, Robert: *Information Visualization. Design for Interaction*, Harlow u.a. ²2007; Ware, Colin: *Information Visualization. Perception for Design*, Amsterdam u.a. ²2004.

non by transforming abstract data into visual-spatial forms“¹³. Die entstehenden Bilder nutzen somit die genuine Funktion graphischer Displays zur Sichtbarmachung von Daten und sind deshalb nicht als Abbilder zu betrachten. Eine Reflexion auf die Methoden und Verfahren der Informationsvisualisierung hilft deshalb diese Funktion genauer zu fassen.

Die Informationsvisualisierung umfasst ein umfangreiches Feld, das von der einfachen Präsentation weniger Daten (z.B. Infographiken¹⁴) bis hin zur interaktiven visuellen Erkundung großer Datenmengen (z.B. Kartographierung wissenschaftlicher Disziplinen¹⁵) reicht. Ihr Ansatz gewinnt vor allem dort an Bedeutung, wo man es mit äußerst komplexen Datenmengen zu tun hat, die anders kaum mehr auszuwerten sind: „visualization helps one see patterns that cannot be readily absorbed in any other way“¹⁶. Wo statistische und numerische Verfahren versagen, setzt die Informationsvisualisierung somit auf ein geschultes Auge und erstellt daher Bilder, die in äußerst unterschiedlichen, konkreten wie abstrakten Formen gestaltet sind. Seit Beginn der 1990er Jahre hat die Anzahl der Publikationen in diesem Feld rasant zugenommen,¹⁷ weshalb in der Informationsvisualisierung durchaus ein Trend zu sehen ist, der sich in den USA auf Grund der finanziellen Förderung im Rahmen der politischen Agenda der *National Home Security* wohl noch verstärken wird.¹⁸ Jenseits spezialisierter Anwendungen in staatlichen und industriellen Rechenzentren konnte sich die Informationsvisualisierung bislang aber kaum durchsetzen.

Das Ziel der Visualisierung ist stets, innere Zusammenhänge von Daten in informativen Bildern einsichtig zu machen, damit der Nutzer mit Überblick entscheiden kann, wo die Berechnungen des Computers keine befriedigende Lösung

-
- 13 So definiert der Herausgeber der 2002 gegründeten Zeitschrift *Information Visualization* im Editorial der ersten Ausgabe. Vgl. Chen, Chaomei: „Editorial: Information Visualization“, in: *Information Visualization*, Jg. 1, Nr. 1, 2002, S. 1-4, hier S. 1.
 - 14 Als jüngeres Beispiel sei hier auf die mittlerweile verbreiteten *tag clouds* verwiesen, z.B. zu den „Themen des Tages“ auf der Eingangsseite von Spiegel-Online. Vgl. auch die Erläuterungen: „Die Wolke mit dem Wichtigsten“, <http://www.spiegel.de/spiegeldigital/0,1518,444591,00.html>, 09.07.2007.
 - 15 Vgl. z.B. Chen, Chaomei: *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*, London u.a. 2003 und die Internetpräsenz der von der Informatikerin Katy Börner kuratierten Ausstellung *Places & Spaces: Mapping Science*, <http://scimaps.org/>, 09.07.2007.
 - 16 Buzydowski, Jan W. u.a.: „Term Co-occurrence Analysis as an Interface for Digital Libraries“, in: Börner, Katy/Chen, Chaomei (Hrsg.): *Visual Interfaces to Digital Libraries*, Berlin u.a. 2002, S. 133-144, hier S. 133.
 - 17 Chen lässt es sich in einem weiteren Editorial nicht nehmen, dies durch eine Anwendung der Informationsvisualisierung auf sich selbst zu zeigen. Vgl. Chen, Chaomei: „Editorial: Information Visualization is Growing“, in: *Information Visualization*, Jg. 1, Nr. 3-4, 2002, S. 159-164.
 - 18 Vgl. hinsichtlich dieser politischen Bedeutung der Informationsvisualisierung bzw. ihrer erweiterten und programmatischen Formulierung als *visual analytics* z.B. die Homepage des *National Visualization and Analytics Center*, <http://nvac.pnl.gov/>, 09.07.2007.

versprechen. Dies wird beispielhaft an der Visualisierung von Texten deutlich. Texte sind heute in einem ungeahnten Umfang gespeichert und prinzipiell leicht verfügbar. Zugleich ist es aber schwieriger geworden, den Text zu finden, der dem eigenen Interesse entspricht. Klassische Lösungen vertrauen auf ein *information retrieval*: Hierzu müssen wir unser Interesse durch wenige Worte und ihre logische Verknüpfung formulieren; der Computer durchsucht dann Internetseiten und Datenbanken und berechnet die Relevanz der Dokumente, um sie schließlich in einer Rangliste auszugeben. Aufgrund der Mängel dieses Verfahrens – das die innere Struktur der Texte ebenso wenig wie ihre inhaltlichen oder formalen Verknüpfungen berücksichtigt und zudem an der hartnäckigen Unfähigkeit der Nutzer scheitert, ihre diffusen und assoziativen Erwartungen formal zu fassen – schlug Lauren B. Doyle schon 1961 ein anderes Vorgehen vor. Angeregt von der Display-Technik des SAGE, für dessen Softwareschmiede SDC (*System Development Corporation*) er arbeitete, kam Doyle auf den Gedanken, der Computer könnte die inneren Zusammenhänge der Texte analysieren, um sie dem Nutzer selbst vor Augen zu führen und interaktiv erkunden zu lassen.¹⁹ Eine „semantic road map“ sollte Überblick verschaffen über eine Unmenge von Texten, die kaum je durchzuarbeiten wären, damit der Nutzer mit wenigen Blicken selbst entscheiden kann, welche Texte für ihn relevant sind.²⁰

Die gegenwärtige Forschung zur Visualisierung von Texten folgt diesem Ansatz und hat vielerlei Methoden und Verfahren entwickelt, die im Folgenden zumindest teilweise kurz vorgestellt werden.²¹ Wie ich zeigen möchte, ist der Übergang von den Texten zu ihrer Visualisierung jedoch notwendigerweise problematisch, weil eine Visualisierung nur insofern hilfreich sein kann, als sie die Komplexität der Texte reduziert, und daher eine prinzipielle Unsichtbarkeit mit sich bringt. Es ist diese Unsichtbarkeit, die einen differenzierteren Blick auf das Display erlaubt.

-
- 19 Doyle, Lauren B.: „Semantic Road Maps for Literature Searchers“, in: *Journal of the ACM*, Jg. 8, Nr. 4, 1961, S. 553-578, hier S. 570.
- 20 Diesen Ansatz bringt auch Robert R. Korfhage – allerdings 30 Jahre später – auf den Punkt: „In this paper we present an argument for a new retrieval paradigm, one that focuses on the organized display of *all* documents, rather than on the linear display of the best.“ (Korfhage, Robert R.: „To See, or Not to See – Is That the Query?“, in: Bookstein, Abraham u.a. (Hrsg.): *Proceedings of the 14th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval (SIGIR '91, Chicago, Illinois, USA, October 13-16, 1991)*, New York 1991, S. 134-141, hier S. 134.)
- 21 Dieser Überblick beschränkt sich auf etablierte Techniken, so dass es kaum sinnvoll ist, ihn mit Verweisen auf eine viel detailliertere Forschungsliteratur zu überfrachten. Es sei stattdessen summarisch verwiesen auf Chen (wie Anm. 15) und Börner, Katy u.a.: „Visualizing Knowledge Domains“, in: *Annual Review of Information Science and Technology*, Jg. 37, 2003, S. 179-255.

2 SICHTBARMACHUNG VON DATEN

Zur Visualisierung von Texten muss in einem ersten Schritt definiert werden, wie der Computer mit den Texten umgehen soll. Es bedarf eines mathematischen Modells, um aus ihnen Daten zu berechnen und ihre inhaltlichen Zusammenhänge formal zu analysieren, wobei die einzelnen Texte oder die AutorInnen als Einheiten der Analyse und damit der Visualisierung betrachtet werden. Zwei ältere bibliometrische Verfahren haben sich in der Informationsvisualisierung etabliert.

Die Kookkurrenzanalyse setzt am Wortgebrauch an und zielt damit auf die Semantik. Es werden in einer statistischen Analyse zuerst Worte bestimmt, die in allen Texten häufig vorkommen und nicht zu den ubiquitären Worten wie Pronomina, Hilfsverben usw. gehören. Für jeden Text wird anschließend die Häufigkeit der so genannten Indexwörter berechnet. Jeder Text wird dadurch aufgefasst als ein Wortprofil, das aus diesen Häufigkeiten besteht und mathematisch gesehen ein Punkt in einem hochdimensionalen Raum darstellt. Das gesamte Textkorpus bildet somit eine hochdimensionale Punktmenge, in der Texte mit ähnlichem Wortprofil nahe beieinander zu liegen kommen. Ein zweites bibliometrisches Verfahren analysiert anhand von Zitationen die Beziehungen zwischen den Texten detaillierter.²² Es hat sich in erster Linie für wissenschaftliche Texte durchgesetzt und wird auch weitgehend auf sie beschränkt bleiben, weil es die Praxis des wissenschaftlichen Zitierens voraussetzt. Zitationsanalysen werten aus, wie Texte Texte zitieren. Es kann zum einen berechnet werden, inwiefern die darzustellenden Texte dieselben Texte zitieren. Ihr Zusammenhang wird so im Hinblick auf die gemeinsamen intellektuellen Bezugspunkte beschrieben. Es kann aber auch berechnet werden, wie oft Texte gemeinsam zitiert werden. Der analysierte Zusammenhang gründet in diesem Fall darin, dass die Texte von späteren Texten in Verbindung gebracht werden. Auf dem ersten Weg wird somit beschrieben, wie sich ein Text situiert, auf dem zweiten Weg, wie er situiert wird. Im letzteren Verfahren, der Kozitationsanalyse, wird daher eine Ordnung dargestellt, die der Diskurs auf frühere Texte zurückwirft, und damit ein dynamisches Bild wissenschaftlicher Diskurse entworfen.²³ Deshalb wird der Kozitationsanalyse bei der Visualisierung von wissenschaftlichen Disziplinen oft der Vorzug gegeben.

Beide bibliometrische Verfahren setzen die Texte in komplexe numerische Daten um, die zur Visualisierung vereinfacht werden müssen. Es können zunächst mehrere Algorithmen angewandt werden, die eine Menge in einem hochdimensionalen Raum in einen Raum von niedrigerer Dimension abbilden. Ein gebräuch-

22 Vgl. den wegweisenden Artikel Small, Henry: „Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents“, in: *Journal of the American Society for Information Science*, Jg. 24, Nr. 4, 1973, S. 265-269.

23 White, Howard D./McCain, Katherine: „Visualizing a Discipline: An Author Co-Citation Analysis of Information Science, 1972-1995“, in: *Journal of the American Society for Information Science*, Jg. 49, Nr. 4, 1998, S. 327-355, hier S. 329 formulieren prägnant: „What is actually mapped is an author's citation image.“

licher Algorithmus ist das so genannte *multidimensional scaling* (MDS): Es rechnet die Punktemenge auf die zwei, drei auf dem Display darstellbaren Dimensionen herunter und erhält ihre in einer zu wählenden Metrik bestimmte Struktur möglichst getreu.²⁴ Das Ergebnis kann in abstrakter Form als Punktemenge in der Fläche oder im dreidimensionalen Raum visualisiert werden, entsprechend einer ‚visuellen Metapher‘ aber auch in Form einer Galaxie oder einer Landschaft.²⁵ Stets werden die Texte vermittels ihrer algorithmischen Berechnung aber als eine räumliche Konfiguration visualisiert.

Ein weiterer Algorithmus wird häufig eingesetzt, um die hochdimensionalen Datenmengen auf drei oder zwei Dimensionen möglichst strukturerhaltend abzubilden, nämlich auf die Knotenpunkte eines Gitters von vorgegebener Größe.²⁶ Diese Visualisierung wird als *self organizing map* (SOM) bezeichnet, da sie mit Karten gemein hat, Territorien darzustellen und jeden Text einer thematischen Region einzuschreiben (vgl. Abb. 1). Allerdings ist hier die Karte wie in anderen Fällen Landschaften oder Galaxien lediglich eine visuelle Metapher zur räumlichen Darstellung abstrakter Daten.²⁷ Sie beschreibt nämlich die Weise der Darstellung und nicht den Prozess der Herstellung, da *self organizing maps* nicht in einem üblichen Sinne kartieren: Sie entstehen nicht durch geometrische Abbildungen oder Projektionen, sondern werden in einem iterativen und konvergenten Prozess errechnet.²⁸

-
- 24 Vgl. Cox, Trevor F./Cox, Michael A. A.: *Multidimensional Scaling*, Boca Raton u.a. ²2001.
- 25 Vgl. zu diesen „visual metaphors“ bspw. Chalmers, Matthew: „Using a Landscape Metaphor to Represent a Corpus of Documents“, in: Frank, Andrew U./ Caspari, Irene (Hrsg.): *Spatial Information Theory*, Berlin u.a. 1993, S. 377-390. Vgl. auch Wise, James A. u.a.: „Visualizing the Non-visual: Spatial Analysis and Interaction with Information from Text Documents“, in: Gershon, Nahum (Hrsg.): *Information Visualization Proceedings, October 30-31, 1995, Atlanta, Georgia, USA*, Los Alamitos, Calif. u.a. 1995, S. 51-58, hier S. 54ff.
- 26 Vgl. Kohonen, Teuvo: *Self-organizing Maps*, Berlin ³2001.
- 27 Diese Metapher kann jedoch auch als Ansatz zur methodischen Bereicherung mit Hilfe der Kartographie dienen. Vgl. Skupin, André: „From Metaphor to Method: Cartographic Perspectives on Information Visualization“, in: Roth, Steven F./Keim, Daniel E. (Hrsg.): *Proceedings IEEE Symposium on Information Visualization (InfoVis2000), 9-10 October, Salt Lake City, Utah*, Los Alamitos, Calif. u.a. 2000, S. 91-97.
- 28 Vgl. Lin, Xia: „Map Displays for Information Retrieval“, in: *Journal of the American Society for Information Science*, Jg. 48, Nr. 1, 1997, S. 40-54, hier S. 51: „The algorithm generates an associative network as the output, rather than the direct mapping of the input.“

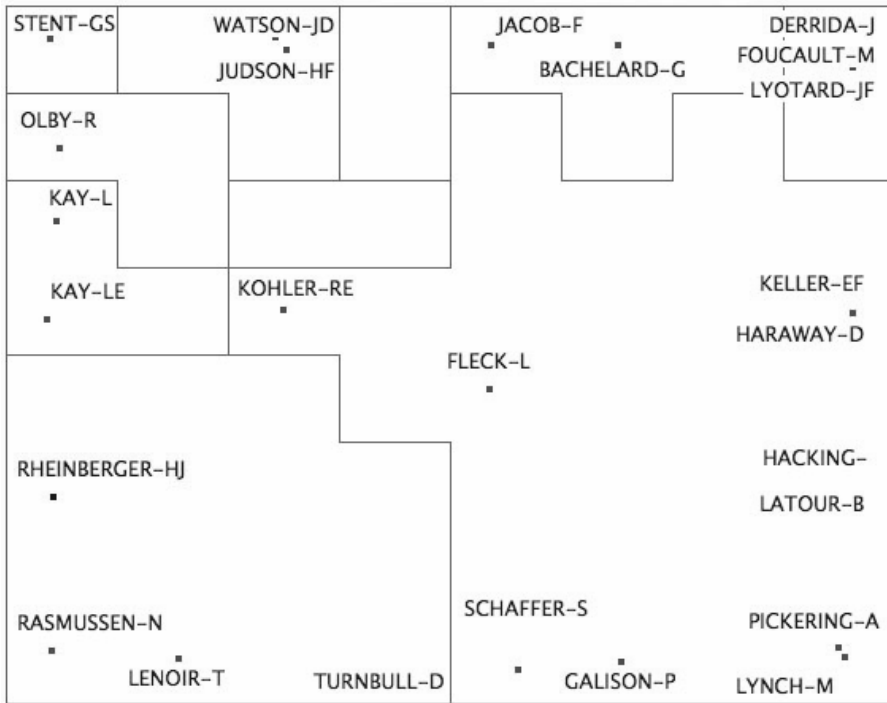


Abb. 1: Self organizing map der 24 mit Hans-Jörg Rheinberger am häufigsten kozierten AutorInnen, berechnet auf der Grundlage des Arts & Humanities Citation Index 1988-1997 und erstellt mit Xia Lins AuthorMap.²⁹

Die bislang vorgestellten Darstellungsverfahren beschränken sich auf die Visualisierung von Texten als einer Menge von Punkten, die von selbstorganisierten Karten zudem in bestimmte thematische Regionen eingeteilt wird. Zusammenhänge zwischen den Texten spiegeln sich in ihrer räumlichen Nähe oder Ferne auf dem Display wider. Die Relationen zwischen den einzelnen Texten oder ihren AutorInnen, die die Koziationsanalyse erfasst, werden dagegen nicht dargestellt. Daher sind in diesem Fall auch Visualisierungen in der Form von Netzwerken üblich, in denen Texte als Punkte visualisiert und zudem alle kozierten Texte miteinander verbunden werden.

Die graphische Realisierung stößt jedoch auf Probleme. Die Punkte müssen auf dem Bildschirm derart positioniert werden, dass sich ein möglichst übersichtliches Bild ergibt, weshalb Überschneidungen der Verbindungen vermieden werden sollten. Einige leistungsfähige Algorithmen lösen dieses Problem recht gut. Ein größeres Problem ist darin zu sehen, dass das Bild oft überladen ist und keine Strukturen zu erkennen gibt. Es ist daher notwendig, das Netzwerk zu vereinfachen.

²⁹ Das Programm ist frei zugänglich unter <http://project.cis.drexel.edu/authorlink/>, 09.07.2007.

chen, indem z.B. nicht alle Verbindungen zwischen den Texten dargestellt werden. Auch hier gibt es wieder verschiedene Algorithmen, unter anderem die so genannten *pathfinder networks*, die von dem Psychologen Roger W. Schvaneveldt eingeführt wurden (vgl. Abb. 2).³⁰

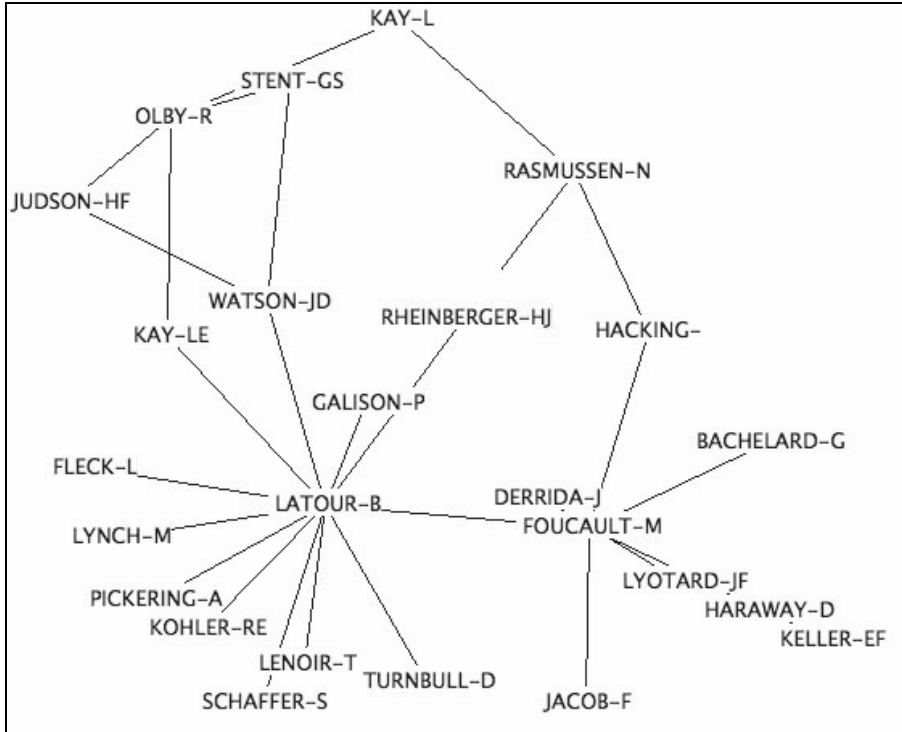


Abb. 2: Pathfinder network zu Abb. 1, erstellt mit Xia Lins AuthorMap.

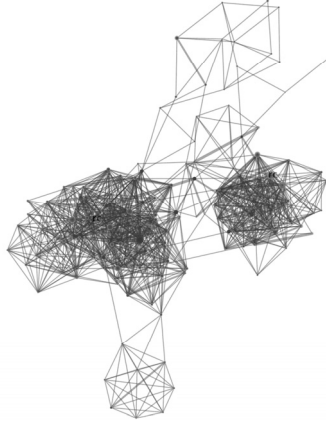
Pathfinder networks werden berechnet, indem die möglichen Wege zwischen allen Punkten eines Netzwerks genauer analysiert und in Abhängigkeit der Häufigkeit von Kozitationen reduziert werden.³¹ Es entstehen auf diese Weise klare und einsichtige Visualisierungen, die den Zusammenhang der Texte aufs Wesentliche zu reduzieren scheinen (vgl. Abb. 3). Je klarer die Strukturen sichtbar sind, desto un-

30 Vgl. Dearholt, Donald W./Schvaneveldt, Roger W.: „Properties of Pathfinder Networks“, in: Schvaneveldt, Roger W. (Hrsg.): *Pathfinder Associative Networks: Studies in Knowledge Organization*, Norwood, N.J. 1990, S. 1-30.

31 Dabei wird jede Verbindung zwischen zwei Texten als umso leichter zu überwinden betrachtet, je häufiger sie kuzitiert werden. Verbindungen zwischen den Texten werden dann einer Minimalitätsforderung unterworfen: Sie werden nur dargestellt, wenn es keinen Weg über andere Texte gibt, der weniger Aufwand erfordert, wenn sie folglich sehr häufig miteinander zitiert werden, häufiger als die Summe der Häufigkeiten von kuzitierten Texten, die sie miteinander umwegig verbinden könnten.

klarer wird aber, inwieweit sie nicht zuallererst durch Vereinfachungen hergestellt wurden. Es ist nämlich sehr viel einfacher, ein prägnantes Bild zu erstellen, als zu gewährleisten, dass es keine falschen Evidenzen erzeugt.

CiteSpace, v. 2.0, Release 10b
July 24, 2007 10:09:11 AM CEST
/Users/arnoschubbach/Documents/IV-Systeme/CiteSpace.2.0.R10b/Examples/Data/Terrorism1990-2003
Timespan: 2001-2003 (Slice Length = 3)
Threshold (c, cc, ccv): 3, 3, 15; 3, 3, 20; 3, 3, 20
Network: N=164, E=1064



CiteSpace, v. 2.0, Release 10b
July 24, 2007 10:09:11 AM CEST
/Users/arnoschubbach/Documents/IV-Systeme/CiteSpace.2.0.R10b/Examples/Data/Terrorism1990-2003
Timespan: 2001-2003 (Slice Length = 3)
Threshold (c, cc, ccv): 3, 3, 15; 3, 3, 20; 3, 3, 20
Network: N=164, E=210

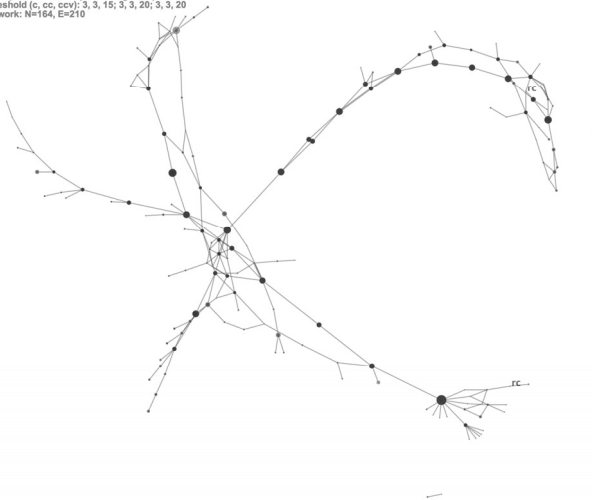


Abb. 3: Übergang zu einem pathfinder network, visualisiert mit Chaomei Chens CiteSpaceII.³²

32 Diese Visualisierungen wurden vom Verfasser auf der Grundlage eines Beispieldatensatzes erstellt, der die Daten zu denjenigen Texten enthält, die eine Suche nach dem

In der Informationsvisualisierung ist dieses Problem wohl bekannt, so dass der unhaltbare Anspruch auf eine simple Widerspiegelung der zugrunde liegenden Daten nicht verfochten wird. Zugleich muss aber die Behauptung aufrechterhalten werden, etwas an oder von den Daten in einer bestimmten Hinsicht darzustellen. Ansonsten verlöre die Informationsvisualisierung jeden Ansatz zu ihrer eigenen Begründung. Xia Lin begreift daher jede Visualisierung von Texten als eine mögliche Darstellung neben anderen, die sich in diesem Sinne als ein „display“ ausweisen soll, statt sich den trügerischen Anschein einer vollständigen Repräsentation zu geben:

Instead, the map for the document space, if it can be created, should emphasize the ‚display‘: The map should be a display (not a representation) of the document space; it should only be a snapshot of the document space with its angles and views defined by the selected indexing and mapping algorithms.³³

Lins Rede von ‚display‘ meint keine technischen Apparate und Dispositive. Jedoch markiert sie mit dem zentralen methodischen Problem der Informationsvisualisierung zugleich den Ort von Bildern, die auf Displays in einem engeren Sinn Daten mittels vielfältiger komplexer Algorithmen sichtbar machen. Die Daten werden zwar weder insgesamt repräsentiert noch als solche sichtbar;³⁴ jede Visualisierung stellt sie aber in einer spezifischen Hinsicht in ihren inneren Zusammenhängen dar und hat jede vermeintliche Kluft zwischen Daten und Bild schon übersprungen. Um die vielfältigen Bildtechniken zu beschreiben, die wie die Informationsvisualisierung zwischen Daten und Bild operieren, ist das Display deshalb nicht als ein Schirm aufzufassen, der zwischen den Blick des Nutzers auf das sichtbare Bild und die per se unsichtbaren Daten und Prozesse im Computer tritt. Vielmehr muss es als produktiver Ort der algorithmischen Sichtbarmachungen von Daten im Bild begriffen werden. Daher ist nicht von der „Entkopplung von Daten und Display“³⁵ auszugehen, sondern von den vielfältigen und stets fragilen Versuchen ihrer Kopplung.

Dies gilt nicht nur für Informationsvisualisierungen, sondern auch für digitale Bilder aus Astronomie oder Nanotechnik und für bildgebende Verfahren in der

Schlagwort ‚terrorism‘ in der medizinischen Literaturdatenbank *PubMed* liefert. Der Datensatz ist wie auch die Java-Applikation *CiteSpace* und weiteres Material verfügbar unter <http://cluster.cis.drexel.edu/~cchen/citespace/>, 09.07.2007.

33 Lin (wie Anm. 28), S. 44.

34 Vgl. zur „Unsichtbarkeit“ der Daten als solcher am Beispiel von SAGE Claus Pias: *Computer Spiel Welten*, München 2002, S. 67ff. und ders.: „Punkt und Linie zum Raster – Zur Genealogie der Computergrafik“, in: Brüderlin, Markus (Hrsg.): *Ornament und Abstraktion. Kunst der Kulturen, Moderne und Gegenwart im Dialog*, Köln 2001, S. 64-69.

35 Vgl. Pias 2002 (wie Anm. 34), S. 80.

Medizin.³⁶ Wissenschaftliche Visualisierungen inszenieren zwar oft den Anschein einer Referenz, als zeigten sie unmittelbar reale Gegenstände in Raum und Zeit. Es ist aber die visuelle Form der Darstellung, die diesen Anschein erweckt, während die komplexen algorithmischen Prozesse zur Sichtbarmachung von Daten im Hintergrund bleiben. Denn diese Bilder werden ebenso berechnet wie Informationsvisualisierungen und sind Resultate der mathematischen Auswertung von Messdaten.³⁷ Sie sind deshalb zuallererst als Visualisierungen von Daten zu betrachten, da sich ihre produktive Funktion in den Wissenschaften in der regelhaften und standardisierten Entstehung von Bildern begründet und nicht in deren vorgeblicher Referenz.³⁸ Statt von einer angeblichen Abbildlichkeit auszugehen, um sie dann zu kritisieren, wäre vielmehr zu untersuchen, wie Daten zu Bildern verarbeitet werden. Dadurch würde die Produktivität algorithmischer Kopplungen von Daten und Bildern in den Mittelpunkt der Diskussion bildgebender Verfahren rücken. Ausgehend von Bildern der Informationsvisualisierung wird somit ein Blick auf das Display möglich, der sich von der sehnsüchtigen oder kritischen Fixierung auf das Abbild löst und sich stattdessen auf den Umgang mit Daten und ihren Strukturen konzentriert.

3 (UN-)SICHTBARKEIT VON DATEN

Der skizzierte Gedanke, dass die Informationsvisualisierung das Display als Ort der Sichtbarmachung von Daten begreift, läuft nun keineswegs darauf hinaus, dass Displays einen unmittelbaren Blick auf die Daten freigeben würden (was ebenso wenig nützlich wie interessant wäre). Denn eine Visualisierung zeigt weder die Daten selbst noch repräsentiert es sie in ihrer vollen Komplexität, wie an der Visualisierung von Texten exemplarisch deutlich wird. Texte werden gemäß einem mathematischen Modell zunächst lediglich in einer bestimmten Hinsicht charakterisiert, nämlich der Semantik des Textkorpus, der intellektuellen Bezugspunkte oder der Zitationen durch spätere Texte. Stets ist der Verlust an Komplexität erheblich. Dies gilt umso mehr, als in der Kookkurrenzanalyse nur wenige Index-

36 Vgl. z.B. Benz, Arnold: „Das Bild als Bühne der Mustererkennung: Ein Beispiel aus der Astrophysik“, in: Heintz/Huber (wie Anm. 4), S. 65-78; Hennig, Jochen: „Changes in the Design of Scanning Tunneling Microscopic Images from 1980 to 1990“, in: *Techné*, Jg. 8, Nr. 2, 2004, S. 36-55; Crelieu, Gérard/Järmann, Thomas: „Abbildung von Wahrnehmung und Denken. Die funktionelle Magnetresonanz-Bildgebung in der Hirnforschung“, in Heintz/Huber (wie Anm. 4), S. 95-108; Angenent, Sigurd u.a.: „Mathematical Imaging in Medical Image Processing“, in: *Bulletin of the American Mathematical Society*, Jg. 43, Nr. 3, 2006, S. 365-396.

37 In diese Richtung weisen auch die Vorschläge von Heintz, Bettina/Huber, Jörg: „Der führerische Blick. Formen und Folgen wissenschaftlicher Visualisierungsstrategien“, in: dies. (wie Anm. 4), S. 9-37, hier S. 20ff.

38 Auf die Referenz setzen irrtümlicherweise Ufologen, wenn sie durch einzelne Fotografien die Existenz von Ufos beweisen wollen, so Latour, Bruno: *Der Berliner Schlüssel. Erkundungen eines Liebhabers der Wissenschaften*, Berlin 1996, S. 183ff.

worte und in der Koziationsanalyse nur viel zitierte AutorInnen oder Texte erfasst werden.³⁹ Zur graphischen Realisierung der Darstellung muss die hochdimensionale Menge dann auf zwei oder drei Dimensionen gebracht werden, wobei die Komplexität weiter reduziert wird und jeder Algorithmus spezifische Vor- und Nachteile hat. Dasselbe gilt für Verfahren, die Verbindungen innerhalb von Netzwerken minimieren, um deren Strukturen sichtbar zu machen. Dieser Verlust an Komplexität ist jedoch zugleich ein Gewinn. Denn eine Visualisierung ist nur insofern produktiv, als sie durch ein übersichtliches und informatives Bild Einsichten gewährt, die anhand der Daten allein nicht gewonnen werden könnten.

Die algorithmische Sichtbarmachung stellt jedoch Daten dar, die sie zuallererst strukturiert. Selbst wenn Daten anders als im Fall von Texten schon in mathematischer Form gegeben sind, müssen z.B. Metriken für den hochdimensionalen Datenraum gesetzt werden, um Nähen zu berechnen und in niederdimensionale Räume möglichst getreu abzubilden; oder es müssen Verbindungen zwischen Elementen gezogen werden, um ein Netzwerk zu erhalten, das vor der Darstellung auf dem Display oft noch zu vereinfachen ist. Die Strukturen werden durch diese Berechnungen konstitutiv bestimmt, wobei die Gefahr unvermeidbar ist, dass eine Struktur erst geschaffen wird, die das Display schließlich offenbar zu machen scheint.

Die Visualisierung kann prinzipiell nur darstellen, was der Strukturierung der Daten entspricht, die von den gewählten Algorithmen, Parametern und Schwellenwerten berechnet wird. Diese Strukturierungen entwerfen jeweils einen Raum des Sichtbaren. Die Sichtbarmachung von Daten ist daher nicht, wie Xia Lin im obigen Zitat andeutet, in Analogie zur Wahl eines Standpunkts in einem gegebenen geometrischen Raum und der Berechnung der entsprechenden Perspektive zu beschreiben. Eine solche Perspektive ergänzt sich mit anderen, im Moment nicht zu sehenden, aber doch sichtbaren, in Husserls Begrifflichkeit „unsichtigen“ Perspektiven zu einem Raum von prinzipiell vollkommener Allsichtigkeit.⁴⁰ Bei Sichtbarmachungen von Daten dagegen werden Räume entworfen, in denen konkrete darzustellende ‚Perspektiven‘ überhaupt erst zu wählen sind. Sie selbst müssen sich daher keineswegs in Analogie zur Perspektive wechselseitig ergänzen.⁴¹

39 Die Visualisierungen haben so zumindest in dieser einfachen Form die Tendenz, etablierte Positionen hervorzuheben, abweichende Positionen und neuere Trends bleiben dagegen unsichtbar. Es gibt aber Ansätze, die sich gerade der Visualisierung solcher Trends widmen. Vgl. z.B. Chen (wie Anm. 15), S. 191ff.

40 Vgl. zur „Unsichtigkeit“ Husserl, Edmund: *Erste Philosophie (1923/24). Zweiter Teil: Theorie der phänomenologischen Reduktion*, hrsg. v. Rudolf Boehm, (Husserliana VIII), Den Haag 1959, S. 146.

41 Einen Eindruck dieser komplexen Verhältnisse gibt auch die große Variabilität der Darstellungen. Wie stark die sichtbaren Strukturen von der Wahl der Algorithmen und selbst der vermeintlich harmlosen Setzung von Parametern oder Schwellenwerten abhängen, wird beim Experimentieren mit komplexeren Visualisierungssystemen schnell deutlich.

Jede Visualisierung von Daten impliziert somit eine Unsichtbarkeit, die nicht die Daten als solche, sondern ihre konkrete Sichtbarkeit auf den Displays und die Sichtbarmachung in ihrer Produktivität, aber auch in ihrer Fragilität charakterisiert.⁴² Diese Unsichtbarkeit resultiert nicht in erster Linie daraus, dass eine Grenze durch die gegebenen Daten gezogen würde, um die einen darzustellen und die anderen zu unterschlagen; sie hat ihren Grund darin, dass von der Strukturierung der Daten abhängt, was überhaupt sichtbar werden kann. Daher verfehlt jede Kritik ihr Ziel, die sich gegen eine Sichtbarmachung wendet, weil diese zugleich unsichtbar macht. Denn es kann, wie Bruno Latour feststellt, jedenfalls bei „wissenschaftlichen Bildern“ nicht darum gehen, „zu wählen zwischen dem Sichtbaren und dem Unsichtbaren“.⁴³ Vielmehr beruhen Visualisierungen auf Verschränkungen von Sichtbarkeit und Unsichtbarkeit. Deshalb sollte aber, wie Xia Lin mit Bezug auf die Visualisierung von Texten fordert, stets kenntlich gemacht werden, dass sich die Sichtbarkeit des Displays der algorithmischen Sichtbarmachung von Daten verdankt, die sehen lässt und zugleich blenden kann.

Die skizzierte Entwicklung von den Daten zum Bild hat jedoch eine notwendige Reflexion übergangen. Das Bild erscheint *auf* dem Display und gewinnt dort eine eigene visuelle Komplexität, die die Komplexität der Daten und ihre algorithmische Verarbeitung nicht einfach widerspiegelt. So wie die zentralperspektivische Konstruktion zum einen den Blick in die Tiefe des Bildraums gewährt, diesen Blick zum anderen aber auch irritieren und auf die Konstellationen auf der Fläche lenken kann, eröffnet jede Visualisierung den Blick in die dargestellten Daten und ihre Strukturen und zugleich auf die reflektierende Oberfläche des Displays.⁴⁴ Es wäre daher nach dem verfolgten Weg der Sichtbarmachung nochmals neu anzusetzen und von der Sichtbarkeit des Bildes auszugehen, um den Ort von Displays zu bestimmen, in denen sich Komplexitäten verschränken, die kaum aufeinander abzubilden sind.

42 Diese Unsichtbarkeit ist über die produktive Strukturierung des Sichtbaren bestimmt, so dass sich wenig ändert, wenn nicht vom einzelnen Bild, sondern von solchen interaktiven Visualisierungen ausgegangen würde, die in jedem einzelnen Bild vieles nicht darstellen, was sie aber im Hintergrund verfügbar halten und auf Anfrage erscheinen lassen. Diese zeitweilige ‚Unsichtbarkeit‘ wäre mit Husserl als aktuelle Unsichtigkeit zu beschreiben, die aber ihrerseits eine prinzipielle Unsichtbarkeit impliziert.

43 Latour, Bruno: *Iconoclasm oder Gibt es eine Welt jenseits des Bilderkrieges?*, Berlin 2002, S. 69.

44 Vgl. zur Fläche des Displays die Beobachtungen von Nake 2004 (wie Anm. 10), S. 340, die er allerdings nicht weiter verfolgt. Mit bildtheoretischem Bezug auf die Malerei Boehm, Gottfried: „Die Wiederkehr der Bilder“, in: ders. (Hrsg.): *Was ist ein Bild?*, München 1994, S. 11-38, hier S. 32ff. sowie Foucault, Michel: *Die Malerei von Manet*, Berlin 1999.

